

LA INFORMACIÓN EN LÍNEA COMO MÉTODO DE ANÁLISIS DE LA ACTIVIDAD
HOTELERA Y PERCEPCIÓN DE LOS TURISTAS

Febrero, 2015

Pedro Manuel Lichtle Fragoso
Secretaría de Turismo*

Juan Carlos Sánchez Salinas
Secretaría de Turismo**

Elaborado con la colaboración de la consultora internacional BeOnPrice. La información y base de datos utilizadas son resultado de un análisis preliminar a nivel nacional e internacional dentro del contexto de una prueba piloto.

Resumen

El objetivo del presente documento es describir una prueba piloto que utiliza dos elementos que surgen a partir del uso de sistemas de extracción de información en línea y en tiempo real: i) monitorear características de la actividad hotelera en destinos turísticos; y ii) consultar opiniones de los consumidores de servicios turísticos.

Considerando las proyecciones para el año 2015 en el comportamiento de los precios promedio por ocupación y la ocupación promedio por destino, se observa la actividad hotelera en dos destinos nacionales: Cancún, Quintana Roo, y México, Distrito Federal. También se observa la actividad hotelera en dos destinos internacionales: Punta Cana, República Dominicana y Honolulu, Estados Unidos. Lo anterior es un método para entender la estacionalidad y la posible rivalidad en precios que presentan los destinos seleccionados.

Los destinos seleccionados presentan un comportamiento positivo en el nivel de satisfacción de los turistas, alcanzando niveles de entre 7.5 y 9 puntos, considerando información en retrospectiva para todo el año 2014. En consecuencia, consiguen una percepción positiva en sus visitantes utilizando datos de información en línea.

La serie de Documentos de Investigación Estadística y Económica presenta resultados preliminares de investigación realizados en la Secretaría de Turismo con el propósito de generar intercambio y debate de ideas para el desarrollo del sector turismo. El contenido de los Documentos de Investigación Estadística y Económica, así como los argumentos vertidos, son responsabilidad exclusiva de los autores y no reflejan necesariamente los de la Secretaría de Turismo.

* Director General de Integración de Información Sectorial. Email: plichtlef@sectur.gob.mx

** Director de Análisis Regional. Email: jsanchezs@sectur.gob.mx

1. Introducción.

La información en línea es valiosa porque puede utilizarse para fines distintos de aquéllos para los que en un inicio fue creada. En el sector turismo se pueden procesar diferentes características de la oferta expresada por los hoteles para la atracción de turistas y utilizarse para generar políticas públicas en beneficio del destino turístico. Por ejemplo, Heerschap (2014) explica que se pueden recolectar y procesar datos en línea del número de habitaciones en los hoteles, el número de camas, su precio por noche, impuestos, amenidades (wifi gratuito), evaluación de los consumidores, así como ofertas de empleo.

El uso de tecnología para explotar la información en línea y en tiempo real, minería de datos, representa un instrumento novedoso para apoyar la planeación y el diseño de la política turística. La información se encuentra en todas partes: en los servidores de las empresas privadas y públicas, en las aplicaciones que circulan por los dispositivos móviles y en las redes sociales con un crecimiento exponencial en los últimos años. Debido a lo anterior, el análisis y la visualización de datos en tiempo real se vuelven un potente instrumento para examinar la información digital de la actividad hotelera.

Al existir mayor oferta de dispositivos móviles y un flujo de información dinámico, los turistas tienen mayor oportunidad de mantenerse conectados por más tiempo y en más lugares, así como compartir sus expectativas y experiencias de viaje. Si bien es cierto que los datos en línea conviven con otros medios tradicionales como la televisión y los medios impresos, como las revistas y periódicos, los turistas se caracterizan por su alta movilidad y la búsqueda de información de manera remota en servicios como sitios web y aplicaciones.¹ En este caso, la extracción de datos en tiempo real también es un instrumento novedoso para examinar la información digital en la demanda por servicios turísticos.

