



UNIVERSIDAD  
**SAN IGNACIO  
DE LOYOLA**

**FACULTAD DE ADMINISTRACIÓN HOTELERA, TURISMO Y  
GASTRONOMÍA**

**Carrera de Administración en Turismo**

**ESTUDIO EXPLORATORIO DEL USO DE  
INDICADORES DE BIG DATA DE GOOGLE COMO  
VARIABLE EN EL MODELAMIENTO DEL FLUJO  
MENSUAL DE VISITANTES AL COMPLEJO  
ARQUEOLÓGICO DE KUÉLAP EN EL PERIODO 2011-  
2016**

**Tesis para optar el Título Profesional de Licenciado en  
Administración en Turismo**

**JHOCELINE PAOLA LEÓN GUTIÉRREZ**

**BRENDA STEPHANIE VILLAR ORTIZ**

**Asesor:**

**José Francisco Melenez Roman**

**Lima – Perú**

**2018**

## ÍNDICE DE CONTENIDOS

CAPÍTULO I: INTRODUCCIÓN .....	5
Problema de Investigación .....	7
Planteamiento del Problema. ....	7
Formulación del Problema. ....	8
Justificación de la Investigación. ....	9
Marco Referencial .....	11
Antecedentes. ....	11
Marco teórico. ....	13
Objetivos e Hipótesis .....	30
Objetivos .....	30
Hipótesis. ....	30
CAPÍTULO II: MÉTODO.....	33
Tipo y Diseño de Investigación.....	33
Tipo de Investigación. ....	33
Diseño de Investigación. ....	33
Variables. ....	36
Muestra. ....	39
Instrumento de Investigación. ....	39
Procedimiento para Recolección de Datos.....	40
Plan de Análisis .....	43
CAPÍTULO III: RESULTADOS.....	49
Presentación de Resultados.....	49
Presentación de Resultados Cualitativos. ....	49
Presentación de Resultados Cuantitativos. ....	60
Discusión.....	73
Conclusiones.....	81
Recomendaciones .....	83
REFERENCIAS .....	85
ANEXOS .....	94
Anexo 1 .....	94
Anexo 2 .....	95
Anexo 3 .....	95
Anexo 4 .....	96
Anexo 5 .....	96
Anexo 6 .....	97

Anexo 7 .....	97
Anexo 8 .....	98
Anexo 9 .....	99
Anexo 10 .....	102
Anexo 11 .....	104
Anexo 12 .....	105
Anexo 13 .....	106
Anexo 14 .....	106
Anexo 15 .....	106
Anexo 16 .....	107
Anexo 17 .....	107
Anexo 18 .....	108
Anexo 19 .....	114
Anexo 20 .....	117
Anexo 21 .....	118
Anexo 22 .....	119

## ÍNDICE DE TABLAS

Tabla 01: Categorías Temáticas Análisis Cualitativo .....	36
Tabla 02: Variables Análisis Cuantitativo .....	37
Tabla 03: Información de los entrevistados .....	39
Tabla 04: Información de ejecución de entrevistas a profundidad .....	42
Tabla 05: Términos Asociados a los Términos Clave de las Entrevistas .....	49
Tabla 06: Parámetros de la primera ecuación .....	60
Tabla 07: Coeficientes y Probabilidad de la primera ecuación .....	60
Tabla 08: Parámetros de la segunda ecuación .....	61
Tabla 09: Coeficientes y Probabilidad de la segunda ecuación .....	62
Tabla 10: Parámetros de la tercera ecuación .....	63
Tabla 11: Coeficientes y Probabilidad de la tercera ecuación .....	63
Tabla 12: Parámetros de la cuarta ecuación .....	65
Tabla 13: Coeficientes y Probabilidad de la cuarta ecuación .....	65
Tabla 14: Parámetros de la quinta ecuación .....	66
Tabla 15: Coeficientes y Probabilidad de la quinta ecuación .....	66
Tabla 16: Parámetros de la sexta ecuación .....	67
Tabla 17: Coeficientes y Probabilidad de la sexta ecuación .....	67
Tabla 18: Parámetros de la séptima ecuación .....	68
Tabla 19: Coeficientes y Probabilidad de la séptima ecuación .....	68
Tabla 20: Parámetros de la octava ecuación .....	69
Tabla 21: Coeficientes y Probabilidad de la octava ecuación .....	69
Tabla 22: Parámetros de la novena ecuación .....	70
Tabla 23: Coeficientes y Probabilidad de la novena ecuación .....	70
Tabla 24: Resumen de los resultados cuantitativos .....	72

## CAPÍTULO I: INTRODUCCIÓN

En la actualidad, se puede observar un crecimiento en cuanto a las búsquedas online que los viajeros potenciales realizan para obtener información sobre el destino turístico al cual desean dirigirse (Instituto Valenciano de Tecnologías Turísticas, 2015); siendo notorio que con el transcurso de los años, el internet se ha posicionado como la principal herramienta de búsqueda de los diferentes mercados a nivel mundial (StatCounter, 2017); por ejemplo, el 69% de los turistas extranjeros que llegaron a Perú en 2016 afirmaron que internet influyó en la elección de su destino (PROMPERÚ, 2017). Por su parte, el 86% de los vacacionistas nacionales en el 2016 usaron internet para buscar información turística, especialmente el buscador Google y el portal “Y Tu Que Planes?” (PROMPERÚ, 2017).

Es posible encontrar una gran cantidad de data relacionada al turismo disponible en la web, pero es necesario un análisis profundo para encontrar el sentido de la información que se obtiene. Un consumidor potencial planeando un viaje, probablemente requiera más que un poco de ayuda del internet cuando busque fuentes de inspiración, compre boletos, reserve alojamiento o busque atractivos. Es por ello que, en la actualidad, se está incrementando la cantidad de actores de la industria turística que adoptan el análisis de grandes cantidades de información para descubrir nuevas formas de tomar decisiones y encontrar oportunidades (Song & Liu, 2016). En ese sentido, las búsquedas en internet proporcionan patrones que pueden ser aprovechados por la industria turística, por ejemplo, para pronosticar el número de llegadas hacia un destino turístico.

Existen diversos métodos para pronosticar el número de llegadas hacia un destino turístico. El tipo autorregresivo es aquel que emplea una serie de datos históricos para realizar proyecciones (Shareef & McAleer, 2005). A este método tradicional, se ha incorporado en recientes investigaciones un indicador del tipo electrónico: *Google Trends*; como en los estudios sobre destinos turísticos de Onder y Gunter (2016) o Park, Lee y Song (2015). Este indicador electrónico refleja el volumen relativo de búsquedas de un término clave en la web y forma parte de lo que actualmente se conoce como *Big Data*.

El presente estudio ha seleccionado al Complejo Arqueológico de Kuélap como atractivo de investigación. Su elección obedece a que es el atractivo turístico más visitado de la región Amazonas y viene siendo promocionado tradicional y digitalmente tanto a nivel nacional como internacional por la entidad encargada de la promoción de la exportación y el turismo (PROMPERÚ) y el Ministerio de Comercio Exterior y Turismo (MINCETUR), lo cual, ha tomado mayor énfasis debido al inicio de operación del sistema de telecabinas.

Cabe mencionar que el atractivo ha experimentado un incremento en el flujo anual de visitantes nacionales y extranjeros entre 2015 y 2016 de 53.94%, llegando a un 15.84% de crecimiento acumulado del 2011 al 2016 (MINCETUR, 2017).

El presente documento tiene por objetivo determinar si los datos digitales obtenidos desde Google Trends pueden ser empleados como variables para el modelamiento predictivo del flujo de visitantes nacionales y extranjeros al Complejo Arqueológico de Kuélap. Para ello, se ha utilizado la ecuación propuesta por Choi y Varian (2009), quienes emplearon el volumen de búsquedas para construir modelos que permitieron realizar pronósticos mensuales de corto plazo. Entre otros hallazgos, demostraron que el flujo mensual de turistas que llegan a Canadá y Estados Unidos puede modelarse con mayor precisión si se incluye los indicadores electrónicos provistos por Google. La mencionada investigación ha servido como referencia para que los indicadores de Google Trends se conviertan en variables de estudio a ser consideradas por profesionales del turismo especializados en el modelado del flujo de turistas que arriban a un espacio turístico determinado (Gawlik, Kabaria, & Kaur, 2011).

Para este estudio en particular, las variables a emplear son los registros históricos mensuales de llegadas y los indicadores de Google Trends: web, imágenes, YouTube y noticias. Asimismo, la investigación busca dar a conocer la potencialidad de un modelo de análisis predictivo para las decisiones de marketing online; ya sea como un ratio de respuesta a campañas pasadas; para conocer la viabilidad de campañas promocionales o para elaborar calendarios de contenidos en base a índices de estacionalidad.

El método de investigación en este estudio contempla información cualitativa y cuantitativa. La investigación cualitativa brinda información descriptiva obtenida mediante entrevistas a profundidad a expertos en materia de marketing digital, prospectiva turística y Big Data para recabar su percepción en relación al planteamiento del problema. Los resultados de la parte cualitativa sirven para el reforzamiento del método cuantitativo, en donde se busca la medición de las variables establecidas a partir de la data histórica y los indicadores de Google Trends, de manera que se pueda cumplir con los objetivos planteados en el estudio.

## Problema de Investigación

### Planteamiento del Problema.

En esta época donde el internet forma parte del día a día, las personas suelen revelar de forma inconsciente, información sobre sus patrones de compra o adquisición de servicios reales o potenciales mientras navegan en la web (Vosen, 2011). Esta navegación genera registros digitales de búsquedas en línea, almacenando información y formando una parte de lo que se conoce como Big Data. Esta última puede ser definida como una herramienta utilizada para describir enormes cantidades de datos que no pueden ser procesados o analizados utilizando herramientas tradicionales, pues demandaría mucho tiempo (IBM, 2016).

La Big Data es capaz de generar registros por medios, por ejemplo, como tarjetas de crédito, lo cual, permite a las empresas analizar la gran cantidad de datos que recopilan durante las transacciones financieras. La enorme colección de datos que las personas generan independientemente a través de ordenadores y del sistema automatizado llega a ser útil para los investigadores. Por ejemplo, Priyanka y Shivaputra (2017) señalaron que “la Big Data es útil en el manejo del tráfico, pues proporciona soluciones para el tráfico en términos de densidad, señalización, navegación inteligente y monitoreo de contaminación de vehículos”. Para ello, usaron la técnica Smart Cluster Heads (SCH), la cual, recopila datos mediante señales de sensores inalámbricos en puntos clave. Los resultados de esta investigación ayudan a disminuir el tiempo de retención en las señales de actividad y disminuir las calles congestionadas (Priyanka & Shivaputra, 2017).

Estas herramientas para el análisis de Big Data suelen generar costos adicionales o requieren conocimientos especializados de los investigadores, porque en muchos casos su análisis es complejo. No obstante, existe una herramienta de fácil uso y gratuita capaz de generar registros valiosos para la investigación en diferentes sectores económicos. Google ha desarrollado una aplicación llamada Google Trends que proporciona índices de búsqueda, la cual, se han convertido en una fuente de datos muy valiosa para diversos campos de estudio: salud pública, ventas, finanzas, ciencias políticas, macroeconomía, entre otros (Borba & Akemi, 2014). En el campo de interés del estudio, el turismo, se utilizan los indicadores generados por Google Trends, por ejemplo, para estimar flujos de visitantes (Onder & Gunter, 2016). Askitas y Zimmermann (2009) señalaron que “la información que los internautas buscan hoy, resulta útil para predecir lo que realizarán en un futuro próximo en términos de viaje”. Dicha información es útil para modelar series de datos que son volátiles (Smith, 2016).

En líneas generales, las autoridades de turismo de los países publican periódicamente indicadores relacionados al sector turismo. Sin embargo, estas publicaciones están generalmente disponibles meses después del recojo de la información. Choi y Varian (2009) señalan que “el internet es comúnmente usado para la planificación de viajes, por lo que la data de Google Trends puede ser útil para predecir las visitas a un destino en el corto plazo”. Esto sucede si se encuentra una correlación entre los arribos en un determinado periodo y las búsquedas en Google. Por tanto, se podrían tener los datos requeridos con una ventaja de algunas semanas respecto a la publicación oficial que suele tener un retraso de aproximadamente un mes (Choi & Varian, 2009). Esta demora en la publicación de la data oficial en turismo se observa en el Perú, por ejemplo, hasta el mes de diciembre del 2017 sólo se podían observar los arribos turísticos al Complejo Arqueológico de Kuélap hasta el mes de octubre del año 2017 (MINCETUR, 2017). Cabe mencionar que, este atractivo resulta relevante como tema de estudio debido al incremento acumulado en su flujo de visitantes desde el año 2011 al 2016 (15.84%), convirtiéndose en el atractivo turístico de más llegadas a la Región Amazonas (MINCETUR, 2017); además de ser uno de los atractivos priorizados por el Estado Peruano para la promoción del turismo.

Este campo de estudio no es común en las investigaciones nacionales, y aún menos para el sector turismo. Es por este motivo que la presente investigación buscó realizar un estudio exploratorio que permita evaluar si la incorporación de los indicadores generados por Google Trends contribuyen como variables significativas para el modelamiento del flujo mensual de visitas de turistas nacionales y extranjeros al Complejo Arqueológico de Kuélap en el periodo 2011-2016; tomando como modelo base la ecuación propuesta por Choi y Varian (2009), y proponiendo la adición de otras variables electrónicas con el objetivo de evaluar el impacto que tienen en la predicción de la demanda. De esta manera, se encontró el nivel de correlación entre las llegadas al Complejo Arqueológico de Kuélap y las variables electrónicas, evaluando el interés que las personas tienen hacia el atractivo y concluyendo en qué medida este interés es significativo en términos de predicción de visitas a Kuélap.

### **Formulación del Problema.**

Esta investigación tiene como problema de investigación el siguiente: “¿Puede el uso de indicadores de Big Data de Google Trends mejorar el modelamiento del flujo mensual de arribos de turistas nacionales y extranjeros al Complejo Arqueológico de Kuélap en el periodo 2011-2016?”

### **Justificación de la Investigación.**

Los registros digitales provistos por Google Trends son una fuente casi ilimitada de información que puede ser de provecho para entidades a cargo de la administración pública de espacios turísticos y por profesionales de turismo. Los indicadores de Google Trends son de fácil y rápido acceso, y no requieren un desembolso financiero (Fotios, 2016). La información obtenida de dicha herramienta se puede dividir en meses, semanas, días y horas; así como por ciudad, región y país; lo que implica que la disponibilidad de datos ya no es un limitante (Bangwayo-Skeete & Skeete, 2015).

Esta investigación se centra en las búsquedas en internet de un atractivo antes de realizar un viaje. La información obtenida de Google Trends podría ser utilizada para realizar un proceso llamado *nowcasting*, que se basa en el uso de datos para intentar predecir el presente o el futuro más cercano (Banco Central Europeo, 2013); en este caso, los datos son las búsquedas en Google de aquellos que indagan en internet sobre el atractivo escogido.

Encontrar una relación directa entre tráfico web y visitas a un determinado sitio turístico contribuiría a la posible predicción de las tendencias del turismo (Lutz, 2013), especialmente de escenarios de muy corto plazo (Saidi, 2010), es decir, futuros muy cercanos. Por tanto, esta información puede ser útil para optimizar los recursos de las agencias de viaje y hoteles, que en su conjunto, componen la oferta turística (RocaSalvatella & Telefónica, 2014). Asimismo, se podrían desarrollar estrategias de marketing online y comunicación más efectivas mediante el modelamiento de la data histórica de visitas y los indicadores de búsquedas en Google; determinando el mejor momento para realizar campañas y conocer con mayor exactitud el impacto de las mismas; así como evaluar las campañas y canales de marketing que han estado funcionando. Además, resultaría posible estimar la estacionalidad mensual del interés (Zeynalov, 2014); pudiéndose realizar estrategias mediante el conocimiento de los momentos en que el atractivo es más buscado y analizando los patrones de estacionalidad para crear un calendario de contenidos para las campañas de marketing online.

Las búsquedas de Google pueden ser empleadas por los profesionales del turismo como un indicador para medir los efectos de una campaña de marketing online; respecto al interés que genera hacia un destino. En principio, el objetivo de dichas iniciativas apunta a crear un mayor interés del visitante potencial hacia un destino; indicador que se traduce en un incremento del número de búsquedas (Sung, 2015). En este estudio, se obtuvo una relación matemática entre las búsquedas y llegadas turísticas, lo que

permitió generar una noción objetiva del ratio de respuesta hacia las campañas digitales. Saidi (2010) manifestó que “de existir relación entre las búsquedas en Google y los arribos de visitantes, significaría que los datos del buscador son en parte representativos de la demanda real” (Saidi, 2010); demostrando así el interés hacia el destino.

El estudio se centra en el Complejo Arqueológico de Kuélap, que es uno de los ejes turísticos de la zona nor-amazónica del Perú al recibir más visitas que cualquier otro atractivo del lugar (MINCETUR, 2017). Kuélap tuvo un importante proyecto de inversión aprobado en el 2014 y finalizado en 2017, el sistema de telecabinas de Kuélap, el cual, representa una mejora para el acceso e infraestructura del lugar. Esta mejora de infraestructura motivó a la Comisión de Promoción del Perú para la Exportación y el Turismo (PROMPERÚ) y el Ministerio de Comercio Exterior y Turismo (MINCETUR) ha priorizar la promoción del atractivo y el plan de trabajo relacionado al mismo. Estos factores están siendo tomados positivamente por el sector privado, que ha comenzado a invertir en las zonas cercanas. Por ejemplo, en 2016, LATAM inauguró su vuelo Lima – Jaén, haciendo más corto el trayecto de los turista hasta Chachapoyas; en 2017, la aerolínea ATSA inauguró un vuelo directo Lima – Chachapoyas (T NEWS, 2017), y grandes tour operadores están iniciando operaciones en la zona (CANATUR, 2017).

La investigación ha utilizado nueve modelos para evidenciar cual presenta una mejor correlación para las visitas a Kuélap: AR-1, Choi y Varian (2009) y siete posibles combinaciones de la ecuación de Choi y Varian, que adicionan otras tres variables electrónicas provenientes de Google Trends (imágenes, videos y noticias) para el periodo 2011-2016. La razón de incorporar estas variables adicionales al modelo, se debe a que se ha observado un interés de parte de las entidades estatales y privadas del sector turismo hacia Kuélap; por lo que se ha pretendido conocer en qué medida la gestión y promoción se traduce en indicadores que permiten tomar decisiones; ello, a través de una prueba empírica que involucra diferentes variables relacionadas al marketing online, con el objetivo de concluir si hasta el momento existe un ratio de respuesta positivo hacia las estrategias de marketing que incluyen al atractivo.

Cabe mencionar que, la investigación está aportando un cuerpo de conocimiento que usa variables digitales para modelos predictivos; las cuales pueden ser más rápidas y más precisas. Asimismo, se enriquece con el aporte de la percepción de expertos en materia de prospectiva turística, marketing turístico y Big Data, que refuerzan el estudio. Y, además, pretende evaluar distintos modelos que cualifiquen a las búsquedas electrónicas como variables independientes, y que a su vez, permitan futuras investigaciones que involucren a estas variables en otros sitios turísticos.

## Marco Referencial

### Antecedentes.

Choi y Varian (2009) modelaron el arribo de turistas extranjeros que llegaron a Hong Kong procedentes de Estados Unidos, Canadá, Gran Bretaña, Francia, Italia, Japón e India; utilizando un modelo de regresión múltiple como el siguiente:

$$\text{Log}X(t) = b_1 + b_2 * \text{Log}X(t - 12) + b_3 * \text{Log}X(t - 1) + c * G(t)$$

Donde la variable dependiente son los visitantes provenientes de un país a Hong Kong en el mes "T". Las variables independientes:  $X(t-12)$ , corresponde a las visitas recibidas para el mismo mes, pero correspondientes al año anterior;  $X(t-1)$  corresponde al número de arribos, pero para el mes inmediato anterior; y  $G(t)$  es el índice de búsquedas mensual en Google Trends; mientras que  $b_2$ ,  $b_3$  y  $c$  son coeficientes que afectan a las variables. Con este modelo estimaron la correlación de las llegadas de cada país y compararon los resultados con el modelo que no hace uso de la variable electrónica. Los investigadores determinaron que las llegadas estaban en función de la data histórica registrada hace un año y un mes, y que la incorporación del indicador de Google Trends mejoraba el valor de R-cuadrado (coeficiente de correlación) en comparación a los modelos que no incluían la variable de búsquedas web. De esta manera, sustentaron el propósito de su estudio, el cual era mostrar cómo se puede utilizar datos de Google Trends para realizar métodos de pronóstico simples en diferentes sectores económicos y fomentar a que otros investigadores realicen sus propios análisis.

Por su parte, Onder y Gunter (2016) investigaron si el uso de los indicadores de Google Trends mejora la precisión de la previsión de la demanda turística de Viena en relación con un modelo de referencia autorregresivo, es decir, un proceso que calcula una variable para un periodo determinado en base a observaciones del mismo indicador en periodos anteriores. Los resultados confirmaron que la precisión de las predicciones mensuales mejora cuando se incluyen los datos de las búsquedas de Google Trends en cada uno de los mercados de origen (Alemania, Italia, Rusia, Reino Unido y Estados Unidos).

Otra investigación que usó Google Trends es la de Park, Lee y Song (2015), donde la herramienta sirvió como método para obtener proyecciones de corto plazo sobre la afluencia de turistas japoneses a Corea del Sur. La variable electrónica se construyó de forma sistemática combinando palabras clave para minimizar errores. Se incorporó el indicador a los modelos estándar de predicción basado en series de tiempo, y se comparó el nivel de precisión en las predicciones, lográndose encontrar que los modelos

que incluyen al indicador digital se desempeñan mucho mejor que los modelos estándar de series de tiempo, en cuanto a la precisión de proyección a corto plazo.

En el estudio de Saidi (2010), se usó el modelo ARIMA, una combinación de métodos autorregresivos y media móvil, e incorporó datos de Google Trends a fin de construir una ecuación que permitiese determinar el número de llegadas mensuales de turistas que pernoctaban en Dubai entre junio del 2004 y octubre del 2010. El modelo ARIMA que se utilizó para esta investigación según las pernoctaciones de los turistas es:

$$d(\text{Log}(\text{GuestNight}_t)) = C + \beta\varepsilon_{t-1} + \gamma\varepsilon_{t-2} + \delta\varepsilon_{t-12}$$

Donde el número de noches de visitantes está en función de los datos registrados de hace un mes, hace dos meses y hace un año. La investigación resaltó que el índice de búsquedas resulta útil para ajustar datos estacionales. Por otro lado, el modelo sirvió para determinar el número de arribos por eventos deportivos, los cuales no poseían estacionalidad identificable, ya que su realización no tenía un patrón definido en cuanto a fechas.

De manera similar a los anteriores autores, Zeynalov (2014) usó Google Trends, llegando a demostrar que los datos provistos por la herramienta resultaron elegibles para la construcción de un modelo econométrico que hacía posible la predicción de corto plazo con respecto al flujo mensual de turistas que llegó a Praga, República Checa, entre los años 2008 y 2013. Estableció la necesidad de que estas búsquedas estuvieran relacionadas a la motivación de viajar y no con otros fines. Esto se debe a que las búsquedas acerca de Praga no necesariamente son todas con fines turísticos, sino, también con otras motivaciones, tales como conocer sobre moda o para fines académicos. Para el modelado fue necesario emplear cálculo vectorial y modelos autorregresivos. La metodología que se utilizó para evaluar y comparar comienza con tres estructuras univariadas del ARIMA que son: el Sistema de estructuras univariadas, el modelo de series temporales (STSM) y el modelo de SVR (autorregresivo).

Por su parte, Fotios (2016) examinó si el volumen de búsquedas para el término *Holidays in Cyprus* resultaba un indicador predictivo de los arribos a Chipre producidos entre enero del 2012 y mayo del 2015. El modelo que se utilizó en esta investigación es *Stationary Test for Seasonality*, la cual, consiste en que la hipótesis nula es igual a una raíz unitaria, la que es más fuerte cuanto más negativo sea el estadístico-t. Entonces, si el valor del estadístico-t es menor que los valores críticos de la prueba, se rechaza la hipótesis nula y la variable no es estacionaria. El estudio demostró que dicha relación existía y que la estacionalidad de las llegadas estaban en sincronía con la estacionalidad

de las búsquedas. El principal indicador tomado en cuenta fue el incremento del coeficiente de determinación (R-cuadrado) igual que Choi y Varian (2009).

En resumen, los estudios encontrados brindan evidencia de que los datos provistos por Google resultan útiles como indicadores válidos y que están relacionados con el flujo de arribos turísticos. Estos indicadores de Google guardan relación con el flujo de llegadas, pero en escenarios en vigencia o de muy corto plazo, es decir futuros muy cercanos. Por otro lado, los modelos que han sido empleados guardan relación con las series de tiempos y modelos de cálculo vectorial. Resulta posible, además, emplear más de un término clave asociado al destino turístico. El principal indicador a tomar en cuenta es el incremento del coeficiente de determinación (R-cuadrado); demostrando que existe una relación entre las variables utilizadas en los diferentes estudios, lo cual denota un grado de asociación, sentido o cercanía (Lahura , 2003). Cabe mencionar que, Onder y Gunter (2016) indicaron que “no se busca construir modelos que usen a Google Trends como variable principal; se trata de tomar modelos que empleen variables reales como elementos base (ej. cantidad de visitantes) y combinarlas con los índices de búsqueda, a fin de obtener ecuaciones con un mayor nivel de precisión”.

En cuanto al aporte al conocimiento, estos modelos que usan variables digitales pueden ser más accesibles, rápidos y precisos. Mientras que, su beneficio aplicativo para el sector turismo se refleja, entre otras aplicaciones, en facilitar la elaboración de estrategias de marketing online.

### **Marco teórico.**

#### ***Big Data.***

La historia del término Big Data, desde el punto de vista académico, se remonta a 1984, y desde el punto de vista empresarial, a 1987. Durante el periodo 2008 y 2010, se publicaron diferentes artículos e informes del fenómeno en The Economist y Forbes; siendo a partir de esas fechas que se considera el inicio de la Big Data como fenómeno social, tecnológico, económico y empresarial (Joyanes, 2013).

El McKinsey Global Institute (2011) ha definido a la Big Data como “un conjunto de datos cuyo tamaño va más allá de las capacidades de las herramientas clásicas de software para capturar, almacenar, gestionar y analizar información”.

Laney (2001) hace énfasis en que la definición pasa por las tres 'Vs' de gestión de datos: volumen, velocidad y variedad. El volumen se refiere a la gran cantidad de datos a ser administrados, la velocidad hace referencia al incremento de la rapidez del ritmo en la generación de datos, y la variedad se refiere a los diversos formatos

de datos que aparentemente no guardan compatibilidad. Sobre esto último, es necesario mencionar que existen datos con estructura definida como los indicadores de Google Trends; pero también existe aquella información que se presenta como datos no estructurados como los comentarios en redes sociales como *Twitter*, *TripAdvisor*, entre otros (Godnov & Redeck, 2016).

Según IBM (2016) existen 2 V's adicionales a las características de la Big Data: veracidad y valor de los datos. La veracidad se refiere a la incertidumbre de los datos; por lo que es necesario invertir tiempo para conseguir datos de calidad. Por su parte, el valor de los datos es la información de utilidad que recolecta la Big Data para obtener los resultados que se fijaron (IBM, 2016).

En la actualidad existe una diversidad de aplicaciones reales de la Big Data, como la predicción de resultados en la Liga Nacional de Fútbol de Estados Unidos (IBM, 2015). En la medicina puede emplearse para reducir el tiempo de los ensayos clínicos mediante recopilación de información en la nube (Bristol Myers Squibb, 2017). Y puede, además, emplearse Big Data para estimar la cantidad de alquileres de automóviles (Avis Budget Group, 2017).

Para el sector de interés de la investigación, la aplicación del presente estudio, que hace uso de la Big Data, representaría un beneficio para la gestión y toma de decisiones más rápida para el alcance de los objetivos en materia turística. Por ejemplo, en el sector privado, RocaSalvatella y Telefónica (2014) elaboraron, en conjunto, un estudio que cruza información de BBVA y Telefónica usando información de tarjetas de crédito y locación GPS de celulares con el objetivo de contribuir a la mejora en la toma de decisiones de las empresas privadas en el sector turismo, especialmente para los hoteles. En cuanto al Estado, en Perú, PROMPERÚ viene implementando un proyecto de Big Data en conjunto con Movistar para obtener hallazgos de traslados internos por turismo (W. Vizarreta, comunicación personal, 01 de agosto, 2017).

### ***Big Data y Turismo.***

El potencial de la Big Data en turismo es grande; puede ser usada para personalizar campañas de marketing y encontrar modelos de negocio ineficientes, es decir; negocios que no son capaces de obtener beneficios razonables de su actividad (Instituto Valenciano de Tecnologías Turísticas, 2015). El análisis de Big Data puede proporcionar ideas de negocio necesarias y puede ser fuente de innovación para las organizaciones de turismo y la industria en general. (Song & Liu, 2016).

La aplicación de Big Data en temas turísticos, por ejemplo, ha logrado determinar los elementos asociados a los niveles de satisfacción e insatisfacción del turista que arribó a hoteles en Japón. Se basaron en los comentarios que hicieron los huéspedes en las páginas web de reservas de viaje, revelando sus impresiones y observaciones sobre la estancia en los hoteles de Japón (Tsujii & Takahashi, 2014). También, se ha determinado la ruta con un mayor nivel de probabilidad de recorrido que realizaría un visitante en Cataluña haciendo uso los dispositivos GPS de celulares (RocaSalvatella & Telefónica, 2014). Por su lado, Fuchs y Lexhagen (2015) orientaron parte de su investigación a la segmentación de turistas que arribaron a Suecia haciendo uso del análisis de comentarios en redes sociales. El gasto de las tarjetas de crédito también ha servido como indicador para estudios relacionados a Big Data. Las compañías Telefónica y RocaSalvatella (2014) utilizaron información de tarjetas de crédito extranjeras para analizar desde una perspectiva de negocios, ciertos aspectos del comportamiento de los visitantes extranjeros a Madrid y Barcelona durante un periodo específico. Otro ejemplo es obtener y analizar información a través de las imágenes en redes sociales, permitiendo a los analistas de datos conocer los intereses de los turistas (Song & Han, 2017). Todos estos hallazgos sumados a los ya mencionados usos de Google Trends, sugieren la versatilidad de investigaciones que siguen la orientación de la Big Data (Sung, 2015).

Lancis (2014) menciona que “la Big Data puede ser empleada para la gestión de destinos turísticos inteligentes a través de guías turísticas digitales en formato multilinguaje que incorpora una serie de funcionalidades como monitorización, avisos, recomendaciones *pull* y botones para compartir contenido multimedia en redes sociales”. Las guías pueden adaptarse para segmentos y productos como el segmento RICE (Reuniones, Incentivos, Congresos y Eventos) que es un tipo de turismo en el que grandes grupos de personas se reúnen para un propósito en particular (Lancis, 2014).

En este sentido, el análisis de la Big Data se encuentra relacionado estrechamente con el reciente surgimiento del diseño turístico y los destinos inteligentes, resultado de los avances en las tecnologías de la información que permiten comprender mejor el mercado de consumo e impulsa a la innovación de la industria de turismo (Zheng Xiang, 2016). La Big Data brinda la posibilidad al sector turismo de realizar una mejor gestión mediante volúmenes de datos e información de visitantes, empresas turísticas y administraciones públicas (Lancis, 2014). Por ejemplo, Wladyka (2017) mencionó que “el incremento del interés puede estar asociado a un incremento en la cantidad de personas que se desplazan hacia un espacio geográfico específico,

mientras que un descenso en el nivel de interés puede vincularse a una disminución en los flujos”. Los indicadores del tipo Big Data, además, se vuelven más representativos conforme se incrementa la proporción de personas que usan internet (Wladyka, 2017).

Se desarrollan nuevas métricas de destino que se conceptualizan en cuatro redes de creación de valor centradas en el visitante. Estas redes de creación se ven representadas en (i) los procesos de marketing y promoción, (ii) ventas y distribución, (iii) diseño de experiencia y (iv) coordinación de asociaciones. La estructura de cada red de creación de valor proporciona información valiosa sobre el rendimiento de un destino para cada proceso; no obstante, cuando las redes de creación de valor se consideran un sistema integrado, pueden utilizarse para conceptualizar un enfoque estratégico de gestión de red para mejorar la competitividad y rendimiento del destino (Stienmetz & Fesenmaier, 2015). Una de las limitaciones de la Big Data es que no resulta un recurso del todo atractivo para investigadores del campo turístico, pues requiere de técnicas de análisis modernas que involucran conocimiento de software especializado y administración de bases de datos, los cuales no son muy populares entre los estudiantes y profesionales de turismo (Baggio, 2016). Dicha brecha debe ser cerrada en base a una formación de carácter interdisciplinaria (Dolnicar & Ring, 2014). Otra limitación es la depuración de datos, la cual no siempre resulta del todo perfecta (RocaSalvatella & Telefónica, 2014). Por ejemplo, las personas pueden buscar información no necesariamente desde donde residen, puede que lo hagan desde un punto totalmente distinto, una zona de tránsito temporal; lo cual complicaría, por ejemplo, conocer los mercados de origen (Bangwayo-Skeete & Skeete, 2015).

### ***Google Trends.***

Google Trends es una aplicación que fue lanzada en 2012, la cual proporciona información semanal sobre las búsquedas en Google con características avanzadas disponibles libremente en <https://www.google.com/trends/>. Cualquier usuario con una cuenta en algún servicio del buscador (*Gmail*, *Blogspot* u otros) puede observar y descargar los índices para uno o más términos de búsqueda.

Google Trends tiene disponible una función de geolocalización y categoría (Fotios, 2016). Ello permite realizar un filtro para delimitar solamente los volúmenes de búsquedas bajo la categoría de viajes en territorio peruano o a nivel mundial. Asimismo, permite conocer los otros lugares de interés que los internautas asocian al destino de búsqueda principal bajo esta categoría de viajes (Google Trends, 2016). La disponibilidad de datos a partir del 2004 permite la existencia de una serie de

tiempo que se construye en relación con determinados términos de búsqueda (Office of National Statistics, 2012). Ello, en base a que una serie de tiempo es el resultado de observar los valores de una variable  $X$  a lo largo de un periodo determinado.

Google Trends no sólo proporciona datos sobre búsquedas de páginas web en Google, también provee indicadores de búsquedas en YouTube, imágenes y noticias; tres opciones de búsqueda que el usuario puede seleccionar al ingresar a Google. Estos tres tipos de búsqueda, al igual que las búsquedas web, hacen uso de algoritmos; que son fórmulas y procesos informáticos que buscan pistas para darle al usuario exactamente lo que busca (Google, s.f.).

Google adquirió a YouTube desde el 2006 y posee sus estadísticas en Google Trends desde el 2008. Esta aplicación permite encontrar las tendencias en YouTube y la popularidad de un término de búsqueda a lo largo del tiempo, lo que puede ayudar a las empresas con el contenido de sus estrategias de video. Cada mes, 1500 millones de usuarios inician sesión y visitan YouTube. Los usuarios dedican más de una hora al día a ver contenido de YouTube en sus dispositivos móviles. Además, durante los dos últimos años, el número de pequeñas y medianas empresas que se anuncian en la página se ha duplicado (YouTube, 2017).

En cuanto a las búsquedas de imágenes, es una opción de Google para hallar imágenes relacionadas a una palabra clave dentro del contenido de diferentes sitios web, basándose en el nombre de la imagen.

En las búsquedas de noticias se encuentran temas y noticias relacionadas a un término específico, el cual puede ser filtrado por fecha y país. El contenido inserto en Google Noticias se compila mediante un algoritmo informático que explora todos los sitios web incluidos en Google (Google, s.f.).

Existen indicadores que son excluidos de Google Trends, como es el caso de las búsquedas no significativas, es decir, aquellas que no alcanzan un mínimo de consultas en un espacio geográfico determinado; la duplicidad de búsquedas de una sola persona en un periodo reducido de tiempo; y los caracteres especiales como apóstrofes, tildes u otros (Google, s.f.).

Los datos proporcionados por Google no son absolutos, pero sí relativos (Baggio, 2016), lo cual puede ser considerado una limitación. Los resultados de búsqueda son proporcionales a un periodo de referencia y la ubicación de una consulta; cada punto de datos se divide por el número total de búsquedas de la región geográfica y el lapso de tiempo que representa para comparar su popularidad relativa, escalándose las

búsquedas en un intervalo del 0 al 100. Si esto no se hiciera, los lugares con mayor volumen de búsquedas siempre aparecerían en los puestos más altos. (Google Trends, 2016).

Otra limitación pertinente radica en la homogenización de términos (Borba & Akemi, 2014). Existen individuos, por ejemplo, que pueden buscar información turística acerca de “Cusco” también con el nombre de “Cuzco”. También puede darse el caso que un individuo busque información acerca de las Islas Ballestas y durante la digitación escriba el nombre del atractivo con faltas ortográficas. De ahí que se afirme que las consultas vienen relacionadas a factores como el grado de instrucción o edad (Zeynalov, 2014). Otro suceso es que el usuario busque un término referido a alguna categoría y encuentre un destino turístico, como es el caso de lima (fruta) y Lima (capital del Perú). Para evitar estos errores semánticos, la investigación hizo uso de cinco palabra clave definidas; contemplando el término “kuelap” y otros cuatro términos asociados por la herramienta Google Trends en base a su algoritmo de búsquedas: “kuelap peru”, “kuelap amazonas”, “kuelap tours” y “fortaleza de kuelap”.

A pesar de las limitaciones, como se observa en los antecedentes, Google Trends ha sido tema de estudios relacionados a la mejora en la precisión de la predicción de las llegadas a un destino turístico usando los datos de búsqueda que se registran en Google, lo cual, permite entender de cierta forma su comportamiento futuro de viaje y hacia dónde direccionan su interés, además de evaluar la correlación entre las búsquedas y las llegadas para un periodo específico.

En este sentido, los indicadores y datos de búsqueda de Google Trends pueden ser utilizados para hallar una correlación con los arribos históricos, tal como concluyeron Choi y Varian (2009). Por su parte, Onder y Gunter (2016) y Saidi (2010) demostraron que Google Trends aporta una mejor precisión a los modelos que utilizan variables históricas. Asimismo, Park, Lee y Song (2015) consideraron que combinando palabras clave se pueden minimizar errores en los resultados de la herramienta digital. Mientras que Zeynalov (2014) argumentó que filtrando en la categoría viaje, los resultados son más segmentados y se construye un mejor modelo econométrico.

### ***Proyección Turística – “Tourism Forecasting”***

Proyectar es un proceso que permite obtener un pronóstico de la magnitud de la evolución de los valores en el futuro. Este proceso ayuda a las compañías a estar mejor preparadas para las condiciones futuras del entorno del mercado (Hart , Rasner, & Lukoszová, 2012). En otras palabras, se organiza la información pasada

para predecir el futuro. Por ejemplo, en el caso de arribos por turismo, la proyección arrojará valores estimados del número de visitantes a un destino turístico concreto.

La importancia económica del sector turismo ha estimulado al incremento del interés en la proyección de demanda turística debido a que facilita la toma de decisiones, reduciendo los riesgos y brindando una mejor y más clara percepción de escenarios futuros (Frechtling, 2001). Por ejemplo, quienes laboran en marketing turístico pueden utilizarla para establecer metas de marketing; explorar la factibilidad de mercados potenciales o simular el impacto de eventos futuros en la demanda, incluyendo planes de marketing alternativos.