Los actores presentes en el sector, por ejemplo, cadenas hoteleras y oficinas gubernamentales tienen un interés natural en los niveles generales de precios, en la actividad hotelera y en la percepción de los turistas sobre los destinos turísticos visitados. Con el fin de tomar una decisión óptima sobre inversión o política pública, se exige que los tomadores de decisiones tengan una noción con claridad sobre el cambio en el nivel de precios pagados por los huéspedes, el costo de oportunidad de incrementar infraestructura hotelera y el mayor acercamiento sobre la percepción de los servicios turísticos ofrecidos.

Por un lado, las cadenas de hoteles y alianzas hoteleras participan en sistemas de reservas internacionales que muestran habitaciones vacías con sus etiquetas de precios y también hay varios canales (online) para reservar una habitación, muy a menudo a diferentes tarifas para cada uno de los canales.² Por otro lado, la respuesta inmediata a preguntas como ¿cuál es el precio por una noche de hotel para dos personas en el centro de Morelia? o ¿qué lugares son recomendables para comer en Tlaxcala?, se encuentran disponibles incluso de manera remota y con un amplio tiempo de anticipación porque el Internet es un

¹ Existen aplicaciones que utilizan la información generada por los usuarios como las condiciones de tráfico en las ciudades y de información sobre lugares de alimentos, recreación y entretenimiento en destinos turísticos.

² Ferrara, 2014.

medio para realizar y complementar muchas de las actividades dentro de cualquier destino turístico en México y en todo el mundo.

La oportunidad de obtener e intercambiar información de manera instantánea hace que la adquisición de datos en línea se convierta en un insumo para complementar la generación de productos y servicios turísticos de mayor calidad, así como del conocimiento de la opinión del turista y aproximarse a ofrecer una experiencia personalizada. Debido a lo anterior, el objetivo del presente documento es describir dos elementos oportunos que surgen a partir del uso de sistemas de extracción de información en línea y en tiempo real: i) monitorear características de la actividad hotelera en destinos turísticos; y ii) consultar opiniones de los consumidores de servicios turísticos.

En cada caso se utiliza una fuente de información con una temporalidad distinta. Para propósitos de predicción en la actividad hotelera (información prospectiva) se utiliza un algoritmo de búsqueda y extracción de información sobre reservas: las reservaciones en línea se realizan para el uso futuro de las instalaciones del hotel. Para propósitos de valoración de opinión sobre los servicios (información retrospectiva) se utiliza un algoritmo de búsqueda y extracción de información sobre opiniones en el pasado. La diferencia temporal para la extracción de la información corresponde al momento en el cual acontece la acción por parte del turista: primero realiza la reservación, después ejerce la ocupación, y por último, expresa su opinión sobre los servicios obtenidos.

2. Extracción de información y datos en línea.

Recientemente, algunos autores se han encargado de desarrollar sistemas de extracción de datos de texto sin procesar y después generan una posible percepción positiva, negativa o neutral del destino. Un ejemplo, son las búsquedas con términos como “Torre Eiffel” y los resultados son utilizados para construir una base de datos para realizar inferencias sobre las opiniones y el comportamiento de los turistas. De acuerdo con Ferrara (2014), este procedimiento implementa clases diferentes de algoritmos, que realizan una búsqueda de fuentes no estructuradas y logran unificar información para utilizarla posteriormente en mejorar la calidad de los servicios en las zonas frecuentadas por turistas.

A nivel internacional, Jean Herget (2014) realizó en República Checa una investigación sobre la posibilidad de describir cómo se distribuyen los turistas por nacionalidad en todo el territorio de aquel país. Utilizó datos entre abril y mayo de 2014, con un total de 1,059 hoteles revisados y extrayendo información de los sitios web a nivel global de 6,837 perfiles de consumidores. Los resultados mostraron que el destino turístico principal en República Checa fueron las montañas Krkonose, en Praga, visitadas en su mayoría por daneses; y Karlovy Vary, visitado por alemanes. Además, el lugar menos visitado en este periodo fue la ciudad de Pilsen.