Existen dos subcategorías principales para los métodos cuantitativos de predicción: extrapolativos y causales (Frechtling, 2001). Los métodos extrapolativos, también son llamados métodos de series de tiempo, y asumen que los patrones del pasado son la clave para predecir el futuro; suponiendo que estos pueden repetirse. Por su parte, los métodos causales intentan simular matemáticamente las relaciones de causa y efecto; determinando las variables causales o variables explicativas que afectan a la variable de predicción y la expresión matemática apropiada para esta relación. Cuando se trata de proyección, ningún método es mejor que otro o se destaca por ser más confiable (Frechtling, 2001). No obstante, en cuanto a resultados, los autores concuerdan en que resulta más exacto predecir el número de personas que su gasto en el destino (Frechtling, 2001). Asimismo, la combinación de modelos de proyección sugiere una mejora para reducir errores en la predicción (Song, Witt, Wong, & Wu, 2007)

Para la proyección turística, se puede hacer uso de diferentes variables; entre ellas, la variable de arribos de turistas sigue siendo la medida de demanda turística más popular (Song & Li, 2008). Esta variable es medida por el total de arribos turísticos desde el punto de origen del viajero al destino turístico. Se puede descomponer en arribos turísticos por vacaciones, por negocios, para visitar a familiares y/o amigos, entre otros. Para medir esta variable, se recomienda utilizar datos provenientes de diferentes fuentes de manera integrada; estos datos pueden ser obtenidos a partir de encuestas y controles administrativos en puntos de ingreso como fronteras y aeropuertos; o en el caso de regiones que no cuentan con puntos de ingreso, utilizar alojamientos o sitios turísticos populares (OMT, 2010).

Además de los arribos, algunos estudios han usado el gasto de turistas en el destino como una variable de demanda, pudiendo ser medido a través de la inclusión de una

sección de gastos específica en las encuestas de recopilación de información de los turistas o a través de métodos alternativos como sistemas de registros bancarios, informes de tarjetas de crédito, o gastos de viaje proporcionados por agencias de viajes, empresas, o las autoridades reguladoras del transporte (OMT , 2010).

Según la Organización Mundial del Turismo (2010), adicional al flujo turístico y el gasto; el empleo, los aspectos demográficos, la duración del viaje y la motivación también son variables a considerar para las estadísticas en turismo. La recopilación de datos sobre el empleo en las industrias turísticas es complejo; y aunque es improbable que se pueda obtener una visión completa del empleo en turismo a partir de una única fuente estadística, se podrían utilizar fuentes como: (i) encuestas por muestreo a los hogares; (ii) encuestas por muestreo a los establecimientos, y (iii) registros administrativos. En cuanto a los aspectos demográficos, las características personales de los visitantes deberían recopilarse a través de información de control recabada en los establecimientos de alojamiento o mediante encuestas realizadas en las fronteras o en lugares específicos. La duración del viaje, es un factor importante para evaluar el nivel de demanda de servicios turísticos, como el alojamiento; por lo que determinar este factor podría servir para estimar el gasto asociado con un viaje. Respecto a la motivación del viaje, este es un factor útil para caracterizar los perfiles del gasto turístico, además de servir para identificar segmentos clave de la demanda turística con enfoque en la planificación y promoción.

Algunas de las dificultades que se encuentran al realizar predicciones turísticas son que, a veces se carece de data histórica; la demanda turística es volátil dependiendo de la estación; los turistas son sensibles a los desastres naturales o desastres ocasionados por el hombre; y el comportamiento del turista es complejo dependiendo de sus propósitos o motivos de viaje (Frechtling, 2001).

Es importante mencionar en este punto que las mediciones y estadísticas en turismo aportan puntos de vista sobre las actividades asociadas con el turismo en un territorio geográfico determinado, ya sea a nivel nacional como regional; sin embargo, es complejo medir las variables en las estadísticas del sector turismo debido a su naturaleza transversal. Las estadísticas de turismo son el producto final de un complejo proceso que incluye numerosas fases que inician en la recopilación y tratamiento de la información en bruto y finalizan en la difusión de los datos tratados para un uso previsto (OMT , 2010).

### **Modelo de Análisis Predictivo.**

El análisis predictivo o modelado predictivo; es un método que permite el análisis de la relación existente entre una unidad de muestra y uno o más atributos o características. Estos modelos implican el análisis de datos y son ampliamente utilizados en negocios como mecanismo para la toma de decisiones; siendo usados en muchas áreas como el marketing, donde se buscan patrones de datos para obtener la respuesta a preguntas sobre el desempeño del cliente (Espino, 2017).

La elección de un modelo suele estar en función a los datos que se tengan disponibles y a las destrezas del investigador (Baggio, 2016). Las investigaciones en turismo disponibles en bases de datos científicas que utilizan modelos de análisis predictivo (Onder & Gunter, 2016 o Park, Lee, & Song, 2015), citan el modelo de Choi y Varian (2009) como el origen de otros modelos de predicción con variables electrónicas. El modelo de Choi y Varian (2009) usa datos históricos y emplea estadística inferencial. Lutz (2013) y Hamia y Heiden (2015) señalan que la importancia de dicho modelo radica en ser el primero que evalúa la incorporación de un indicador electrónico como variable independiente para alcanzar un mayor nivel de precisión en otros indicadores económicos, por ejemplo, arribos de turistas (Park, Lee, & Song, 2015).

Los métodos computacionales resultan ser los más precisos debido a que emplean métodos más sofisticados como la inteligencia artificial y *machine learning* (Kim & Moosa, 2001). Dichos métodos permiten que la obtención de resultados sea más rápida, fácil y accesible. Este estudio maneja grandes cantidades de datos obtenidos por Google Trends, los cuales podrán ser tratados de forma más organizada y entendible a través del uso de estos métodos computacionales; que también permitirán obtener los coeficientes a analizar de manera más rápida y clara.

Choi y Varian (2009) hacen referencia a un modelo externo del tipo AR-1 aplicable al turismo, el cual refiere que las llegadas a un destino están relacionadas con sus datos predecesores registrados hace un año y hace un mes. Este modelo es conveniente debido a que requiere de pocas variables: los arribos que ha registrado un sitio turístico y las búsquedas en línea sobre el mismo. Este modelo, a pesar de su sencillez, resulta suficientemente robusto (De la Oz, 2014). El modelo de referencia AR-1 señala que la cantidad de arribos mensuales está determinada por la siguiente ecuación:

$$\text{Log}X(t) = b_1 + b_2 * \text{Log}X(t - 12) + b_3 * \text{Log}X(t - 1) \quad \text{Modelo AR - 1}$$

En esta ecuación:  $X(t)$  representa la cantidad de llegadas registrada en el mes “T” (junio del 2015, por ejemplo);  $X(t-12)$  corresponde a las visitas recibidas para el mismo mes, pero correspondientes al año anterior (junio del 2014 de acuerdo al ejemplo);  $X(t-1)$  es también un flujo que corresponde al número de arribos, pero para el mes inmediato anterior (mayo del 2015 según la situación ejemplificada).

Choi y Varian (2009) incorporaron el indicador de Google Trends al modelo AR-1 y construyeron la siguiente relación:

$$\text{Log}X(t) = b_1 + b_2 * \text{Log}X(t - 12) + b_3 * \text{Log}X(t - 1) + c * G(t) \text{ Modelo C. \& V. 2009}$$

La ecuación anterior incorpora el índice de búsquedas mensual  $G(t)$ , el cual será adicionado en forma lineal y no sobre la base de un logaritmo como los anteriores. En cuanto a los demás términos:  $b_1$  es un término independiente;  $b_2$ ,  $b_3$  y  $c$  son coeficientes que afectan a las variables  $X(t-12)$ ;  $X(t-1)$  y  $G(t)$  respectivamente.

La ecuación de Choi y Varian (2009) demostró que los indicadores proporcionados por Google Trends, una herramienta que brinda indicadores en tiempo real, se pueden correlacionar con diferentes indicadores económicos, entre ellos, los arribos turísticos.

Los datos de visitas disponibles y la bibliografía en mencion han direccionado a que el modelo de Choi y Varian (2009) sea el más adecuado para la presente investigación, debido a su accesible aplicación y análisis.

### ***Marketing Digital/Online.***

El marketing es el proceso de planificación y ejecución de la concepción, fijación de precios, promoción y distribución de ideas, bienes y servicios para crear un intercambio y satisfacer a los individuos y/o organizaciones; siendo éste una parte integral del proceso de gestión de una empresa comercial (Rice, 1997).

El modelo de análisis predictivo explicado en el marco teórico hace uso del modelamiento de la data histórica de arribos y los índices de búsquedas en Google; lo cual, podría servir para encontrar una correlación, que de ser positiva, podría reflejar el interés hacia un atractivo específico en términos digitales.

Este estudio contempla el marketing online, el cual, según Kotler & Armstrong (2012), se considera la forma de marketing directo con mayor y más rápido crecimiento debido al impacto en el uso del internet por parte de la oferta y la demanda; compartiendo enormes cantidades de información, lo que puede ser aprovechado por los profesionales en marketing para crear valor y establecer relaciones con los

clientes a través de un sitio web; anuncios y promociones en línea; creando o participando en comunidades web y redes sociales; o utilizando correo electrónico.

El marketing online se dirige a un público más activo y que recibe la información bajo sus propias condiciones, a diferencia del marketing tradicional que se dirige a un público hasta cierto punto pasivo y menos segmentado que recibe mensajes de manera unidireccional. Este tipo de marketing permite relacionarse con el cliente de una manera más profunda y significativa a través de una comunicación bidireccional e interactiva gracias a las nuevas tecnologías (Kotler & Armstrong, 2012).

La tecnología ha ido mejorando día a día, causando que inevitablemente, las compañías se vean afectadas e incorporen las tecnologías de información y comunicación a sus estrategias. Los departamentos de marketing ya integran estrategias en internet a su repertorio previo de estrategias tradicionales, para capturar la atención de su público objetivo, quienes pasan más y más tiempo online (Guðbjörg, 2014). Debe haber una mezcla de acciones, tomando en cuenta que los medios tradicionales son “ganchos de alcance” o métodos *call to action* que buscan una respuesta inmediata, y los digitales son medios para mantener la frecuencia. Además, es importante que toda la organización esté involucrada en el proceso de marketing, no sólo el área encargada (Fonseca, 2017).

Kotler, Kartajaya y Setiawan (2010) mencionan que el marketing ha ido evolucionando en etapas: marketing 1.0, 2.0 y 3.0. El inicio del marketing, la versión 1.0, tenía como objetivo vender productos, estandarizando y reduciendo los costos de producción para que el precio de los bienes sea menor y así, se pueda vender más en base a un precio más asequible para los consumidores. Luego, el marketing 2.0, apareció en la “época de la información tecnológica”, teniendo como objetivo la diferenciación, satisfacción y retención de los clientes, centrándose en ellos y segmentándolos para producir mejores productos acorde a sus necesidades. Ahora, ha llegado el marketing 3.0 con el objetivo de “hacer del planeta un lugar mejor”, centrandó el concepto del marketing en los valores que se ofrecen a los consumidores y realizando marketing de colaboración, pasando a ser “prosumidores” (combinación entre consumidores y productores), permitiendo a los clientes dar su opinión sobre la empresa y participando activamente con ella.

El marketing para destinos debe considerar la forma correcta para integrar cada decisión de viaje en sus campañas, de manera que se optimice el conocimiento y la actitud positiva hacia el destino, y por consiguiente, se incremente el gasto de los turistas (Fesenmaier & Stienmetz, 2013). Un destino necesita del marketing

electrónico dado que hoy en día el uso del internet es bastante alto por parte de los viajeros, por lo cual, el objetivo es que todo el itinerario pre-viaje (inspiración, planificación, reservas) se realice online (European Travel Commission & Organización Mundial del Turismo, 2008).

Según Think with Google (2016), el proceso de compra de un viaje pasa por cuatro etapas: inspiración, planificación, reserva y experiencia en el destino; no obstante, con el aumento de la tecnología, se identificó una quinta etapa: compartir; la que se basa en revivir la experiencia a través de imágenes, fotos y opiniones, y a su vez influenciar a otros. Cabe mencionar que esta quinta etapa, en la actualidad, se incorpora dentro de las demás. Así, es en la etapa de inspiración donde se debe crear el deseo e interés de viajar a un destino; en la etapa de planeamiento donde se comunica el valor; en la etapa de reserva donde se brindan las facilidades necesarias para que este proceso no sea complejo y se ofrece lo mejor del destino; y la etapa de experiencia del viaje donde se fomenta el retorno e inspira a un nuevo viaje (Think with Google, 2016).

En este sentido, cada Organización de Gestión de Destino (DMO por sus siglas en inglés) es la llamada a planificar sus estrategias de marketing, y con los nuevos avances tecnológicos, considerar diferentes modelos para gestionar su marketing electrónico (European Travel Commission & Organización Mundial del Turismo, 2008). En el caso de Perú, se podría mencionar que se trabaja desde el “momento cero de la verdad” o ZMOT por sus siglas en inglés, que es ese momento previo a la compra, donde las primeras impresiones ocurren y el camino hacia la compra inicia (Google, 2011). En Perú, se proporciona información actualizada online a los viajeros potenciales en el ZMOT, por ejemplo, a inicios de noviembre del 2017 se ha lanzado la nueva campaña de PROMPERÚ, nombrada: “Perú, el país más rico del mundo”, para la cual, se ha creado una página web ([www.perupaismasrico.com](http://www.perupaismasrico.com)) donde se brinda información y se visualiza una serie de videos que pueden ser compartidos en redes sociales (PROMPERÚ, 2017), influenciando de esta manera, en la toma de decisiones en las etapas de búsqueda, planificación y selección del destino (Google, 2011); desarrollando y proyectando la marca; y proporcionando ideas e inspiración para que el viajero seleccione a Perú como su destino de viaje.

Actualmente, a nivel electrónico, se maneja el llamado *inbound marketing* o marketing de atracción. El marketing de atracción es un enfoque holístico orientado a atraer individuos a la marca y convertirlos en clientes duraderos. Este tipo de marketing es una estrategia que incorpora técnicas reconocidas para atraer y retener

a los clientes, tales como SEO, marketing en redes sociales, e-mail marketing, relaciones públicas, CRM, analítica web, minería de datos, etc. La idea central del marketing de atracción es el utilizar estrategias capaces de ganar el interés del cliente y atraerlo a la empresa en lugar de enviar un mensaje publicitario y esperar su reacción (Bezhovski, 2015).

Una de las herramientas utilizadas para las estrategias de inbound marketing, es el posicionamiento en buscadores, el cual sirve para atraer a quienes navegan en la red. Un destino que recién se está dando a conocer, encuentra en los buscadores web, como Google, una herramienta útil para captar mayor visibilidad ante el público extranjero y la industria turística. No obstante, el sector turismo en Latinoamérica aún presenta un bajo uso de prácticas inbound; a pesar de que la experiencia ha demostrado que estas técnicas dan muy buenos resultados y brindan una gran oportunidad para la industria (Samsing, 2015).

La DMO, como parte integral de la industria, debe adoptar las nuevas estrategias de marketing online para continuar eficientemente con el logro de sus objetivos planteados y la promoción del destino en el mundo (European Travel Commission & Organización Mundial del Turismo, 2008). Por ejemplo, se podría utilizar la herramienta Google Trends para observar y comparar el nivel de búsqueda sobre una determinada palabra durante un tiempo específico; tomar decisiones en base al conocimiento de aquellos momentos en que el destino es más buscado y analizar los patrones de estacionalidad para crear un calendario de contenidos para las campañas de marketing online. También los indicadores de la herramienta podrían ser empleados para medir los efectos de una campaña de marketing online, en base a la obtención de una relación matemática entre las búsquedas y las llegadas, generando una noción objetiva del ratio de respuesta.

### ***Marketing en Buscadores Web.***

Algunos de los más famosos buscadores web son Google, Yahoo y Bing. Google es el buscador número uno hasta el momento; aunque en algunos países predominan buscadores locales como Baidu en China, Yandex en Rusia y Timway en Hong Kong (Durmaz & Hali, 2016). Según Dumaz y Hali (2016) “el marketing con buscadores web es una parte del marketing online que se enfoca en proveer una posición privilegiada dentro de los buscadores; para lo cual, se debe tener una apropiada mezcla entre la organización de tu sitio web y el análisis de palabras clave”. El estar bien posicionado en los buscadores web, a diferencia de otras técnicas de marketing,

es una acción que ayuda a los usuarios a encontrar lo que buscan en lugar de sólo captar su atención y tratar de influir en su decisión de compra.

Google es el buscador más utilizado hoy en día, y desde 1998 viene reinventado a los motores de búsqueda; de forma que la palabra Google se encuentra en el diccionario definido como una palabra para búsquedas en internet. También ha redefinido la publicidad en línea a través de la provisión de una plataforma de publicidad dentro de un motor de búsqueda (Kotler, Kartajaya, & Setiawan, 2010).

Dentro del marketing online en buscadores web, se encuentran las estrategias SEO (Search Engine Optimization) & SEM (Search Engine Marketing), consideradas las dos bases más importantes del marketing online (Durmaz & Hali, 2016).

Las estrategias SEO consisten en desarrollar e implementar una serie de procedimientos para fortalecer la notoriedad de un sitio web con el objetivo de aparecer en las primeras posiciones de la lista de resultados. Por su parte, las estrategias SEM son complementarias a las SEO, centrándose en enlaces de pago y caracterizándose por segmentar el mercado y ofrecer lo que el usuario está buscando (Sustaeta, 2014). De esta manera, las estrategias SEO tratan de posicionar un sitio web en los buscadores de forma orgánica, es decir, se posiciona en la web en base al contenido de la misma; mientras que las estrategias SEM tratan de fortalecer la presencia de un sitio web en base a un posicionamiento de pago y segmentado, es decir, se realiza un pago para que la web se ubique de forma preferencial en la parte superior o derecha del buscador, siendo lo primero que observa el internauta cuando realiza la búsqueda de un tema en específico.

Para lograr un buen posicionamiento a través de estrategias SEO, el sitio web debe contener notoriedad en internet, es decir, tener enlaces de otros sitios webs y viceversa; debe ser activo y con contenido actualizado; permitir la participación activa de las personas (ej. a través de comentarios); estar habilitado en diferentes idiomas; y evitar el *spam* y el *copyright*. Por su parte, el posicionamiento SEM se realiza a través de pago, siendo su herramienta más conocida *Google Adwords*. Para hacer uso de esta herramienta es necesario crear una cuenta gratuita en Google Adwords, escribir el anuncio que se mostrará, seleccionar las palabras clave para el anuncio y realizar el pago (Sustaeta, 2014). Se debe considerar que cada palabra clave tiene un costo diferente, por ejemplo, la palabra Perú costará mucho más que la palabra Kuélap, por ser más utilizada y conocida. Es así que, el costo de esta herramienta depende de la disponibilidad de pago y las palabras clave que se desee utilizar.

Como se ha mencionado, Google es el buscador más utilizado, dominando más del 90% de las búsquedas a nivel mundial (StatCounter, 2017), y siendo de esta manera, el mejor buscador web para planificar y aplicar estrategias SEO y SEM. Este buscador cuenta con Google Trends, una aplicación que proporciona información semanal o mensual sobre las búsquedas que se realizan en Google. Los buscadores web son usados para el 85% de las búsquedas de viajes (Solimar International, s.f), por lo que los datos de Google Trends pueden proporcionar *insights* sobre las intenciones, intereses y deseos de las personas, lo cual, puede permitir a los gestores tomar decisiones informadas. Dinis, Costa y Pacheco (2016) mencionaron que “la información de Google Trends puede ser utilizada como un aproximado a futuros movimientos a un destino y es de gran interés para los gestores debido a que están disponibles tiempo antes del viaje y de las estadísticas turísticas oficiales de un país”.

### ***Complejo Arqueológico de Kuélap.***

El Complejo Arqueológico de Kuélap es considerado un atractivo turístico porque se pueden realizar actividades de turismo cultural, naturaleza y aventura que, además de contar con el recurso turístico de la Zona Arqueológica Monumental Kuélap; posee instalaciones turísticas y equipamiento como el sistema de telecabinas para el acceso, la zona de ventas de artesanías o el museo de sitio (MINCETUR, 2011). Este atractivo se encuentra situado en los andes amazónicos nororientales del Perú, en la provincia de Luya, región Amazonas. Se ubica a 3 000 msnm y su temperatura oscila entre los 13°C y 23°C. Perteneciente a la cultura pre inca Chachapoyas, Kuélap, se localiza en la cima de una montaña y goza de una impresionante flora y fauna que lo vuelven un sitio propicio para practicar turismo de aventura, cultural y de naturaleza (PROMPERÚ, 2017).

Kuélap fue declarado Zona Arqueológica Monumental Kuélap, mediante Resolución Directoral Nacional RDN N°694-2007/INC; además, se encuentra en la lista tentativa de la UNESCO para declararse Patrimonio Cultural de la Humanidad. Es el atractivo de mayor afluencia turística dentro de las referencias estadísticas que presenta la región Amazonas: Complejo Arqueológico de Kuélap, Catarata de Gocta y Museo de Leymebamba, con 56 010 visitas entre nacionales y extranjeros en el 2016 y un crecimiento de 53.94% respecto al año 2015; esto según revisión de la Base de Datos de Turismo del Ministerio de Comercio Exterior y Turismo. (MINCETUR, 2017).

Según el Perfil del Vacacionista Extranjero que Visita Amazonas, el 75% viaja por cuenta propia, 89% busca información turística sobre el destino en internet; 38% toma

más de un mes para planificar su viaje; 89% ha visitado Kuélap y 94% presenta recordación sobre Kuélap como atractivo turístico de Amazonas (PROMPERÚ, 2015). Por su parte, el Perfil del Vacacionista Nacional que Visita Amazonas señala que el 84% viaja por cuenta propia, 81% busca información turística sobre el destino en internet; 36% toma una semana o menos para planificar su viaje; 86% ha visitado el Kuélap y 97% presenta recordación sobre Kuélap como atractivo turístico de Amazonas (PROMPERÚ, 2015).

El Complejo Arqueológico de Kuélap se encuentra bajo la administración del Ministerio de Cultura - Dirección Regional de Cultura Amazonas. Organizaciones como el MINCETUR, Plan COPESCO y el Gobierno Regional Amazonas vienen interviniendo con proyectos desde el año 2003, con la finalidad de investigar, conservar y mantener el atractivo. El complejo, es un espacio que permite realizar observación de aves, observación de flora y fauna, trekking, fotografía, entre otras actividades. Se ofrece servicio de guiado y en los alrededores del complejo se encuentran kioscos de venta de comida/bebida, venta de artesanías y alquiler de caballos. Un poco más alejado, en el distrito de Tingo, se accede a servicios higiénicos, hostales, y restaurantes (MINCETUR, 2013).

En cuanto a los objetivos referentes a turismo planteados para este atractivo, se puede tomar como referencia el Plan Estratégico de Desarrollo Concertado de Amazonas 2008 – 2021, donde se contempla poner en valor y formar circuitos o corredores turísticos e implementar infraestructura y servicios turísticos en la región Amazonas con inversión pública y privada; tomando en cuenta la presencia de Kuélap que ha sido determinante para la instalación de diversos prestadores de servicios turísticos, principalmente en la ciudad de Chachapoyas. También se considera promover el posicionamiento de productos y recursos turísticos de la Región Amazonas en el mercado nacional e internacional; poniendo énfasis en el desarrollo de un concepto y una marca que proporcione identidad y potencie la inserción de los productos y servicios turísticos regionales. Asimismo, se busca promover el desarrollo turístico en la región conservando el patrimonio cultural, arqueológico y natural, enfocándose en la sostenibilidad del patrimonio como base del desarrollo turístico regional (Gobierno Regional de Amazonas, 2007).

El más grande proyecto de inversión realizado es el Sistema de Telecabinas de Kuélap, que comenzó su construcción en 2015 e inicio operaciones en 2017. Las telecabinas trasladan en aproximadamente 20 minutos a los visitantes hasta el complejo, a comparación de años anteriores cuando se debía caminar durante 90

minutos aproximadamente hasta el atractivo (El Comercio, 2017). Esto significa una gran facilidad para el acceso a esta importante zona turística del oriente peruano, además de la mejora de la presentación del atractivo turístico. Según el Plan Copesco Nacional, esta obra busca mejorar la accesibilidad a la zona arqueológica monumental Kuélap; adecuar las condiciones para el recorrido interno del atractivo; incrementar la difusión y promoción del atractivo; lograr una eficiente capacidad de gestión turística cultural de las autoridades; y optimizar los conocimientos de calidad y cultura turística de los agentes involucrados (Plan Copesco Nacional, s.f).

Las Telecabinas de Kuélap fueron construidas bajo la modalidad de Asociación Público – Privada por el Consorcio Telecabinas Kuélap, integrado por la empresa francesa Pomagalski S.A.S y la peruana Ingenieros Civiles y Contratistas Generales S.A. Para usarlo, el visitante debe llegar a la estación de embarque en Tingo Nuevo y abordar los buses turísticos del concesionario para llegar a la estación de salida de las telecabinas. Una vez en la estación de salida, se abordarán las cabinas hasta llegar al Parador de La Malca, donde se inicia el recorrido peatonal hacia la fortaleza.

Actualmente, el Complejo Arqueológico de Kuélap tiene presencia web con llegada a turistas extranjeros en el portal Peru.Travel, donde se contemplan datos e información sobre el atractivo. Asimismo, el público nacional puede acceder a información y paquetes turísticos sobre este atractivo a través del portal Y Tú Que Planes. Ambos portales son el segundo y tercer sitio web sugerido por el buscador cuando se digita el término “Kuélap”, sólo después de su referencia en Wikipedia. Cabe mencionar que en la primera página del buscador no aparecen ofertas de paquetes turísticos, solamente información sobre el atractivo.

## Objetivos e Hipótesis

### Objetivos

**Objetivo General:** Determinar si los indicadores de Big Data de Google Trends son estadísticamente significativos<sup>1</sup> para mejorar el modelamiento del flujo de visitantes nacionales y extranjeros al Complejo Arqueológico de Kuélap en el periodo 2011-2016.

**Objetivo específico 1:** Explorar si los indicadores de Big Data de Google Trends contribuyen al mejoramiento de estrategias de marketing online del Complejo Arqueológico de Kuélap.

**Objetivo específico 2:** Determinar si el modelo AR-1 es estadísticamente significativo para pronosticar las llegadas en el periodo de 2011-2016 para el caso del Complejo Arqueológico de Kuélap (Este modelo no hace uso de variables electrónicas).

**Objetivo específico 3:** Determinar si la inclusión de la variable G(web) es estadísticamente significativa en comparación con el modelo AR-1 que sólo emplea variables históricas en el periodo de 2011-2016 para el caso del Complejo Arqueológico de Kuélap.

**Objetivo específico 4:** Determinar si la inclusión de las combinaciones de las variables G(YouTube), G(Imágenes) y/o G(Noticias) son estadísticamente significativas en comparación con el modelo que sólo emplea datos históricos y G(web) en el periodo de 2011-2016 para el caso del Complejo Arqueológico de Kuélap.

### Hipótesis

H<sub>10</sub> - El modelo AR-1 no es estadísticamente significativo para pronosticar las llegadas en el periodo de 2011-2016 para el caso del Complejo Arqueológico de Kuélap.

H<sub>11</sub> - El modelo AR-1 es estadísticamente significativo para pronosticar las llegadas en el periodo de 2011-2016 para el caso del Complejo Arqueológico de Kuélap.

H<sub>20</sub> - La inclusión de la variable G(web) no es estadísticamente significativa en comparación con el modelo AR-1 que sólo emplea variables históricas en el periodo de 2011-2016 para el caso del Complejo Arqueológico de Kuélap.

H<sub>21</sub> - La inclusión de la variable G(web) es estadísticamente significativa en comparación con el modelo AR-1 que sólo emplea variables históricas en el periodo de 2011-2016 para el caso del Complejo Arqueológico de Kuélap.

---

<sup>1</sup> Es un adjetivo que se emplea para describir la validez, respaldada por la estadística, para demostrar que la asociación no ha ocurrido en virtud del azar (Soporte Minitab, s.f.).

H3<sub>0</sub> - La inclusión de la combinación de las variables G(Youtube), G(Imágenes) y G(Noticias) no es estadísticamente significativa en comparación con el modelo que sólo emplea datos históricos y G(web) en el periodo de 2011-2016 para el caso del Complejo Arqueológico de Kuélap.

H3<sub>1</sub> - La inclusión de la combinación de las variables G(Youtube), G(Imágenes) y G(Noticias) es estadísticamente significativa en comparación con el modelo que sólo emplea datos históricos y G(web) en el periodo de 2011-2016 para el caso del Complejo Arqueológico de Kuélap.

H4<sub>0</sub> - La inclusión de la combinación de las variables G(Youtube) y G(Imágenes) no es estadísticamente significativa en comparación con el modelo que sólo emplea datos históricos y G(web) en el periodo de 2011-2016 para el caso del Complejo Arqueológico de Kuélap.

H4<sub>1</sub> - La inclusión de la combinación de las variables G(Youtube) y G(Imágenes) es estadísticamente significativa en comparación con el modelo que sólo emplea datos históricos y G(web) en el periodo de 2011-2016 para el caso del Complejo Arqueológico de Kuélap.

H5<sub>0</sub> - La inclusión de la combinación de las variables G(Imágenes) y G(Noticias) no es estadísticamente significativa en comparación con el modelo que sólo emplea datos históricos y G(web) en el periodo de 2011-2016 para el caso del Complejo Arqueológico de Kuélap.

H5<sub>1</sub> - La inclusión de la combinación de las variables G(Imágenes) y G(Noticias) es estadísticamente significativa en comparación con el modelo que sólo emplea datos históricos y G(web) en el periodo de 2011-2016 para el caso del Complejo Arqueológico de Kuélap.

H6<sub>0</sub> - La inclusión de la combinación de las variables G(Youtube) y G(Noticias) no es estadísticamente significativa en comparación con el modelo que sólo emplea datos históricos y G(web) en el periodo de 2011-2016 para el caso del Complejo Arqueológico de Kuélap.

H6<sub>1</sub> - La inclusión de la combinación de las variables G(Youtube) y G(Noticias) es estadísticamente significativa en comparación con el modelo que sólo emplea datos históricos y G(web) en el periodo de 2011-2016 para el caso del Complejo Arqueológico de Kuélap.

H7<sub>0</sub> - La inclusión de la variable G(Imágenes) no es estadísticamente significativa en comparación con el modelo que sólo emplea datos históricos y G(web) en el periodo de 2011-2016 para el caso del Complejo Arqueológico de Kuélap.

H7<sub>1</sub> - La inclusión de la variable G(Imágenes) es estadísticamente significativa en comparación con el modelo que sólo emplea datos históricos y G(web) en el periodo de 2011-2016 para el caso del Complejo Arqueológico de Kuélap.

H8<sub>0</sub> - La inclusión de la variable G(Youtube) no es estadísticamente significativa en comparación con el modelo que sólo emplea datos históricos y G(web) en el periodo de 2011-2016 para el caso del Complejo Arqueológico de Kuélap.

H8<sub>1</sub> - La inclusión de la variable G(Youtube) es estadísticamente significativa en comparación con el modelo que sólo emplea datos históricos y G(web) en el periodo de 2011-2016 para el caso del Complejo Arqueológico de Kuélap.

H9<sub>0</sub> - La inclusión de la variable G(Noticias) no es estadísticamente significativa en comparación con el modelo que sólo emplea datos históricos y G(web) en el periodo de 2011-2016 para el caso del Complejo Arqueológico de Kuélap.

H9<sub>1</sub> - La inclusión de la variable G(Noticias) es estadísticamente significativa en comparación con el modelo que sólo emplea datos históricos y G(web) en el periodo de 2011-2016 para el caso del Complejo Arqueológico de Kuélap.

## CAPÍTULO II: MÉTODO

### Tipo y Diseño de Investigación

#### **Tipo de Investigación.**

El presente estudio tiene una orientación mixta, es decir, integra datos cuantitativos y cualitativos. Los métodos mixtos se caracterizan por producir datos más variados mediante la diversidad de observaciones y perspectivas, considerándose varias fuentes y tipos de datos, contextos o ambientes para el análisis de la información. El investigador es quien define el enfoque que tiene mayor peso – cuantitativo o cualitativo –, o si se les otorga la misma prioridad a ambos (Hernández, Fernández, & Baptista, 2014). Este tipo de investigación reforzó la credibilidad de los procedimientos y dio mayor solidez al enfoque y resultados del estudio.

La primera etapa, cualitativa, consistió en recolectar y analizar, mediante entrevistas a profundidad, la percepción de expertos en materia de Big Data, prospectiva turística y marketing turístico sobre el problema de investigación y el modelo planteado. Los resultados de esta etapa sirvieron de soporte para la etapa cuantitativa, dándole mayor consistencia al planteamiento de este estudio, así como para evaluar la utilidad de las variables electrónicas en un modelo predictivo que también sirva para plantear estrategias de marketing online en un atractivo en crecimiento como el Complejo Arqueológico de Kuélap. La segunda etapa, cuantitativa, buscó evaluar variables de un modelo de análisis predictivo para conocer si los resultados que relacionan a estas variables de arribos históricos y electrónicas podrían servir para el modelamiento predictivo de llegadas de turistas al Complejo Arqueológico de Kuélap en escenarios de corto plazo.

#### **Diseño de Investigación.**

Esta investigación mixta posee un diseño anidado o incrustado concurrente de modelo dominante (DIAC), llamado así porque el método de menor prioridad, en este caso el cualitativo, es anidado o incrustado dentro del que se considera central, el cuantitativo. Esta acción se debe a que el método considerado secundario responde a diferentes preguntas de investigación respecto al método principal (Hernández, Fernández, & Baptista, 2014). Los datos recolectados en ambos métodos han sido conectados en la sección de Discusión para que se proporcione una visión más amplia del problema de investigación.

El método dominante del estudio, el cuantitativo, tiene un diseño de investigación del tipo correlacional, ya que busca medir la relación entre dos o más variables,

estableciendo un grado de correlación correspondiente. Este estudio no pretende explicar el fenómeno que es objeto de investigación, pues se limita a investigar grados de correlación (Hernández, Fernández, & Baptista, 2014). Esto significa analizar si un aumento o disminución en una variable coincide con un aumento o disminución en la otra variable.

A su vez, el estudio es de carácter no experimental. Este tipo de estudios son conocidos como investigaciones *Ex Post Facto*, es decir, “después de los hechos ocurridos”. Asimismo, estas investigaciones se realizan sin la manipulación intencional de variables, observando los fenómenos en su ambiente natural para analizarlos (Hernández, Fernández, & Baptista, 2014). Esto quiere decir que no se puede influir directamente sobre las variables por tratarse de situaciones que ya se dieron, sucediendo lo mismo sobre sus efectos. Por ejemplo, para el caso de este estudio se puede hablar de los hechos históricos pasados como los arribos a Kuélap o las búsquedas en línea sobre el atractivo.

Para el presente estudio, se probó un análisis de correlación y significancia, utilizando las variables del modelo AR-1, del modelo de Choi y Varian (2009) y siete posibles combinaciones del modelo de Choi y Varian (2009) adicionando tres indicadores obtenidos desde Google Trends. El modelo precedente sólo contempla búsquedas web; no obstante, las búsquedas de videos y búsquedas de imágenes representan también factores importantes hoy en día para el marketing, pues podrían complementar la medición del interés de aquellos viajeros que se encuentran en la fase de inspiración y planeamiento de viaje. Según Forbes (2014) “las imágenes complementan y mejoran el contenido, captan el interés de los usuarios y aumentan las posibilidades de ser encontrados a través de búsquedas orgánicas en internet”. Por su parte, “los videos significan una parte importante del marketing de contenidos; son extremadamente visuales y auditivos, haciendo que el usuario recuerde con mayor facilidad el contenido y mensaje del mismo, permitiendo la mejora de estrategias SEO y aumentando los clics hacia la página web principal” (Forbes, 2017). Se incluyó también la variable noticias, debido al impacto periodístico desde finales del 2014 que ha tenido la noticia de la construcción y puesta en marcha del sistema de telecabinas hacia Kuélap. Conociendo los modelos, se evaluó la correlación, nivel de significancia y estadístico-t de las variables de los dos modelos revisados en la literatura de análisis predictivo y las combinaciones derivadas de los indicadores agregados.

Cabe mencionar que el análisis de correlación es usado para examinar la relación entre una o más variables, dónde el resultado más cercano a +1.0 significa una correlación

positiva más estrecha entre las variables (Veal, 2006). En este estudio, esta prueba se analiza mediante el R-cuadrado y el R-cuadrado ajustado. En cuanto al nivel de significancia, representando por el valor de probabilidad o valor-P, muestra que una relación significativa es aquella que no ha ocurrido por casualidad (Veal, 2006), siendo mayor el error mientras más cerca se encuentre este valor a 1 (Guajarati, 2006). Los indicadores analizados serán más detallados en la sección Plan de Análisis.

El primer modelo evaluado es el AR-1, que es el modelo base citado por Choi y Varian (2009):

$$\text{Log}X(t) = b_1 + b_2 * \text{Log}X(t - 12) + b_3 * \text{Log}X(t - 1) \text{ Modelo AR - 1}$$

El segundo, el modelo elaborado por Choi y Varian (2009):

$$\text{Log}X(t) = b_1 + b_2 * \text{Log}X(t - 12) + b_3 * \text{Log}X(t - 1) + c * G(t) \text{ Modelo C. \& V. 2009}$$

Finalmente, se probaron todas las combinaciones posibles con las variables agregadas:

Ecuación de Choi y Varian (2009) + Gimagen (t) + GYouTube(t) + Gnoticias(t)

$$\text{Log}X(t) = b_1 + b_2 * \text{Log}X(t - 12) + b_3 * \text{Log}X(t - 1) + c_1 * Gweb(t) + c_2 * Gimagen(t) + c_3 * Gyoutube(t) + c_4 * Gnoticias(t)$$

Ecuación de Choi y Varian (2009) + Gimagen (t) + GYouTube(t)

$$\text{Log}X(t) = b_1 + b_2 * \text{Log}X(t - 12) + b_3 * \text{Log}X(t - 1) + c_1 * Gweb(t) + c_2 * Gimagen(t) + c_3 * Gyoutube(t)$$

Ecuación de Choi y Varian (2009) + Gimagen + Gnoticias(t)

$$\text{Log}X(t) = b_1 + b_2 * \text{Log}X(t - 12) + b_3 * \text{Log}X(t - 1) + c_1 * Gweb(t) + c_2 * Gimagen(t) + c_3 * Gnoticias(t)$$

Ecuación de Choi y Varian (2009) + GYouTube(t) + Gnoticias(t)

$$\text{Log}X(t) = b_1 + b_2 * \text{Log}X(t - 12) + b_3 * \text{Log}X(t - 1) + c_1 * Gweb(t) + c_2 * Gyoutube(t) + c_3 * Gnoticias(t)$$

Ecuación de Choi y Varian (2009) + Gimagen (t)

$$\text{Log}X(t) = b_1 + b_2 * \text{Log}X(t - 12) + b_3 * \text{Log}X(t - 1) + c_1 * Gweb(t) + c_2 * Gimagen(t)$$

Ecuación de Choi y Varian (2009) + GYouTube(t)

$$\text{Log}X(t) = b_1 + b_2 * \text{Log}X(t - 12) + b_3 * \text{Log}X(t - 1) + c_1 * Gweb(t) + c_2 * Gyoutube(t)$$

## Ecuación de Choi y Varian (2009) + Gnoticias (t)

$$\text{Log}X(t) = b_1 + b_2 * \text{Log}X(t - 12) + b_3 * \text{Log}X(t - 1) + c_1 * Gweb(t) + c_2 * GNoticias(t)$$

Se sustenta la adición de nuevas variables al modelo con la literatura de la regresión lineal múltiple (base del modelo de Choi y Varian); la cual, menciona que este método busca esclarecer la relación de una variable dependiente y varias independientes, estableciendo los símbolos matemáticos ( + ó - ) de acuerdo a una variable positiva o negativa (Universidad Nacional Mayor de San Marcos, 2006); es decir, las variables independientes se sumarán o restarán en base a su número resultante (coeficiente), ya sea positivo o negativo para cada una de ellas.

### **Variables.**

En la etapa cualitativa, las entrevistas a profundidad cuentan con variables deductivas porque son derivadas del marco teórico previamente descrito en este documento. Las categorías temáticas consideradas que engloban el marco teórico fueron las siguientes:

Tabla 01.