La Oficina de Turismo de Austria (Haller, 2000) adoptó un sistema en su portal en línea para extraer información de viajes a través de Internet, aplicación que se convertiría en el portal turístico líder no sólo en Austria, sino también en otros países como Alemania y Suiza. La funcionalidad que ofrece el citado portal en línea se clasifica en tres componentes:

- i) Internet: pone al alcance del público diferentes módulos, entre los que destacan el Atlas y el de Reservaciones; mientras que el Atlas permite al cliente navegar a través de todo tipo de información turística, el de Reserva realiza una búsqueda estructurada precisa basada en un subconjunto de la información del turismo presentado por el Atlas (como pueblos, hoteles, habitaciones disponibles, eventos y sitios de campamentos), dando lugar inclusive a la reservación en línea desde estos módulos.
- ii) Extranet: permite a los proveedores de información turística autorizados, actualizar y ampliar su información y los productos del turismo ofrecidos directamente.
- iii) Intranet: permite configurar el sistema de varias formas para enfrentar estrategias definidas por el usuario.

La discusión del presente documento tiene como base los programas que generan automáticamente datos de salida en formatos de base de datos. De acuerdo con Fiumara (2007), hay una serie de aplicaciones donde las herramientas de extracción pueden aprovechar plenamente su poder, donde una de las aplicaciones más prometedora es la comparación automática de productos, por ejemplo, la posibilidad de que un consumidor compare diferentes ofertas de un mismo bien y sus características en tiempo real.

Para nuestro caso, generalmente es el sector privado el que genera un mayor flujo de información en línea, por ejemplo, en la actualidad se observa el crecimiento en las reservaciones de hotel con anticipación a través de las agencias de viaje en línea (OTA's por sus siglas en inglés). También existe una nueva tendencia en la valoración y emisión de opinión sobre los destinos en redes sociales y foros de discusión.

Adicionalmente, existen instrumentos de extracción que generan información sobre las tarifas en rutas aéreas de las aerolíneas y se puede observar que su comportamiento depende del tiempo en la anticipación de compra, de la distancia de la ruta y del tipo de asiento seleccionado (clase). La extracción de datos también ocurre sobre las tarifas de hoteles y su dependencia en la anticipación de compra, de la distancia hacia las atracciones turísticas del destino o centro de la ciudad y del tipo de cuarto seleccionado o categoría del hotel en cuestión.

En este sentido, el dinamismo del turismo puede tener una representación a partir de la extracción de información del segundo tipo: precios promedio de hoteles en México y percepción de los turistas. Por un lado, en un destino en particular la tarifa de una noche de hotel es una variable que influye directamente en la decisión de los turistas sobre el viaje a realizar, por lo tanto, la posibilidad de comparar diferentes ofertas y sus características en tiempo real puede ser de gran utilidad para conocer el comportamiento de la oferta a la que se enfrentan los consumidores. Por otro lado, el reconocimiento de la percepción del turista sobre un destino es una aplicación de la extracción de datos que todavía se encuentra en desarrollo, pero su relevancia puede observarse en la mejora continua de los servicios.

De acuerdo con Sarawagi (2008), en la extracción de datos sobre percepción es necesario asociar una connotación determinada con el valor de un adjetivo que describa la connotación. El valor de este adjetivo requiere la combinación de pistas que se encuentran repartidas en muchas palabras diferentes alrededor de la información extraída.

Por ejemplo, dada una palabra o una frase, se requiere extraer partes de una página web que puede representar una crítica sobre el tema. Entonces, para inferir si la crítica es positiva o negativa se requiere la asociación con la connotación previamente definida, esto también se llama extracción de opinión y hoy día es un tema de interés para la investigación del comportamiento del turista con un desarrollo potencial.

Sin embargo, la extracción de información sobre opiniones presenta mayores dificultades que el seguimiento de variables como las tarifas de rutas aéreas y hoteleras. Lo anterior, debido a que no solamente son variables cuantitativas sino también conllevan un análisis cualitativo que debe considerarse previamente a la generación de conclusiones sobre el comportamiento de las variables y su posible tendencia.

Crandall (2009) señala varios problemas en la determinación del comportamiento de la opinión de un consumidor, tales como determinar si cada segmento de texto (frase, párrafo o sección) es una "opinión" sobre el tema referido; identificar la opinión del titular (la persona u organización que genera la información); la determinación de la polaridad de la opinión (positiva o negativa); y el tema o asunto con exactitud. Y con mayor dificultad se puede inferir la transacción que describe, como el servicio o la calidad del producto.

Las empresas tienen el objetivo final de maximizar utilidades. Para conseguir dicho objetivo requieren de un análisis exhaustivo sobre productos, clientes, competidores y también de la estructura del mercado objetivo; comúnmente a este tipo de análisis se le conoce como Inteligencia de Mercados, la cual tiene una relación estrecha con la extracción de datos en línea. En este tipo de escenario, las técnicas de extracción de datos sirven para satisfacer dos requisitos principales: i) estrategias de planificación para llegar a los clientes potenciales; y ii) utilizar la menor cantidad de recursos para acceder al cliente.

Walchhofer (2010) sugiere que los mercados electrónicos son más volátiles y sensibles debido a la velocidad en el cambio del comportamiento de los consumidores, también señala que los métodos de segmentación de mercado y discriminación de precios se vuelven más complicados con la velocidad de transmisión de la información. En consecuencia, se requiere una imagen clara, reciente y completa de los mercados en línea para orientar estrategias de lanzamiento de productos, incluyendo en particular las decisiones de precios. Lo mismo ocurre en el sector turismo con las ofertas de vuelos, operadores de planes turísticos, renta de automóviles y tarifas en los hoteles.

Las herramientas de extracción de datos generan reportes sobre el mercado del turismo considerando el nivel de precios, la variación y su distribución territorial. Siguiendo a Walchhofer (2010) los extractores usan atributos como: i) tiempo, estacionalidad, periodos vacacionales, días laborales y días que corresponden a fines de semana; ii) descomposición geográfica (país, estado o ciudad) y subdivisiones como destinos turísticos; y iii) otros atributos de la oferta, como los rankings de hoteles, número de camas en el cuarto y el tipo de hotel, entre otras.

3. Estrategias de procesamiento de datos.

Siguiendo lo expresado por Ronald Jansen (*International Trade Statistics, United Nations Statistics Division*), el mayor auge en procesamiento de datos se ha presentado en los proveedores y en las fuentes de datos, porque son los primeros agentes que intervienen en el proceso de la cadena de generación de información. El ejercicio que aquí se expone, surge de una prueba piloto con una empresa dedicada a desarrollar estrategias de extracción de información adaptadas a las necesidades de sus clientes a nivel internacional: es un acercamiento entre la Secretaría de Turismo y la posibilidad de generar utilidad para el sector turismo a partir de información con la estrategia de procesamiento de datos de BeOnPrice.³

La información extraída de las agencias de viaje en línea no es exclusiva, empresas como Google (*Google Hotel Finder*) tienen sus propias herramientas. De acuerdo con BeOnPrice⁴, el atractivo de *Google Hotel Finder* no tiene que ver con el número de usuarios que retoma de las OTAs, sino por la visibilidad que proporciona. Es un caso muy similar al de la red social *Google Plus*, que pese a poseer una menor popularidad que sus rivales Facebook y Twitter, posee una mayor relevancia que éstos para el posicionamiento en las páginas de resultados de búsqueda.

Por un lado, la popularidad de los hoteles tiene una mayor dificultad de medición a partir de encuestas de satisfacción por lo que resulta una buena idea utilizar motores de búsqueda y extracción para identificar si los usuarios en línea (turistas) tienen una mayor preferencia por un hotel, definiendo una localización específica y un determinado tiempo de ocurrencia.

La popularidad de un hotel es un activo estrechamente relacionado con el éxito en sus reservaciones porque es una característica que genera valor adicional como el nombre o la marca a nivel internacional. Aún más, la popularidad de un hotel puede ser una estrategia de monitoreo de la actividad hotelera y, además, es un activo relacionado con la consulta de opiniones de los consumidores, por lo tanto, su seguimiento podría apoyar la planeación y el diseño de la política turística.