*Categorías Temáticas Análisis Cualitativo*

<b>Nombre de la Categoría</b>	<b>Descripción</b>
Big Data	<p>Se ha tomado en consideración la categoría de Big Data debido a que el enfoque del estudio utiliza esta gran cantidad de información para hallar una correlación entre las visitas a Kuélap y las búsquedas en Google.</p> <p>Esta categoría se trabajó en la entrevista mediante cinco preguntas abiertas referentes a los conocimientos del entrevistado sobre el concepto de Big Data; como afecta este concepto a nivel empresarial; sus limitaciones; si dentro de su empresa u organización hacen uso de esta tecnología y en que beneficiaría la Big Data a la gestión turística en el país.</p>
Proyección Turística	<p>La proyección en turismo permite organizar la información pasada para tener una mejor percepción de escenarios futuros que faciliten la toma de decisiones y reduzcan riesgos.</p> <p>Para las entrevistas resultó importante conocer su perspectiva en cuanto a proyección turística y en qué sentido esto es relevante para la gestión de un atractivo turístico. Es así que se exploraron tres aspectos en esta categoría: los beneficios de los modelos de proyección para la gestión turística; las desventajas de estos modelos y que variables considera relevantes para un modelo de este tipo.</p>

Google Trends	<p>La última categoría es Google Trends, una herramienta de Big Data que permite observar y comparar el nivel de búsqueda en Google sobre una palabra clave durante un periodo específico.</p> <p>Google Trends es la herramienta de Big Data que se utilizará en el estudio, por lo que se consideró importante conocer la perspectiva de los entrevistados respecto a su conocimiento sobre la herramienta; como su aplicación beneficiaría al marketing online en un destino turístico peruano y a Kuélap; cuales serían las desventajas para ambos casos y su percepción respecto a adicionar variables de Google Trends (web, YouTube, imágenes y noticias) a un modelo predictivo. Asimismo, se exploró sobre las ventajas y desventajas de estos indicadores para las estrategias de marketing online en un atractivo y los beneficios y desventajas que aporta a un modelo predictivo.</p>
---------------	--

En cuanto a la etapa cuantitativa, se tuvo en consideración las variables presentadas en la Tabla 02; las mismas que han sido aplicadas para los nueve modelos de análisis.

Tabla 02.

*Variables Análisis Cuantitativo*

Nombre de la Variable	Descripción
Flujo Mensual de Visitantes que Arriban al Complejo Arqueológico de Kuélap	Esta variable incluye el total de turistas y excursionistas nacionales y extranjeros que visitaron Kuélap. Es por ello que se generaliza y sólo se hará uso del indicador asociado a las visitas en general.
LogX(t), LogX(t-12) y LogX(t-1)	Los datos de la variable provienen de la entidad a cargo de la administración del complejo: el Ministerio de Cultura – Dirección Desconcertada de Cultura, Amazonas, y es publicada por el MINCETUR. Hace inclusión tanto de visitantes nacionales como extranjeros que han arribado al atractivo entre los años 2011 al 2016 y se distingue entre las cantidades correspondientes a cada grupo (total, nacionales y extranjeros). Es necesario mencionar que se añadieron las llegadas de diciembre 2010 (ver anexo 1) debido a que el modelo requiere los datos del mes inmediato anterior.
Donde:	Se utiliza el plazo de un año en base a la teoría de los modelos de serie de tiempo, donde se asume que los patrones del pasado son la clave para predecir el futuro; suponiendo que estos pueden repetirse (Frechtling, 2001). Asimismo, la data utilizada no registra una estacionalidad mensual definida; por tanto, se aconseja observar lo registrado durante todos los meses. En relación al plazo de un mes, es necesario conocer como se está moviendo el mercado respecto a la planificación de viaje.
Índice Mensual de Búsquedas Web de Google Trends referido a Kuélap	Se obtiene de una fuente secundaria (Google). Se ingresa a la dirección <a href="https://www.google.com/trends/">https://www.google.com/trends/</a> y se procede a digitar los términos que se desea buscar. Para el caso de Kuélap, Google ya tiene predeterminado el término global. Los términos clave a ingresar fueron “kuelap”, “kuelap peru”, “kuelap amazonas” “kuelap tours” y “fortaleza de kuelap” (ver anexo 2), todos los términos son asociados por la misma herramienta Google Trends al término clave Kuélap y fueron seleccionados a criterio en base a su índice de búsquedas. También se consideró pertinente seleccionar términos compuestos en base a lo planteado por Park, Lee y Song (2015), quienes consideraron que la combinación de palabras clave minimiza errores en los resultados. La base de datos de estos
G(t) / Gweb(t)	

	<p>indicadores está disponible por semanas o por meses, dependiendo del periodo indicado (ver anexo 3). Luego, en el caso de indicadores semanales, se agrupan en meses y así se obtienen los datos mensuales (ver anexo 4).</p> <p>El aplicativo permite seleccionar si se desea delimitar las búsquedas en un espacio geográfico determinado. Resulta posible obtener el interés global y del Perú hacia Kuélap. No obstante, el buscador señala que en el extranjero no se ha realizado un volumen significativo de búsquedas bajo el término “kuelap” (ver anexo 5). Esto no significa que absolutamente ninguna persona haya realizado búsqueda alguna acerca del atractivo, sino que en otros países las búsquedas no alcanzan una masa crítica necesaria para generar un indicador (Baggio, 2016). Otra posible explicación es que los turistas extranjeros realicen las búsquedas sobre el atractivo una vez se encuentren en territorio peruano.</p>
<p>Índice Mensual de Búsquedas de imágenes de Google Trends referido a Kuélap.</p> <p>Gimagen(t)</p>	<p>La percepción de un destino ya no sólo se relaciona con la información, sino por medio de imágenes. Muchas agencias de viajes utilizan este medio para promocionar los destinos que ofrecen. La imagen de un destino de manera visual es de vital importancia para la toma de decisiones del turista, ya que refleja lo que se espera ver del destino (Fayeke &amp; Crompton, 1991).</p> <p>Google Trends brinda al usuario la posibilidad de elegir datos desde diferentes tipos de búsquedas, uno de ellos es el buscador de Google Imágenes. Para obtener esta información, se ingresa a la dirección <a href="https://www.google.com/trends/">https://www.google.com/trends/</a>, se redacta el término, por ejemplo, “kuelap”, para finalmente filtrar por búsqueda de imágenes (ver anexo 6). Esta variable posee el mismo tratamiento de datos que la del Índice de Búsquedas Web de Google Trends.</p>
<p>Índice Mensual de Búsquedas en YouTube de Google Trends referido a Kuélap.</p> <p>GYoutube(t)</p>	<p>Los índices de búsquedas en YouTube representan otra variable obtenida desde Google Trends, que es aplicable al modelo. Cabe mencionar que los videos juegan un rol muy importante en la fase de planificación; 64% de las personas que ven videos relacionados a viajes están pensando en realizar uno (Google, 2017).</p> <p>Esta variable tuvo el mismo tratamiento para extraer índices como las búsquedas web y de imágenes. No obstante, se filtró la información seleccionando la opción Búsquedas en YouTube (ver anexo 7).</p>
<p>Índice Mensual de Búsquedas en Noticias de Google Trends referido a Kuélap.</p> <p>GNoticias(t)</p>	<p>Los índices de búsquedas de noticias son otra variable obtenida desde Google Trends para el modelo. Se ha tomado en cuenta esta variable, debido a la coyuntura digital que ha tenido la noticia de la construcción y puesta en marcha de las telecabinas de Kuélap, y su posible influencia al turismo de la región.</p> <p>Esta variable tuvo el mismo tratamiento de datos que las tres anteriores. Los datos fueron filtrados seleccionando la opción búsquedas de noticias (ver anexo 8).</p>
<p>Término independiente</p> <p>b1</p>	<p>Es el término que no está afecto a ninguna variable y que pertenece a la ecuación de regresión como parte del ajuste.</p>
<p>Coeficientes</p> <p>b2, b3, c1, c2 y c3</p>	<p>En las variaciones por unidad que se produzcan en <math>\text{LogX}(t-1)</math>, <math>\text{LogX}(t-12)</math>, <math>G_{\text{web}}(t)</math>, <math>G_{\text{imagen}}(t)</math>, <math>G_{\text{youtube}}(t)</math>, <math>G_{\text{noticias}}(t)</math>; la variable <math>\text{LogX}(t)</math> registrará un cambio equivalente en los coeficientes <math>b_2, b_3, b_4, c_1, c_2, c_3</math> y <math>c_4</math>, que afectan a cada variable independiente (ver anexo 20).</p>

## Muestra.

La muestra para la etapa cualitativa de esta investigación fue intencional y estuvo conformada por cinco unidades de análisis. Los participantes fueron conocedores de marketing turístico, prospectiva turística y Big Data de entidades públicas y privadas del sector turismo. Cada participante accedió de manera voluntaria a la realización de las entrevistas y al uso de sus datos para fines del estudio. Los participantes dieron su autorización para mencionar sus datos en este estudio; los mismos que se presentan a continuación:

Tabla 03.

*Información de los entrevistados*

Nombre del Entrevistado	Cargo	Entidad	Especialidad
Andrea Martínez Bertramini	Consultora en Turismo, Mercadeo y Gestión Pública	Consultora KM CERO	Marketing Turístico
Walter Vizarreta	Subdirector de Inteligencia y Prospectiva Turística	PROMPERÚ	Prospectiva Turística
Guillermo Graglia	Director	Graglia Consulting	Marketing
Renato Peña	Director/Fundador	Brandia	Big Data
Felipe Arentsen	Especialista Técnico de Big Data del Perú	IBM	Big Data

En cuanto a la etapa cuantitativa, no se cuenta con una muestra concreta. Los datos se obtuvieron a partir de Google Trends tomando en cuenta el tiempo (2011 – 2016) y los parámetros de búsqueda (región, categoría, palabras clave).

## Instrumento de Investigación.

En la etapa cualitativa, se realizaron cinco entrevistas a profundidad de carácter semi estructuradas a expertos en materia de marketing turístico, prospectiva turística y Big Data. Se tuvo una muestra selectiva utilizada para complementar la información recabada en el marco teórico y obtener una mejor comprensión del objeto de estudio.

Las entrevistas sirvieron para explorar en profundidad el punto de vista de los expertos sobre la viabilidad del uso de variables electrónicas en el análisis predictivo y que, a su vez, sirvan para las estrategias de marketing online de un atractivo turístico peruano; específicamente el Complejo Arqueológico de Kuélap.

Para este estudio, se elaboró una guía de pautas dividida en tres categorías seleccionadas de acuerdo al marco teórico: Big Data, proyección turística y Google Trends; contando con un total de 21 preguntas. Se buscó realizar una entrevista semi estructurada de manera que se tuviera libertad para introducir preguntas adicionales, no

contempladas en un inicio, para precisar y obtener mayor información. Asimismo, se preparó un formato de consentimiento informado (Anexo 9) para comunicar al entrevistado sobre el propósito y plan del estudio, la confidencialidad y sus derechos como participante. Este formato de consentimiento informado fue adaptado de un estudio de la Universidad de Georgetown en Estados Unidos (Georgetown University, s.f). La validación de ambos instrumentos se realizó con Jaime Briceño, Coordinador Académico e Investigador de Mercados de la Carrera de Marketing de la Universidad San Ignacio de Loyola, el día 15 de julio del 2017. En esta validación se encontró que eran necesarios algunos ajustes semánticos en ambos instrumentos, de manera que sean entendibles y puedan ser utilizados en las entrevistas. Las fichas técnicas del instrumento para la etapa cualitativa son el formato de consentimiento informado y la guía de pautas (ver anexos 9 y 10 respectivamente).

Respecto a la etapa cuantitativa, se usó *Microsoft Excel* y *SPSS Statistics* para poder trabajar con la data obtenida de Google Trends y evaluar la correlación de los indicadores para los modelos. Microsoft Excel es un programa informático desarrollado por Microsoft desde su primera versión en 1985 hasta la última diseñada en el 2016. Contiene una serie de paquetes estadísticos disponibles bajo el nombre de “herramienta para análisis”. Estas aplicaciones permiten realizar análisis de datos complejos (Microsoft, 2016). Para el presente estudio, este programa se utilizó para ordenar y almacenar los datos históricos y electrónicos para luego transformar el total de arribos mensuales en logaritmos. Para este proceso se hizo uso de una plantilla elaborada por las investigadoras (ver anexo 11). Por su parte, SPSS Statistics sirvió para realizar la función de regresión lineal y evaluar el R-cuadrado, R-cuadrado ajustado, nivel de significancia y distribución *t-student* de todos los posibles modelos.

### **Procedimiento para Recolección de Datos.**

Para la primera parte de la etapa cualitativa, se seleccionaron a los posibles entrevistados y se les enviaron cartas vía correo electrónico para solicitar una entrevista a nombre de la Universidad San Ignacio de Loyola (ver anexo 12). Una vez cada entrevistado aceptó, se realizaron las coordinaciones para determinar el día y el lugar de la entrevista con cada uno.

Después de las coordinaciones respectivas con cada entrevistado, se logró entrevistar a una muestra de cinco expertos: Andrea Martínez y Guillermo Graglia, expertos en marketing; Walter Vizarreta, experto en prospectiva turística; y Renato Peña y Felipe Arentsen, expertos en Big Data.

Antes de iniciar la entrevista, se presentó el formato de consentimiento informado a cada entrevistado, comentándole sobre el propósito del estudio y porque se le seleccionó para participar del mismo. Después de informarles los puntos en el formato de consentimiento informado, los entrevistados marcaron “SÍ” en los casilleros referentes a que: (i) entendían toda la información del formato; (ii) estaban voluntariamente de acuerdo en participar de la investigación; (iii) estaban de acuerdo en que se use su nombre como aporte al estudio; y, (iv) estaban de acuerdo en ser grabados (audio). Luego de ello, firmaron el formato y se iniciaron las entrevistas y grabación de las mismas.

Cabe mencionar, que solamente el experto en prospectiva turística, Walter Vizarreta, no permitió que se graben las primeras secciones sobre Big Data y proyección turística. El entrevistado solicitó que la primera parte no fuera grabada debido a que dio detalles de su trabajo en estos temas, y la política de confidencialidad del estado no permite tener registro de ello. Sin embargo, accedió a que se apunten detalles y que esta información sea utilizada con fines académicos. Algunas de las preguntas contempladas en la guía de pautas fueron omitidas cuando los entrevistados las respondían en alguna pregunta anterior, de forma que no se redunde.

Las entrevistas fueron aplicadas entre el 27 de julio y 18 de agosto del 2017 en la locación previamente pactada con cada uno y teniendo una duración entre 25 a 55 minutos dependiendo del entrevistado. La duración de las entrevistas fue corta ya que se tuvieron que realizar preguntas concretas y directas para responder el objetivo de la investigación cualitativa, ello, debido a que los expertos no contaban con mucho tiempo disponible.

A continuación se presenta la información de ejecución de cada entrevista según la fecha en que fueron realizadas.

Tabla 04.

*Información de ejecución de entrevistas a profundidad*

<b>Nombre del Entrevistado</b>	<b>Cargo</b>	<b>Entidad</b>	<b>Fecha de la Entrevista</b>	<b>Lugar de la Entrevista</b>	<b>Duración de la Entrevista</b>
Andrea Martínez Bertramini	Consultora en Turismo, Mercadeo y Gestión Pública	Consultora KM CERO	27/07/2017	Cafetería Don Mamino	55 minutos
Walter Vizarreta	Subdirector de Inteligencia y Prospectiva Turística	PROMPERÚ	01/08/2017	Oficina en PROMPERÚ	40 minutos
Guillermo Graglia	Director	Graglia Consulting	08/08/2017	Starbucks La Fontana	30 minutos
Renato Peña	Director/Fundador	Brandia	10/08/2017	Pastelería María Almenara	25 minutos
Felipe Arentsen	Especialista Técnico de Big Data del Perú	IBM	18/08/2017	Starbucks Parque Sur	35 minutos

Posterior a la realización de las cinco entrevistas, se procedió a transcribir una a una sin omitir palabras, de forma que puedan ser analizadas con los programas Voyant Tools y TextRazor. Se analizaron las entrevistas en español con el programa TextRazor; no obstante, para su correcto análisis en Voyant Tools, se tuvieron que traducir al inglés debido a que la herramienta trabaja mejor en su idioma nativo (Voyant Tools, s.f.), pudiendo hacer uso de su aplicativo interno de auto limpieza de conectores, preposiciones, etc.

En cuanto a la etapa cuantitativa, las variables ya descritas provienen de fuentes secundarias. Las llegadas son provistas por el Ministerio de Cultura – Dirección Desconcertada de Cultura, Amazonas, y se distingue a visitantes nacionales de extranjeros. En el caso de los datos provistos por Google, se usaron como base las búsquedas a nivel global (ver anexo 13), las cuales incluyen búsquedas en territorio nacional. Cabe mencionar que con la observación y uso de la herramienta, se ha detectado que la mayoría de estas búsquedas globales se realizan en Perú. El previo testeado de la herramienta ha facilitado la definición de las palabras clave a utilizar, seleccionando cinco términos que tienen un significativo índice de búsquedas: “kuelap”, “kuelap peru”, “kuelap amazonas” “kuelap tours” y “fortaleza de kuelap”. Se optó por esos términos porque la herramienta Google Trends los sugería como palabras asociadas, además de seguir lo hallado por Park, Lee, y Song (2015) referente a la combinación de palabras clave para minimizar errores.

Es necesario mencionar la depuración de los datos. Esta fue posible gracias a las funciones especiales del buscador. El primer filtro del buscador es la “función autocompletar”; la cual permite relacionar el término de búsqueda a algún significado. Por ejemplo, la palabra clave “kuelap” está referida a “sitio arqueológico en Perú” (ver anexo 14), por lo que el primer filtro sería de carácter semántico. Además, para el caso de la palabra “kuelap”, la aplicación filtra los caracteres especiales como la tilde, omitiéndolas automáticamente en las búsquedas.

Luego, se delimitó el periodo de tiempo desde enero del 2011 hasta diciembre del 2016, usando el filtro de intervalo de tiempo. Se consideraron seis años porque dicha cantidad de tiempo permite tener una base de datos robusta y con mayores fluctuaciones, además que coincide con el mismo periodo de tiempo extraído para los arribos a Kuélap. Posteriormente, se filtró el interés hacia el Complejo Arqueológico de Kuélap como destino de viaje y no con otro fin tal como buscar reportes académicos del complejo (ver anexo 15). Este filtro permitió incrementar la precisión de los resultados (Gawlik, Kabaria, & Kaur, 2011). Una vez depurados los datos, se obtuvieron los índices de búsquedas globales más óptimos en una gráfica (ver anexo 16).

Para aplicar el modelo se descargó la versión CSV por cada tipo de búsqueda: web, imágenes, YouTube y noticias, las cuales arrojan una hoja de cálculo en Excel con toda la data que se ha filtrado (ver anexo 17).

Toda la data obtenida fue ordenada y almacenada en la plantilla de Excel mencionada en la sección Instrumentos de Investigación, calculándose el logaritmo de los arribos mensuales de turistas para el periodo 2011-2016 y el mes inmediato anterior (diciembre 2010). Luego de este proceso, se exportó la plantilla al programa SPSS, con el cual se halló la función de correlación y evaluó los diferentes coeficientes para las variables de los modelos: AR-1, Choi y Varian (2009) y las siete posibles combinaciones con el modelo de Choi y Varian (2009).

### **Plan de Análisis**

Se transcribieron las entrevistas aplicadas durante el periodo del 27 de julio al 18 de agosto 2017. El análisis de la etapa cualitativa fue de carácter descriptivo porque buscó darle un significado o sentido a la información obtenida a partir de la recolección de los datos. Para el hallazgo de los resultados se hizo uso de los programas Voyant Tools y TextRazor, aplicándose y analizando ambas herramientas entre el 09 y 24 de setiembre de 2017.

La herramienta Voyant Tools es una web basada en lectura de texto y análisis de contexto. Está diseñada para hacer más sencillo el trabajo de colección de texto para una variedad de formatos HTML, XML, PDF, RTF y MS Word (Voyant Tools, s.f.). Este programa permitió encontrar relaciones entre palabras comunes dentro de las entrevistas hechas y armar un cuadro de doble entrada para su mejor análisis.

Para utilizar Voyant Tools, se agruparon todas las entrevistas traducidas en un documento que fue subido al programa. Voyant Tools mostró un mapa de palabras clave que asocio distintos términos incluidos en las entrevistas, comenzando por los términos clave de los cuales se desligan términos asociados. Se debe mencionar que se dejaron de lado los conectores repetidos en las entrevistas y solamente se tomaron en cuenta las palabras que mantenían sentido para el estudio. Los términos clave hallados y sus asociados obtenidos por la herramienta Voyant Tools, se seleccionaron en base a las variables y conceptos clave del estudio que se tocaron en las entrevistas: Big Data, Google Trends, Google Búsquedas, Google Imágenes, YouTube, Google Noticias, Modelo Predictivo y Marketing Online; los mismos que fueron colocados en un cuadro de doble entrada junto a sus términos asociados para su mejor análisis.

Por su parte, TextRazor es una herramienta online que identifica algorítmicamente palabras potenciales en base al contexto de la oración para evitar diferencias semánticas. Es capaz de procesar miles de palabras por segundo utilizando técnicas avanzadas de procesamiento de lenguaje e inteligencia artificial para analizar gramaticalmente y extraer información semántica del contenido original (TextRazor, s.f.). Esta herramienta se usó para extraer las citas textuales más resaltantes según el algoritmo del programa, haciendo más objetiva la presentación de resultados. Cabe mencionar que, tomando en cuenta las citas textuales del programa, se seleccionaron aquellas ideas, que a criterio, representan un mayor aporte a la investigación.

En el programa TextRazor se cargaron una a una cada entrevista a profundidad; de forma que la herramienta arroje las citas textuales más relevantes encontradas haciendo uso de su algoritmo. Las investigadoras revisaron las ideas con la mayor cantidad de palabras relevantes en cada entrevista, y se seleccionó a criterio aquellas que contienen mayores aportes para la investigación. Estas citas textuales fueron analizadas y presentadas mediante su transcripción y análisis; dividiéndolas en las tres categorías enmarcadas en la entrevista a profundidad, las mismas que engloban los principales conceptos abordados en el marco teórico: Big Data, proyección turística y Google Trends.

Cabe mencionar que para disminuir sesgos, se validó a criterio de las investigadoras la consistencia y relevancia de la información obtenida de los softwares en base a los conceptos que guardan relación con el foco del estudio (big data, análisis predictivo y marketing turístico), que responden el problema de investigación y enriquecen la discusión. Esto se hizo en base al análisis de los resultados obtenidos.

Para la etapa cuantitativa, se descargaron los datos en Excel con los indicadores obtenidos por cada término clave: “kuelap”, “kuelap peru”, “kuelap amazonas” “kuelap tours” y “fortaleza de kuelap”, para los periodos entre enero 2011 y diciembre del 2016. Estos resultados se agruparon en un solo documento dividiéndose por periodo y los cinco términos clave. Posterior a ello, se realizó la sumatoria de los datos por cada mes para obtener un indicador total de las búsquedas relacionadas al atractivo por motivos turísticos, y de esta forma ordenar la información obtenida. Este proceso se realizó para todo el periodo de estudio y por cada tipo de búsqueda: web, imágenes, YouTube y noticias (ver anexo 18). Se utilizó el total por cada tipo de búsqueda y se colocaron los resultados junto a los logaritmos de los arribos en la plantilla armada (ver anexo 19).

La plantilla se exportó al programa SPSS, el cual, presenta funciones que permiten analizar datos estadísticos de principio a fin. Es un programa de fácil uso que facilita trabajar con grandes bases de datos, por lo que cualquier investigador puede acceder a usarlo. El programa estadístico SPSS fue creado en 1968, y desde entonces ha contado hasta la fecha con 24 versiones; pasando en su versión 17.0.3, en el 2009, a ser parte de la compañía IBM (IBM, 2017). Con este programa se halló la función de correlación y analizaron los indicadores que observaron Choi y Varian (2009): el R-cuadrado y el valor-P o nivel de significancia. Asimismo, se evaluó también el R-cuadrado ajustado y la distribución t-student en el modelo propuesto en el estudio; ello, con el objetivo de evaluar si la inclusión de las variables mejoran el modelo. Cabe mencionar que el R-cuadrado incrementa en función a la adición de una variable al modelo, mientras que el R-cuadrado ajustado corrige alguna estimación excesiva, adecuándose a la cantidad de variables que se han considerado e incrementándose solamente si lo adicionado mejora la precisión del modelo (Minitab, 2013).

De acuerdo a la teoría estadística, el valor de R-cuadrado (coeficiente de determinación) está asociado a un valor R, llamado coeficiente de correlación, elevado al exponente dos. La Facultad de Ciencias de la Universidad Tecnológica de Pereira en sus investigaciones considera a 0.8 como un valor de R muy alto; en consecuencia, un R-cuadrado mínimo esperable sería de  $0.8^2$  o 0.64. Cabe mencionar, que este R-cuadrado indica cuánta variación en una variable se puede explicar por la variación

de otra (Universidad Tecnológica de Pereira, s.f). Dicho de otra forma, el R-cuadrado permite conocer el grado de relación, reflejado en un índice de ajuste, entre una variable dependiente y las variables independientes en un modelo; siendo ésta, una medida de la precisión general del modelo.

En un caso extremo, podría darse la situación de que R tenga un valor negativo, lo cual significaría una relación inversa entre las variables (Schneider, Hommel, & Blettner, 2010). La relación directa o inversa se hace evidente en los signos de los coeficientes de las variables: número de arribos mensuales y los indicadores provistos por Google Trends: búsquedas web, imágenes, YouTube y noticias.

El R-cuadrado ajustado es igual o menor que el R-cuadrado; posicionándose entre los valores de 1, cuando el modelo predice perfectamente los valores, y 0, cuando el modelo no tiene ningún valor predictivo. Si el valor se sitúa más cerca a 1, mayor será el ajuste del modelo (IBM, s.f.).

Con respecto al valor-P (valor de probabilidad), conocido también como nivel de significancia; debe asociarse a un valor alfa. El valor de alfa puede ir desde 0,01 hasta 1; pero mientras más cerca a 1 esté, se corre mayor riesgo a equivocación. Por ejemplo, si se acepta un valor alfa de 0.1 significaría que en la prueba matemática hay una probabilidad de equivocación de 10% (Guajarati, 2006). Es por ello que Fitcher (1980) asocia el valor-P a 0.05 para sus pruebas matemáticas, considerándolo el valor apropiado para el límite de significancia (Gómez, s.f). Cabe destacar que el indicador de 0.05 fue el empleado por Choi y Varian (2009).

También, se considera la prueba del estadístico-t, que permite analizar y comparar dos medias para conocer en qué medida son significativamente diferentes. Esta prueba grafica una curva donde, si sus resultados se encuentran dentro de la región de aceptación, se aprueba la hipótesis nula (Veal, 2006); que para esta investigación representaría que las variables involucradas no son significativas para el modelo.

Teniendo los datos ordenados, a mediados del mes de octubre de 2017, se procedió a correlacionar las variables de los nueve modelos y evaluar los indicadores: AR-1, Choi y Varian (2009) y las siete combinaciones posibles que involucran al modelo de Choi y Varian, sumándole tres indicadores obtenidos desde Google Trends. Con ello se quiso evaluar si la inclusión de la variable búsquedas web de Google Trends ha logrado mejorar el valor del coeficiente de determinación (R-cuadrado) y el R-cuadrado ajustado del modelo AR-1; y si la adición de las variables google imágenes, google YouTube y google noticias mejora aun más el valor de los indicadores (Ver anexo 20).

## Limitaciones

Se debe tener en cuenta que a pesar de aplicarse una metodología validada previamente (Choi y Varian en 2009; Onder y Gunter en 2016 y; Park, Lee y Song en 2015), los autores de las referencias han utilizado el modelo para evaluar ciudades, y la presente investigación pretendió aplicar la metodología en un atractivo, agregando además, nuevas variables no examinadas por otros autores (búsquedas de imágenes, búsquedas de noticias y búsquedas en YouTube). Por tanto, los resultados pueden reflejar valores negativos en algunas de las evaluaciones de la prueba matemática; no obstante, ello no significa que se está dejando de realizar un aporte al conocimiento. Las pruebas negativas también aportan valor al conocimiento, logrando evidenciar deficiencias en alguno de los campos evaluados: web, multimedia (imágenes o YouTube) o cobertura mediática (noticias).

Los entrevistados consideraron que una limitación de la herramienta Google Trends dentro del campo de estudio del turismo es que la información obtenida es insuficiente para realizar análisis más profundos sobre los turistas potenciales que reflejan interés en un atractivo turístico. Esto se debe a que la herramienta solamente manifiesta el interés, aspectos geográficos y con que otros conceptos se relaciona al atractivo, mas no aspectos demográficos que podrían servir para elaborar un perfil más complejo del usuario que busca Kuélap en Google. Para el caso de la presente investigación, no es necesario obtener información más profunda de Google Trends, pues no se pretendió elaborar un perfil. Los datos obtenidos han sido suficientes para conocer la relación entre las llegadas a Kuélap y las búsquedas en internet sobre el destino.

Los datos proporcionados por Google no representan el número de búsquedas en un periodo de tiempo determinado; es decir, no se cuenta con valores absolutos; lo cual dificulta el pronóstico de arribos. La presente investigación no tiene por objetivo construir un modelo para realizar predicciones; sin embargo, sí busca obtener un mejor indicador ajustado para el flujo mensual gracias a la incorporación de la variable de Google Trends.

Google Trends sólo puede mostrar el volumen de búsquedas, pero no puede identificar si el interés hacia un término en específico es de naturaleza positiva o negativa. Por ejemplo, en los atentados de París ocurridos en el 2015, se intensificaron las consultas web hacia la capital francesa como destino turístico (ver anexo 21); no obstante, ello coincidió con una disminución en los arribos turísticos a la ciudad de París de -2.2% al año siguiente (Dirección General de Empresas del Gobierno Francés, 2017). Por tanto, el motivo por el cual un viajero consulta información en Google acerca de un destino

turístico que se encuentra bajo amenaza de peligro, se debería más al hecho de, por ejemplo, cancelar una reserva o buscar más información al respecto a fin de aminorar riesgos. Por este suceso, se contaminarían los datos provistos por Google Trends y ya no podrían servir para el pronóstico acertado de arribos a París. Si se quisiese realizar, se tendría que esperar otro periodo de cinco años mínimo para tener data histórica, partiendo por ejemplo del 2017, cuando se asume que las búsquedas ya se han normalizado. A la fecha, no han sucedido acontecimientos similares en la región Amazonas o en el Perú que puedan provocar este tipo de fenómeno; al contrario, los arribos por turismo han incrementado en el año 2016, tanto al país (+8%) como a la región (+53.94%).

En cuanto a las palabras clave que se han usado, éstas han derivado de la herramienta Google Trends en relación a aquellas palabras buscadas en Google que son asociadas al término Kuélap. Éstas podrían considerarse arbitrarias debido a que aquellos términos que ahora son relacionados a Kuélap, en unos meses podrían cambiar. No obstante, en el periodo que se realizó la investigación, las palabras seleccionadas fueron las que representaron una mayor relevancia e índice de búsqueda según el algoritmo del programa Google Trends para el término Kuélap, por lo cual, el modelo es considerado válido.

Otras limitaciones pertinentes se han tocado a lo largo del documento. Existe la posibilidad de que un porcentaje de búsquedas no se hayan realizado por viajeros potenciales, sino por estudiantes recopilando información para un trabajo relacionado al turismo. También, se puede haber digitado mal algún término y esto ya no suma a la data de Google Trends para el término de búsqueda. Estas limitaciones, a la fecha, son inevitables.

## CAPÍTULO III: RESULTADOS

### Presentación de Resultados

#### Presentación de Resultados Cualitativos.

El análisis de las entrevistas se realizó haciendo uso de las dos herramientas explicadas en el Capítulo II: Método; siendo éstas Voyant Tools y TextRazor.

El objetivo de la etapa cualitativa es dar solidez al tema de investigación, siendo este novedoso y poco estudiado. Asimismo, permitió dar mayor consistencia al planteamiento del estudio, así como evaluar la viabilidad de las variables electrónicas para un modelo predictivo y las estrategias de marketing online en un atractivo turístico. Para su mejor análisis, los resultados obtenidos de Voyant Tools, se presentan un cuadro de doble entrada, mostrando los términos clave y sus respectivas palabras relacionadas. Por su parte, la información recabada de TextRazor, se está presentando en las tres categorías que dividieron la entrevista a profundidad, las mismas que engloban los conceptos abordados en el marco teórico: Big Data, Proyección Turística y Google Trends.

A continuación se presentan los resultados obtenidos a través de Voyant Tools.

Tabla 05.

*Términos Asociados a los Términos Clave de las Entrevistas*

<b>Término Clave</b>	<b>Términos Asociados</b>
Big Data	Novedoso, útil, complejo, grande, costoso, tecnológico, moderno, entrenamiento, especialización
Google Trends	útil, gratuito, insuficiente, necesario, solución, capacidad, interesante, preciso, herramienta
Google Búsquedas	Excelente, diferente, información
Google Imágenes	sentimiento, complementario
YouTube	filmografía, complementario
Google Noticias	Entorno, positivo
Modelo Predictivo	útil, demanda, rentabilidad, infraestructura, budget, optimización, procesos
Marketing Online	instrumento, indicador, resultado

En la Tabla 05 se aprecian los términos principales relacionados al estudio, los cuales han sido ordenados de modo vertical de acuerdo a su relevancia y de manera horizontal según su nivel de confianza, ello, expresado de forma gráfica en el programa Voyant Tools.

En la tabla, se puede observar que los entrevistados asociaron mayormente el término Big Data a un concepto: “novedoso”, “útil”, “complejo”, “grande”, “costoso”, “tecnológico”, “moderno”, “para entrenamiento” y “especialización”. Todas estas palabras asociadas reflejan lo que comprende la Big Data; no obstante, la herramienta utilizada para este estudio (Google Trends) no presenta costos adicionales y su complejidad de uso y análisis no representa mayores dificultades para los investigadores. Por tanto, significaría una herramienta apropiada para iniciarse en el mundo de la Big Data.

La herramienta principal del estudio (Google Trends), ha sido relacionada a palabras como: “útil”, “gratuito”, “insuficiente”, “necesario”, “solución”, “capacidad”, “interesante”, “preciso” y “herramienta”. Si bien Google Trends es útil, es gratuita a comparación de otras herramientas de Big Data y significaría una solución al objeto de este estudio; la información obtenida aun es insuficiente para realizar análisis más profundos sobre los turistas potenciales que reflejan interés en un atractivo turístico, de manera que se pueda realizar, por ejemplo, un perfil completo de la persona que busca Kuélap.

Las variables a utilizar en el modelo propuesto también han sido asociadas a diversos términos. Google Búsquedas se relacionó a: “excelencia”, “diferente” e “información”. Esta asociación puede reflejar que se trata de una variable diferente a las usuales en los modelos predictivos, que es capaz de brindar un tipo de información asociada a las tendencias actuales; hágase del hecho que ahora la mayoría de búsquedas sobre algún destino turístico inician en la web. Por ejemplo, para el caso de Perú en 2016, el 69% de los turistas extranjeros y 86% de vacacionistas nacionales afirmaron que internet influyó en la elección de su destino (PROMPERÚ, 2017).

La siguiente variable, Google Imágenes, ha sido relacionada a: “sentimiento” y “complementaria”. Los entrevistados la consideraron complementaria debido a que no resultaría una variable determinante, desde su punto de vista, para reflejar el interés hacia un destino; ello a consecuencia de que las imágenes también pueden ser encontradas en los sitios web. No obstante, uno de los entrevistados indicó que la importancia de las imágenes radica más en aspectos relacionados a los sentimientos, es decir, a lo que estas imágenes puedan hacer sentir al visualizarlas, por ejemplo, incentivar a querer viajar al lugar por sus paisajes.

En cuanto a la variable relacionada a YouTube, se le asocia a: “complementaria” y “relacionada a la filmografía”. A pesar de ser considerada una variable complementaria, la mayoría de los entrevistados mencionaban que es una variable interesante a considerar en vista de la influencia que poseen los videos en la decisión de compra de las personas, ello, según los últimos estudios relacionados al turismo. También

añadieron que deberían utilizarse este tipo de variables en estudios relacionados a la filmografía y su influencia para el turismo.

Otro de los conceptos tocados en las entrevistas está relacionado a la variable de Google Noticias del modelo de análisis predictivo. Esta variable ha sido asociada a términos como: “entorno” y “positivo”, sugiriéndose para el modelo, siempre y cuando las búsquedas relacionadas al entorno del atractivo sean por aspectos positivos. Esta sugerencia, resulta una limitación para Google Trends, pues la herramienta no diferencia entre búsquedas relacionadas a aspectos positivos o negativos del atractivo buscado.

El término clave modelo predictivo ha sido relacionado a: “útil”, “demanda”, “rentabilidad”, “infraestructura”, “presupuesto”, “optimización” y “procesos”. Estos modelos podrían ser útiles para reflejar una posible demanda hacia un atractivo turístico, lo cual, serviría para optimizar el presupuesto, los procesos, la infraestructura y la rentabilidad de los involucrados en el sector. También, permiten plantear mejores estrategias, por ejemplo, en el campo del marketing. Sin embargo, es importante tomar estos modelos como herramientas de orientación y no asumir que sus resultados se darán de forma absoluta.

El último concepto abordado ha sido marketing online, el cual ha sido relacionado a: “instrumento”, “indicador” y “resultado”. El marketing online es considerado un instrumento que sirve para tomar decisiones adecuadas en base a información previamente consultada. También, los entrevistados relacionan al concepto con la palabra indicador, esto obtenido a partir de la sección Google Trends de la entrevista, donde señalaron que es una herramienta capaz de validar si las campañas y comunicación están llegando al público objetivo, pudiéndose medir el efecto online a través de las búsquedas. En cuanto al término resultado, este se asoció a marketing online en base a su conexión con la Big Data y la gestión en turismo, refiriéndose a que el resultado de esos campos, trabajando en conjunto, puede beneficiar a próximas campañas online y hacer mejoras en el tipo de comunicación de las mismas; además de redirigir el esfuerzo de manera mucho más eficiente y granular. Asimismo, centrándose en Google Trends como herramienta de Big Data, se asocia al marketing online para la obtención de resultados que permitan realizar un mejor manejo de contenidos para las estrategias SEO y para la mejora de la calidad de información en el ámbito online del atractivo de interés.

La segunda herramienta utilizada para el análisis de la etapa cualitativa fue TextRazor, la cual, permitió obtener los resultados presentados a continuación; los mismos que han sido divididos en las tres secciones abarcadas en las entrevistas a profundidad.

Cabe indicar que para la presentación de resultados en esta sección, se están utilizando los nombres de los entrevistados, quienes en el formato de consentimiento informado aplicado en cada entrevista, aceptaron y firmaron a favor de utilizar sus identidades para la presente investigación.

## **Big Data**

A. Martínez definió a la Big Data como “abundante información, que se ha formulado producto de la interacción de las personas con la internet y el mundo digital; y que se va, almacenando y puede ser usada de distintos modos” (A. Martínez, comunicación personal, 27 de julio, 2017). Esta información debe ser almacenada y gestionada para facilitar el entendimiento y toma de decisiones de una compañía u organización; además, esta información se puede encontrar en una gran variedad de formatos. F. Arentsen refirió que “no solamente es información estructurada, tabulada y bien ordenada, sino que tenemos otro tipo de información como videos, formatos semi estructurados, audios, archivos, Word, PDF, fotos, de todo” (F. Arentsen, comunicación personal, 18 de agosto, 2017). Entonces, la información de la Big Data no solamente se encuentra en abundancia sino también es muy variada.

La tendencia del uso de Big Data se está acercando a todos los sectores económicos ya sea a nivel privado como estatal. R. Peña mencionó que “se ha estado hablando mucho de que la Big Data solamente es para grandes compañías, sin embargo; no es así. Estamos viendo organismos gubernamentales, organismos del estado que están utilizando la Big Data para democratizarla.” (R. Peña, comunicación personal, 10 de agosto, 2017). Esta es una idea que coincide con lo afirmado por Vizarreta durante su entrevista.

En cuanto a la relación de la Big Data con el turismo, A. Martínez señaló que

Hoy en día una de las grandes tendencias es que, cada vez más, el usuario está esperando que nos adelantemos a sus necesidades; ni siquiera que le propongamos cosas, sino que nos adelantemos un poco a cuales van a ser sus deseos y sus necesidades; y la única forma de hacerlo es en realidad obteniendo información de Big Data que me permita ver cuál es su comportamiento pasado, hacia donde está yendo y como yo puedo predecir qué es lo que necesita a través de modelos de programas, digamos en la red, en el mundo digital (A. Martínez, comunicación personal, 27 de julio, 2017).