Por otro lado, el RevPAR (Ingreso por Habitación Disponible) es el indicador de competitividad en la industria hotelera, pues permite a los hoteles comparar su éxito a la hora de optimizar el uso del inventario disponible en un entorno competitivo. Sin embargo, siguiendo a BeOnPrice⁵ hay otra métrica que está ganando terreno en el sector hotelero recientemente, el TRevPAR (Total de Ingresos por Habitación Disponible).

Fundamentalmente el TRevPAR no sólo mide los flujos de ingresos principales, sino que toma en cuenta las oportunidades para generar ingresos suplementarios. En este sentido, considera los servicios e infraestructuras con potencial para generar ingresos (como

³ La información y base de datos utilizadas son resultado de un análisis preliminar de una muestra a nivel nacional e internacional dentro del contexto de una prueba piloto. Los derechos de propiedad sobre los algoritmos de extracción de información corresponden a la consultora internacional BeOnPrice. El acceso y procesamiento de la información utilizada ocurrió el día 16 de enero de 2015.

⁴ BeOnPrice, 19 de febrero de 2014.

⁵ BeOnPrice, 24 de abril de 2014.

espacios para banquetes o conferencias). El TRevPAR sirve para conocer en qué medida las distintas áreas del hotel juegan su papel en la generación de ingresos.

Los productos complementarios que presentan costos de oportunidad para desarrollar infraestructura también son difíciles de medir con encuestas hacia los hoteles, por cuestiones de medición y estimación. Por consiguiente, el cálculo de este tipo de variables por un proveedor de servicios externos facilita la exploración de alternativas para la comparación de ofertas en un destino turístico.

Por lo antes descrito, la extracción de información en tiempo real sobre el RevPAR y TRevPar son complementos de Inteligencia de Mercados para conocer el comportamiento competitivo apegado a la realidad, ya que es información de datos publicados por los propios portales de reservaciones de servicios de hospedaje. A diferencia de otras fuentes de información, las OTAs proveen directamente los datos, sin necesidad de preguntar a los hoteleros por la misma.

Por último, BeOnPrice⁶ expresa la necesidad en el sector turístico de ser capaces de medir las señales que emite el mercado con el ejemplo de Expedia. Expedia muestra los tres principios para el uso de *Big Data en Turismo*: i) los resultados se mejoran probando y aprendiendo, realizando pruebas de hipótesis y la aplicación de los resultados, así como monitoreando la experiencia de viaje a través de la personalización; ii) usar los datos y análisis para entender mejor al cliente y proporcionarle una mejor experiencia de viaje, desde el momento en que el cliente emprende las primeras búsquedas online; y iii) entregar datos útiles a los socios, ayudándolos a ser más eficaces en el mercado.

3.1 Estrategias comerciales de hoteles en línea.

De acuerdo con Conde (2011), la importancia de la utilización del comercio electrónico en la industria turística es que ayuda a mejorar las operaciones tradicionales practicadas en los negocios, tanto en la estrategia empresarial como en la reingeniería del negocio turístico. También existen diferentes clases de estrategias comerciales y las utilizadas en la prueba piloto del presente documento son: i) Negocio a Negocio: extranet entre los hoteleros y las OTAs (por ejemplo Booking); y ii) Cliente a Negocio: clientes registrando sus preferencias sobre la opinión sobre un hotel o destino.

Asimismo, Conde (2011) considera que en el sector turismo destaca la utilización de distintas estrategias comerciales dentro de las OTAs, por ejemplo, un modelo de subasta: en el cual los vendedores (hoteles) colocan el precio al cual están dispuestos a vender un producto o servicio (noche en una habitación) e invitan a los compradores a hacer su oferta. Considerando que la prueba piloto extrae información de Booking como insumo de la información en línea, se puede sugerir que Booking utiliza una estrategia de aparador electrónico donde colecciona la información que proveen los hoteles y la muestra en su plataforma comercial.⁷ Por lo que se podría afirmar que la información contenida en Booking no presenta modificaciones para el cliente final sobre precios y reservaciones en línea.

⁶ BeOnPrice, 30 de junio de 2014.

⁷ En el Anexo se presentan las estrategias desarrolladas por Conde (2011).

De acuerdo con Schaal (2012), las ventajas competitivas de Booking han sido su actuación independiente y la capacidad de promover una amplia gama de oferta hotelera. A ello, agrega que su modelo comercial privilegió una estructura organizativa sencilla y reemplazó a las gestiones complejas por trámites técnicamente simples y eficientes con sus clientes directos (hoteles).

El modelo de Booking contrasta con los modelos comerciales seguidos por sus principales competidores: modelos complejos y con mayores comisiones. Sin embargo, Booking no considera la totalidad de los hoteles, en ocasiones, sus clientes potenciales no están dispuestos a cumplir con las condiciones y con el pago de comisiones. Además, existe la posibilidad de que los hoteles de categorías bajas no presenten recursos suficientes para el desarrollo de un micro-sitio en la plataforma de Booking.

Debido a lo anterior, la información contenida en la plataforma de Booking podría ofrecer solamente tendencias sobre el comportamiento de las unidades de alojamiento con información en línea y no necesariamente una réplica de la actividad real, principalmente por la escasa participación de pequeños hoteles con escasos recursos y con poco interés de ofrecer sus servicios en línea.

4. Prueba Piloto: algoritmos de BeOnPrice aplicados a la información de Booking.

En la prueba piloto realizada se muestra información agregada sobre las principales variables turísticas de interés para el sector, como pueden ser los precios promedio por habitación, la ocupación promedio en un destino, la demanda o la reputación online, entre otros; con base en un portal de venta de habitaciones en Internet: Booking. Este tipo de pruebas piloto tienen la ventaja de proveer un fácil acceso y comprensión para los agentes interesados en realizar proyectos de investigación, análisis y medición del sector.

4.1 Monitoreo de características de la actividad hotelera.

Considerando los retos de la generación de estadísticas oficiales como los cambios en el entorno, restricciones presupuestales y la dificultad de generar confianza en el turista, se están explorando herramientas para generar información. De ésta forma, la extracción de información puede servir para la generación de tendencias en algunos destinos seleccionados a nivel nacional y su comparación con la competencia en otros destinos turísticos a nivel internacional.

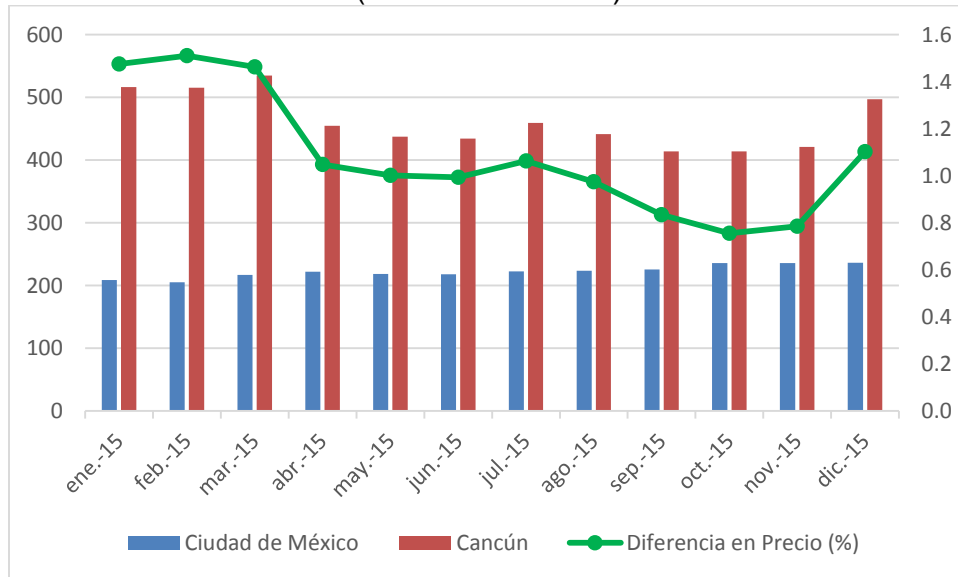
Aquí es importante resaltar que la información es prospectiva (año 2015) para los precios promedio y para la ocupación promedio; y retrospectiva (año 2014) para obtener la opinión de los turistas, considerando el comportamiento de las unidades de alojamiento que ofrecen información en línea. De ésta manera, los algoritmos de extracción trabajan con información agregada de unidades de alojamiento (hoteles). Para definir los hoteles se utilizan dos tipos de filtros: i) zona: permite seleccionar aquellos hoteles que pertenecen a una región; y ii) hoteles: permite seleccionar aquellos hoteles que son de un tipo, tienen una cierta categoría y/o un número de habitaciones.