Asimismo, G. Graglia relacionó el turismo y la Big Data mencionando que

Es una excelente oportunidad para las empresas de tener realmente información muy detallada de sus clientes y poder personalizar ya sea sus productos, servicios en el caso de turismo, la experiencia en sí, es un tema que afecta todas las etapas del consumo, si vamos hablar del turismo, porque transforma básicamente las bases de cómo se hace marketing desde el antes; de poder anticipar preferencias en base a la información que tenemos y pronosticar el comportamiento de la gente lo más probable (G. Graglia, comunicación personal, 08 de agosto, 2017).

Existen diversas dificultades definidas en este estudio; F. Arentsen considera que

Se necesita ser minucioso en el tratamiento de los datos, pues no se puede estar totalmente seguros de que la información, especialmente en redes sociales, sea del todo verídica. Por ejemplo, alguien podría inscribirse en Facebook diciendo que es Juanito Pérez, pero en realidad es Felipe Arentsen. Entonces si esa persona quiere manejar la información de Juanito Pérez, no es 100% confiable (F. Arentsen, comunicación personal, 18 de agosto, 2017).

La presente investigación no hace uso de redes sociales como variables de investigación; y si bien los indicadores de Google Trends tienen limitantes, los índices de búsqueda de un término clave en la herramienta son más confiables debido al algoritmo que utiliza Google para filtrar las búsquedas.

### **Proyección Turística**

La planificación de un destino turístico es una acción que puede tomar como referente el análisis de proyecciones, por ejemplo, de arribos de turistas. Esta acción de proyectar para planificar puede alimentarse de información obtenida a partir de la Big Data; la cual, podría mejorar la precisión del modelo predictivo y generar *insights* más cercanos a la realidad; facilitando así, la toma de decisiones y evitando recaer en decisiones erradas pasadas.

A. Martínez se refirió a lo mencionado indicando que

En realidad, no sólo el flujo sino el movimiento de éstos, es súper importante para un crecimiento organizado, a futuro, que permita que sea sostenible. Estamos hablando de la sostenibilidad de un destino que no se hizo, por ejemplo con Machu Picchu. Se sabía la gente que llegaba y entraba, pero no se hizo nada con esta información; el ritmo al que estaba creciendo; la necesidad de abrir otras rutas de entrada; la misma administración del visitante en el sitio; las puertas de entrada que son súper importantes; si vienen del norte, de Cajamarca, en el caso de Kuélap, si están viniendo por Jaén, por Tarapoto; y que me conviene a mi desde el punto de vista de flujos para tomar una decisión sobre qué aeropuerto abro; qué carretera construyo. De alguna manera orientar una experiencia que sea sostenible en el tiempo y que sea exitosa para que no pase lo de Machu Picchu (A. Martínez, comunicación personal, 27 de julio, 2017).

Por su parte, G. Graglia se refiere a la importancia de estos modelos para anticiparse a diversos acontecimientos partiendo de la base inicial que cuanto más lejos se avance en el horizonte temporal, más inexactos son los modelos. “Son modelos que cuanto más corto es el plazo, más exacto son para llegar al otro extremo. Para mí son importantes; primero, para determinar capacidad de servicios, podría anticiparte a temporadas altas” (G. Graglia, comunicación personal, 08 de agosto, 2017). La importancia de los modelos de *nowcasting* como los evaluados en esta investigación, se centran en proporcionar un pronóstico del ahora. Resultan útiles, por ejemplo, para pronosticar la demanda del sistema de telecabinas de Kuélap en un escenario de muy corto plazo. Dado que la capacidad física del mecanismo no es ilimitada, es mejor conocer con mayor certidumbre los momentos en los que la demanda presenta mayores niveles de fluctuación.

Estos modelos de proyección benefician a la gestión turística para una toma de decisiones más real. F. Arentsen menciona que “se podría hacer un proceso que constantemente esté entregando información en tiempo real para tomar también decisiones en base al tiempo real y lo que está pasando” (F. Arentsen, comunicación personal, 18 de agosto, 2017). Estos modelos necesitan de variables para cumplir su función. R. Peña sugiere que “las variables engloben todos los aspectos posibles, lo

cual depende mucho de la categoría, y para el turismo es importantísimo un análisis con variables casi 360, un análisis de PEST (Político, Económico, Social y Tecnológico).” (R. Peña, comunicación personal, 10 de agosto, 2017). Es así que, además de las variables propuestas en el modelo; se requiere, desde luego, los elementos externos. Elementos como la conflictividad social, el tipo de cambio (en especial para el turismo receptivo) y otros que afectan el flujo de llegadas. Para el caso del presente estudio, se ha utilizado la herramienta Google Trends, que, por una característica intrínseca, está más asociada a la dimensión tecnológica.

Como cualquier tema de discusión, la proyección turística está predispuesta a tener desventajas. R. Peña considera que

Hay una desventaja, porque hay un desconocimiento del tema y hay un esfuerzo que hay que realizar para capacitarte al respecto. No es complicado, pero hay que meter en la cabeza aspectos de estadística disyuntiva y un poco más de análisis; cosa que es muy complicada porque hay personas que realmente nunca han sabido cómo medir sus cosas, y para ellos es un lenguaje totalmente distinto el tener que comenzar a ver y recabar información para tomar decisiones a futuro (R. Peña, comunicación personal, 10 de agosto, 2017).

En este sentido, la Big Data no tiene una desventaja intrínseca, pero sí requiere de competencias estadístico-matemáticas. Una desventaja a considerar es que es complicado transmitir este tipo de conocimientos a terceros no especialistas. Quienes tengan competencias básicas relacionadas a la Big Data deben orientarse a buscar la forma como transmitir sus hallazgos en un entorno donde no todos están acostumbrados a usar métricas, como en el mercado peruano.

Asimismo, es importante resaltar que estos modelos tienen un margen de error y pueden fallar; por lo que deberían ser considerados como una herramienta de orientación. G. Graglia señaló que

Los modelos son modelos al fin. Sí, algunos serán muy exactos o no, pero todos tienen un margen de error, y el principal peligro que tiene planificar en base a modelos es no prever alternativas. Los modelos econométricos, van a proyectar un número y tienen que ser de orientación (G. Graglia, comunicación personal, 08 de agosto, 2017).

## Google Trends

G. Graglia señala que “hoy estamos en un mundo de hipersegmentación, donde estos indicadores de Google te permiten llegar a ese hipersegmento, en un modelo más detallado que antes no se podía. Era muy difícil y ahora se hace relativamente fácil, el tema es saber gestionar esta información” (G. Graglia, comunicación personal, 08 de agosto, 2017). La Big Data en parte está relacionada a la hipersegmentación. Gracias a Google Trends se conoce en qué regiones existe más interés hacia Kuélap; así como que otros destinos están asociados al atractivo. Antes esto resultaba imposible; pero hoy resulta mucho más fácil obtener y administrar dicha información.

Uno de los atributos que más se resalta es que no requiere de un desembolso de dinero. La Big Data, en general, es vista como una solución más económica que el uso de encuestas tradicionales. W. Vizarreta mencionó que “algunos distritos peruanos han recurrido a su trabajo porque quieren usar información de Big Data, ya que no tienen para hacer las encuestas que hace PROMPERÚ” (W. Vizarreta, comunicación personal, 01 de agosto, 2017). Al menos, esto es cierto para el caso de Google Trends, ya que no requiere un costo adicional. Su uso está recomendado para la gestión de sitios turísticos que no cuentan con el capital necesario para costear investigaciones de mercado. No obstante, la comprensión de la Big Data no resulta del todo sencilla, ya que a veces quienes están interesados en aplicarla, no tienen una idea clara de por dónde empezar.

Google Trends permite al usuario obtener información de los índices de búsqueda sobre un término clave; pero también ofrece información de palabras asociadas al término principal. A este contexto se refirió A. Martínez mencionando que

Google Trends también te permite, si estás viendo que te encuentras muy vinculado a Machu Picchu o a algo que es un ícono fuerte, pero que ya no lo quieres desarrollar más; la posibilidad de desarrollar otras formas de introducir un nuevo concepto o direccionar tu contenido a lo que las personas lo están asociando (A. Martínez, comunicación personal, 27 de julio, 2017).

En Google se encuentra prácticamente toda la información de búsquedas que se realizan en internet, pudiéndose encontrar a través de Google Trends, los índices de búsqueda de Kuélap.

F. Arentsen asocia estas búsquedas al interés del usuario y al marketing digital mencionando que

Ahí está toda la información donde están buscando las personas actualmente. Va a tener toda la información de las personas que están mostrando interés en venir. Si la gente está relacionando la búsqueda a un destino de vacaciones: con el marketing digital se puede hacer que, cuando una persona busque algo relacionado a vacaciones, aparezca Kuélap por defecto (F. Arentsen, comunicación personal, 18 de agosto, 2017).

Es útil considerar como variable las búsquedas sobre el término Kuélap debido a que demuestra que ya existe un interés en el atractivo. A. Martínez considera que

Definitivamente si alguien ya busca específicamente un atractivo como ese, es porque ya hay un interés bastante más definido. Si sólo busca Perú o Amazonas, ahí no, porque quizás está todavía en una etapa de exploración. Ya si busca Kuélap, creo que ya su valor como un elemento más para predecir las llegadas es útil para considerar” (A. Martínez, comunicación personal, 27 de julio, 2017).

No obstante, como señala W. Vizarreta, “el interés y la concreción es bastante compleja de correlacionar, pero hay que hacerlo definitivamente” (W. Vizarreta, comunicación personal, 01 de agosto, 2017). Si bien es cierto, los indicadores de Google Trends sirven para correlacionar llegadas, pero todavía resulta insuficiente debido a su propia naturaleza. Los factores externos tales como el clima o estabilidad política afectan al interés potencial hasta tal punto, que puede hacer que un turista desista de viajar al destino por el cual se mostraba interesado. Google Trends como herramienta no es absoluta, pero sirve para una comprensión inicial de un problema.

Es importante mencionar las limitaciones que existen en el uso de esta herramienta. W. Vizarreta menciona que “no todos los que buscan un destino en Google cumplen con la definición de lo que es ser un turista” (W. Vizarreta, comunicación personal, 01 de agosto, 2017). Pese a que, no todos los que usan Google cumplen con la definición de turistas, el algoritmo de Google Trends está diseñado para clasificar una búsqueda bajo la categoría viaje; pero no necesariamente se trata de un potencial viajero, también puede tratarse de un estudiante universitario recopilando información para un trabajo relacionado al turismo. Otro aspecto que puede considerarse una limitación, pero en el

campo de la investigación, tal como indica G. Graglia, es que “la mayoría en nuestro país no tiene capacidades todavía para gestionar esta información, no saben cómo usarla” (G. Graglia, comunicación personal, 08 de agosto, 2017).

Aunque las búsquedas web representan un variable útil para el modelo del estudio, las otras tres variables: Google Imágenes, Google YouTube y Google Noticias, deben ser exploradas en una metodología de prueba y error. En relación a esta acotación, F. Arentsen menciona que

Uno se va a dar cuenta de que variables funcionan cuando se correlacionen. El tema ahora no es una restricción tecnológica porque podemos almacenar y guardar con todo lo que hay disponible. Se debería ir más allá y buscar información más detallada, porque mientras más detalle se tenga, mejor; porque posiblemente la segmentación que hace Google es súper buena para ellos, pero no para tu caso de uso” (F. Arentsen, comunicación personal, 18 de agosto, 2017).

R. Peña señala que “las variables de Google YouTube y noticias son útiles siempre y cuando haya contenido positivo, tanto artículos; cantidad de artículos que tengan esos aspectos, idiomas en distintos portales; nivel alto en videos, archivos, todo influye” (R. Peña, comunicación personal, 10 de agosto, 2017). La posible existencia entre el contenido en YouTube y el contenido en noticias resulta favorable para Kuélap y otros sitios turísticos siempre y cuando, se trate de contenido positivo; aunque es posible inferir que la información con contenido negativo puede afectar a las llegadas, disminuyendo su cantidad. En cuanto a la variable imágenes, R. Peña mencionó que “cada vez que él ha buscado un destino, lo ha buscado porque se ha encontrado una foto y se ha enamorado, por lo que sí se puede llegar a medir el nivel de las búsquedas de imágenes de Kuélap, definitivamente tiene que haber un porcentaje en atribución a ventas” (R. Peña, comunicación personal, 10 de agosto, 2017). De esta manera, se podría observar un primer acercamiento entre el contenido en formato de imágenes y las llegadas a Kuélap.

Los entrevistados consideraron que las variables electrónicas proporcionadas por Google Trends son útiles para el modelo predictivo planteado en el estudio, pero las reforzarían con variables externas al ámbito electrónico. A. Martínez señala que

Debería considerarse el tema de accesibilidad aérea, disponibilidad de asientos; incluso por ejemplo, el paso de frontera de ecuatorianos. En cuanto a redes sociales; para los extranjeros quizás Instagram, porque hay personas apasionadas por los viajes que publican muchas fotos en Instagram para el público de afuera (A. Martínez, comunicación personal, 27 de julio, 2017).

Por su parte, F. Arentsen considera importante el clima y las redes sociales (F. Arentsen, comunicación personal, 18 de agosto, 2017). G. Graglia, opta por la accesibilidad, fortaleza del atractivo, temporadas y momentos de vacaciones, motivación de viaje, seguridad del destino, plazas de aviones y capacidad (G. Graglia, comunicación personal, 08 de agosto, 2017). Finalmente, R. Peña, se inclina por las variables macro como aspectos políticos, económicos, sociales y tecnológicos (R. Peña, comunicación personal, 10 de agosto, 2017).

En líneas generales, los resultados cualitativos reflejan que las variables de Google Trends pueden ser utilizadas para modelar el flujo de llegadas a un atractivo turístico, partiendo de que las mismas son relevantes para el proceso de decisión de viaje; no obstante, se deben realizar las pruebas respectivas para afirmar su utilidad respecto al caso concreto de Kuélap. Asimismo, estos datos bien podrían ser complementados con información obtenida de redes sociales o aspectos externos al entorno digital, como capacidad de aviones, clima, seguridad, entre otros. Es importante también resaltar que los resultados deben ser tomados a manera de orientación y no como valores determinantes; es decir, muestran un camino a seguir, pero no el absoluto.

## Presentación de Resultados Cuantitativos.

Para esta sección, se han evaluado los cuatro coeficientes descritos en el Capítulo II: Método (R-cuadrado, valor-P, R-cuadrado ajustado y distribución t-student).

### Ecuación 01: Modelo AR-1

Las primeras variables consideradas provienen del modelo AR-1:

$$\text{Log}X(t) = b_1 + b_2 * \text{Log}X(t - 12) + b_3 * \text{Log}X(t - 1)$$

Leyenda:

- $\text{Log}X(t)$  es el Logaritmo de los arribos a Kuélap en un mes X
- $\text{Log}Y(t - 12)$  es el Logaritmo de los arribos a Kuélap de hace un año
- $\text{Log}X(t - 1)$  es el Logaritmo de los arribos a Kuélap de hace un mes
- $b_1$  es un término independiente
- $b_2$  y  $b_3$  son coeficientes que afectan a las variables

Correlacionando los tres logaritmos en SPSS; donde la variable dependiente es  $\text{Log}X(t)$  y las variables independientes son  $\text{Log}X(t - 12)$  y  $\text{Log}X(t - 1)$ , se obtuvo lo siguiente:

Tabla 06.

#### *Parámetros de la primera ecuación*

Estadísticas de la regresión	
Coefficiente de correlación múltiple	<b>.838</b>
Coefficiente de determinación R <sup>2</sup>	<b>.702</b>
R <sup>2</sup> ajustado	<b>.694</b>
Observaciones	72

Tabla 07.

#### *Coefficientes y Probabilidad de la primera ecuación*

	Coefficientes	Estadístico t	Probabilidad
CONSTANTE	<b>.144</b>	.545	.588
LogY(t-12)	<b>.703</b>	8.598	<b>.000</b>
LogY(t-1)	<b>.269</b>	3.448	<b>.001</b>

En relación a los resultados, se observa una correlación positiva entre las variables de +0.838 (coeficiente de correlación múltiple), encontrándose cerca de una correlación perfecta de +1.0 y superando el criterio de +0.8 considerado en el Plan de Análisis. El R-cuadrado muestra que el +0.702 de la variabilidad de la variable dependiente queda explicada por su relación con las variables independientes.

De acuerdo al valor-P (valor de probabilidad) o nivel de significancia, las variables asociadas al periodo t-12 y t-1 se deben incluir en el modelo debido a que los resultados obtenidos se encuentran a un nivel  $p < 0.05$  mencionado en el Plan de Análisis.

Adicionalmente, la prueba del estadístico-t para las variables  $\text{LogY}(t-12)$  y  $\text{LogY}(t-1)$ , permitieron encontrar que las puntuaciones de ambas variables se encuentran fuera de la región de aceptación en la gráfica t-student (ver anexo 22, tabla A6), demostrando que el valor del estadístico-t es menor que los valores críticos de la prueba, rechazándose la hipótesis nula.

En este sentido, se puede mencionar que existe relación entre el flujo de visitas turísticas que recibe el Complejo Arqueológico de Kuélap en un mes cualquiera con la cantidad registrada en un año anterior y en un mes anterior.

Teniendo en cuenta que la correlación fue positiva, con la inclusión de los dos indicadores, la ecuación debería quedar de la siguiente manera:

$$\text{LogX}(t) = 0.144 + 0.703 * \text{LogX}(t - 12) + 0.269 * \text{LogX}(t - 1)$$

#### Ecuación 02: Choi y Varian (2009)

El segundo modelo es el elaborado por Choi y Varian (2009). Este modelo agrega la variable  $G(t)$  o índice de búsquedas web, a la ecuación anterior:

$$\text{LogX}(t) = b_1 + b_2 * \text{LogX}(t - 12) + b_3 * \text{LogX}(t - 1) + c * G(t)$$

Leyenda:

- $\text{LogX}(t)$  es el Logaritmo de los arribos a Kuélap en un mes X
- $\text{LogX}(t - 12)$  es el Logaritmo de los arribos a Kuélap de hace un año
- $\text{LogX}(t - 1)$  es el Logaritmo de los arribos a Kuélap de hace un mes.
- $G(t)$  es el índice de búsquedas web sobre Kuélap en un mes X
- $b_1$  es un término independiente
- $b_2, b_3$  y  $c$  son coeficientes que afectan a las variables

Tabla 08.

#### *Parámetros de la segunda ecuación*

Estadísticas de la regresión	
Coefficiente de correlación múltiple	<b>.884</b>
Coefficiente de determinación $R^2$	<b>.782</b>
$R^2$ ajustado	<b>.773</b>
Observaciones	72

Tabla 09.

*Coefficientes y Probabilidad de la segunda ecuación*

	Coeficientes	Estadístico t	Probabilidad
CONSTANTE	<b>.232</b>	1.019	.312
LogY (t-12)	<b>.557</b>	7.315	<b>.000</b>
LogY(t-1)	<b>.309</b>	4.572	<b>.000</b>
G(t)	<b>.004</b>	4.995	<b>.000</b>

Con este modelo, se puede observar una correlación positiva de +0.884, más cercana a la correlación perfecta de +1.0 en comparación con el modelo precedente (AR-1). Asimismo, el R-cuadrado alcanza un valor de +0.782, superando al anterior de +0.702 y al +0.064 considerado aceptable por la Universidad de Pereira en el Plan de Análisis. Se demuestra una correlación más significativa entre las variables y una mejora en la precisión del modelo. Con respecto al valor de R-cuadrado ajustado; el anterior modelo alcanzó un valor de +0.694, mientras el actual registró +0.773. Este resultado en el R-cuadrado ajustado significa que la inclusión de la variable G(t) determina un nivel de correlación de forma consistente y no únicamente por ser una variable adicional.

Los valores del estadístico-t de las tres variables involucradas: LogY(t-12), LogY(t-1) y G(t), se encuentran fuera de la región de aceptación en la gráfica t-student (ver anexo 22, tabla A7), resultando una prueba más para evidenciar que estas variables son estadísticamente significativas para el modelo. Se evidencia que el valor del estadístico-t es menor que los valores críticos de la prueba y se rechaza la hipótesis nula.

Además, el nivel de significancia o valor-P de todas las variables independientes incluidas LogX (t-12), LogX (t-1) y G(t) se encuentran a un nivel  $p < 0.05$ , lo cual significa que las variables son estadísticamente significativas para el modelo. Esta relación muestra que existe correlación entre la cantidad de turistas que llegan al Complejo Arqueológico de Kuélap en un mes cualquiera con la cantidad registrada en un año anterior, en un mes anterior y las cifras provistas de las búsquedas web en Google.

En este sentido, el modelo podría considerar los coeficientes y quedar de la siguiente manera:

$$\text{LogX}(t) = 0.232 + 0.557 * \text{LogX}(t - 12) + 0.309 * \text{LogX}(t - 1) + 0.004 * G(t)$$

Para el presente estudio, se adicionaron otros indicadores electrónicos proporcionados por la herramienta Google Trends: Google Imágenes, Google Youtube y Google Noticias. El objetivo de esta prueba fue evaluar si alguno de éstos indicadores podría mejorar la correlación en la predicción de visitas al Complejo Arqueológico de Kuelap.

En este contexto, se realizaron siete posibles combinaciones para evaluar los indicadores adicionales; encontrándose los siguientes resultados:

Ecuación 03: Choi y Varian (2009) + GImagen (t) + GYouTube(t) + GNoticias(t)

$$LogX(t) = b_1 + b_2 * LogX(t - 12) + b_3 * LogX(t - 1) + c_1 * Gweb(t) + c_2 * Gimagen(t) + c_3 * Gyoutube(t) + c_4 * Gnoticias(t)$$

Leyenda:

- $LogX(t)$  es el Logaritmo de los arribos a Kuélap en un mes X
- $LogX(t - 12)$  es el Logaritmo de los arribos a Kuélap de hace un año
- $LogX(t - 1)$  es el Logaritmo de los arribos a Kuélap de hace un mes
- $Gweb(t)$  es el índice de búsquedas w ebsobre Kuélap en un mes X
- $Gimagen(t)$  es el índice de búsquedas de imágenes sobre Kuélap en un mes X
- $Gyoutube(t)$  es el índice de búsquedas en youtube sobre Kuélap en un mes X
- $Gnoticias(t)$  es el índice de búsquedas de noticias sobre Kuélap en un mes X
- $b_1$  es un término independiente
- $b_2, b_3, c_1, c_2, c_3$  y  $c_4$  son coeficientes que afectan a las variables

Tabla 10.

*Parámetros de la tercera ecuación*

Estadísticas de la regresión	
Coefficiente de correlación múltiple	<b>.891</b>
Coefficiente de determinación R <sup>2</sup>	<b>.795</b>
R <sup>2</sup> ajustado	<b>.776</b>
Observaciones	72

Tabla 11.

*Coefficientes y Probabilidad de la tercera ecuación*

	Coefficientes	Estadístico t	Probabilidad
Intercepción	.333	1.422	.160
LogY(t-12)	.558	7.345	<b>.000</b>
LogY(t-1)	.278	3.997	<b>.000</b>
Gweb(t)	.005	4.978	<b>.000</b>
Gimagen(t)	-.001	-1.899	<b>.062</b>
Gyoutube(t)	.000	.405	<b>.687</b>
Gnoticias(t)	.000	.897	<b>.373</b>

Este modelo, donde se incluyen todas las variables disponibles, muestra una correlación positiva de +0.891, siendo más cercana a una correlación perfecta de +1.0 en comparación con el modelo de Choi y Varian (2009). Asimismo, muestra un valor más alto del R-cuadrado y el R-cuadrado ajustado respecto al modelo anterior.

No obstante, a pesar de determinar un valor más alto en los coeficientes mencionados; lo cual, podría significar que existe una correlación más fuerte entre las variables; el valor de probabilidad (valor-P) de las variables agregadas: GImagen(t), GYoutube(t) y GNoticias(t) evidencia que no existiría significancia para el modelo, y que el incremento de estas variables se ha debido a una casualidad o a la intervención de factores exógenos a la ecuación, tal como menciona Veal (2006). Se entiende ello, en base a las puntuaciones del valor-P para las tres variables adicionales (GImagen[t], GYoutube[t] y GNoticias[t]), los cuales se encuentran a un nivel  $p > 0.05$ .

Respecto a la prueba t-student, los valores del estadístico-t de las variables LogY(t-12), LogY(t-1) y G(t) se encuentran fuera de la región de aceptación en la gráfica, pero los valores del estadístico-t correspondientes a las variables GImagen(t), GYoutube(t) y GNoticias(t) son mayores que los valores críticos de la prueba, encontrándose dentro de la región de aceptación (ver anexo 22, tabla A8), por lo que las tres variables agregadas no tendrían significancia para el modelo. En este sentido, se aconseja el retiro de las variables GImagen(t) GYoutube(t) y GNoticias(t) del modelo.

#### Ecuación 04: Choi y Varian (2009) + GImagen (t) + GYouTube(t)

$$\text{Log}X(t) = b_1 + b_2 * \text{Log}X(t - 12) + b_3 * \text{Log}X(t - 1) + c_1 * Gweb(t) + c_2 * Gimagen(t) + c_3 * Gyoutube(t)$$

Leyenda:

- $\text{Log}X(t)$  es el Logaritmo de los arribos a Kuélap en un mes X
- $\text{Log}X(t - 12)$  es el Logaritmo de los arribos a Kuélap de hace un año
- $\text{Log}X(t - 1)$  es el Logaritmo de los arribos a Kuélap de hace un mes
- $Gweb(t)$  es el índice de búsquedas w eb sobre Kuélap en un mes X
- $Gimagen(t)$  es el índice de búsquedas de imágenes sobre Kuélap en un mes X
- $Gyoutube(t)$  es el índice de búsquedas en youtube sobre Kuélap en un mes X
- $b_1$  es un término independiente
- $b_2, b_3, c_1, c_2$  y  $c_3$  son coeficientes que afectan a las variables

Tabla 12.

*Parámetros de la cuarta ecuación*

Estadísticas de la regresión	
Coefficiente de correlación múltiple	<b>.890</b>
Coefficiente de determinación R <sup>2</sup>	<b>.792</b>
R <sup>2</sup> ajustado	<b>.776</b>
Observaciones	72

Tabla 13.

*Coefficientes y Probabilidad de la cuarta ecuación*

	<i>Coefficientes</i>	<i>Estadístico t</i>	<i>Probabilidad</i>
Intercepción	.304	1.310	.195
LogY(t-12)	.555	7.325	<b>.000</b>
LogY(t-1)	.291	4.267	<b>.000</b>
Gweb(t)	.005	4.945	<b>.000</b>
Gimagen(t)	-.001	-1.764	<b>.082</b>
Gyoutube(t)	8.271E-05	.366	<b>.716</b>

\*8.271E-05 significa:  $8.271 \times 10^{-5}$

La ecuación 04 muestra un coeficiente de correlación múltiple por encima del +0.8 considerado aceptable en el Plan de Análisis. Asimismo, también muestra un valor más alto del coeficiente de correlación R-cuadrado y el R-cuadrado ajustado respecto al modelo de Choi y Varian (2009). No obstante, como en la combinación precedente, este aumento de los coeficientes se podría deber a una casualidad debido a que el valor-P de las variables agregadas: Gimagen(t) y GYoutube(t), no mantendría significancia para el modelo, pues los mismos se encuentran a un nivel  $p > 0.05$ .

Ello también se demuestra a través de los valores obtenidos en el estadístico-t de las variables Gimagen(t) y GYoutube(t) son mayores que los valores críticos de la prueba, encontrándose dentro de la región de aceptación de la prueba t-student, a diferencia de los valores de las variables LogY(t-12), LogY(t-1) y G(t) (ver anexo 22, tabla A9); por lo cual, las variables agregadas no tendrían significancia para el modelo y se aconseja su retiro del mismo.

### Ecuación 05: Choi y Varian (2009) + GImagen + GNoticias(t)

$$\text{Log}X(t) = b_1 + b_2 * \text{Log}X(t - 12) + b_3 * \text{Log}X(t - 1) + c_1 * Gweb(t) + c_2 * Gimagen(t) + c_3 * Gnoticias(t)$$

Leyenda:

- $\text{Log}X(t)$  es el Logaritmo de los arribos a Kuélap en un mes X
- $\text{Log}X(t - 12)$  es el Logaritmo de los arribos a Kuélap de hace un año
- $\text{Log}X(t - 1)$  es el Logaritmo de los arribos a Kuélap de hace un mes
- $Gweb(t)$  es el índice de búsquedas w eb sobre Kuélap en un mes X
- $Gimagen(t)$  es el índice de búsquedas de imágenes sobre Kuélap en un mes X
- $Gnoticias(t)$  es el índice de búsquedas de noticias sobre Kuélap en un mes X
- $b_1$  es un término independiente
- $b_2, b_3, c_1, c_2$  y  $c_3$  son coeficientes que afectan a las variables

Tabla 14.

*Parámetros de la quinta ecuación*

Estadísticas de la regresión	
Coefficiente de correlación múltiple	<b>.891</b>
Coefficiente de determinación R <sup>2</sup>	<b>.794</b>
R <sup>2</sup> ajustado	<b>.779</b>
Observaciones	72

Tabla 15.

*Coefficientes y Probabilidad de la quinta ecuación*

	<i>Coefficientes</i>	<i>Estadístico t</i>	<i>Probabilidad</i>
Intercepción	.318	1.383	.171
LogY(t-12)	.560	7.443	<b>.000</b>
LogY(t-1)	.281	4.095	<b>.000</b>
Gweb(t)	.005	5.028	<b>.000</b>
Gimagen(t)	-.001	-1.868	<b>.066</b>
Gnoticias(t)	.000	.885	<b>.379</b>

En la ecuación 05 también se logra observar el incremento del coeficiente de correlación múltiple, superando el +0.8 considerado aceptable. Además, se visualiza un aumento del R-cuadrado y R-cuadrado ajustado respecto a la ecuación que sólo agrega búsquedas web.

Sin embargo, como se ha podido observar en las anteriores combinaciones, las variables agregadas: GImagen(t) y GNoticias(t) no estarían añadiendo significancia al modelo debido a que el nivel de probabilidad de las mismas se encuentran a un nivel  $p > 0.05$ .

Adicional a ello, otro factor que refuerza la evidencia de que las dos variables adicionadas no presentan significancia para el modelo es que los valores de la gráfica t-student correspondientes a las variables GImagen(t) y GNoticias(t) son mayores que los valores críticos de la prueba, econtrándose dentro de la región de aceptación y aprobándose con ello la hipótesis nula (ver anexo 22, tabla A10), que para el presente estudio indica que las variables consideradas no serían significativas.

#### Ecuación 06: Choi y Varian (2009) + GYouTube(t) + GNoticias(t)

$$\text{Log}X(t) = b_1 + b_2 * \text{Log}X(t - 12) + b_3 * \text{Log}X(t - 1) + c_1 * Gweb(t) + c_2 * Gyoutube(t) + c_3 * Gnoticias(t)$$

Leyenda:

- $\text{Log}X(t)$  es el Logaritmo de los arribos a Kuélap en un mes X
- $\text{Log}X(t - 12)$  es el Logaritmo de los arribos a Kuélap de hace un año
- $\text{Log}X(t - 1)$  es el Logaritmo de los arribos a Kuélap de hace un mes
- $Gweb(t)$  es el índice de búsquedas w eb sobre Kuélap en un mes X
- $Gyoutube(t)$  es el índice de búsquedas en youtube sobre Kuélap en un mes X
- $Gnoticias(t)$  es el índice de búsquedas de noticias sobre Kuélap en un mes X
- $b_1$  es un término independiente
- $b_2, b_3, c_1, c_2$  y  $c_3$  son coeficientes que afectan a las variables

Tabla 16.

#### *Parámetros de la sexta ecuación*

Estadísticas de la regresión	
Coefficiente de correlación múltiple	<b>.885</b>
Coefficiente de determinación R <sup>2</sup>	<b>.783</b>
R <sup>2</sup> ajustado	<b>.777</b>
Observaciones	72

Tabla 17.

#### *Coefficientes y Probabilidad de la sexta ecuación*

	<i>Coefficientes</i>	<i>Estadístico t</i>	<i>Probabilidad</i>
Intercepción	.245	1.047	.299
LogY(t-12)	.559	7.221	<b>.000</b>
LogY(t-1)	.303	4.348	<b>.000</b>
Gweb(t)	.004	4.744	<b>.000</b>
Gyoutube(t)	-5.837E-06	.538	<b>.593</b>
Gnoticias(t)	.000	-.026	<b>.979</b>

\*-5.837E-06 significa:  $-5.837 \times 10^{-6}$

En la ecuación 06 es posible observar el incremento del valor del coeficiente de correlación múltiple a comparación del modelo de Choi y Varian (2009) y el AR-1. Asimismo, se incrementa respecto a los valores del R-cuadrado y el R-cuadrado

ajustado, lo que podría evidenciar que existe correlación en el modelo. No obstante, esta correlación no estaría determinada por las dos variables agregadas: GYoutube(t) y GNoticias(t). Se deduce ello, debido a que los niveles de significancia de ambas variables se encuentran en un nivel  $p > 0.05$ . Se refuerza esta acotación con lo hallado en la prueba t-student (estadístico-t) que evidencia que los valores correspondientes a las variables GYoutube(t) y GNoticias(t) son mayores que los valores críticos de la prueba, encontrándose dentro de la región de aceptación y aprobándose la hipótesis nula (ver anexo 22, tabla A11). De esta manera, se demuestra que las variables adicionadas no son estadísticamente significativas para el modelo.

Ecuación 07: Choi y Varian (2009) + GImagen (t)

$$\text{Log}X(t) = b_1 + b_2 * \text{Log}X(t - 12) + b_3 * \text{Log}X(t - 1) + c_1 * Gweb(t) + c_2 * Gimagen(t)$$

Leyenda:

- $\text{Log}X(t)$  es el Logaritmo de los arribos a Kuélap en un mes X
- $\text{Log}X(t - 12)$  es el Logaritmo de los arribos a Kuélap de hace un año
- $\text{Log}X(t - 1)$  es el Logaritmo de los arribos a Kuélap de hace un mes
- $Gweb(t)$  es el índice de búsquedas w ebsobre Kuélap en un mes X
- $Gimagen(t)$  es el índice de búsquedas de imágenes sobre Kuélap en un mes X
- $b_1$  es un término independiente
- $b_2, b_3, c_1$  y  $c_2$  son coeficientes que afectan a las variables

Tabla 18.

*Parámetros de la séptima ecuación*

Estadísticas de la regresión	
Coefficiente de correlación múltiple	<b>.890</b>
Coefficiente de determinación R <sup>2</sup>	<b>.792</b>
R <sup>2</sup> ajustado	<b>.779</b>
Observaciones	72

Tabla 19.

*Coefficientes y Probabilidad de la séptima ecuación*

	Coefficientes	Estadístico t	Probabilidad
Constante	.290	1.276	.206
LogY(t-12)	.557	7.423	<b>.000</b>
LogY(t-1)	.293	4.359	<b>.000</b>
Gweb(t)	.005	4.996	<b>.000</b>
Gimagen(t)	-.001	-1.737	<b>.087</b>

La ecuación 07 combina el modelo precedente de Choi y Varian (2009) con una variable individual, en este caso, GImagen(t). Esta combinación refleja un coeficiente de

correlación múltiple que supera el +0.8 considerado aceptable en el Plan de Análisis; además de un valor mayor en los otros coeficientes analizados (R-cuadrado y R-cuadrado ajustado). Sin embargo, nuevamente la variable añadida al modelo no guardaría significancia debido al valor obtenido en el nivel de probabilidad, el cual se encuentra en un nivel  $p > 0.05$ .

Además, cabe mencionar que el valor del estadístico-t de esta variable añadida, es mayor que el valor crítico de la prueba, encontrándose dentro de la región de aceptación de la prueba t-student, a diferencia de las tres variables:  $\text{LogY}(t-12)$ ,  $\text{LogY}(t-1)$  y  $G(t)$  (ver anexo 22, tabla A12); y reforzando así, la evidencia de que la variable agregada no tiene significancia para el modelo.

#### Ecuación 08: Choi y Varian (2009) + GYouTube(t)

$$\text{LogX}(t) = b_1 + b_2 * \text{LogX}(t - 12) + b_3 * \text{LogX}(t - 1) + c_1 * Gweb(t) + c_2 * Gyoutube(t)$$

Leyenda:

- $\text{LogX}(t)$  es el Logaritmo de los arribos a Kuélap en un mes X
- $\text{LogX}(t - 12)$  es el Logaritmo de los arribos a Kuélap de hace un año
- $\text{LogX}(t - 1)$  es el Logaritmo de los arribos a Kuélap de hace un mes
- $Gweb(t)$  es el índice de búsquedas w eb sobre Kuélap en un mes X
- $GYoutube(t)$  es el índice de búsquedas de videos en YouTube sobre Kuélap en un mes X
- $b_1$  es un término independiente
- $b_2, b_3, c_1$  y  $c_2$  son coeficientes que afectan a las variables

Tabla 20.

#### Parámetros de la octava ecuación

Estadísticas de la regresión	
Coefficiente de correlación múltiple	<b>.884</b>
Coefficiente de determinación R <sup>2</sup>	<b>.782</b>
R <sup>2</sup> ajustado	<b>.769</b>
Observaciones	72

Tabla 21.

#### Coefficientes y Probabilidad de la octava ecuación

	Coefficientes	Estadístico t	Probabilidad
Constante	.232	1.000	.321
LogY(t-12)	.557	7.243	<b>.000</b>
LogY(t-1)	.309	4.528	<b>.000</b>
Gweb(t)	.004	4.831	<b>.000</b>
Gyoutube(t)	-6.090E-06	-.027	<b>.978</b>

\*-6.090E-06 significa:  $-6.090 \times 10^{-6}$

En esta ecuación, el coeficiente de correlación múltiple y el R-cuadrado se mantienen respecto a la ecuación de Choi y Varian (2009). A pesar de ello, el R-cuadrado ajustado disminuye ligeramente, y si bien esto evidenciaría que la adición de esta variable no determinaría una mejora en el modelo; en un futuro, si el atractivo cuenta con mayor presencia en YouTube, podría haber una mejora en la correlación.

No obstante, lo que resultaría determinante, es que la variable añadida no estaría evidenciando una significancia. Esto se debe a que se encuentra en un nivel  $p > 0.05$ . Cabe mencionar que el valor del estadístico-t de la variable añadida es mayor que el valor crítico de la prueba, encontrándose dentro de la región de aceptación de la prueba t-student (ver anexo 22, tabla A13); y reforzando así, la evidencia de que la variable agregada no tendría significancia para el modelo.

#### Ecuación 09: Choi y Varian (2009) + GNoticias (t)

$$\text{Log}X(t) = b_1 + b_2 * \text{Log}X(t - 12) + b_3 * \text{Log}X(t - 1) + c_1 * Gweb(t) + c_2 * GNoticias(t)$$

Leyenda:

- $\text{Log}X(t)$  es el Logaritmo de los arribos a Kuélap en un mes X
- $\text{Log}X(t - 12)$  es el Logaritmo de los arribos a Kuélap de hace un año
- $\text{Log}X(t - 1)$  es el Logaritmo de los arribos a Kuélap de hace un mes
- $Gweb(t)$  es el índice de búsquedas w e sobre Kuélap en un mes X
- $GNoticias(t)$  es el índice de búsquedas de noticias sobre Kuélap en un mes X
- $b_1$  es un término independiente
- $b_2, b_3, c_1$  y  $c_2$  son coeficientes que afectan a las variables

Tabla 22.

*Parámetros de la novena ecuación*

Estadísticas de la regresión	
Coefficiente de correlación múltiple	<b>.885</b>
Coefficiente de determinación R <sup>2</sup>	<b>.783</b>
R <sup>2</sup> ajustado	<b>.770</b>
Observaciones	72

Tabla 23.

*Coefficientes y Probabilidad de la novena ecuación*

	Coefficientes	Estadístico t	Probabilidad
Constante	.246	1.067	.290
LogY(t-12)	.559	7.293	<b>.000</b>
LogY(t-1)	.303	4.390	<b>.000</b>
Gweb(t)	.004	4.905	<b>.000</b>
Gnoticias(t)	.000	-.542	<b>.590</b>

En la ecuación 09, se aprecia que el coeficiente de correlación múltiple es mayor al +0.8 aceptado en el Plan de Análisis y supera ligeramente al obtenido en la ecuación de Choi y Varian (2009); además, se presenta un ligero incremento del coeficiente R-cuadrado en comparación a la ecuación precedente. Sin embargo, como en el caso de la ecuación 08, el R-cuadrado ajustado disminuye, evidenciando que el indicador se incrementa en base a la adición de variables y no porque esta adición pueda determinar una mejora en el modelo. Se demuestra así, que en estos momentos no existiría correlación entre las variables de la ecuación.