4.1.1 Prueba piloto en destinos nacionales e internacionales seleccionados.

Destinos nacionales: en este apartado se cubrió información referente a la proyección de demanda para un destino y ocupación media del mes en curso para todo el año 2015, en los destinos nacionales seleccionados: México, Distrito Federal, y Cancún, Quintana Roo.

En primer lugar, para el caso de los hoteles en Cancún, la base de la prueba piloto contabiliza más de 120 unidades de alojamiento, de un total de 176 unidades de alojamiento en dicha región correspondientes al sistema DataTur, lo que significa una cobertura del 68%. En segundo lugar, para el caso de la Ciudad de México el acopio de información es de 37 establecimientos en la prueba piloto (sólo para la región del corredor Reforma-Polanco); en donde hay 45 unidades de alojamiento correspondientes en DataTur. Lo anterior significa que la muestra es de aproximadamente 82%.⁸

Gráfico 1. Proyección de precio promedio por habitación para 2015.
(Monto en dólares)



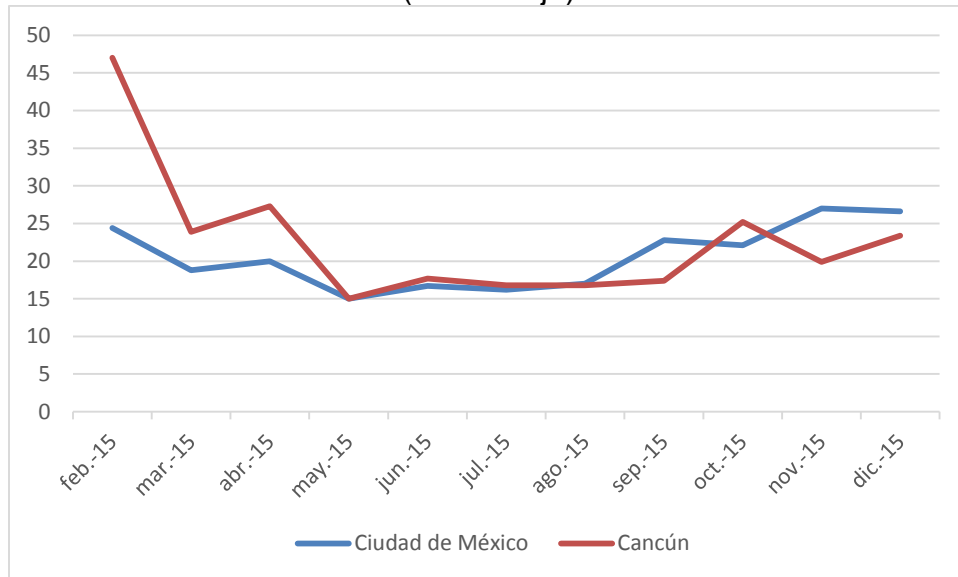
Fuente: Prueba piloto con información de BeOnPrice.

En el Gráfico 1 se puede observar la diferencia en precios en los destinos nacionales: Cancún presentará en todo el año 2015 precios promedio por habitación por arriba de los proyectados para la Ciudad de México y la diferencia es mayor a inicio del año. También se puede observar una tendencia decreciente en la diferencia en precios, sin embargo, al cierre de año se observa un incremento hasta llegar a un 100% de diferencia, por último, la brecha entre ambos precios se mantiene en todo el año 2015.