Asimismo, la variable añadida al modelo, no estaría evidenciando una significancia debido a que se encuentra a un nivel  $p > 0.05$ . Esta no significancia, se reafirma con los resultados del estadístico-t de la variable añadida, la cual, es mayor que el valor crítico de la prueba; encontrándose dentro de la región de aceptación de la prueba t-student y ubicándose dentro de la región de aceptación (ver anexo 22, tabla A14).

Para un mejor entendimiento de los resultados cuantitativos del estudio, se ha realizado una tabla resumen. Esta tabla demuestra que en todas las combinaciones propuestas que agregan nuevas variables al modelo de Choi y Varian (2009), para el caso del Complejo Arqueológico de Kuélap, no se tendría mayor significancia en comparación al modelo precedente. Si bien en algunos casos se muestra un mayor nivel de correlación ( $R^2$  y  $R^2$  ajustado), esto se debería a una coincidencia o factores externos (Veal, 2006) ya que el análisis del nivel de significancia en valor-P es mayor a +0.05.

La tabla se muestra en la página siguiente para que se aprecie en su totalidad.

Tabla 24.

*Resumen de los resultados cuantitativos*

Modelo/ indicador	R2	R2 Ajust.	valor-P					
			(t-12)	(t-1)	Gw(t)	Gi(t)	Gy(t)	Gn(t)
Ecuación 01 (AR-1)	.702	.694	.000	.001				
Ecuación 02 (Choi y Varian 2009)	.782	.773	.000	.000	.000			
Ecuación 03	.795	.776	.000	.000	.000	.062	.687	.373
Ecuación 04	.792	.776	.000	.000	.000	.082	.716	
Ecuación 05	.794	.779	.000	.000	.000	.066		.379
Ecuación 06	.783	.777	.000	.000	.000		.593	.979
Ecuación 07	.792	.779	.000	.000	.000	.087		
Ecuación 08	.782	.769	.000	.000	.000		.978	
Ecuación 09	.783	.770	.000	.000	.000			.590

De izquierda a derecha: pruebas matemáticas que se realizaron en la etapa cuantitativa AR-1, modelo de Choi y Varian (2009) y las siete combinaciones planteadas en el estudio. El cuadro compara los indicadores R-cuadrado, R-cuadrado ajustado y valor-P.

La tabla muestra que los coeficientes analizados para la ecuación 01 (AR-1) son positivos respecto a los valores establecidos en la sección Plan de Análisis (R2 y R2 ajustado >0.064 y valor-P <0.05). Asimismo, la ecuación 02 (Choi y Varian), se incrementó respecto a los coeficientes R2 y R2 ajustado de la ecuación precedente, reflejando lo siguiente: el R2 pasó de +0.702 a +0.782; y el R2 ajustado de +0.694 a +0.773. De esta manera, se demuestra que ambos modelos podrían servir para la predicción de llegadas a Kuélap; pero que el modelo de Choi y Varian (2009) presenta un mejor ajuste que el modelo AR-1.

Por otro lado, se evidencia que las siete combinaciones del modelo de Choi y Varian con las variables imágenes, YouTube y noticias, no serían útiles para pronosticar la demanda a Kuélap a comparación de los modelos que emplean sólo datos históricos y búsquedas web. Ello se evidencia, especialmente, con los resultados del valor-P que superan el <0.05 aceptable; lo cual significa que las variables no son representativas para los modelos.

## Discusión

Las pruebas realizadas han permitido determinar que la hipótesis alternativa  $H1_1$  debe ser aprobada. La prueba matemática demostró que el modelo AR-1 es estadísticamente significativo para pronosticar las llegadas a Kuélap en el periodo 2011-2016.

La hipótesis alternativa  $H2_1$  también se aprueba, debido a que la inclusión de la variable electrónica búsquedas web ha mejorado la correlación en comparación con el modelo AR-1, demostrando tener un mejor ajuste y que la data histórica bien puede ser complementada con este indicador electrónico. Con esto, se valida la prueba realizada por Choi y Varian (2009) al igual que lo hicieron Onder y Gunter (2016), Park, Lee y Song (2015), entre otros.

Se rechaza la hipótesis alternativa  $H3_1$  y se aprueba la hipótesis nula  $H3_0$ , porque la probabilidad (valor-p) que involucra a las variables independientes G(Youtube), G(Imágenes) y G(Noticias) se encuentra a un nivel  $p > 0.05$ ; lo cual, significa que las variables deben ser retiradas del modelo; concluyendo que las tres variables añadidas no son estadísticamente significativas para el mismo; aún cuando el coeficiente R-cuadrado supera a la ecuación base de Choi y Varian (2009).

La hipótesis nula  $H4_0$  referente a que la inclusión de la combinación de las variables G(Youtube) y G(Imágenes) no es estadísticamente significativa para la ecuación, también se aprueba. Esto se debe a que los valores de probabilidad (valor-p) de las variables adicionales demuestran poca significancia para el modelo, considerando que se encuentran a un nivel  $p > 0.05$ .

La hipótesis nula  $H5_0$  también debe ser aprobada, ya que la prueba matemática que involucra a las variables independientes G(Imágenes) y G(Noticias) presenta variables que no son estadísticamente significativas debido a la cercanía de su nivel de significancia (valor-p) a +1. Esto a pesar que el valor del coeficiente R-cuadrado supera el valor +0.064 establecido.

De la misma manera, la hipótesis nula  $H6_0$  es aprobada porque la probabilidad de las variables independientes involucradas G(Youtube) y G(Noticias) supera el valor de +0.05, deduciendo que las variables en mención no son estadísticamente significativas para el modelo. Ello, aún cuando el coeficiente R-cuadrado presenta un valor mayor al establecido.

Respecto a los modelos que añaden una a una las variables adicionales obtenidas desde Google Trends. Si bien, en la ecuación 07, el valor R-cuadrado demostró superar el valor +0.064 establecido; la hipótesis nula  $H7_0$ , que agrega la variable independiente

G(Imágenes) al modelo precedente, debe ser aprobada, pues la variable añadida se encuentra a un nivel  $p > 0.05$ ; evidenciando así, que no es estadísticamente significativa para el modelo.

La hipótesis nula  $H_{80}$ , que considera la inclusión de la variable G(Youtube) se aprueba porque la probabilidad que involucra a la variable se encuentra en un nivel  $p > 0.05$ , esto, a pesar de que el R-cuadrado de la ecuación supera el valor establecido en el Plan de Análisis. Se concluye, entonces, que la variable G(Youtube) no es estadísticamente significativa para el modelo.

Finalmente, la hipótesis nula  $H_{90}$  alusiva a que la inclusión de la variable G(Noticias) no es estadísticamente significativa para el modelo en comparación al precedente de Choi y Varian (2009), se aprueba, debido a que el valor de probabilidad de la variable independiente adicionada se encuentra en un nivel  $p > 0.05$ ; además, el coeficiente R-cuadrado ajustado no supera al valor de +0.773 del modelo base.

En este sentido, la prueba matemática demostró que la inclusión de los indicadores adicionales de Google Trends: imágenes, YouTube y noticias no mejoran la precisión del modelo. Y a pesar que, en la mayoría de casos, la adición de estos indicadores incrementa los valores de los coeficientes R-cuadrado y R-cuadrado ajustado; como variables independientes no cuentan con un nivel de significancia en valor  $p < 0.05$ , evidenciándose con ello, como señala Veal (2006), que el aumento en los coeficientes se debe a una casualidad o a factores exógenos y no refleja significancia.

Los hallazgos de la investigación han dejado aportes relevantes en cuanto a la contribución de conocimientos, permitiendo que se enriquezca la discusión con distintos contrastes provenientes de las diferentes fuentes de información del estudio: literatura revisada, metodología cualitativa y metodología cuantitativa.

De esta manera, los resultados cuantitativos y cualitativos del estudio coinciden con lo señalado por Choi y Varian (2009) respecto a que existe relación entre las búsquedas web y las llegadas a un sitio turístico, en este caso, el Complejo Arqueológico de Kuélap. El modelo que adiciona la variable Google Búsquedas mejora la precisión de aquellos modelos que sólo emplean variables históricas, tal como afirman Onder y Gunter (2016). No obstante, no todos los indicadores provenientes de Google Trends que se han adicionado en el estudio resultan pertinentes para el modelamiento de arribos hacia este atractivo.

Los expertos entrevistados indicaron que el uso de variables electrónicas de la Big Data de Google Trends resulta interesante para el modelamiento predictivo de llegadas a un

atractivo turístico, generando resultados empresariales o de gestión positivos. Esto se contradice con los resultados cuantitativos de la investigación que han demostrado que, por lo menos para Kuélap, la adición de estas variables electrónicas, a excepción de las búsquedas web, no mejora el modelo predictivo con un grado de significancia aceptable.

Park, Lee y Song (2015) concluyeron que era necesaria la combinación de términos clave para obtener resultados apropiados. En el caso de esta investigación, se ha requerido más de un término clave. Los términos clave utilizados: “kuelap”, “kuelap peru”, “kuelap amazonas” “kuelap tours” y “fortaleza de kuelap”, se obtuvieron de la lista de términos con mayor índice de búsquedas que están relacionados y asociados a la palabra “kuelap” proporcionada por la herramienta Google Trends. Esto, como menciona la entrevistada A. Martínez, ofrece una visión más clara de aquellos términos que en el campo del marketing digital pueden servir para mejorar estrategias SEO y SEM, de manera que, si se lanzase alguna campaña, se tiene una línea base de términos clave a utilizar para generar tráfico web. En este sentido, Google Trends permite al usuario obtener información de los índices de búsqueda sobre un término clave; pero también ofrece información de palabras asociadas al término principal, pudiéndose usar estos vocablos como nuevas palabras clave para las estrategias SEO; de manera que el usuario pueda encontrar el contenido que la empresa desea mostrar a través de un posicionamiento orgánico en buscadores.

Aun a pesar de lo señalado por Park, Lee & Song (2015), Google Trends muestra que las frases como “vacaciones en Kuélap”, no tienen volúmenes relevantes de búsquedas; caso contrario sucedió con Fotios (2016), quien empleó el término “holidays in Cyprus” y obtuvo registros de Google que pudieron ser correlacionados con las llegadas a Chipre. Esto evidencia que el uso de frases compuestas no es útil para todos los casos.

Con respecto a lo señalado por Saidi (2010) y Choi y Varian (2009), la investigación coincide en que se podría pronosticar llegadas empleando datos como las visitas producidas hace un mes y hace un año. Se concuerda también con Choi y Varian (2009) en que la incorporación de las búsquedas web de Google mejora los resultados de pronósticos respecto al modelo que sólo utiliza datos históricos (AR-1). Sin embargo, como indicó el entrevistado G. Graglia, no se debe tomar estos resultados rígidamente, sino a manera de orientación y utilizarse en conjunto con la compilación de mayor información.

Asimismo, la investigación coincide con lo señalado por Zeynalov (2014), quien indicó que “para encontrar una relación entre búsquedas y llegadas era necesario que las primeras estén vinculadas a la motivación de viajar”. En este caso, las búsquedas de Kuélap han sido filtradas en la categoría viaje. La correlación entre arribos y consultas en Google puede existir debido a que se ha podido depurar la información. Esto concuerda con lo señalado

por la entrevistada A.Martínez; quien considera que las búsquedas web de la palabra específica Kuélap, reflejan un interés definido hacia el atractivo, siendo útil considerarlas para el modelo y la correlación. Además, este indicador electrónico, podría significar un buen aporte para saber si las campañas y comunicación de marketing online están funcionando.

En relación al turismo, Song y Liu (2016) señalan que en la red se puede hallar una gran cantidad de datos disponibles, pero es necesario un análisis profundo para encontrarle sentido. Los mismos autores mencionaron que el análisis de Big Data es capaz de brindar ideas de negocio y de innovación para las organizaciones de turismo y la industria en general, lo que coincide con lo señalado por el entrevistado R. Peña, quien indicó que se está viendo que los actores privados y estatales relacionados al turismo están utilizando la Big Data para hacerla accesible y utilizarla a su favor. Lancis (2014) considera que la Big Data brinda la posibilidad al sector turismo de realizar una mejor gestión mediante volúmenes de datos e información de visitantes, empresas turísticas y administraciones públicas. Para la entrevistada A. Martínez. el análisis de Big Data proporciona información capaz de brindar una visión de lo que está sucediendo con el comportamiento de los turistas y hacia dónde se dirige. No obstante, cabe mencionar que para este estudio, no todos los formatos de datos disponibles en la red mejoran la correlación de los modelos; evidenciándose ello en la no significancia de las variables relacionadas a Google Youtube, Google Noticias y Google Imágenes para el caso particular de Kuélap.

La Big Data no sólo comprende datos estructurados como las búsquedas web; tal como afirma Laney (2001), también existe aquella información que se presenta como datos no estructurados, tales como los comentarios en redes sociales como *Twitter*, *TripAdvisor*, entre otros. Estos datos han sido utilizados en investigaciones como la de Fuchs y Lexhagen (2015), quienes investigaron la segmentación de turistas que arribaron a Suecia haciendo uso del análisis de comentarios en redes sociales. El análisis de comentarios en redes sociales podría beneficiar y mejorar el modelamiento para el caso de Kuélap, pues se puede identificar palabras clave que permitan conocer quienes están interesados en viajar. Tal como señaló el entrevistado F. Arentsen, aun cuando la herramienta de Big Data utilizada en este estudio, Google Trends, no proporcione características propias de los usuarios; sí permite observar lo que los turistas están buscando en internet, siendo esto una señal del interés que tienen hacia un tópico específico, el cual podría tratarse de un destino para vacacionar o un servicio determinado, lo que colaboraría con la toma de decisiones de quienes trabajan en el sector. Por ejemplo, se podrían definir estrategias de marketing de atracción para un mayor acercamiento al público y que los viajeros potenciales decidan por Kuélap como atractivo de visita.

Baggio (2016) considera que las técnicas de análisis de Big Data requieren de un conocimiento más especializado; sin embargo, afirmó que no dominar métodos especializados no debe ser una limitante para que los profesionales o estudiantes en turismo prueben esta metodología. El entrevistado G. Graglia, coincide con ello, señalando que, en el Perú, la mayoría no cuenta aún con capacidades para gestionar y utilizar esta información. Por su parte, el entrevistado R. Peña menciona que administrar datos de Big Data requiere competencias como la estadística y analítica de datos, pero ello no significa que este proceso sea complicado. Google Trends es una herramienta de fácil uso que proporciona información de Big Data, siendo su análisis menos complicado que el de otras herramientas; por lo que podría considerarse que brinda un primer acercamiento al concepto.

La proyección para destinos turísticos, como bien menciona Hart, Rasner y Lukoszová (2012), es un proceso que prevé las condiciones futuras del entorno del mercado y como señala Frechtling (2001), facilita la toma de decisiones, reduciendo los riesgos y brindando una mejor y más clara percepción de escenarios futuros. Ante esto, se puede resaltar lo indicado por el entrevistado F. Arentsen, quien manifestó que sería interesante el poder tener un modelo que otorgue información para tomar decisiones en tiempo real y en relación al entorno. Google Trends ofrece indicadores que pueden ser divididos en meses o semanas, por lo que sus datos podrían ser utilizados para la evaluación de modelos que generen resultados, por ejemplo, de una semana anterior.

Para el entrevistado G. Graglia, la Big Data y los modelos de análisis predictivo anticipan las preferencias de los viajeros en base a información, pudiéndose pronosticar el comportamiento de una manera más precisa y transformando el marketing. Ante ello, se puede resaltar lo estipulado por Stienmetz y Fesenmaier (2015), quienes consideran que se pueden desarrollar nuevas métricas para un destino conceptualizadas en cuatro redes de creación de valor: marketing y promoción, ventas y distribución, diseño de experiencia y coordinación de asociaciones. Entre éstas métricas podría considerarse el modelo de Choi y Varian (2009), como una evidencia de respuesta ante las estrategias de marketing online de un atractivo; reflejando si efectivamente han generado interés en los usuarios. En el caso de Kuélap, la investigación sugiere que la presencia de búsquedas en Google mediante estrategias de SEO o SEM, sería una técnica exitosa para conseguir más turistas, al demostrarse que estas búsquedas son significativas para la predicción.

Frechtling (2001) afirmó que cuando se trata de proyecciones, ningún método es mejor que otro o se destaca por ser más confiable. Esta idea, podría ser complementada con lo señalado por el entrevistado G. Graglia, quien indica que los modelos de predicción tienen

un margen de error; por lo cual, siempre debe haber un plan de contingencia y alternativas. No obstante, sería pertinente indicar que, el estudio a través del análisis de los coeficientes (R-cuadrado, R-cuadrado ajustado, nivel de probabilidad y distribución t-student), ha evidenciado que la adición de la variable Google Búsquedas reduce el error para el caso de Kuélap.

Según Kim y Moosa (2001), para probar un modelo de análisis predictivo, los métodos computacionales son los más precisos e indicados, pues permiten que los resultados sean obtenidos de forma más rápida, fácil y accesible. Esto significa una ventaja, como señala el entrevistado W. Vizarreta, para quienes no cuentan con suficientes recursos económicos, analíticos y disponibilidad de tiempo. Esto es cierto para el caso de Google Trends, ya que no requiere un costo adicional y su uso está recomendado para la gestión de destinos turísticos que no cuentan con el capital necesario para costear investigaciones de mercado más extensas.

Los expertos entrevistados consideraron que las variables adicionales provenientes de Google Trends podrían mejorar el modelo para un atractivo turístico; sin embargo, la prueba matemática evidenció que este no sería el caso del Complejo Arqueológico de Kuélap, salvo para la variable Google Búsquedas. Se afirma entonces, como mencionó el entrevistado F. Arentsen, que solamente se va a poder conocer si las variables adicionadas funcionan, si se realiza la prueba matemática correspondiente. Asimismo, esto podría ser el reflejo de lo mencionado por el entrevistado R. Peña, en relación a que existe una falta de promoción con contenido relevante, coherente, sostenible y sustentable en la promoción de los sitios turísticos peruanos. Esto probablemente no aplique para todos los atractivos; por ejemplo, en Machu Picchu se podría tener diferentes resultados y funcionar las combinaciones del modelo; sin embargo, ello sólo se podría comprobar con la prueba matemática.

Aparentemente, para que exista una correlación entre las variables electrónicas y las llegadas a un atractivo turístico, éste debería tener algunas condiciones adicionales relacionadas a la gestión del atractivo; como el uso de variables electrónicas dentro de las estrategias de marketing online, lo cual, puede ser ejecutado a través de notas de prensa electrónicas, videos en YouTube con etiquetas adecuadas, entre otros. Esto podría beneficiar al modelo de predicción que utiliza todos los indicadores disponibles de Google Trends, siendo verificado por una prueba matemática como la propuesta en el estudio.

Wladyka (2017) mencionó que el interés a lo largo del tiempo representado por el volumen de búsquedas en internet sobre un sitio turístico puede estar asociado a un incremento en la cantidad de personas que viajan a este lugar. Los resultados numéricos evidencian que

Kuélap aún no tiene una fuerte presencia web, tomando en cuenta la influencia de las llegadas hacia el atractivo de las búsquedas de imágenes, videos y noticias. Esto refleja que aún existe una labor pendiente en cuanto a las estrategias digitales para el atractivo. Las entidades privadas y estatales encargadas de promover Kuélap como destino turístico pueden tomar ventaja de Google Trends como una métrica en tiempo real para evaluar el interés a lo largo del tiempo respecto al atractivo, ventaja que no poseen la mayoría de indicadores macroeconómicos como ingresos, ventas de paquetes turísticos, etc., porque carecen de inmediatez en la disponibilidad de datos.

A pesar de que los resultados cuantitativos reflejaron que los modelos que combinan variables no son significativos para el caso de Kuélap; la investigación cualitativa dejó declaraciones interesantes en cuanto a las variables adicionadas. El entrevistado R. Peña mencionó que “la variable Google Noticias es útil siempre y cuando haya contenido y cantidad de artículos en distintos idiomas”. Para este estudio, la variable Google Noticias presentó un nivel de significancia (valor-p) que superaba el +0.05 establecido, lo que refleja una mayor probabilidad de error y evidencia que el indicador aún no es significativo para hallar las visitas a Kuélap. Este hecho podría reflejar que el atractivo todavía no cuenta con gran exposición mediática, por lo que la inclusión de esta variable no afecta para la mejora del modelo. La adición de la variable Google YouTube tampoco mejoro el modelo, esto, probablemente porque el atractivo no presenta una gran cantidad de contenido expuesto en la plataforma YouTube, lo cual, mencionado por el entrevistado W. Vizarreta “sería una buena estrategia para atraer a los turistas potenciales, por ejemplo, a través de un video musical o una película taquillera que muestre parte del destino, aun cuando no se trate de herramientas tradicionales de marketing”.

Si bien es cierto que los indicadores de Google Trends sirven para correlacionar llegadas, todavía resultan insuficientes debido a su propia naturaleza. Los factores externos tales como el clima, estabilidad política o otros factores de competitividad del destino como la conectividad área, afectan el interés potencial de forma que pueden hacer que un turista desista de viajar al destino por el cual se mostraba interesado. Google Trends como herramienta no es absoluta, pero sirve para una comprensión inicial de un problema, demostrando que en la actualidad, las búsquedas en internet significan un factor clave para conocer el interés de las personas hacia un destino o atractivo turístico. Por tanto, resultaría beneficioso para aquellos sitios turísticos que ya presentan este fenómeno, como el caso de Kuélap, que se concentren en mejorar sus estrategias de marketing online relacionadas a buscadores web.

Se puede rescatar como último hallazgo de la investigación que los entrevistados también mencionaron que reforzarían el modelo con variables externas al ámbito electrónico que podrían darle mayor solidez a la prueba matemática y generar nuevas métricas, tal como capacidad de aerolíneas, seguridad, clima, estacionalidad del atractivo, etc. Frechtling (2001) señala que, es importante considerar que la demanda turística es volátil dependiendo de la estación; los turistas son sensibles a los desastres naturales o desastres ocasionados por el hombre; el comportamiento del turista es complejo dependiendo de sus propósitos o motivos de viaje; entre otros aspectos que se deben tener en cuenta al momento de considerar un modelo de predicción. En este punto es importante mencionar que, la finalidad de la investigación se centra en evaluar variables electrónicas para determinar si éstas son significativas para un modelo predictivo; por tanto, no se utilizaron variables como las sugeridas considerando que no responderían los objetivos de la investigación.

## Conclusiones

Los antecedentes académicos relacionan a las búsquedas en internet con destinos y esta investigación ha demostrado que es viable emplear la metodología para un atractivo, una ubicación más específica.

La investigación cualitativa y lectura de bibliografía presentada en el marco teórico han permitido explorar y encontrar que los indicadores proporcionados por Google Trends benefician a las estrategias de marketing online de un atractivo turístico, especialmente aquellas estrategias de SEO & SEM vinculadas a los buscadores web y uso de palabras clave, los cuales, representan un primer acercamiento al marketing de atracción.

Se determinó que el modelo AR-1, que no hace uso de variables electrónicas, hubiera permitido pronosticar las llegadas a Kuélap en el periodo de 2011-2016. Además, se comprobó y concluyó, de la misma manera que Choi y Varian (2009), que la inclusión de la variable G(web) al modelo, mejora la precisión de un modelo que solamente utiliza variables históricas. En esta prueba, los coeficientes analizados se incrementaron respecto al modelo AR-1: el coeficiente de correlación múltiple de +0.838 a +0.884; el R-cuadrado de +0.702 a +0.782; y el R-cuadrado ajustado de +0.694 a +0.773. Por tanto, se demuestra que la variable de llegadas mensuales a Kuélap está correlacionada con la cantidad de arribos alcanzado hace un año, hace un mes y el índice proveniente del tráfico de búsquedas web en Google.

Los expertos entrevistados consideraron que, si bien los indicadores de búsquedas web representan una variable útil para el modelo del estudio; las otras variables: Google Imágenes, Google YouTube y Google Noticias, deben ser exploradas en una metodología de prueba y error para detectar que tan beneficiosas son para el modelamiento del flujo turístico de un atractivo en específico. De acuerdo a ello, en relación a la evaluación de las combinaciones de las variables G(YouTube), G(Imágenes) y/o G(Noticias); los hallazgos de la investigación comprobaron que no son estadísticamente significativas en comparación con el modelo que sólo emplea datos históricos y búsquedas web. La prueba matemática evidenció que para el caso específico de Kuélap, las llegadas no guardan relación con los indicadores provenientes de Google imágenes, YouTube y noticias, por lo que no hubieran permitido pronosticar las llegadas a Kuélap en el periodo 2011 - 2016. Estos últimos resultados probablemente se deban a que Kuélap es un atractivo en crecimiento que aún no cuenta con tanta exposición y posicionamiento web.

La investigación ha aplicado una metodología que normalmente no es usada en estudios relacionados al turismo, en la cual existen pocas pruebas empíricas de sus resultados.

Asimismo, ha aportado un cuerpo de conocimiento que usa variables digitales para modelos predictivos que pueden resultar más rápidas y precisas. Y, además, ha contemplado la riqueza del contraste entre la percepción de los expertos sobre el objeto del estudio, los conocimientos generales de los autores citados y los resultados de la prueba matemática para el caso específico de Kuélap.

Los resultados cualitativos demuestran que los estudios que hacen uso de Big Data se están volviendo populares entre los investigadores de diferentes sectores económicos. La herramienta de Big Data utilizada en este estudio, Google Trends, es vista como una solución más económica, especialmente para aquellos actores (operadores, gestores de destino, etc.) que no cuentan con los recursos económicos necesarios para costear investigaciones de mercado. Esta herramienta ha sido reconocida como XXXXX, el cual permite conocer, a través de la búsqueda de un término clave en concreto, el interés de un usuario hacia un sitio turístico específico.

Por el lado del marketing turístico, es posible utilizar los indicadores de Google Trends como una métrica para monitorear y evaluar campañas promocionales orientadas a incrementar las búsquedas en internet de un sitio turístico. Si PROMPERÚ, por ejemplo, hiciese una campaña online promocionando Kuélap, el resultado de ésta, debería reflejarse en un incremento de las búsquedas; caso contrario, significaría que la campaña no cumplió sus objetivos. La herramienta también sirve para tener conocimiento de aquellos momentos en que el destino es más buscado y analizar los patrones de estacionalidad para crear un calendario de contenidos para las campañas de marketing online.

Los hallazgos de la investigación han permitido reafirmar que la presencia web es un punto clave dentro del funnel del inbound marketing, pues permite generar tráfico web. Este tráfico web genera leads, los cuales representan una muestra de interés en el destino. Esto puede ser de utilidad para las DMO's, las cuales, pueden plantear estrategias que generen un mayor tráfico hacia los sitios web del destino o atractivo, y que a partir de ello, se conviertan los leads en visitas al sitio turístico. Para el caso de Kuélap, sería importante iniciar con una mayor presencia en internet para atraer a los potenciales viajeros, tomando en cuenta los videos e imágenes como ventajas para plantear estrategias y acciones de marketing de contenidos y marketing de atracción.

## Recomendaciones

Es recomendable complementar los datos de Google con los de alguna red social para mejorar la precisión de sus resultados. Con el análisis sentimental aplicado a los comentarios, es posible obtener puntuaciones positivas y negativas. Si es que un destino se encuentra sumergido en conflictos sociales internos, esto se reflejará en los comentarios online y perjudicará a la demanda (El Rouby, 2015). Google cuantifica el interés, pero no si el mismo es de naturaleza negativa o positiva. El empleo de datos textuales y su conversión a patrones numéricos, sin embargo, presenta una serie de desafíos; siendo el más importante, el conjunto de procedimientos para depurar los datos textuales.

Sería viable replicar el estudio a otros atractivos turísticos. Se recomienda para futuras investigaciones incorporar las variables Google Imágenes, Google YouTube y Google Noticias para el modelamiento del flujo mensual a sitios turísticos que cuenten con una mayor exposición y posicionamiento web; o realizar nuevamente el estudio para el Complejo Arqueológico de Kuélap en caso se incremente la presencia en internet.

En el caso de variables como búsquedas de YouTube, es recomendable cruzar las mismas con llegadas a lugares que hayan sido objeto de filmaciones de videos musicales, películas o series que hayan tenido un impacto considerable en la audiencia. Aun cuando para el caso de Kuélap no exista correlación entre las reproducciones de YouTube y los arribos, esto no significa que ocurra lo mismo para otros sitios turísticos. Además, a diferencia de Google Trends, YouTube cuenta con un indicador absoluto, que es la cantidad de reproducciones. Este aspecto de YouTube podría servir para futuras investigaciones que se centren en la correlación de, por ejemplo, la reproducción de un video musical que muestre sitios turísticos y su impacto en las llegadas al país de referencia.

Para el caso de imágenes, específicamente en extranjeros, se podría realizar mediciones que empleen Instagram. Esta red social permite realizar un conteo de imágenes e interacciones; las mismas que pueden ser correlacionadas con las visitas a destinos o atractivos peruanos realizadas por viajeros foráneos. Las imágenes también pueden ser procesadas y de ellas pueden obtenerse sentimientos y emociones (Gajarla & Gupta, s.f). Con ambas es posible medir cómo los viajeros sintieron su experiencia durante el viaje.

Otras futuras investigaciones podrían estar relacionados a los modelos predictivos que utilizan variables electrónicas ajustadas con variables no electrónicas, o también a estudios que usen variables electrónicas para proyectar variables no electrónicas como ingresos o empleo.

Una de las limitaciones más importantes del estudio es que Google Trends brinda indicadores relativos. Estos indicadores pueden ser complementados en investigaciones futuras con herramientas virtuales que muestren valores absolutos (cantidad de búsquedas); de manera que se ejemplifiquen proyecciones sobre un sitio o producto turístico. Por ejemplo, el portal KWFinder es una herramienta de investigación y análisis de palabras clave que muestra el volumen de búsquedas que se han realizado hacia un término clave en los últimos 12 meses (Kortis, 2016). No obstante, si bien esta herramienta brinda las búsquedas absolutas que podrían servir para proyectar; la data no está limpia, es decir, no está filtrada debido a que incluye búsquedas sobre viajes, así como contenido histórico o motivos académicos.

Los académicos deben interpretar estos resultados desde una perspectiva correlacional: el volumen de búsquedas acerca de Kuélap está correlacionado con los arribos a dicho atractivo. No es relación causal, es decir, no significa que una de las variables es causa de la otra. Por tanto, la variación en una variable es sistemática con respecto a la otra. En esta investigación, se ha demostrado que el indicador de búsquedas web de Google Trends está correlacionado y contribuye a fortalecer un modelo predictivo a Kuélap. Asimismo, en cuanto a la gestión comercial, Google puede ser empleado para medir el interés hacia un destino o atractivo turístico.

La investigación ha demostrado que, de cierta manera, las búsquedas web en Google influyen en las llegadas a Kuélap; por lo cual, sería pertinente que las entidades involucradas dirijan acciones de promoción online para el atractivo. El Gobierno Local podría iniciar con la creación de una página web donde se de la información necesaria sobre los servicios turísticos ofrecidos en Kuélap o la adición de una sección relacionada al turismo en el sitio web de la Región. Luego de ello, podría realizarse un calendario de contenidos para campañas de marketing online más especializadas, utilizando como base la estacionalidad de búsquedas del atractivo, las cuales, involucren otras estrategias multimedia como videos e imágenes.

Finalmente, sería aconsejable para los gestores en turismo del Gobierno Local que utilicen la información obtenida a partir de la investigación, como base para centrar estrategias digitales que permitan optimizar los ingresos y arribos hacia el atractivo; que son las dos metas de gestión de turismo en el Estado. Se debería realizar un trabajo en conjunto a los gestores de turismo privados de la región, demostrándoles los resultados y motivándolos a elaborar estrategias de SEO o SEM dirigidas a promocionar su oferta turística vía online.

## REFERENCIAS

- Askitas, N., & Zimmermann, K. (2009). *Google Econometrics and Unemployment Forecasting*. Bonn, Alemania. Recuperado el 23 de diciembre de 2016, de <http://ftp.iza.org/dp4201.pdf>
- Avis Budget Group. (2017). *Annual Reports*. Recuperado el 20 de setiembre de 2017, de <http://ir.avisbudgetgroup.com/annuals.cfm>
- Baggio, R. (2016). Big Data, Business Intelligence and Tourism. Recuperado el 23 de enero de 2017, de [http://www.iby.it/turismo/papers/baggio\\_BigDataSurvey.pdf](http://www.iby.it/turismo/papers/baggio_BigDataSurvey.pdf)
- Banco Central Europeo. (2013). *Now-Casting and The Real-Time Data Flow*. Frankfurt, Alemania. Recuperado el 25 de abril de 2017, de <https://www.ecb.europa.eu/pub/pdf/scpwps/ecbwp1564.pdf?74129250f82f3a27d5c798fc79f10dbb>
- Bangwayo-Skeete, P., & Skeete, R. (Febrero de 2015). Can Google data improve the forecasting performance of tourist arrivals? Mixed-data sampling approach. *Tourism Management*, 46, 454-464. Recuperado el 23 de diciembre de 2016, de <http://www.sciencedirect.com/science/article/pii/S0261517714001460>
- Bezhovski, Z. (2015). Inbound Marketing - A New Concept in Digital Business. Recuperado el 22 de junio de 2017, de [https://www.researchgate.net/publication/291833697\\_Inbound\\_Marketing\\_-\\_A\\_New\\_Concept\\_in\\_Digital\\_Business](https://www.researchgate.net/publication/291833697_Inbound_Marketing_-_A_New_Concept_in_Digital_Business)
- Borba, A., & Akemi, A. (2014). Big Data Usage in the Marketing Information. *Journal of Data Analysis and Information Processing*, 2, 77-85. Recuperado el 13 de enero de 2017, de [http://file.scirp.org/pdf/JDAIP\\_2014082816562983.pdf](http://file.scirp.org/pdf/JDAIP_2014082816562983.pdf)
- Bristol Myers Squibb. (2017). *Publications*. Recuperado el 20 de setiembre de 2017, de <https://www.bms.com/media/publications.html>
- CANATUR. (2017). *Viajes Pacífico proyecta alcanzar 60% de ocupación en nuevo hotel Achamaqui de Chachapoyas*. Recuperado el 30 de diciembre de 2017, de <http://www.portaldeturismo.pe/index.php/noticia/viajes-pacifico-proyecta-alcanzar-60-de-ocupacion-en-nuevo-hotel-achamaqui-de-chachapoyas>
- Choi, H., & Varian, H. (2009). *Predicting the Present with Google Trends*. Recuperado el 23 de diciembre de 2016, de [https://static.googleusercontent.com/media/www.google.com/es//googleblogs/pdfs/google\\_predicting\\_the\\_present.pdf](https://static.googleusercontent.com/media/www.google.com/es//googleblogs/pdfs/google_predicting_the_present.pdf)
- De la Oz, M. (2014). *Predicting Tourist Inflows to Punta Cana, Dominican Republic, Using Google Trends*. Tesis de maestría, Utah State University. Recuperado el 22 de febrero de 2017, de <https://digitalcommons.usu.edu/cgi/viewcontent.cgi?referer=https://www.google.com.pe/&httpsredir=1&article=1362&context=gradreports>
- Dinis, G., Costa, C., & Pacheco, O. (2016). *The Use of Google Trends Data as Proxy of Foreign Tourist Inflows to Portugal*. Recuperado el 10 de setiembre de 2017, de [http://iacudit.org/journal/volumes/v3n1/v3n1\\_66-75.pdf](http://iacudit.org/journal/volumes/v3n1/v3n1_66-75.pdf)

- Dirección General de Empresas del Gobierno Francés. (junio de 2017). *Dirección General de Empresas del Gobierno Francés (DGE)*. Recuperado el 29 de setiembre de 2017, de [https://www.entreprises.gouv.fr/files/files/directions\\_services/etudes-et-statistiques/4p-DGE/2017-06-4p-N71-EVE-ENGLISH.pdf](https://www.entreprises.gouv.fr/files/files/directions_services/etudes-et-statistiques/4p-DGE/2017-06-4p-N71-EVE-ENGLISH.pdf)
- Dolnicar, S., & Ring, A. (2014). Tourism marketing research: Past, present and future. *Annals of Tourism Research*. Recuperado el 13 de enero de 2017, de <http://www.sciencedirect.com/science/article/pii/S0160738314000358>
- Durmaz, Y., & Hali, I. (Junio de 2016). *Travel from Traditional Marketing to Digital Marketing*. Recuperado el 10 de setiembre de 2017, de [https://www.researchgate.net/profile/Yakup\\_Durmaz/publication/303940843\\_Travel\\_from\\_Traditional\\_Marketing\\_to\\_Digital\\_Marketing/links/575f9dc708aed884621bb59/Travel-from-Traditional-Marketing-to-Digital-Marketing.pdf](https://www.researchgate.net/profile/Yakup_Durmaz/publication/303940843_Travel_from_Traditional_Marketing_to_Digital_Marketing/links/575f9dc708aed884621bb59/Travel-from-Traditional-Marketing-to-Digital-Marketing.pdf)
- Editorial Elearning. (2010). *Marketing Digital*. Malaga, España: Publicaciones Vertice. Obtenido de [https://books.google.com.pe/books?hl=es&lr=&id=qdGjYHkGB1UC&oi=fnd&pg=PP2&dq=marketing+digital&ots=P\\_A4HtrmFY&sig=AGKI8ulTofaA37CsZnC1UJC8TL0#v=onepage&q=marketing%20digital&f=false](https://books.google.com.pe/books?hl=es&lr=&id=qdGjYHkGB1UC&oi=fnd&pg=PP2&dq=marketing+digital&ots=P_A4HtrmFY&sig=AGKI8ulTofaA37CsZnC1UJC8TL0#v=onepage&q=marketing%20digital&f=false)
- El Comercio. (marzo de 2017). Teleférico de Kuélap: la inversión y su impacto económico. Recuperado el 12 de mayo de 2017, de <http://elcomercio.pe/economia/peru/teleferico-kuelap-inversion-y-su-impacto-economico-noticia-1972840>
- El Roubly, N. (2015). The Use of Text mining to examine the Effect of the Egyptian Revolution on Tourists' Sentiments towards visiting the Country. *Journal of Tourism Research*. Recuperado el 29 de setiembre de 2017, de <http://jotr.eu/index.php/volume10/89-sentiments>
- Espino, C. (2017). *Análisis predictivo: técnicas y modelos utilizados y aplicaciones del mismo - herramientas Open Source que permiten su uso*. Recuperado el noviembre de 02 de 2017, de <http://openaccess.uoc.edu/mwg-internal/de5fs23hu73ds/progress?id=GHot8xknVcny0gyMU7I4bga4t7z7YX1NMRpZLXmP55Q>,
- European Travel Commission & Organización Mundial del Turismo. (2008). *Handbook on E-marketing for Tourism Destinations*. Recuperado el 26 de mayo de 2017, de [http://www.edenbulgaria.eu/.../1212\\_attachment\\_10\\_\\_handbook\\_20on\\_20e\\_marketing\\_2](http://www.edenbulgaria.eu/.../1212_attachment_10__handbook_20on_20e_marketing_2)
- Fayeke, C., & Crompton, J. (1991). Image differences between prospective, first-time, and repeat visitors to the Lower Rio Grande Valley. Recuperado el 09 de mayo de 2017, de <http://journals.sagepub.com/doi/abs/10.1177/004728759103000202>
- Fesenmaier, D., & Stienmetz, J. (2013). *Effects of Channel, Timing, and Bundling on Destination Advertisement Response*. Recuperado el 27 de octubre de 2017, de <http://scholarworks.umass.edu/cgi/viewcontent.cgi?article=1840&context=ttra>
- Fonseca, E. (08 de mayo de 2017). Taller de Marketing Digital Turístico en PROMPERÚ. Lima. Recuperado el 08 de mayo de 2017

- Forbes. (2014). *Your Guide To Using Images In Your Content Marketing Strategy*. Recuperado el 02 de noviembre de 2017, de <https://www.forbes.com/sites/jaysondemers/2014/09/29/your-guide-to-using-images-in-your-content-marketing-strategy/#535cb35f7b8a>
- Forbes. (2017). *Video Marketing: The Future Of Content Marketing*. Recuperado el 02 de noviembre de 2017, de <https://www.forbes.com/sites/forbesagencycouncil/2017/02/03/video-marketing-the-future-of-content-marketing/#470a657b6b53>
- Fotios, K. (2016). *Web Search activity: A Forecast Tool for Tourist Arrivals in Cyprus*. Thessaloniki, Grecia. Recuperado el 16 de enero de 2017, de <https://repository.ihu.edu.gr/xmlui/bitstream/handle/11544/14480/Kallasidis-Fotis-Dissertation.pdf?sequence=1>
- Frechtling, D. (2001). *Forecasting Tourism Demand: Methods and Strategies*. Recuperado el 21 de junio de 2017, de <https://books.google.com.pe/books?id=AWAeunPTiYEC&printsec=frontcover#v=onepage&q&f=false>
- Fuchs, M., & Lexhagen, M. (2015). Applying business intelligence for Generation in Tourism Destinations – A Case Study from Sweden. *Tourism and Leisure*. Recuperado el 06 de enero de 2017, de [http://link.springer.com/chapter/10.1007%2F978-3-658-06660-4\\_11](http://link.springer.com/chapter/10.1007%2F978-3-658-06660-4_11)
- Gajarla, V., & Gupta, A. (s.f). *Emotion Detection and Sentiment Analysis of Images*. Georgua Tech, College of Computing. Recuperado el 20 de setiembre de 2017, de [https://www.cc.gatech.edu/~hays/7476/projects/Aditi\\_Vasavi.pdf](https://www.cc.gatech.edu/~hays/7476/projects/Aditi_Vasavi.pdf)
- Gawlik, E., Kabaria, H., & Kaur, S. (2011). *Predicting tourism trends with Google Insights*. Recuperado el 15 de enero de 2017, de <http://cs229.stanford.edu/proj2011/GawlikKaurKabaria-PredictingTourismTrendsWithGoogleInsights.pdf>
- Georgetown University. (s.f). *Georgetown University Medical Center*. Recuperado el 28 de junio de 2017, de <https://ora.georgetown.edu/irb/irbICFtemplates>
- Gobierno Regional de Amazonas. (2007). *Plan Regional de Desarrollo Concertado de Amazonas 2008 - 2021*. Recuperado el 27 de octubre de 2017, de [http://www.mesadeconcertacion.org.pe/sites/default/files/archivos/2015/documentos/11/1\\_prdc\\_amazonas\\_2008\\_2021.pdf](http://www.mesadeconcertacion.org.pe/sites/default/files/archivos/2015/documentos/11/1_prdc_amazonas_2008_2021.pdf)
- Godnov, U., & Redeck, T. (2016). Application of text mining in tourism: Case of Croatia. *Annals of Tourism Research*, 58, 162-166. Recuperado el 13 de enero de 2017, de <http://www.sciencedirect.com/science/article/pii/S0160738316300159>
- Gómez. (s.f). *SOBRE EL CONCEPTO DEL P-VALOR*. Recuperado el 30 de diciembre de 2017, de [http://www.ahepe.es/VICongreso/descargas/MA\\_Gomez\\_Villegas.pdf](http://www.ahepe.es/VICongreso/descargas/MA_Gomez_Villegas.pdf)
- Google. (2011). *Winning the Zero Moment of Truth* . Recuperado el 21 de junio de 2017, de <https://www.thinkwithgoogle.com/marketing-resources/micro-moments/2011-winning-zmot-ebook/>

- Google. (2017). *Think With Google*. Recuperado el 08 de mayo de 2017, de I-Want-to-Get-Away Moments: What They Mean for Travel Marketing: <https://www.thinkwithgoogle.com/articles/get-away-moments-travel-marketing.html>
- Google. (s.f.). *Ayuda de Noticias*. Recuperado el 13 de setiembre de 2017, de <https://support.google.com/news/publisher#topic=4359865>
- Google. (s.f.). *Google Dentro de Google*. Recuperado el 13 de setiembre de 2017, de Algoritmos: <https://www.google.com/intl/es-419/insidesearch/howsearchworks/algorithms.html>
- Google. (s.f.). *Google Digital Unlocked*. Recuperado el 13 de setiembre de 2017, de Trends Help: [https://support.google.com/trends/answer/4355213?hl=en&ref\\_topic=6248052](https://support.google.com/trends/answer/4355213?hl=en&ref_topic=6248052)
- Google Trends. (2016). *Google Trends*. Recuperado el 14 de enero de 2017, de <https://www.google.com/trends/explore?cat=67&geo=TH&q=%2Fm%2F016wzw>
- Guajarati, D. (2006). Principios de econometría. Madrid: McGraw-Hill Interamericana de España. Recuperado el 02 de noviembre de 2017
- Guðbjörg, A. (2014). *Effect of social media marketing on traditional marketing campaigns in young icelandic companies*. Recuperado el 27 de octubre de 2017, de [https://skemman.is/bitstream/1946/18659/1/MIB0614\\_Thesis\\_Anna\\_Cowden\\_EffectofSocialMediaMarketingonTraditionalMarketingCampaignsinYoungIcelandicCompanies.pdf](https://skemman.is/bitstream/1946/18659/1/MIB0614_Thesis_Anna_Cowden_EffectofSocialMediaMarketingonTraditionalMarketingCampaignsinYoungIcelandicCompanies.pdf)
- Hart, M., Rasner, J., & Lukoszová, X. (2012). *Demand Forecasting Significance for Contemporary Process Management of Logistics Systems*. Recuperado el 06 de Abril de 2017, de <http://clc2012.tanger.cz/files/proceedings/09/reports/836.pdf>
- Hernández, R., Fernández, C., & Baptista, P. (2014). *Metodología de la Investigación* (6ta edición ed.). Recuperado el 20 de febrero de 2017, de <http://upla.edu.pe/portal/wp-content/uploads/2017/01/Hern%C3%A1ndez-R.-2014-Metodologia-de-la-Investigacion.pdf.pdf>
- IBM. (2016). *A Smarter Planer - Big Data*. Recuperado el 03 de diciembre de 2016, de <https://www.ibm.com/big-data/us/en/>
- IBM. (2016). *The 5 Vs of Big Data*. Recuperado el 11 de mayo de 2017, de <https://www.ibm.com/blogs/watson-health/the-5-vs-of-big-data/>
- IBM. (2017). *IBM SPSS Statistics Base*. Recuperado el 14 de mayo de 2017, de <http://www-03.ibm.com/software/products/es/spss-stats-base>
- IBM. (s.f.). *IBM Knowledge Center*. Recuperado el 29 de noviembre de 2017, de R2 Ajustado: [https://www.ibm.com/support/knowledgecenter/es/SS4QC9/com.ibm.solutions.wa\\_an\\_overview.2.0.0.doc/rsquared\\_adjusted.html](https://www.ibm.com/support/knowledgecenter/es/SS4QC9/com.ibm.solutions.wa_an_overview.2.0.0.doc/rsquared_adjusted.html)
- IBM, I. B. (2015). *Analyze NFL data for the big game*. Recuperado el 20 de setiembre de 2017, de <https://www.ibm.com/communities/analytics/watson-analytics-blog/analyze-nfl-data-for-the-big-game/>

- Instituto Valenciano de Tecnologías Turísticas. (2015). *BIG DATA: retos y oportunidades para el turismo*. Recuperado el 03 de diciembre de 2016, de <http://www.thinktur.org/media/Big-Data.-Retos-y-oportunidades-para-el-turismo.pdf>
- Joyanes, L. (2013). *Big Data: Análisis de grandes volúmenes de datos en organizaciones* (Vol. 1era edición). Alfaomega Grupo. Recuperado el noviembre de 02 de 2017, de [https://books.google.com.pe/books?id=1GywDAAAQBAJ&pg=PT39&lpg=PT39&dq=big+data+aparicion&source=bl&ots=\\_WO6Q64cZM&sig=F1yYJpKturhmwi1qUxdgVEdlLUs&hl=es-419&sa=X&ved=0ahUKEwj697fChvbXAhVkt8KHR\\_GAA8Q6AEIWzAl#v=onepage&q=big%20data%20aparicion&f=false](https://books.google.com.pe/books?id=1GywDAAAQBAJ&pg=PT39&lpg=PT39&dq=big+data+aparicion&source=bl&ots=_WO6Q64cZM&sig=F1yYJpKturhmwi1qUxdgVEdlLUs&hl=es-419&sa=X&ved=0ahUKEwj697fChvbXAhVkt8KHR_GAA8Q6AEIWzAl#v=onepage&q=big%20data%20aparicion&f=false)
- Kim, J., & Moosa, I. (2001). Seasonal behaviour of monthly international tourist flows: specification and implications for forecasting models. *Tourism Economics*. Recuperado el 13 de enero de 2017, de <http://journals.sagepub.com/mwg-internal/de5fs23hu73ds/progress?id=stOMWdeYcD6Y9isEd6pxTcHZozM5zymP9bp38iSJS4,&dl>
- Kortis, M. (2016). *KWFinder guide*. Recuperado el 30 de diciembre de 2017, de [https://mangools.com/blog/kwfinder-guide/?\\_ga=2.153367032.113725893.1514754909-252290730.1514754909](https://mangools.com/blog/kwfinder-guide/?_ga=2.153367032.113725893.1514754909-252290730.1514754909)
- Kotler, P., & Armstrong, G. (2012). *Marketing* (Vol. decimocuarta edición). Recuperado el 03 de noviembre de 2017, de [https://profdariomarketing.files.wordpress.com/2014/03/marketing\\_kotler-armstrong.pdf](https://profdariomarketing.files.wordpress.com/2014/03/marketing_kotler-armstrong.pdf)
- Kotler, P., Kartajaya, H., & Setiawan, I. (2010). *Marketing 3.0*. New Jersey: John Wiley & Sons, Inc. Recuperado el 18 de mayo de 2017, de [http://manajemen-pemasaran.com/katalogmanajemen/2010%20-%20\(EBOOK\)%20Marketing%2030-Hermawan%20KertajayaPhilip%20Kotler%20Hermawan%20Kartajaya%20Iwan%20Setiawan.pdf](http://manajemen-pemasaran.com/katalogmanajemen/2010%20-%20(EBOOK)%20Marketing%2030-Hermawan%20KertajayaPhilip%20Kotler%20Hermawan%20Kartajaya%20Iwan%20Setiawan.pdf)
- Kumar, S., Morstatter, F., & Liu, H. (2013). *Twitter Data Analytics*. Springer. Recuperado el 29 de setiembre de 2017, de <https://www.google.com.pe/url?sa=t&rct=j&q=&esrc=s&source=web&cd=2&cad=rja&uact=8&ved=0ahUKEwiYy4zSnNLWAhWH5iYKHR5FCJUQFggwMAE&url=http%3A%2F%2Ftweetracker.fulton.asu.edu%2Ftda%2FTwitterDataAnalytics.pdf&usg=AOvVaw1jax8O-ljQUOz3dC2DZ1Kv>
- Lahura, E. (2003). *El Coeficiente de Correlación y Correlaciones Espúreas*. Recuperado el 31 de mayo de 2017, de <http://files.pucp.edu.pe/departamento/economia/DDD218.pdf>
- Lancis, E. (Abril de 2014). *Big Data aplicada al sector turístico*. Recuperado el 20 de junio de 2017, de <http://www.segittur.es/es/sala-de-prensa/detalle-documento/Presentacin-Big-Data-aplicada-al-sector-turstico-/#.WgMtOFvWzcd>
- Lancis, E. (2014). *Turismo y Big Data: Construyendo el Futuro*. Recuperado el 30 de diciembre de 2017, de <http://www.segittur.es/es/sala-de-prensa/detalle-documento/Presentacin-Big-Data-aplicada-al-sector-turstico-/#.WkkZNVWWbIV>

- Lutz, F. (2013). Trending in Africa Using Google Trends to Nowcast Consumer Confidence in South Africa. Recuperado el 17 de diciembre de 2016, de [www.essa2013.org.za/fullpaper/essa2013\\_2610.pdf](http://www.essa2013.org.za/fullpaper/essa2013_2610.pdf)
- McKinsey Global Institute. (2011). *Big Data: The next frontier for innovation, competition, and productivity*. Recuperado el 03 de diciembre de 2016, de <http://www.mckinsey.com/business-functions/digital-mckinsey/our-insights/big-data-the-next-frontier-for-innovation>
- Microsoft. (14 de diciembre de 2016). *Usar las Herramientas para análisis para realizar análisis de datos complejos*. Recuperado el 03 de marzo de 2017, de <https://support.office.com/es-es/article/Usar-las-Herramientas-para-an%C3%A1lisis-para-realizar-an%C3%A1lisis-de-datos-complejos-6c67ccf0-f4a9-487c-8dec-bdb5a2cefab6>
- MINCETUR. (2011). *Turismo: Guía para la Formulación de Proyectos de Inversión Exitosos*. Recuperado el 10 de setiembre de 2017, de [https://www.mincetur.gob.pe/wp-content/uploads/documentos/turismo/proyectos\\_inversion\\_publica/instrumentos\\_metodologicos/Guia\\_Turismo\\_capitulo1.pdf](https://www.mincetur.gob.pe/wp-content/uploads/documentos/turismo/proyectos_inversion_publica/instrumentos_metodologicos/Guia_Turismo_capitulo1.pdf)
- MINCETUR. (2013). *Zona Arqueológica Monumental Kuélap*. Recuperado el 11 de mayo de 2017, de [http://ficha.sigmincetur.mincetur.gob.pe/index.aspx?cod\\_Ficha=754](http://ficha.sigmincetur.mincetur.gob.pe/index.aspx?cod_Ficha=754)
- MINCETUR. (2017). *Sistema de Información: Estadística de Turismo*. Recuperado el 10 de febrero de 2017, de <http://datosturismo.mincetur.gob.pe/appdatosTurismo/Content2.html>
- Minitab. (2013). *Multiple Regression Analysis: Use Adjusted R-Squared and Predicted R-Squared to Include the Correct Number of Variables*. Recuperado el 29 de noviembre de 2017, de <http://blog.minitab.com/blog/adventures-in-statistics-2/multiple-regression-analysis-use-adjusted-r-squared-and-predicted-r-squared-to-include-the-correct-number-of-variables>
- Office of National Statistics. (2012). Recuperado el 23 de diciembre de 2016, de <https://www.ons.gov.uk>
- OMT. (2010). *Recomendaciones internacionales para estadísticas de turismo 2008*. Naciones Unidas. Recuperado el 27 de octubre de 2017, de [https://unstats.un.org/unsd/publication/Seriesm/SeriesM\\_83rev1s.pdf](https://unstats.un.org/unsd/publication/Seriesm/SeriesM_83rev1s.pdf)
- Onder, I., & Gunter, U. (2016). Forecasting Tourism Demand with Google Trends: The Case of Vienna. *Conference: The 8th International Conference on Computational and Financial Econometrics*. Recuperado el 10 de diciembre de 2016, de [https://www.researchgate.net/publication/274009577\\_Forecasting\\_Tourism\\_Demand\\_with\\_Google\\_Trends\\_The\\_Case\\_of\\_Vienna](https://www.researchgate.net/publication/274009577_Forecasting_Tourism_Demand_with_Google_Trends_The_Case_of_Vienna)
- Park, S., Lee, J., & Song, W. (2015). Short-term forecasting of Japanese tourist inflow to South Korea using Google trends data. *Journal of Travel & Tourism Marketing*. Recuperado el 03 de diciembre de 2016, de <http://www.tandfonline.com/doi/figure/10.1080/10548408.2016.1170651?scroll=top&needAccess=true>

- Plan Copesco Nacional. (s.f). *Mejoramiento de los servicios turísticos públicos zona arqueológica monumental Kuélap, distrito de Tingo, provincia de Luya, Amazonas*. Recuperado el 27 de octubre de 2017, de [http://plancopesconacional.gob.pe/index.php?option=com\\_k2&view=item&id=134](http://plancopesconacional.gob.pe/index.php?option=com_k2&view=item&id=134)
- Priyanka, M., & Shivaputra, S. (2017). Traffic Management in Smart cities Using Big Data. *International Journal of Innovative and Emerging*. Recuperado el 14 de setiembre de 2017, de <http://ijiere.com/FinalPaper/FinalPaperTraffic%20Management%20in%20Smart%20cities%20Using%20Big%20Data171203.pdf>
- PROMPERÚ. (2015). *Turismo In*. Recuperado el 27 de octubre de 2017, de Conociendo al Turista que Visita Amazonas 2015: <https://www.promperu.gob.pe/turismoin/Sitio/VisorDocumentos?titulo=Conociendo%20al%20turista%20que%20visita%20Amazonas&url=~/Uploads/conociendoAITurista/Receptivo-Amazonas.pdf&nombObjeto=conociendoAITurista>
- PROMPERÚ. (2015). *Turismo In*. Recuperado el 27 de octubre de 2017, de Conocinedo al Vacacionista Nacional que Visita Amazonas 2015: <https://www.promperu.gob.pe/turismoin/Sitio/VisorDocumentos?titulo=Conociendo%20al%20turista%20que%20visita%20Amazonas&url=~/Uploads/conociendoAITurista/Interno-Amazonas.pdf&nombObjeto=conociendoAITurista>
- PROMPERÚ. (2017). *"Perú, el país más rico del mundo"*. Recuperado el 27 de octubre de 2017, de <http://www.peru.travel/perupaismasrico/latam/#material-de-campana>
- PROMPERÚ. (2017). *Perfil del Turista Extranjero 2016*. Recuperado el 11 de diciembre de 2017, de [https://www.promperu.gob.pe/TurismoIN/sitio/VisorDocumentos?titulo=Perfil%20del%20Turista%20Extranjero%202016&url=~/Uploads/perfiles\\_extranjeros/39/PUBLICACIÓN%20PERFIL%20TURISTA%20EXTRANJERO%202016%20\(SPREAD\)\\_LC.pdf&nombObjeto=PerfTuristaExt&back=/Turism](https://www.promperu.gob.pe/TurismoIN/sitio/VisorDocumentos?titulo=Perfil%20del%20Turista%20Extranjero%202016&url=~/Uploads/perfiles_extranjeros/39/PUBLICACIÓN%20PERFIL%20TURISTA%20EXTRANJERO%202016%20(SPREAD)_LC.pdf&nombObjeto=PerfTuristaExt&back=/Turism)
- PROMPERÚ. (2017). *Perfil del Vacacionista Nacional 2016*. Recuperado el 11 de diciembre de 2017, de [https://www.promperu.gob.pe/TurismoIN/sitio/VisorDocumentos?titulo=Perfil%20del%20Vacacionista%20Nacional%202016&url=~/Uploads/perfiles\\_vacac\\_nac/1038/PVN%2016%20\(SPREAD\).pdf&nombObjeto=PerfVacacionistaNac&back=/TurismoIN/sitio/PerfVacacionistaNac&issuuid](https://www.promperu.gob.pe/TurismoIN/sitio/VisorDocumentos?titulo=Perfil%20del%20Vacacionista%20Nacional%202016&url=~/Uploads/perfiles_vacac_nac/1038/PVN%2016%20(SPREAD).pdf&nombObjeto=PerfVacacionistaNac&back=/TurismoIN/sitio/PerfVacacionistaNac&issuuid)
- PROMPERÚ. (2017). *Peru.Travel*. Recuperado el 11 de mayo de 2017, de <http://www.peru.travel/es-es/que-hacer/milenario/complejo-arqueologico-kuelap.aspx>
- Rice, C. (1997). *Understanding Customers*. doi:[http://dx.b-ok.org/genesis/751000/738a0e9a5012a204cb6a52ef0d443825/\\_as/\[Chris\\_Rice,\\_Craig\\_Simmons,\\_Chartered\\_Institute\\_of\(b-ok.org\).pdf](http://dx.b-ok.org/genesis/751000/738a0e9a5012a204cb6a52ef0d443825/_as/[Chris_Rice,_Craig_Simmons,_Chartered_Institute_of(b-ok.org).pdf)
- RocaSalvatella & Telefónica. (2014). Big Data and Tourism: New indicators for tourism management. Recuperado el 03 de diciembre de 2016, de <http://telefonicacatalunya.com/wp-content/uploads/2014/05/BIG-DATA-Y-TURISMO-eng-interactivo.pdf>

- Saidi, N. (2010). Forecasting Tourism in Dubai Forecasting Tourism in Duba. Recuperado el 13 de diciembre de 2016, de <http://nassersaidi.com/wp-content/uploads/2012/06/Economic-Note-8.pdf>
- Samsing, C. (2015). *Inbound Marketing y su Estado en Latinoamérica 2015*. Recuperado el 10 de setiembre de 2017, de <https://offers.hubspot.es/estado-inbound-marketing-2015>
- Schneider, A., Hommel, G., & Blettner, M. (2010). Linear Regression Analysis. *Dtsch Arztebl Int*. Recuperado el 17 de enero de 2017, de <https://www.aerzteblatt.de/int/archive/article?id=79009>
- Shareef, R., & McAleer. (2005). *Modelling multivariate tourism demand and Volatility in Small Tourism Economies*. Recuperado el 03 de diciembre de 2016, de <https://core.ac.uk/download/pdf/6408158.pdf>
- Smith, P. (2016). Google's MIDAS Touch: Predicting UK Unemployment with Internet Search Data. *Journal of Forecasting*. Recuperado el 05 de diciembre de 2016, de <http://onlinelibrary.wiley.com/doi/10.1002/for.2391/full>
- Solimar International. (s.f). Recuperado el 10 de setiembre de 2017, de <http://www.solimarinternational.com/resources-page/e-books>
- Song, H., & Han , L. (2017). *Predicting Tourist Demand Using Big Data*. Recuperado el 11 de setiembre de 2017, de [http://www.springer.com/cda/content/document/cda\\_downloaddocument/9783319442624-c1.pdf?SGWID=0-0-45-1592857-p180195151](http://www.springer.com/cda/content/document/cda_downloaddocument/9783319442624-c1.pdf?SGWID=0-0-45-1592857-p180195151).
- Song, H., & Li, G. (2008). Tourism demand modelling and forecasting—A review of recent research. Recuperado el 11 de febrero de 2017, de <https://core.ac.uk/download/pdf/397016.pdf>
- Song, H., & Liu, H. (2016). Predicting Tourist Demand Using Big Data. En *Analytics in Smart Tourism Design* (págs. 13-29). Recuperado el 2017 de marzo de 2017, de [https://link.springer.com/chapter/10.1007%2F978-3-%20319-44263-%201\\_2](https://link.springer.com/chapter/10.1007%2F978-3-%20319-44263-%201_2)
- Song, H., Witt, S., Wong, K., & Wu, D. (2007). Tourism forecasting: To combine or not to combine? 28, 1068-1078. doi:<https://doi.org/10.1016/j.tourman.2006.08.003>
- StatCounter. (2017). *StatCounter Global Stats*. Recuperado el 10 de setiembre de 2017, de <http://gs.statcounter.com/search-engine-market-share>
- Stienmetz & Fesenmaier. (2015). *Traveling the Network: A Proposal for Destination Performance Metrics*. Recuperado el 20 de mayo de 2017, de <http://www.tandfonline.com/doi/abs/10.1080/15980634.2013.11434673>
- Sung, D. (2015). Title: The Effects of Google Trends on Tourism Industry in South Korea. *2nd International Conference on Big Data Analysis and Data Mining*. Recuperado el 20 de diciembre de 2016, de <http://datamining.conferenceseries.com/abstract/2015/the-effects-of-google-trends-on-tourism-industry-in-south-korea>
- Sustaeta, Á. (2014). *La importancia del SEO y el SEM en el emprendimiento*. Recuperado el 21 de mayo de 2017, de <https://repositorio.unican.es/xmlui/bitstream/handle/10902/6459/SUSTAETANAVARRROANGELMANUEL.pdf?sequence=1>

- T NEWS. (2017). *ATSA ya comercializa en su web vuelos Lima - Chachapoyas*. Recuperado el 30 de diciembre de 2017, de <http://tnews.com.pe/atsa-ya-comercializa-en-su-web-vuelos-lima-chachapoyas-desde-us249/>
- TextRazor. (s.f.). *TextRazor*. Recuperado el 13 de setiembre de 2017, de [https://www.textrazor.com/entailments\\_and\\_synonyms](https://www.textrazor.com/entailments_and_synonyms)
- Think with Google. (2016). *How Micro-Moments are Reshaping the Travel Customer Journey*. Recuperado el 11 de diciembre de 2017, de <https://www.thinkwithgoogle.com/marketing-resources/micro-moments/micro-moments-travel-customer-journey/>
- Tsujii, K., & Takahashi, M. (2014). Extraction of Accommodation Evaluation by Foreign Tourists with Text Mining. *International Journal of Trade, Economics and Finance*. Recuperado el 05 de enero de 2017, de <http://www.ijtef.org/papers/355-C00005.pdf>
- Universidad Nacional Mayor de San Marcos. (2006). *Capítulo V: Análisis de regresión y correlación lineal*. Recuperado el 22 de mayo de 2017, de [http://sisbib.unmsm.edu.pe/bibvirtualdata/libros/2006/Estadistica\\_Descrip/05\\_cap5.pdf](http://sisbib.unmsm.edu.pe/bibvirtualdata/libros/2006/Estadistica_Descrip/05_cap5.pdf)
- Universidad Tecnológica de Pereira. (s.f.). *Análisis de Correlaciones*. Recuperado el 26 de mayo de 2017, de <http://academia.utp.edu.co/seminario-investigacion-llfiles/2017/03/06a.An%C3%A1lisisDeCorrelaciones.pdf>
- Veal, A. (2006). *Research Methods for Leisure and Tourism* (Vol. tercera edición). Harlow, Essex, Inglaterra: Pearson Education. Recuperado el 02 de noviembre de 2017
- Vosen, S. (2011). Forecasting private consumption: survey-based indicators vs. Google trends. *Journal of Forecasting*. Recuperado el 10 de diciembre de 2016, de <http://onlinelibrary.wiley.com/doi/10.1002/for.1213/abstract>
- Voyant Tools. (s.f.). *Voyant Tools Documentation*. Recuperado el 11 de setiembre de 2017, de <http://docs.voyant-tools.org/start/>
- Voyant Tools. (s.f.). *Voyant Tools Help*. Recuperado el 02 de noviembre de 2017, de <https://voyant-tools.org/docs/#!/guide/languages>
- Wladyka, D. (2017). Queries to Google Search as Predictors of Migration Flows from Latin America to Spain. *Journal of Population and Social Studies*, 25(4), 312-327. doi:10.25133
- YouTube. (2017). *YouTube*. Recuperado el 14 de setiembre de 2017, de <https://www.youtube.com/intl/es/yt/advertise/>
- Zeynalov, A. (2014). Nowcasting Tourist Arrivals to Prague:. Recuperado el 15 de diciembre de 2016, de [https://mpira.ub.uni-muenchen.de/60945/1/MPRA\\_paper\\_60945.pdf](https://mpira.ub.uni-muenchen.de/60945/1/MPRA_paper_60945.pdf)
- Zheng Xiang, D. R. (2016). *Big Data Analytics, Tourism Design and Smart Tourism*. Recuperado el 22 de febrero de 2017, de [https://link.springer.com/chapter/10.1007/978-3-319-44263-1\\_17](https://link.springer.com/chapter/10.1007/978-3-319-44263-1_17)

## ANEXOS

### Anexo 1

Tabla A1.

*Flujo de vacacionistas nacionales y extranjeros que llegaron al Complejo Arqueológico de Kuélap en 2010-2016*

	2010			2011			2012			2013			2014			2015			2016		
	Nacional	Extranjero	Total	Nacional	Extranjero	Total															
Enero	1 091	357	1 448	1 613	479	2 092	1 909	398	2 307	2 420	572	2 992	2 986	528	3 514	2 827	649	3 476	3 255	526	3 781
Febrero	771	523	1 294	1 082	436	1 518	907	387	1 294	1 907	356	2 263	2 230	453	2 683	1 482	414	1 896	2 569	763	3 332
Marzo	445	443	888	764	424	1 188	533	424	957	1 045	368	1 413	1 145	451	1 596	861	331	1 192	2 342	667	3 009
Abril	1 134	448	1 582	1 162	460	1 622	1 191	552	1 743	622	460	1 082	1 969	520	2 489	1 017	405	1 422	875	587	1 462
Mayo	584	401	985	736	420	1 156	837	506	1 343	1 277	537	1 814	1 906	491	2 397	1 008	557	1 565	2 096	733	2 829
Junio	722	458	1 180	810	404	1 214	1 356	445	1 801	1 453	612	2 065	1 557	537	2 094	1 531	737	2 268	1 745	848	2 593
Julio	2 501	965	3 466	3 659	889	4 548	2 815	774	3 589	3 030	736	3 766	4 597	882	5 479	4 522	837	5 359	6 243	1 997	8 240
Agosto	2 267	735	3 002	2 757	810	3 567	3 612	583	4 195	4 162	824	4 986	3 612	788	4 400	4 902	968	5 870	5 803	1 176	6 979
Septiembre	1 660	593	2 253	2 656	660	3 316	1 701	527	2 228	2 029	526	2 555	1 962	672	2 634	2 761	603	3 364	4 370	724	5 094
Octubre	1 901	471	2 372	2 404	464	2 868	2 023	550	2 573	2 883	521	3 404	3 260	668	3 928	3 593	755	4 348	5 316	1 009	6 325
Noviembre	3 045	372	3 417	3 359	497	3 856	2 910	525	3 435	3 975	416	4 391	5 380	561	5 941	2 105	582	2 687	6 381	827	7 208
Diciembre	1 420	389	1 809	2 053	433	2 486	2 111	384	2 495	2 392	372	2 764	2 438	553	2 991	2 483	455	2 938	4 402	756	5 158
<b>Total</b>	<b>17 541</b>	<b>6 155</b>	<b>23 696</b>	<b>23 055</b>	<b>6 376</b>	<b>29 431</b>	<b>21 905</b>	<b>6 055</b>	<b>27 960</b>	<b>27 195</b>	<b>6 300</b>	<b>33 495</b>	<b>33 042</b>	<b>7 104</b>	<b>40 146</b>	<b>29 092</b>	<b>7 293</b>	<b>36 385</b>	<b>45 397</b>	<b>10 613</b>	<b>56 010</b>

Fuente: Ministerio de Cultura - Dirección Desconcentrada de Cultura - Amazonas

Elaboración: MINCETUR (2017)

## Anexo 2

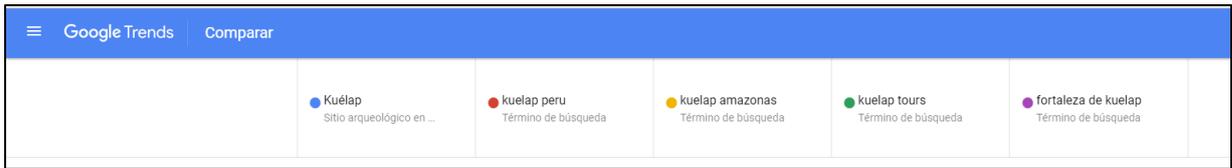


Figura A1. Ejemplo de uso de Google Trends para capturar información bajo los términos “kuelap”, “kuelap peru”, “kuelap amazonas”, “kuelap tours” y “fortaleza de kuelap”.

Fuente: Google Trends

## Anexo 3

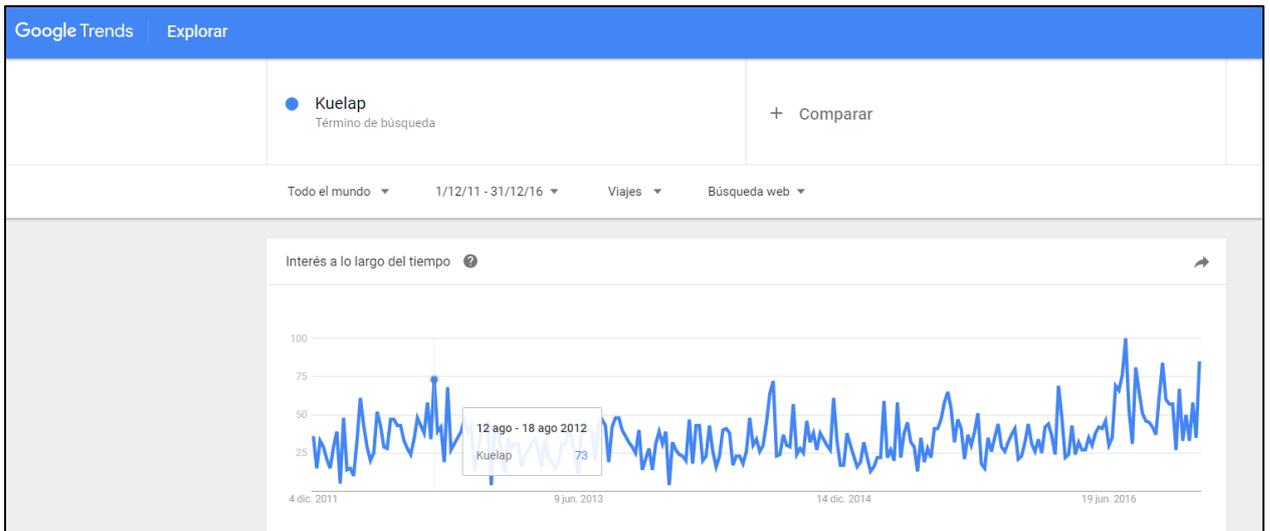
The image shows a screenshot of the Microsoft Excel interface. The title bar reads 'multiTimeline - Excel'. The ribbon includes 'ARCHIVO', 'INICIO', 'INSERTAR', 'DISEÑO DE PÁGINA', 'FÓRMULAS', 'DATOS', 'REVISAR', and 'VISTA'. The 'INICIO' ribbon is active, showing options for font (Calibri, size 11), text alignment (Ajustar texto, Combinar y centrar), and number formatting (Formato condicional, Dar formato como tabla). The spreadsheet area shows a table with the following data:

Semana	Kuelap: (Todo el mundo)
04/12/2011	21
11/12/2011	31
18/12/2011	16
25/12/2011	23
01/01/2012	21
08/01/2012	25
15/01/2012	30
22/01/2012	20
29/01/2012	29
05/02/2012	10
12/02/2012	24
19/02/2012	24
26/02/2012	5
04/03/2012	29
11/03/2012	38

Figura A2. Ejemplo de la vista preliminar de la base de datos semanal generada por Google bajo el término “kuelap”

Fuente: Google Trends

## Anexo 4



*Figura A3.* Resultados por semana en búsquedas web del término “Kuélap”. En la captura del término “kuelap” para la segunda semana de agosto del 2012 se obtuvo un índice de búsquedas de 73. De esta forma es posible adicionar los resultados semana a semana y obtener indicadores mensuales.

Fuente: Google Trends

## Anexo 5



*Figura A4.* Mapa mundial de las búsquedas con el término “kuelap”

Fuente: Google Trends

## Anexo 6



Figura A5. Resultados por semana de búsquedas de imágenes del término “kuelap”. En la captura para la tercera semana de noviembre del 2014, se obtuvo un índice de búsquedas mediante imágenes de 64 bajo el término “kuelap”. De esta forma, es posible adicionar los resultados semana a semana y obtener indicadores mensuales.

Fuente: Google Trends

## Anexo 7

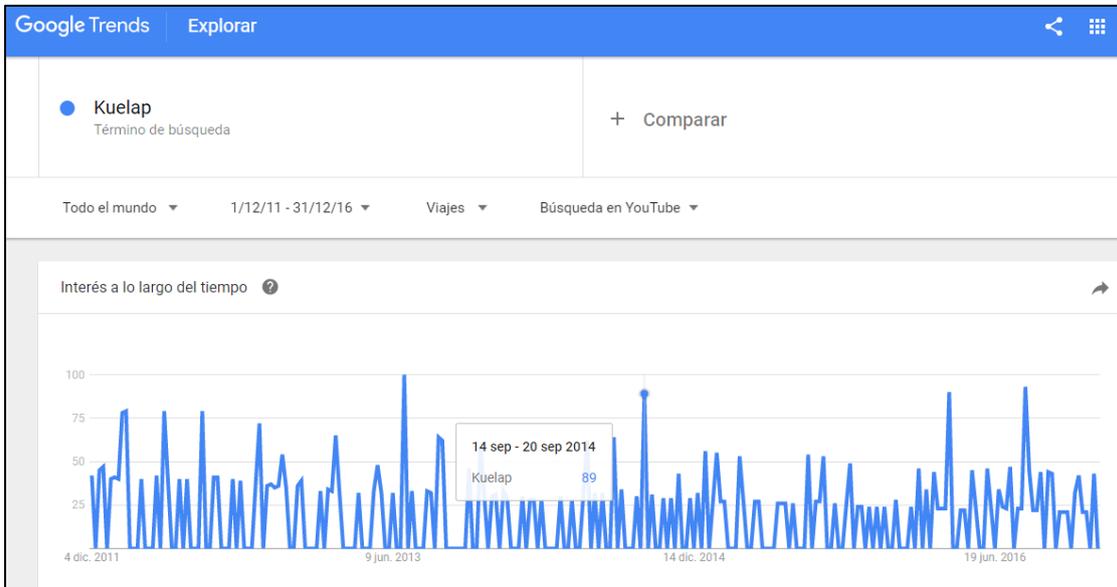
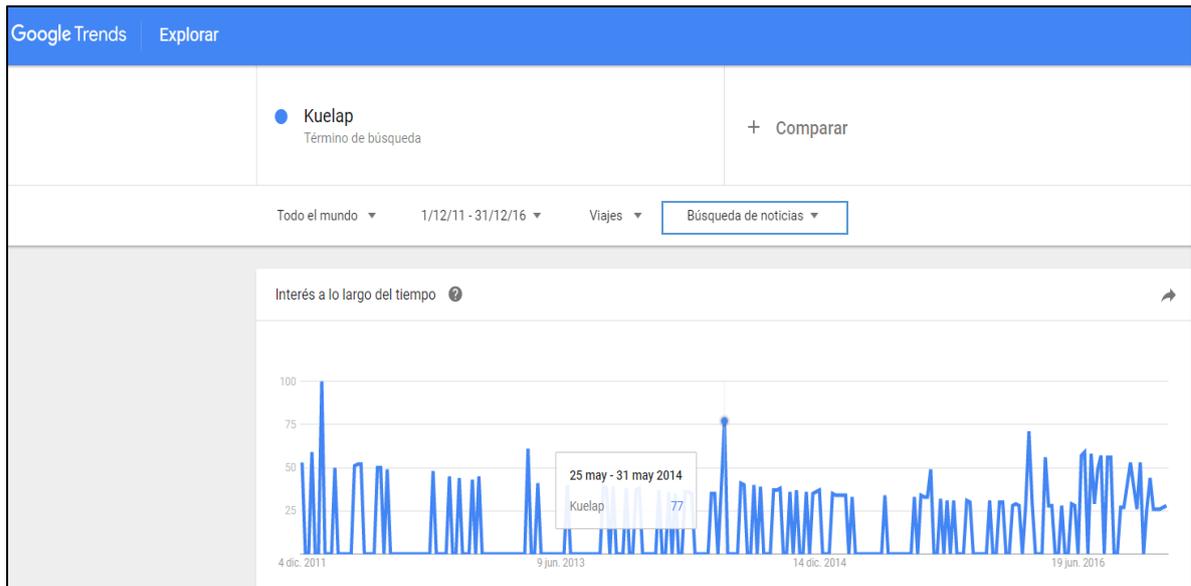


Figura A6. Resultados por semana de búsquedas en YouTube del término “kuelap”. En la captura para la tercera semana de setiembre del 2014 se obtuvo un índice de búsquedas de 89 en YouTube bajo el término “kuelap”

Fuente: Google Trends

## Anexo 8



*Figura A7.* Resultados por semana de búsquedas de noticias del término “kuelap”. En la captura para la última semana de mayo del 2014 se obtuvo un índice de búsquedas de 77 en noticias bajo el término “kuelap”

Fuente: Google Trends

## Anexo 9

### FORMATO DE CONSENTIMIENTO INFORMADO

**Título del Estudio:** Estudio Exploratorio del Uso de Indicadores de Big Data de Google como Variable en el Modelamiento del Flujo Mensual de Visitantes al Complejo Arqueológico de Kuélap en el Periodo 2011-2016

**Investigadores:** Jhoceline Paola León      **E-mail:** paola.leon1107@gmail.com  
Gutiérrez  
Brenda Stephanie Villar                      brendavillarortiz@gmail.com  
Ortiz

**Asesor:** José Francisco Melenez      **E-mail:** jmelenez@usil.edu.pe  
Roman

#### **Introducción**

Buenos días/tardes. Nuestros nombres son Paola León y Brenda Villar, y estamos realizando un estudio exploratorio sobre el uso de la Big Data de Google como variable en el modelamiento del flujo mensual de visitantes a Kuélap, y de esta manera obtener el grado de Licenciadas en Administración en Turismo.

Le invitamos a participar de esta investigación, así que por favor tome todo el tiempo que necesite para decidir si acepta, pues la decisión de participar o no, es completamente suya. Si decide no participar, le pedimos que marque la palabra "NO" y firme al final de este formato.

Si tiene alguna pregunta, siéntase libre de hacérsela a las investigadoras de este estudio.

#### **Propósito del Estudio**

El propósito de este estudio es conocer distintas opiniones para determinar si los indicadores de Big Data de Google son válidos para la mejora en la precisión de un modelo de proyección turística, y éste a su vez, pueda aportar beneficios al diseño de estrategias de marketing online de un atractivo en crecimiento como el Complejo Arqueológico de Kuélap.

#### **Plan de Estudio**

Le hemos pedido participar de este estudio debido a que lo consideramos un experto en materia de prospectiva turística/marketing turístico/Big Data; por lo que los aspectos a tratar en esta entrevista serán relacionados a dicho tema.

Si usted decide formar parte de esta investigación, solamente le realizaremos una entrevista de aproximadamente 30 minutos en la locación donde gentilmente puede atendernos. Durante la entrevista, se le harán preguntas relacionadas al Big Data, proyección turística y marketing, las cuales serán grabadas si usted lo consiente.

Usted es libre de detener su participación en cualquier momento. Sin embargo, le pedimos avise a las investigadoras primero.

### **Beneficios**

Si está de acuerdo en formar parte de este estudio, no habrá un beneficio directo para usted. Sin embargo, la información reunida en esta investigación nos ayudará a validar diferentes aspectos de nuestro trabajo. Con gusto compartiremos los resultados de nuestro estudio con usted.

### **Confidencialidad**

Toda la información que se recoja será confidencial y no se usará para ningún otro propósito fuera de los de esta investigación académica; siendo las dos investigadoras del estudio las únicas personas autorizadas para manejar la información. Asimismo, sus respuestas serán reunidas y utilizadas junto a las opiniones de otros entrevistados.

Nos gustaría incluir su nombre u otra información identificable en los resultados de este estudio; sin embargo, si usted desea estos datos se pueden mantener anónimos.

### **Derechos como Participante de la Investigación**

La participación en esta investigación es completamente voluntaria en cualquier momento. Usted puede elegir no participar del estudio o dejar de hacerlo en algún punto. Si decide no participar o dejar de hacerlo, no habrá ninguna consecuencia negativa.

Si decide no seguir participando del estudio durante la entrevista, debe informar a las investigadoras de su decisión. La información obtenida a través de su participación será incluida en el análisis y reporte final de este estudio.

### **Preguntas o Preocupaciones**

Siéntase libre de realizar las preguntas que considere necesarias durante la entrevista. Asimismo, si usted tuviera preguntas adicionales posteriores a la entrevista, puede contactar a las investigadoras y/o asesor del estudio a los e-mails proporcionados al inicio del formato de consentimiento informado, cuya copia se le entregará una vez firmada.

### **Declaración del Investigador que Obtiene el Consentimiento Informado**

Hemos explicado completamente el estudio al participante. Se explicó el propósito, los posibles beneficios, la confidencialidad y que la participación es completamente voluntaria.

Además, hemos invitado al participante a que realice preguntas y se dará respuesta a todas sus consultas durante la entrevista.

\_\_\_\_\_  
Jhoceline Paola León Gutiérrez

\_\_\_\_\_  
Brenda Stephanie Villar Ortiz

Lima, \_\_\_\_ de \_\_\_\_\_ del 2017

### **Derechos como Participante**

Entiendo toda la información incluida en este formato de consentimiento informado.

- SI**  
 **NO**

Estoy voluntariamente de acuerdo en participar de esta investigación.

- SI**  
 **NO**

Estoy de acuerdo en que se utilice mi nombre como aporte para este estudio.

- SI**  
 **NO**

Estoy de acuerdo en que la entrevista será grabada en audio como parte de la investigación.

- SI**  
 **NO**

Firma: \_\_\_\_\_

Nombre:

Lima, \_\_\_\_ de \_\_\_\_\_ del 2017

*Una vez firme este formato, recibirá una copia del mismo y las investigadoras guardarán el original para su expediente de investigación.*

## Anexo 10

### GUÍA PARA ENTREVISTA A PROFUNDIDAD

#### Presentación

Buenos días/tardes. Mi nombre es..... y estamos realizando un estudio exploratorio sobre el uso de la Big Data de Google como variable en el modelamiento del flujo mensual de visitantes al Complejo Arqueológico de Kuélap, y de esta manera obtener el grado de Licenciadas en Administración en Turismo.

El propósito de este estudio es conocer distintas opiniones para determinar si los indicadores de Big Data de Google son válidos para la mejora en la precisión de un modelo de proyección turística, y este a su vez, aporta beneficios a las estrategias de marketing online de un destino en crecimiento como Kuélap. En este sentido, siéntase libre de compartir sus ideas en este espacio, lo importante es conocer su opinión sincera. ¡Desde ya muchas gracias por su tiempo!

#### Categoría I: Big Data

- ¿Qué se le viene a la mente cuando menciono el término Big Data?
- Según su perspectiva ¿En qué sentido afecta el uso de la Big Data a nivel empresarial/organizacional?
- ¿Qué limitaciones le encuentra al uso de la Big Data?
- ¿Usted o su equipo utilizan herramientas y/o análisis de Big Data para el desarrollo de sus tareas laborales? ¿Cómo?
- ¿En qué podría beneficiar la aplicación de Big Data a la gestión del turismo en Perú?

#### Categoría II: Proyección Turística/Modelos de Análisis Predictivo

Tenemos conocimiento que existen diferentes métodos para estimar el número de llegadas hacia un destino turístico en futuros a corto o largo plazo dependiendo del método. Este tipo de métodos también son llamados modelos de análisis predictivo. En este sentido:

- ¿En qué podría beneficiar a la gestión turística el uso de los modelos para proyectar demanda?
- ¿Cuáles serían las desventajas?
- ¿Qué variables considera que se deben usar para desarrollar estos modelos de proyección?

#### Categoría III: Google Trends

- ¿Ha escuchado de la herramienta Google Trends del buscador Google?

*(Mencionar el concepto sólo si es necesario, de lo contrario, pasar a las preguntas)*

Google Trends es una herramienta de Big Data, la cual permite observar y comparar el nivel de búsqueda en Google sobre un determinado término o keyword durante un periodo de tiempo específico, utilizando para ello, valores relativos en una escala de 0 a 100.

Conociendo el concepto de esta herramienta:

- ¿En qué podría beneficiar el uso de Google Trends a la gestión de marketing online de un atractivo turístico peruano?
- ¿Cuáles serían las desventajas de su uso?
- ¿Cómo beneficiaría el uso de esta herramienta en un atractivo como Kuélap?
- ¿Y cuáles serían las desventajas?

Google Trends proporciona una serie de índices relacionados a las búsquedas en la web de Google, búsquedas de imágenes, búsquedas en Youtube y búsquedas de noticias. Algunos autores han utilizado los índices de búsquedas web proporcionados por esta herramienta para desarrollar modelos de predicción a corto plazo en un atractivo turístico. Conociendo estas facilidades que brinda Google Trends:

- ¿Usted considera que el índice de las búsquedas web sobre Kuélap representa una variable relevante para mejorar la predicción de las llegadas de turistas? ¿Por qué?
  - ¿Usted considera que el índice de las búsquedas de imágenes sobre Kuélap representa una variable relevante para mejorar la predicción de las llegadas de turistas? ¿Por qué?
  - ¿Usted considera que el índice de las búsquedas en YouTube sobre Kuélap representa una variable relevante para mejorar la predicción de las llegadas de turistas? ¿Por qué?
  - ¿Usted considera que el índice de las búsquedas de noticias sobre Kuélap representa una variable relevante para mejorar la predicción de las llegadas de turistas? ¿Por qué?
- ¿Encuentra algún beneficio en estos indicadores para mejorar las estrategias de marketing online en un atractivo turístico peruano?
  - ¿Cuáles serían las desventajas?

Ahora bien, para terminar, nos interesa conocer la relación entre los modelos de predicción y los posibles impactos que podrían tener en el desarrollo e implementación de estrategias de marketing online:

- ¿Qué beneficios aporta un modelo de proyección que utiliza variables electrónicas al marketing online?
- ¿Qué desventajas le encuentra?

¿Tiene algún otro comentario adicional?

¡Muchas Gracias por su valioso tiempo!

## Anexo 11

Tabla A2.

Plantilla en Excel para procesar los datos y exportar a SPSS

TOTAL DE ARRIBOS ACTUAL	TOTAL DE ARRIBOS HACE UN MES	TOTAL DE ARRIBOS HACE UN AÑO	LOGARITMO - TOTAL DE ARRIBOS ACTUAL	LOGARITMO - TOTAL DE ARRIBOS HACE UN MES	LOGARITMO - TOTAL DE ARRIBOS HACE UN AÑO	ÍNDICE BÚSQUEDAS GOOGLE WEB	ÍNDICE BÚSQUEDAS GOOGLE IMÁGENES	ÍNDICE BÚSQUEDAS GOOGLE YOUTUBE	ÍNDICE BÚSQUEDAS GOOGLE YOUTUBE
			LogY (t)	LogY(t-1)	LogY(t-12)	G(t)	N(t)	I(t)	Y(t)

## Anexo 12



UNIVERSIDAD  
**SAN IGNACIO  
DE LOYOLA**

Av. La Fontana 750  
La Molina, Lima - Perú  
tel (511) 317 1000  
www.usil.edu.pe

**N° 073-2017/TURISMO**

Lima, 20 de julio de 2017

Señora  
ANDREA MARTÍNEZ BERTRAMINI  
Consultor Turismo, Mercadeo y Gestión Pública  
**KM CERO**  
Presente.-

Estimada señora Martínez:

Tengo el agrado de dirigirme a usted para presentarle a las Señoritas Bachilleres de la Carrera de Administración en Turismo de la Universidad San Ignacio de Loyola:

**LEÓN GUTIÉRREZ, Jhoceline Paola**  
**VILLAR ORTIZ, Brenda Stephanie**

Las Srtas. León y Villar están realizando su tesis de grado sobre el Uso de Big Data de Google para el Modelamiento del Flujo Mensual de Visitantes al Complejo Arqueológico de Kuélap. Por ello, le solicitaríamos concederles una cita para que le puedan realizar una entrevista.

Los contactos para realizar las coordinaciones del caso son:  
jmelenez@usil.edu.pe (José Francisco Melenez) / paola.leon1107@gmail.com (Paola León Gutiérrez) / brendavillarortiz16@gmail.com (Brenda Villar Ortiz).

Sin otro particular, aprovecho la ocasión para enviarle mis cordiales saludos.

Atentamente,

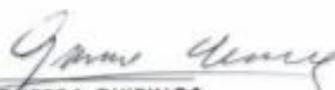
  
**DIANA GUERRA CHIRINOS**  
Directora  
Carrera de Administración en Turismo



Figura A8. Ejemplo de carta enviada para solicitud de entrevista a profundidad

## Anexo 13



Figura A9. Expansión de los resultados a nivel mundial y por periodo de tiempo (enero 2011 – diciembre 2016) .

## Anexo 14

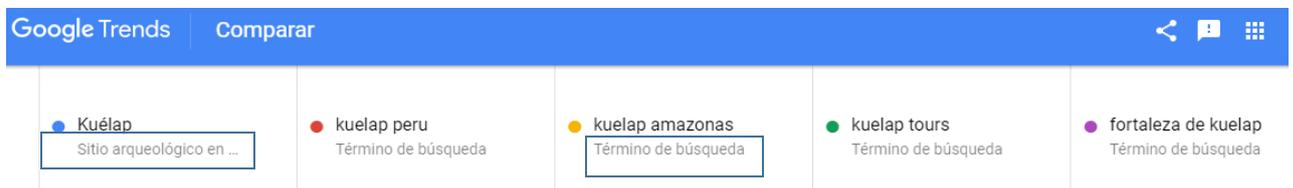


Figura A10. Función Autocompletar de Google Trends para los términos “kuelap”, “kuelap peru”, “kuelap amazonas”, “kuelap tours” y “fortaleza de kuelap”.

## Anexo 15

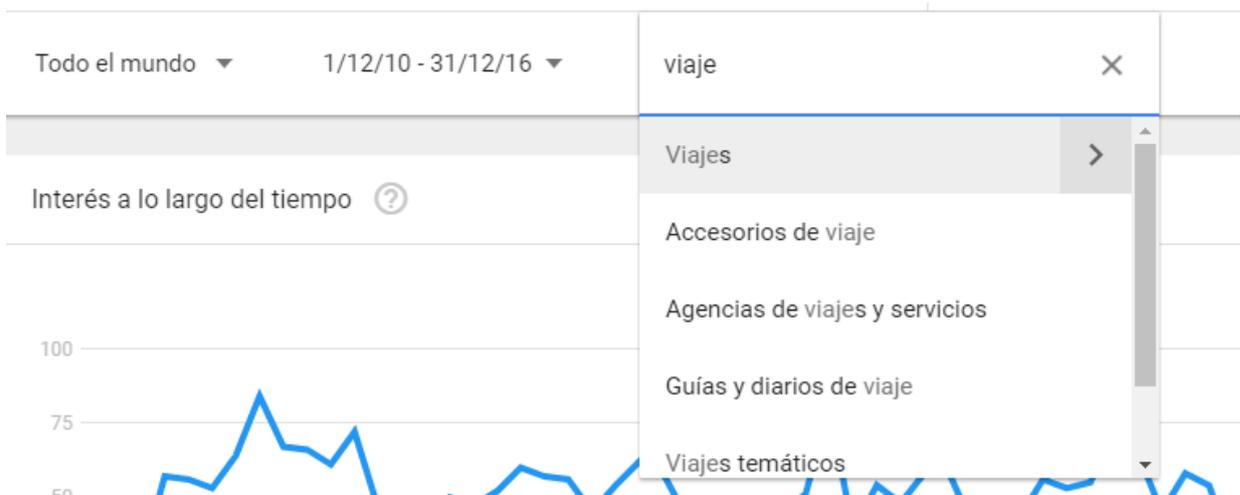


Figura A11. Restricción de los resultados a la categoría Viajes

## Anexo 16

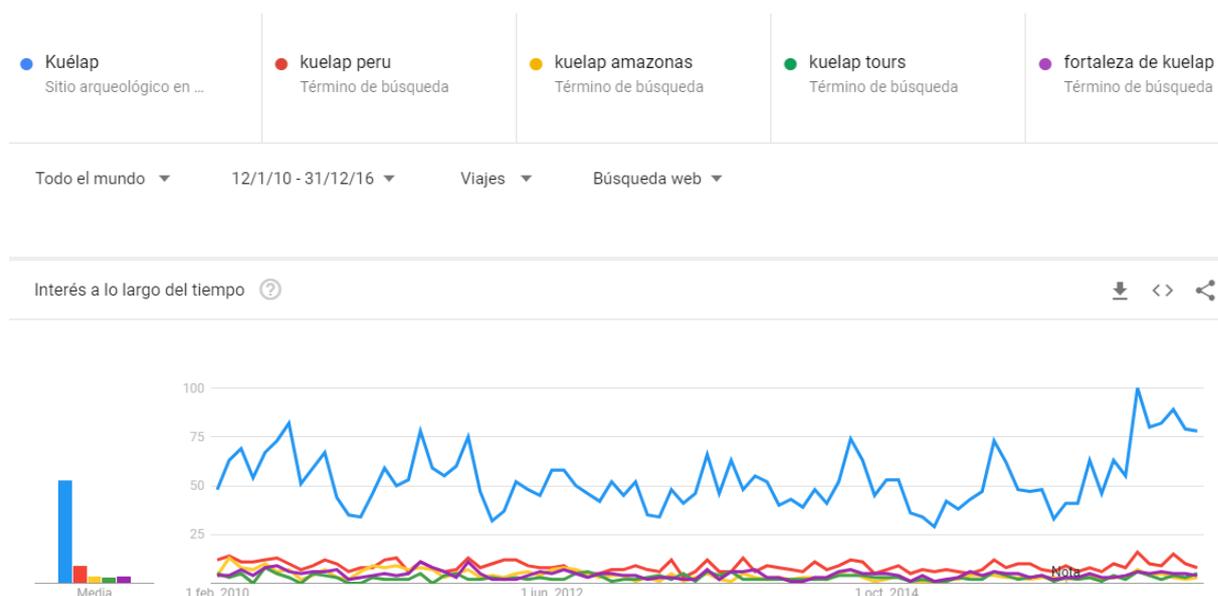


Figura A12. Gráfica de Resultados en Google Trends – términos “kuelap”, “kuelap peru”, “kuelap amazonas”, “kuelap tours” y “fortaleza de kuelap”. Donde el término “kuelap” está representado por la línea azul, el término “kuelap peru” por la línea roja, el término “kuelap amazonas” por la línea amarilla, el término “kuelap tours” por la línea verde y el término “fortaleza de kuelap” por la línea morada.

## Anexo 17

Tabla A3.

Ejemplo de Índices de Búsqueda Web arrojado por Google Trends para los términos “kuelap”, “kuelap peru”, “kuelap amazonas”, “kuelap tours” y “fortaleza de kuelap”.

Categoría: Viajes					
Mes	kuelap: (Todo el mundo)	kuelap peru: (Todo el mundo)	kuelap amazonas: (Todo el mundo)	kuelap tours: (Todo el mundo)	fortaleza de kuelap: (Todo el mundo)
2010-12	49	13	3	5	7
2011-01	36	6	2	0	4
2011-02	33	8	4	0	3
2011-03	49	9	8	4	3
2011-04	62	10	8	3	5
2011-05	55	13	8	3	5
2011-06	50	7	6	2	6
2011-07	78	12	10	4	8
2011-08	58	10	7	6	7
2011-09	54	6	6	4	6
2011-10	62	8	6	5	3
2011-11	74	14	6	2	

## Anexo 18

Tabla A4.

Sumatoria de términos “kuelap”, “kuelap peru”, “kuelap amazonas”, “kuelap tours” y “fortaleza de kuelap” para el periodo 2011-2016

Categoría: Viajes

Locación: Todo el Mundo

BUSQUEDAS WEB							BUSQUEDA DE NOTICIAS						
Mes	kuelap	kuelap peru	kuelap amazonas	kuelap tours	fortaleza de kuelap	TOTAL	Mes	kuelap	kuelap peru	kuelap amazonas	kuelap tours	fortaleza de kuelap	TOTAL
01/11	42	8	3	2	5	60	01/11	57	0	0	0	0	57
02/11	29	8	5	7	4	53	02/11	0	0	0	0	0	0
03/11	55	9	10	6	4	84	03/11	0	0	0	0	0	0
04/11	59	8	5	0	4	76	04/11	0	0	0	0	0	0
05/11	56	10	11	4	4	85	05/11	51	0	0	0	0	51
06/11	56	11	6	0	6	79	06/11	0	0	0	0	0	0
07/11	88	13	10	9	9	129	07/11	0	0	0	0	0	0
08/11	64	11	4	4	4	87	08/11	0	0	0	0	0	0
09/11	59	11	6	3	6	85	09/11	0	0	0	0	0	0
10/11	59	5	5	3	6	78	10/11	0	0	0	0	0	0
11/11	65	14	5	0	9	93	11/11	0	0	0	0	0	0
12/11	47	7	3	2	6	65	12/11	50	0	0	0	0	50
01/12	34	11	2	2	3	52	01/12	0	0	0	0	0	0
02/12	39	10	2	4	4	59	02/12	0	0	0	0	0	0
03/12	51	13	6	2	2	74	03/12	0	0	0	0	0	0
04/12	54	9	7	3	3	76	04/12	0	0	0	0	0	0
05/12	56	11	6	3	3	79	05/12	0	0	0	0	0	0
06/12	58	8	6	2	6	80	06/12	0	0	0	0	0	0
07/12	57	8	6	3	8	82	07/12	0	0	0	0	0	0

08/12	49	8	4	5	5	71	08/12	0	0	0	0	0	0
09/12	48	7	2	5	4	66	09/12	0	0	0	0	0	0
10/12	46	6	4	3	4	63	10/12	0	0	0	0	0	0
11/12	56	9	6	0	4	75	11/12	29	0	0	0	0	29
12/12	51	7	5	2	4	69	12/12	0	0	0	0	0	0
01/13	46	8	3	3	3	63	01/13	0	0	0	0	0	0
02/13	38	6	4	4	3	55	02/13	0	0	0	0	0	0
03/13	40	5	2	4	4	55	03/13	31	0	0	0	0	31
04/13	41	10	4	3	6	64	04/13	0	0	0	0	0	0
05/13	37	5	3	1	3	49	05/13	0	0	0	0	0	0
06/13	43	10	5	1	5	64	06/13	0	0	0	0	0	0
07/13	63	9	4	4	6	86	07/13	0	0	0	0	0	0
08/13	49	6	1	4	2	62	08/13	0	0	0	0	0	0
09/13	62	8	3	2	4	79	09/13	0	0	0	0	0	0
10/13	45	11	4	2	4	66	10/13	26	0	0	0	0	26
11/13	54	7	4	2	6	73	11/13	0	0	0	0	0	0
12/13	54	8	4	1	4	71	12/13	29	0	0	0	0	29
01/14	37	6	4	3	3	53	01/14	26	0	0	0	0	26
02/14	47	7	2	3	2	61	02/14	27	0	0	0	0	27
03/14	43	8	2	2	2	57	03/14	0	0	0	0	0	0
04/14	58	13	3	2	4	80	04/14	0	0	0	0	0	0
05/14	43	6	1	2	3	55	05/14	22	0	0	0	0	22
06/14	51	8	4	3	4	70	06/14	0	0	0	0	0	0
07/14	74	11	6	2	6	99	07/14	23	0	0	0	0	23
08/14	61	14	3	3	6	87	08/14	23	0	0	0	0	23
09/14	53	7	3	4	4	71	09/14	0	0	0	0	0	0

10/14	51	7	1	2	5	66	10/14	44	0	0	0	22	66
11/14	50	8	2	3	4	67	11/14	0	0	0	0	0	0
12/14	38	6	2	2	2	50	12/14	0	0	0	0	0	0
01/15	37	5	1	3	3	49	01/15	22	0	0	0	0	22
02/15	37	5	2	1	3	48	02/15	0	0	0	0	0	0
03/15	45	10	2	3	4	64	03/15	0	0	0	0	0	0
04/15	42	6	1	4	3	56	04/15	24	0	0	0	0	24
05/15	42	8	2	3	4	59	05/15	24	0	0	0	0	24
06/15	50	7	3	2	3	65	06/15	0	0	0	0	0	0
07/15	75	11	4	4	7	101	07/15	0	0	0	0	0	0
08/15	56	5	2	4	5	72	08/15	0	0	0	0	0	0
09/15	52	9	1	2	5	69	09/15	21	0	0	0	0	21
10/15	46	8	2	1	3	60	10/15	21	0	0	0	0	21
11/15	47	7	2	3	4	63	11/15	0	0	0	0	0	0
12/15	34	5	4	2	3	48	12/15	0	0	0	0	0	0
01/16	40	6	3	2	3	54	01/16	0	0	0	0	0	0
02/16	46	9	2	3	3	63	02/16	0	0	0	0	0	0
03/16	59	6	3	2	2	72	03/16	17	0	0	0	0	17
04/16	45	9	3	2	1	60	04/16	20	0	0	0	0	20
05/16	60	8	4	3	4	79	05/16	0	0	0	0	0	0
06/16	60	10	4	3	5	82	06/16	28	0	0	0	0	28
07/16	97	17	10	6	7	137	07/16	26	0	0	0	0	26
08/16	89	12	5	3	5	114	08/16	19	0	0	0	0	19
09/16	69	7	2	2	6	86	09/16	19	19	0	0	0	38
10/16	83	12	3	3	6	107	10/16	32	18	0	0	0	50
11/16	78	10	3	6	5	102	11/16	32	0	0	0	0	32
12/16	84	10	6	5	5	110	12/16	47	19	0	0	0	66

Categoría: Viajes

Locación: Todo el Mundo

BUSQUEDAS IMAGENES							BUSQUEDA DE YOUTUBE						
Mes	kuelap	kuelap peru	kuelap amazonas	kuelap tours	fortaleza de kuelap	TOTAL	Mes	kuelap	kuelap peru	kuelap amazonas	kuelap tours	fortaleza de kuelap	TOTAL
01/11	29	7	7	0	0	43	01/11	31	22	0	0	0	53
02/11	39	14	20	0	7	80	02/11	41	22	23	0	0	86
03/11	59	6	6	0	16	87	03/11	19	0	0	0	0	19
04/11	62	7	13	0	17	99	04/11	27	0	0	0	0	27
05/11	46	9	14	0	7	76	05/11	43	0	20	0	20	83
06/11	72	21	0	0	9	102	06/11	66	27	53	0	0	146
07/11	76	18	21	0	22	137	07/11	59	27	0	0	0	86
08/11	85	11	13	0	14	123	08/11	49	19	19	0	19	106
09/11	71	12	9	0	15	107	09/11	36	11	0	0	11	58
10/11	38	5	15	0	5	63	10/11	61	19	14	0	14	108
11/11	69	0	13	0	7	89	11/11	36	18	0	0	18	72
12/11	42	14	13	0	0	69	12/11	38	0	17	0	0	55
01/12	17	6	5	0	6	34	01/12	34	17	0	0	0	51
02/12	24	7	6	0	5	42	02/12	26	0	0	0	0	26
03/12	36	12	9	0	10	67	03/12	33	0	0	0	0	33
04/12	33	6	5	0	7	51	04/12	44	17	0	0	0	61
05/12	36	7	7	0	9	59	05/12	23	0	0	0	0	23
06/12	32	11	10	0	10	63	06/12	23	17	0	0	0	40
07/12	72	18	16	0	11	117	07/12	29	17	0	0	0	46
08/12	57	12	9	0	9	87	08/12	18	0	17	0	0	35

09/12	38	6	14	0	0	58	09/12	46	17	17	0	17	97
10/12	32	12	0	0	6	50	10/12	64	21	21	0	21	127
11/12	49	11	8	0	9	77	11/12	37	0	16	0	0	53
12/12	51	6	8	0	6	71	12/12	26	32	0	0	0	58
01/13	21	4	5	0	4	34	01/13	18	16	0	0	21	55
02/13	28	15	7	0	6	56	02/13	23	0	0	0	0	23
03/13	30	4	0	0	19	53	03/13	15	0	0	0	0	15
04/13	30	7	6	0	7	50	04/13	23	0	0	0	0	23
05/13	30	7	5	0	4	46	05/13	26	0	16	0	16	58
06/13	40	12	7	0	8	67	06/13	27	0	0	0	0	27
07/13	53	11	8	4	8	84	07/13	25	16	16	0	21	78
08/13	37	6	7	0	4	54	08/13	34	16	21	0	0	71
09/13	46	9	7	0	7	69	09/13	35	16	0	0	0	51
10/13	38	13	8	0	6	65	10/13	26	16	0	0	16	58
11/13	43	5	5	0	5	58	11/13	34	0	0	0	0	34
12/13	23	3	0	0	5	31	12/13	26	0	16	0	16	58
01/14	14	0	0	0	3	17	01/14	17	16	21	0	0	54
02/14	21	4	4	0	5	34	02/14	43	17	0	0	17	77
03/14	24	7	0	0	0	31	03/14	35	18	16	0	16	85
04/14	33	11	9	0	5	58	04/14	24	16	0	0	0	40
05/14	32	7	4	0	3	46	05/14	24	17	0	0	0	41
06/14	48	10	7	0	0	65	06/14	39	23	0	0	23	85
07/14	53	6	13	0	6	78	07/14	17	0	0	0	0	17
08/14	36	4	8	0	7	55	08/14	39	17	17	0	17	90
09/14	41	6	5	0	7	59	09/14	48	0	17	0	22	87
10/14	35	5	8	0	3	51	10/14	33	0	0	0	0	33
11/14	50	8	6	0	4	68	11/14	45	17	17	0	17	96
12/14	22	5	8	0	5	40	12/14	16	0	0	0	17	33

01/15	8	0	4	0	4	16	01/15	39	34	0	0	0	73
02/15	17	3	3	3	3	29	02/15	33	0	0	0	0	33
03/15	24	6	7	0	3	40	03/15	17	0	0	0	16	33
04/15	21	7	4	0	0	32	04/15	46	0	0	0	0	46
05/15	11	0	4	0	0	15	05/15	27	0	0	0	0	27
06/15	28	9	5	0	4	46	06/15	17	17	0	0	0	34
07/15	53	10	10	0	9	82	07/15	54	17	0	17	0	88
08/15	38	11	3	0	4	56	08/15	56	17	0	0	0	73
09/15	34	5	7	0	7	53	09/15	19	0	0	0	0	19
10/15	38	9	6	4	9	66	10/15	41	0	0	0	0	41
11/15	41	6	12	0	4	63	11/15	39	22	0	0	16	77
12/15	18	0	3	0	0	21	12/15	16	0	0	0	0	16
01/16	20	15	0	0	3	38	01/16	24	0	16	0	0	40
02/16	18	10	3	0	5	36	02/16	19	23	0	0	0	42
03/16	31	9	5	6	5	56	03/16	60	0	0	0	0	60
04/16	36	7	10	0	3	56	04/16	33	16	16	0	0	65
05/16	28	8	8	0	3	47	05/16	35	19	49	0	0	103
06/16	45	10	8	3	4	70	06/16	61	0	0	0	0	61
07/16	64	11	10	0	13	98	07/16	100	35	18	0	18	171
08/16	39	8	3	4	5	59	08/16	65	17	0	0	0	82
09/16	46	7	7	3	6	69	09/16	62	27	0	0	18	107
10/16	29	8	3	0	3	43	10/16	63	17	0	0	0	80
11/16	29	3	4	0	4	40	11/16	46	23	0	0	0	69
12/16	30	6	3	0	3	42	12/16	44	17	0	0	0	61

## Anexo 19

Tabla A5.

Plantilla en Excel para procesar los datos y exportar a SPSS con la información incluida

TOTAL DE ARRIBOS ACTUAL	TOTAL DE ARRIBOS HACE UN MES	TOTAL DE ARRIBOS HACE UN AÑO	LOGARITMO - TOTAL DE ARRIBOS ACTUAL LogY (t)	LOGARITMO - TOTAL DE ARRIBOS HACE UN MES LogY(t-1)	LOGARITMO - TOTAL DE ARRIBOS HACE UN AÑO LogY(t-12)	ÍNDICE BÚSQUEDAS GOOGLE WEB G(t)	ÍNDICE BÚSQUEDAS GOOGLE IMÁGENES I(t)	ÍNDICE BÚSQUEDAS GOOGLE YOUTUBE Y(t)	ÍNDICE BÚSQUEDAS GOOGLE NOTICIAS N(t)
2092	1809	1448	3.3205616801952400000	3.2574385668598100000	3.1607685618611300000	60	43	53	57
1518	2092	1294	3.1812717715594600000	3.3205616801952400000	3.1119342763326800000	53	80	86	0
1188	1518	888	3.0748164406451700000	3.1812717715594600000	2.9484129657786000000	84	87	19	0
1622	1188	1582	3.2100508498751400000	3.0748164406451700000	3.1992064791616600000	76	99	27	0
1156	1622	985	3.0629578340845100000	3.2100508498751400000	2.9934362304976100000	85	76	83	51
1214	1156	1180	3.0842186867392400000	3.0629578340845100000	3.0718820073061300000	79	102	146	0
4548	1214	3466	3.6578204560157000000	3.0842186867392400000	3.5398285583779000000	129	137	86	0
3567	4548	3002	3.5523031093383500000	3.6578204560157000000	3.4774106879072500000	87	123	106	0
3316	3567	2253	3.5206145218782400000	3.5523031093383500000	3.3527611917238300000	85	107	58	0
2868	3316	2372	3.4575791469957600000	3.5206145218782400000	3.3751146846922300000	78	63	108	0
3856	2868	3417	3.5861370252307900000	3.4575791469957600000	3.5336449787987600000	93	89	72	0
2486	3856	1809	3.3955011243056300000	3.5861370252307900000	3.2574385668598100000	65	69	55	50
2307	2486	2092	3.3630475945210900000	3.3955011243056300000	3.3205616801952400000	52	34	51	0
1294	2307	1518	3.1119342763326800000	3.3630475945210900000	3.1812717715594600000	59	42	26	0
957	1294	1188	2.9809119377768400000	3.1119342763326800000	3.0748164406451700000	74	67	33	0
1743	957	1622	3.2412973871099900000	2.9809119377768400000	3.2100508498751400000	76	51	61	0
1343	1743	1156	3.1280760126687200000	3.2412973871099900000	3.0629578340845100000	79	59	23	0
1801	1343	1214	3.2555137128195300000	3.1280760126687200000	3.0842186867392400000	80	63	40	0
3589	1801	4548	3.5549734583332400000	3.2555137128195300000	3.6578204560157000000	82	117	46	0
4195	3589	3567	3.6227319651647200000	3.5549734583332400000	3.5523031093383500000	71	87	35	0
2228	4195	3316	3.3479151865016900000	3.6227319651647200000	3.5206145218782400000	66	58	97	0
2573	2228	2868	3.4104397862103500000	3.3479151865016900000	3.4575791469957600000	63	50	127	0
3435	2573	3856	3.5359267413955700000	3.4104397862103500000	3.5861370252307900000	75	77	53	29
2495	3435	2486	3.3970705499594100000	3.5359267413955700000	3.3955011243056300000	69	71	58	0

2992	2495	2307	3.4759615891924200000	3.3970705499594100000	3.3630475945210900000	63	34	55	0
2263	2992	1294	3.3546845539547300000	3.4759615891924200000	3.1119342763326800000	55	56	23	0
1413	2263	957	3.1501421618485600000	3.3546845539547300000	2.9809119377768400000	55	53	15	31
1082	1413	1743	3.0342272607705500000	3.1501421618485600000	3.2412973871099900000	64	50	23	0
1814	1082	1343	3.2586372827240800000	3.0342272607705500000	3.1280760126687200000	49	46	58	0
2065	1814	1801	3.3149200559924200000	3.2586372827240800000	3.2555137128195300000	64	67	27	0
3766	2065	3589	3.5758803156806500000	3.3149200559924200000	3.5549734583332400000	86	84	78	0
4986	3766	4195	3.6977522741677500000	3.5758803156806500000	3.6227319651647200000	62	54	71	0
2555	4986	2228	3.4073909044707300000	3.6977522741677500000	3.3479151865016900000	79	69	51	0
3404	2555	2573	3.5319895514125500000	3.4073909044707300000	3.4104397862103500000	66	65	58	26
4391	3404	3435	3.6425634371043900000	3.5319895514125500000	3.5359267413955700000	73	58	34	0
2764	4391	2495	3.4415380387021600000	3.6425634371043900000	3.3970705499594100000	71	31	58	29
3514	2764	2992	3.5458017571592800000	3.4415380387021600000	3.4759615891924200000	53	17	54	26
2683	3514	2263	3.4286206726719400000	3.5458017571592800000	3.3546845539547300000	61	34	77	27
1596	2683	1413	3.2030328870147100000	3.4286206726719400000	3.1501421618485600000	57	31	85	0
2489	1596	1082	3.3960248966085900000	3.2030328870147100000	3.0342272607705500000	80	58	40	0
2397	2489	1814	3.3796680340336500000	3.3960248966085900000	3.2586372827240800000	55	46	41	22
2094	2397	2065	3.3209766773428200000	3.3796680340336500000	3.3149200559924200000	70	65	85	0
5479	2094	3766	3.7387013004347100000	3.3209766773428200000	3.5758803156806500000	99	78	17	23
4400	5479	4986	3.6434526764861900000	3.7387013004347100000	3.6977522741677500000	87	55	90	23
2634	4400	2555	3.4206157706257700000	3.6434526764861900000	3.4073909044707300000	71	59	87	0
3928	2634	3404	3.5941714791149100000	3.4206157706257700000	3.5319895514125500000	66	51	33	66
5941	3928	4391	3.7738595523766900000	3.5941714791149100000	3.6425634371043900000	67	68	96	0
2991	5941	2764	3.4758164130313200000	3.7738595523766900000	3.4415380387021600000	50	40	33	0
3476	2991	3514	3.5410797677766300000	3.4758164130313200000	3.5458017571592800000	49	16	73	22
1896	3476	2683	3.2778383330020500000	3.5410797677766300000	3.4286206726719400000	48	29	33	0
1192	1896	1596	3.0762762554042200000	3.2778383330020500000	3.2030328870147100000	64	40	33	0
1422	1192	2489	3.1528995963937500000	3.0762762554042200000	3.3960248966085900000	56	32	46	24
1565	1422	2397	3.1945143418824700000	3.1528995963937500000	3.3796680340336500000	59	15	27	24
2268	1565	2094	3.3556430502208700000	3.1945143418824700000	3.3209766773428200000	65	46	34	0
5359	2268	5479	3.7290837570436100000	3.3556430502208700000	3.7387013004347100000	101	82	88	0
5870	5359	4400	3.7686381012476100000	3.7290837570436100000	3.6434526764861900000	72	56	73	0
3364	5870	2634	3.5268559871258700000	3.7686381012476100000	3.4206157706257700000	69	53	19	21

4348	3364	3928	3.6382895354142600000	3.5268559871258700000	3.5941714791149100000	60	66	41	21
2687	4348	5941	3.4292676664331700000	3.6382895354142600000	3.7738595523766900000	63	63	77	0
2938	2687	2991	3.4680517914542400000	3.4292676664331700000	3.4758164130313200000	48	21	16	0
3781	2938	3476	3.5776066773625400000	3.4680517914542400000	3.5410797677766300000	54	38	40	0
3332	3781	1896	3.5227049927347500000	3.5776066773625400000	3.2778383330020500000	63	36	42	0
3009	3332	1192	3.4784221877400800000	3.5227049927347500000	3.0762762554042200000	72	56	60	17
1462	3009	1422	3.1649473726218400000	3.4784221877400800000	3.1528995963937500000	60	56	65	20
2829	1462	1565	3.4516329474569900000	3.1649473726218400000	3.1945143418824700000	79	47	103	0
2593	2829	2268	3.4138025167693500000	3.4516329474569900000	3.3556430502208700000	82	70	61	28
8240	2593	5359	3.9159272116971200000	3.4138025167693500000	3.7290837570436100000	137	98	171	26
6979	8240	5870	3.8437931983259100000	3.9159272116971200000	3.7686381012476100000	114	59	82	19
5094	6979	3364	3.7070589406276000000	3.8437931983259100000	3.5268559871258700000	86	69	107	38
6325	5094	4348	3.8010605298478600000	3.7070589406276000000	3.6382895354142600000	107	43	80	50
7208	6325	2687	3.8578147779710100000	3.8010605298478600000	3.4292676664331700000	102	40	69	32
5158	7208	2938	3.7124813378019200000	3.8578147779710100000	3.4680517914542400000	110	42	61	66

## Anexo 20

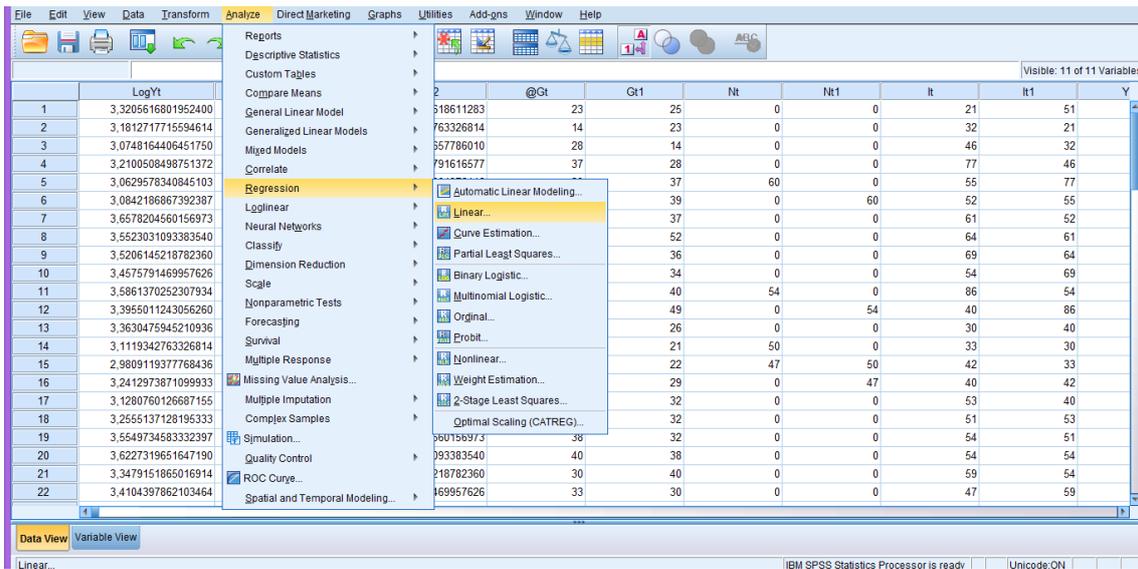


Figura A13. Herramienta SPSS: Analizar – Regresión – Linear para obtener los resultados de la regresión lineal de las ecuaciones.

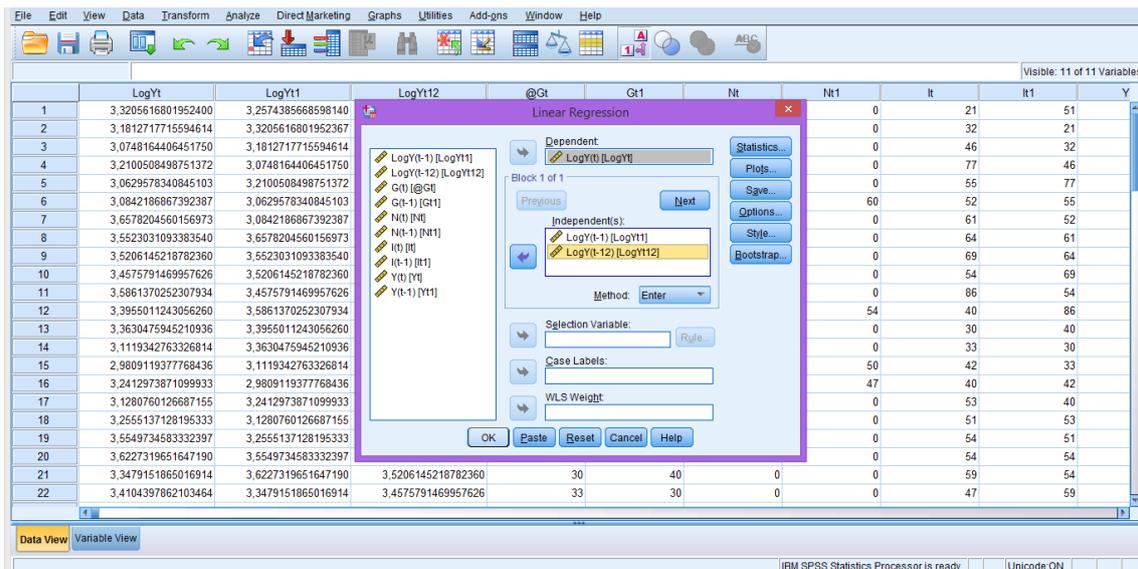


Figura A14. Selección de variables a considerar en SPSS por cada una de las nueve ecuaciones de la investigación.

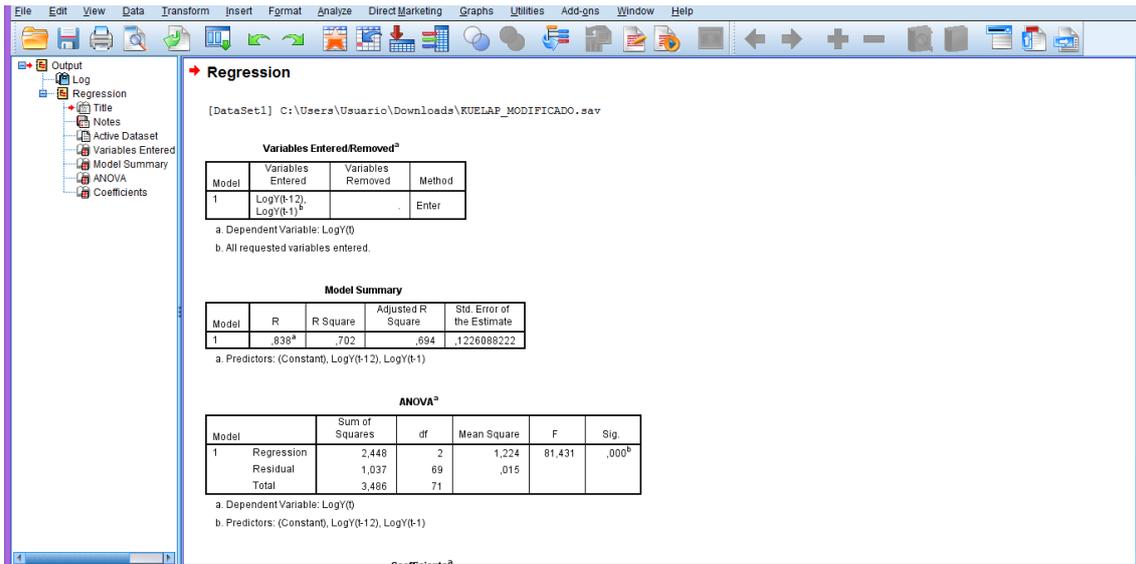


Figura A15. Resultado de la regresión lineal obtenidos del programa SPSS. Se deben seleccionar los resultados del cuadro de coeficientes.

## Anexo 21

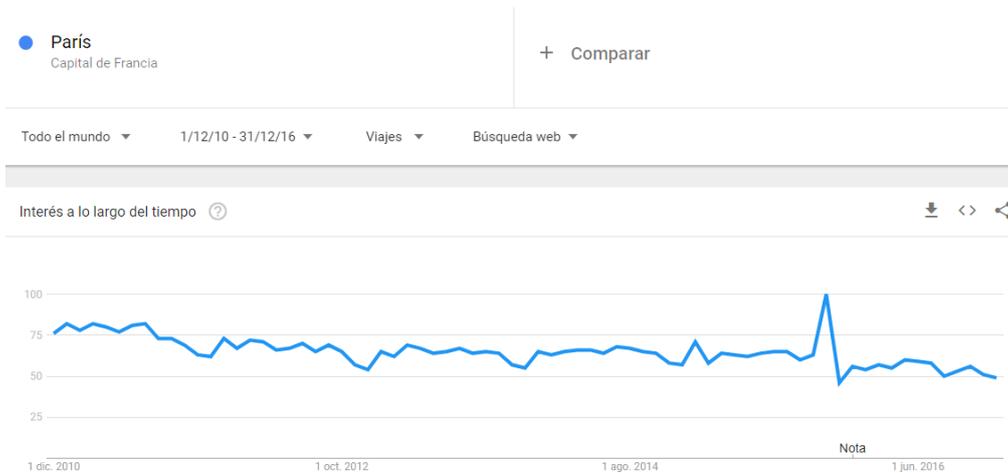


Figura A16. Índice de búsquedas web de Google Trends bajo el término "París". Las búsquedas en YouTube y noticias presentaron igual índice (100) que las búsquedas web en el mes de noviembre del 2015.

## Anexo 22

Tabla A6.

Análisis de varianza de la primera ecuación: Modelo AR-1

	Grados de libertad	Suma de cuadrados	Promedio de los cuadrados	F	Valor crítico de F
Regresión	2	2.448	1.224	81.431	<b>.0001</b>
Residuos	69	1.037	.015		
Total	71	3.486			

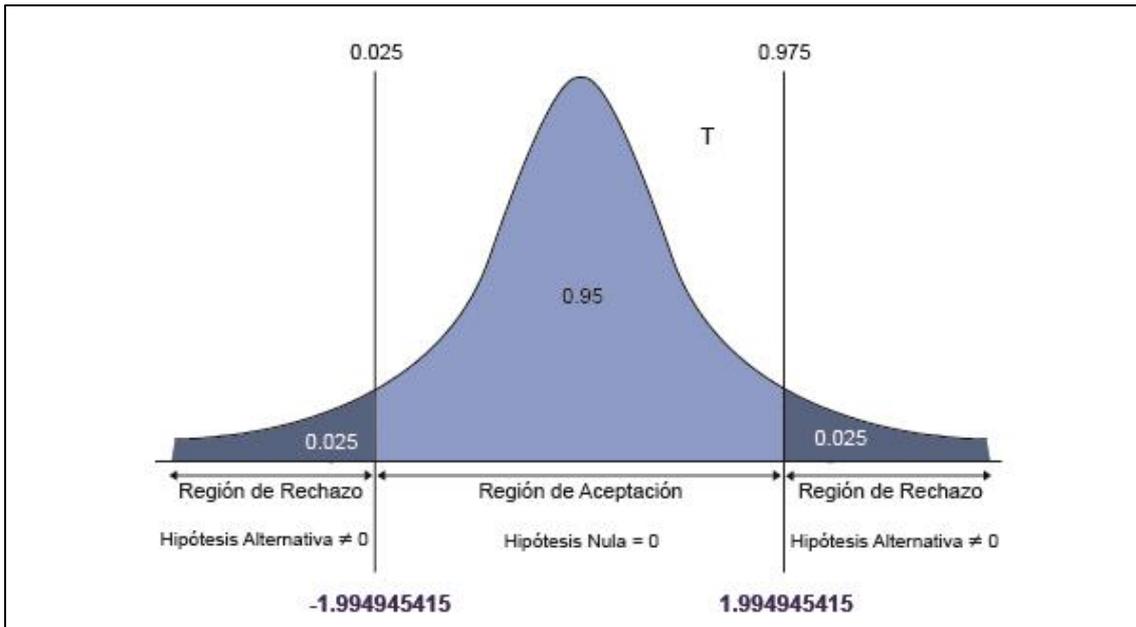


Figura A17. Gráfica t-student de la primera ecuación: Modelo AR-1

Se realizó la prueba matemática t-student para hallar el estadístico-t de la primera ecuación (Modelo AR-1). Para esto, se realizaron las operaciones: =INV.T(0.025, GL) e =INV.T(0.975, GL) en Microsoft Excel; donde GL son los grados de libertad de los residuos obtenidos en la tabla A6. Como resultado se obtuvo  $-1.994945415$  y  $1.994945415$ , que son los rangos limitantes de la región de aceptación. Esto significa que los valores del estadístico-t de la primera ecuación que se encuentran fuera de estos rangos, deberían rechazar la hipótesis nula, y aquellos que se encuentran dentro, deberían aprobarla. Para el caso de la presente investigación, la aprobación de la hipótesis nula denota que las variables involucradas no presentan significancia para el modelo predictivo.

Tabla A7.

Análisis de varianza de la segunda ecuación: Modelo Choi y Varian (2009)

	Grados de libertad	Suma de cuadrados	Promedio de los cuadrados	F	Valor crítico de F
Regresión	3	2.727	.909	81.449	<b>.0001</b>
Residuos	68	.759	.011		
Total	71	3.486			

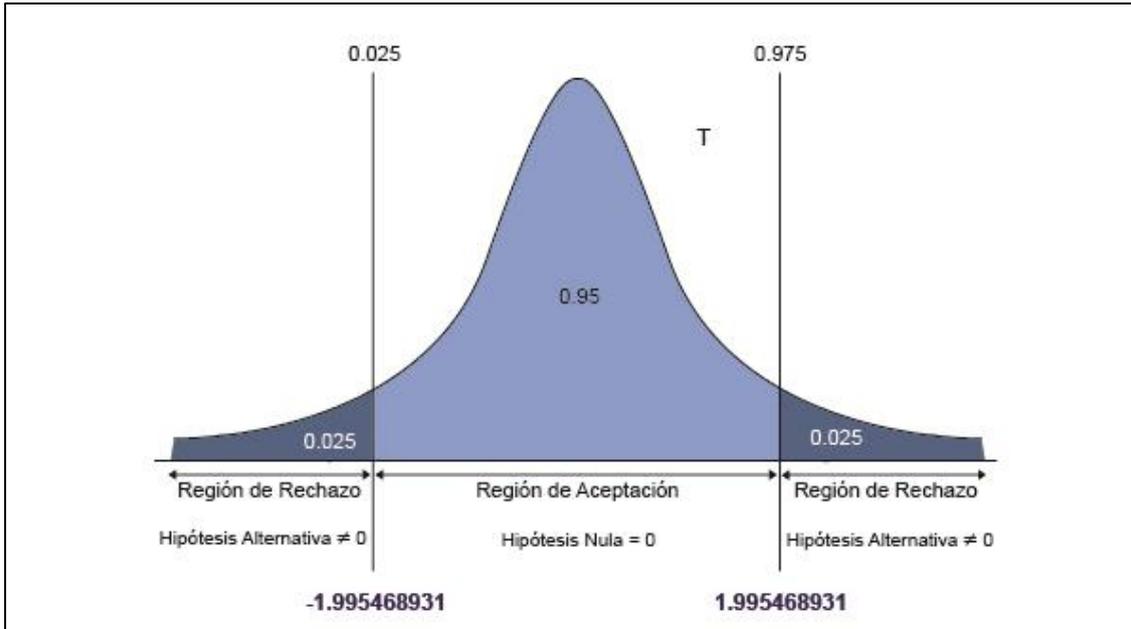


Figura A18. Gráfica t-student de la segunda ecuación: Modelo Choi y Varian (2009)

Se realizó la prueba matemática t-student para hallar el estadístico-t de la segunda ecuación (Modelo Choi & Varian). Para esto, se realizaron las operaciones: =INV.T(0.025, GL) e =INV.T(0.975, GL) en Microsoft Excel; donde GL son los grados de libertad de los residuos obtenidos en la tabla A7. Como resultado se obtuvieron las puntuaciones: -1.995468931 y 1.995468931, que son los rangos limitantes de la región de aceptación. Ello significa que los valores del estadístico-t de la segunda ecuación que se encuentran fuera de estos rangos, deberían rechazar la hipótesis nula, y aquellos que se encuentran dentro, deberían aprobarla. Para el caso de la presente investigación, la aprobación de la hipótesis nula denota que las variables involucradas no presentan significancia para el modelo predictivo.

Tabla A8.

Análisis de varianza de la tercera ecuación: Modelo Choi y Varian (2009) + Gimagen (t) + GYouTube(t) + Gnoticias(t)

	Grados de libertad	Suma de cuadrados	Promedio de los cuadrados	F	Valor crítico de F
Regresión	6	2.770	.462	41.920	<b>.0001</b>
Residuos	65	.716	.011		
Total	71	3.486			

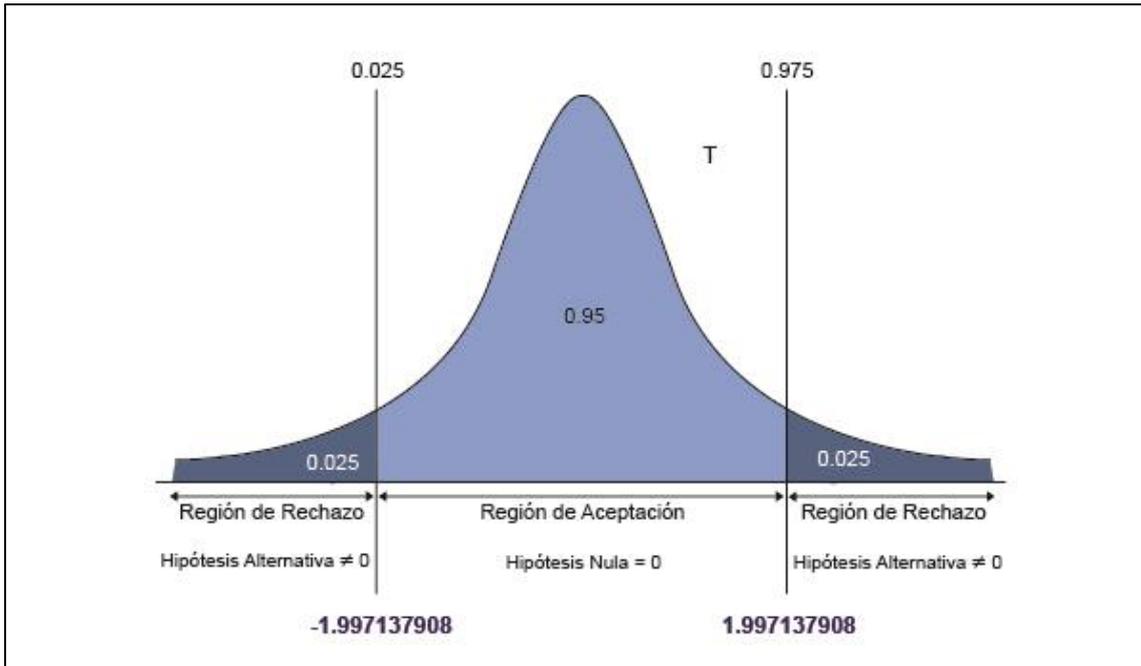


Figura A19. Gráfica t-student de la tercera ecuación: Modelo Choi y Varian (2009) + Gimagen (t) + GYouTube(t) + Gnoticias(t)

Se realizó la prueba matemática t-student para hallar el estadístico-t de la tercera ecuación. Para esto, se realizaron las operaciones: =INV.T(0.025, GL) e =INV.T(0.975, GL) en Microsoft Excel; donde GL son los grados de libertad de los residuos obtenidos en la tabla A8. Como resultado se obtuvo -1.997137908 y 1.997137908, que son los rangos limitantes de la región de aceptación. Estas puntuaciones limitantes significan que los valores del estadístico-t de la tercera ecuación que se encuentran fuera de estos rangos, deberían rechazar la hipótesis nula, y aquellos que se encuentran dentro, deberían aprobarla. Para el caso de la presente investigación, la aprobación de la hipótesis nula denota que las variables involucradas no presentan significancia para el modelo predictivo.

Tabla A9.

Análisis de varianza de la cuarta ecuación: Modelo Choi y Varian (2009) + Gimagen (t) + GYouTube(t)

	Grados de libertad	Suma de cuadrados	Promedio de los cuadrados	F	Valor crítico de F
Regresión	5	2.761	.552	50.292	<b>.0001</b>
Residuos	<b>66</b>	.725	.011		
Total	71	3.486			

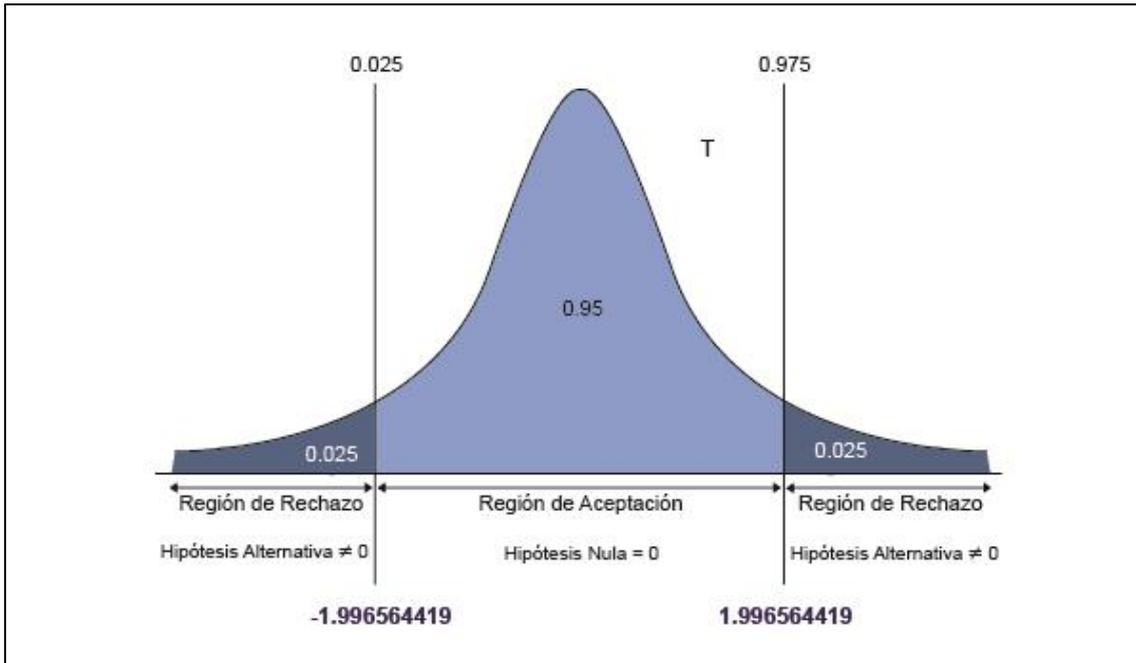


Figura A20. Gráfica t-student de la cuarta ecuación: Modelo Choi y Varian (2009) + Gimagen (t) + GYouTube(t)

Se realizó la prueba matemática t-student para hallar el estadístico-t de la cuarta ecuación. Para esto, se realizaron las operaciones: =INV.T(0.025, GL) e =INV.T(0.975, GL) en Microsoft Excel; donde GL son los grados de libertad de los residuos obtenidos en la tabla A9. Como resultado se obtuvieron las puntuaciones: -1.996564419 y 1.996564419, que son los rangos limitantes de la región de aceptación. Esto significa que los valores del estadístico-t de la cuarta ecuación que se encuentran fuera de estos rangos, deberían rechazar la hipótesis nula, y aquellos que se encuentran dentro, deberían aprobarla. Para el caso de la presente investigación, la aprobación de la hipótesis nula denota que las variables involucradas no presentan significancia para el modelo predictivo.

Tabla A10.

Análisis de varianza de la quinta ecuación: Modelo Choi y Varian (2009) + Gimagen + Gnoticias(t)

	Grados de libertad	Suma de cuadrados	Promedio de los cuadrados	F	Valor crítico de F
Regresión	5	2.768	.554	50.916	<b>.0001</b>
Residuos	66	.718	.011		
Total	71	3.486			

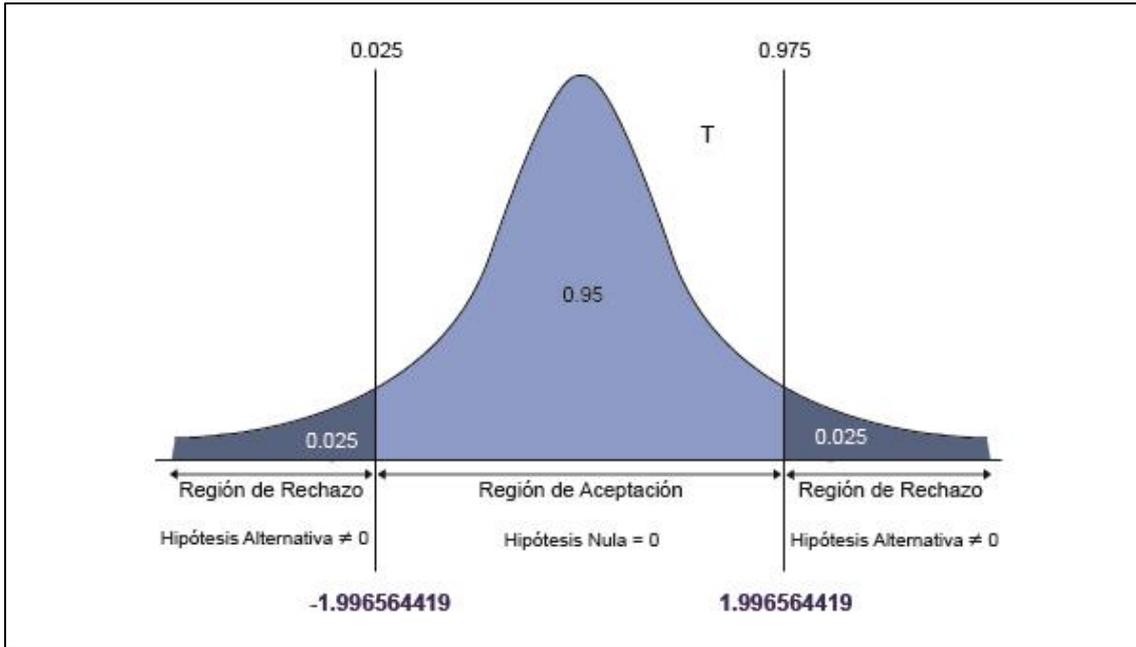


Figura A21. Gráfica t-student de la quinta ecuación: Modelo Choi y Varian (2009) + Gimagen + Gnoticias(t)

Se realizó la prueba matemática t-student para hallar el estadístico-t de la quinta ecuación. Para esto, se realizaron las operaciones: =INV.T(0.025, GL) e =INV.T(0.975, GL) en Microsoft Excel; donde GL son los grados de libertad de los residuos obtenidos en la tabla A10. Como resultado se obtuvo -1.996564419 y 1.996564419, que son los rangos limitantes de la región de aceptación. Estas puntuaciones significan que los valores del estadístico-t de la quinta ecuación que se encuentran fuera de estos rangos, deberían rechazar la hipótesis nula, y aquellos que se encuentran dentro, deberían aprobarla. Para el caso de la presente investigación, la aprobación de la hipótesis nula denota que las variables involucradas no presentan significancia para el modelo predictivo.

Tabla A11.

Análisis de varianza de la sexta ecuación: Modelo Choi y Varian (2009) + GYouTube(t) + Gnoticias(t)

	Grados de libertad	Suma de cuadrados	Promedio de los cuadrados	F	Valor crítico de F
Regresión	5	2.730	.546	47.698	<b>.0001</b>
Residuos	66	.756	.011		
Total	71	3.486			

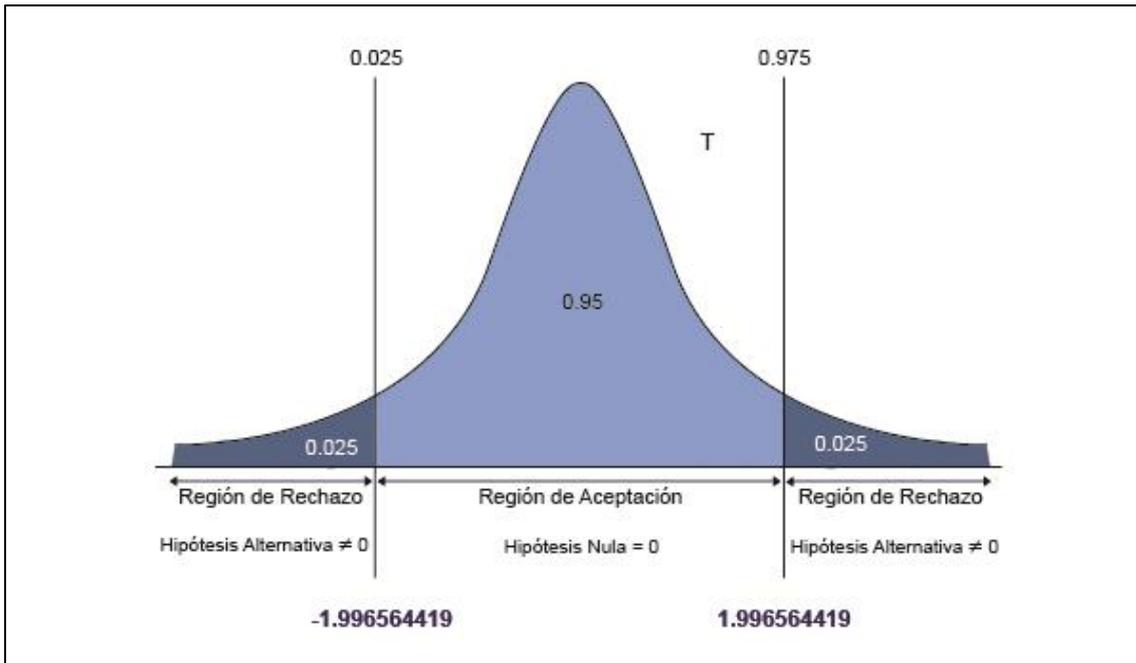


Figura A22. Gráfica t-tudent de la sexta ecuación: Modelo Choi y Varian (2009) + GYouTube(t) + Gnoticias(t)

Se realizó la prueba matemática t-student para hallar el estadístico-t de la sexta ecuación. Para esto, se realizaron las operaciones: =INV.T(0.025, GL) e =INV.T(0.975, GL) en Microsoft Excel; donde GL son los grados de libertad de los residuos obtenidos en la tabla A11. Como resultado se obtuvieron las puntuaciones: -1.996564419 y 1.996564419, que son los rangos limitantes de la región de aceptación. Esto significa que los valores del estadístico-t de la sexta ecuación que se encuentran fuera de estos rangos, deberían rechazar la hipótesis nula, y aquellos que se encuentran dentro, deberían aprobarla. Para el caso de la presente investigación, la aprobación de la hipótesis nula denota que las variables involucradas no presentan significancia para el modelo predictivo.

Tabla A12.

Análisis de varianza de la séptima ecuación: Modelo Choi y Varian (2009) + Gimagen

	Grados de libertad	Suma de cuadrados	Promedio de los cuadrados	F	Valor crítico de F
Regresión	4	2.759	.690	63.655	<b>.0001</b>
Residuos	67	.726	.011		
Total	71	3.486			

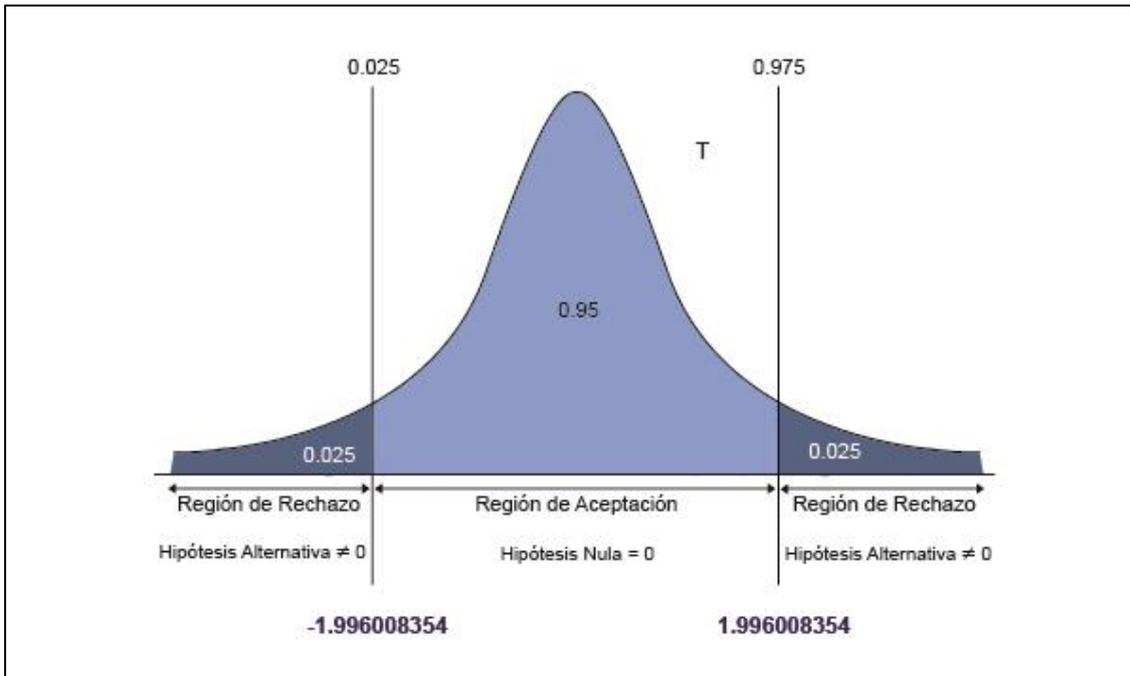


Figura A23. Gráfica t-student de la séptima ecuación: Modelo Choi y Varian (2009) + Gimagen

Se realizó la prueba matemática t-student para hallar el estadístico-t de la séptima ecuación. Para esto, se realizaron las operaciones:  $=INV.T(0.025, GL)$  e  $=INV.T(0.975, GL)$  en Microsoft Excel; donde GL son los grados de libertad de los residuos obtenidos en la tabla A12. Como resultado se obtuvo  $-1.996008354$  y  $1.996008354$ , que son los rangos limitantes de la región de aceptación. Estas puntuaciones significan que los valores del estadístico-t de la séptima ecuación que se encuentran fuera de estos rangos, deberían rechazar la hipótesis nula, y aquellos que se encuentran dentro, deberían aprobarla. Para el caso de la presente investigación, la aprobación de la hipótesis nula denota que las variables involucradas no presentan significancia para el modelo predictivo.

Tabla A13.

Análisis de varianza de la octava ecuación: Modelo Choi y Varian (2009) + GYouTube(t)

	Grados de libertad	Suma de cuadrados	Promedio de los cuadrados	F	Valor crítico de F
Regresión	4	2.727	.682	60.189	<b>.0001</b>
Residuos	67	.759	.011		
Total	71	3.486			

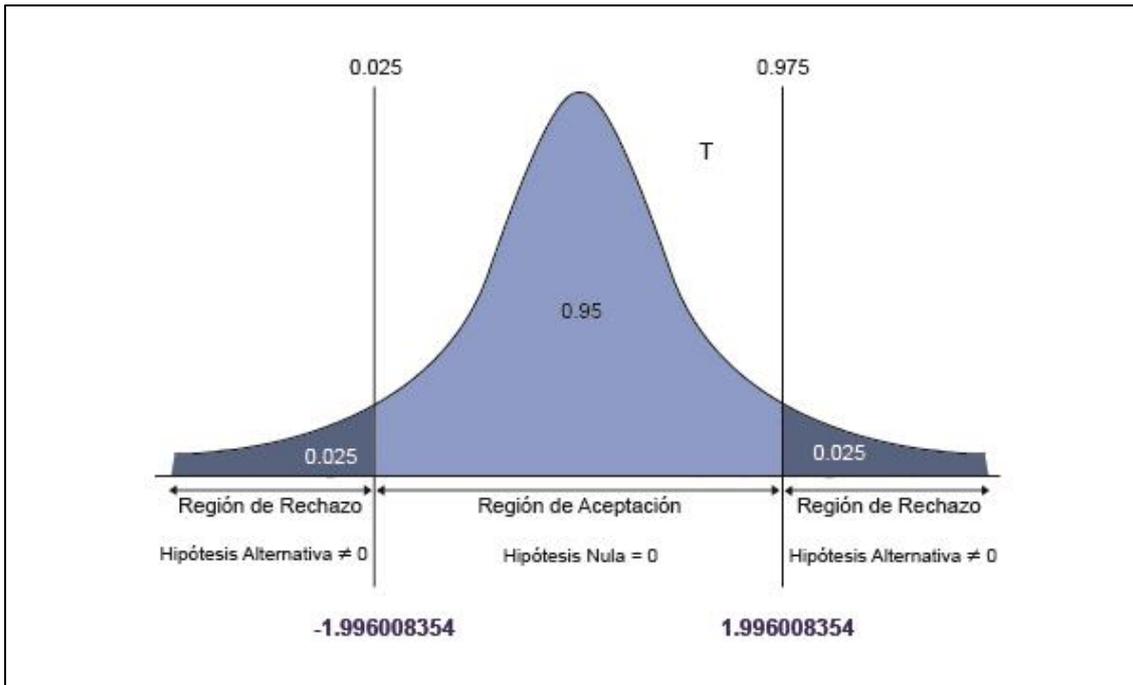


Figura A24. Gráfica t-student de la octava ecuación: Modelo Choi y Varian (2009) + GYouTube(t)

Se realizó la prueba matemática t-student para hallar el estadístico-t de la octava ecuación. Para esto, se realizaron las operaciones: =INV.T(0.025, GL) e =INV.T(0.975, GL) en Microsoft Excel; donde GL son los grados de libertad de los residuos obtenidos en la tabla A13. Como resultado se obtuvieron las puntuaciones: -1.996008354 y 1.996008354, que son los rangos limitantes de la región de aceptación. Esto significa que los valores del estadístico-t de la octava ecuación que se encuentran fuera de estos rangos, deberían rechazar la hipótesis nula, y aquellos que se encuentran dentro, deberían aprobarla. Para el caso de la presente investigación, la aprobación de la hipótesis nula denota que las variables involucradas no presentan significancia para el modelo predictivo.

Tabla A14.

Análisis de varianza de la novena ecuación: Modelo Choi y Varian (2009) + Gnoticias(t)

	Grados de libertad	Suma de cuadrados	Promedio de los cuadrados	F	Valor crítico de F
Regresión	4	2.730	.683	60.525	<b>.0001</b>
Residuos	67	.756	.011		
Total	71	3.486			

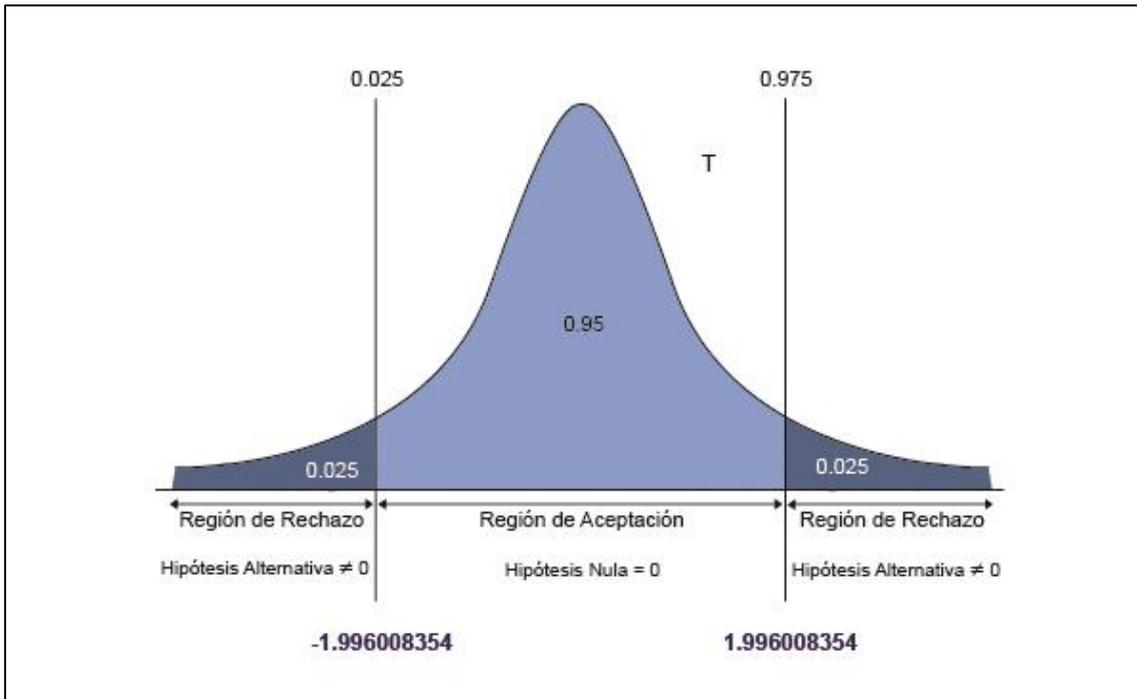


Figura A25. Gráfica t-student de la novena ecuación: Modelo Choi y Varian (2009) + Gnoticias(t)

Se realizó la prueba matemática t-student para hallar el estadístico-t de la novena ecuación. Para esto, se realizaron las operaciones: =INV.T(0.025, GL) e =INV.T(0.975, GL) en Microsoft Excel; donde GL son los grados de libertad de los residuos obtenidos en la tabla A14. Como resultado se obtuvo -1.996564419 y 1.996564419, que son los rangos limitantes de la región de aceptación. Estas puntuaciones significan que los valores del estadístico-t de la novena ecuación que se encuentran fuera de estos rangos, deberían rechazar la hipótesis nula, y aquellos que se encuentran dentro, deberían aprobarla. Para el caso de la presente investigación, la aprobación de la hipótesis nula denota que las variables involucradas no presentan significancia para el modelo predictivo.