⁸ Fuente: BeOnPrice y DataTur. La información puede filtrarse atendiendo a diferentes dimensiones como tipo de habitación, tipo de hotel, categoría del hotel y número de habitaciones. Sin embargo, la prueba piloto contempla solamente valores acumulados de las variables como primer acercamiento al monitoreo de actividad hotelera en destinos nacionales.

El comportamiento de los precios promedio en los destinos nacionales indica que Cancún tiene un nivel de precios más alto que la Ciudad de México (corredor Reforma-Polanco) y además presenta variaciones estacionales que no se observan en la Ciudad de México. En suma, la proyección de los precios promedio para cada destino podría depender del comportamiento de la demanda: en la Ciudad de México se presenta una demanda constante y en Cancún una demanda estacional.

Gráfico 2. Proyección de ocupación promedio para 2015.
(Porcentaje)



Fuente: Prueba piloto con información de BeOnPrice.

El Gráfico 2 señala que no existe una gran diferencia entre la ocupación proyectada para cada destino, también da cuenta de la estacionalidad que podría mejorarse con una periodicidad semanal: la Ciudad de México presenta los mayores niveles de reservas para su ocupación promedio en los meses de noviembre y diciembre; Cancún presenta los mayores niveles de reservas para su ocupación promedio en febrero y abril.

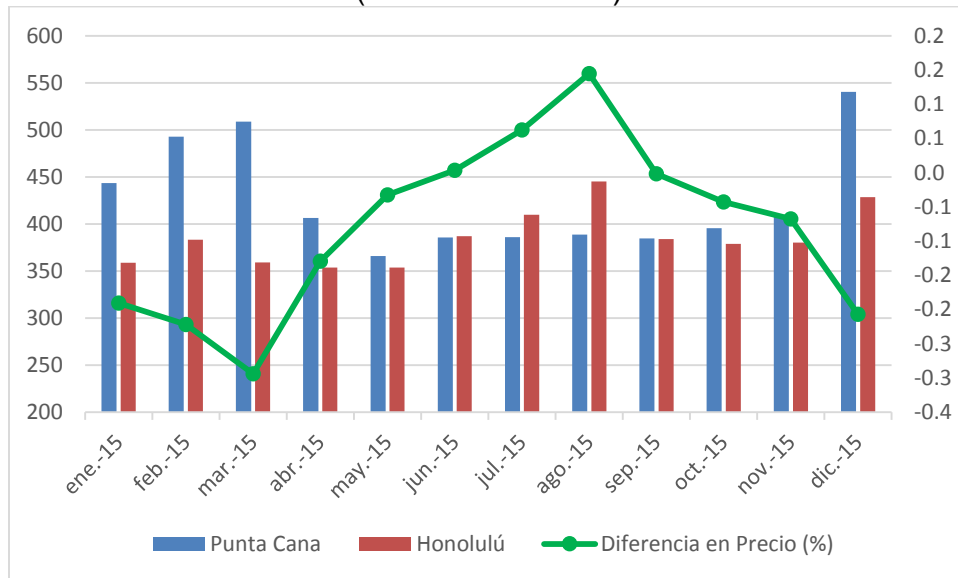
En ambos casos se observa una ocupación por debajo del 30% en todo el año 2015, considerando que es información de programación de viajes en línea que depende del tiempo de anticipación de la reservación y tiene una alta correlación con la ocupación real. Lo anterior podría indicar dos posibles situaciones sobre los datos en línea: i) los hoteles pueden mantener una estrategia comercial en donde se bloquea en línea cierto número de cuartos anualmente para reservas de usuarios con beneficios adicionales; y ii) las reservaciones en línea aseguran un nivel de ocupación promedio a lo largo del año.

Además, la tendencia observada en la proyección de ocupación nos podría confirmar que si bien se pueden observar variaciones a lo largo del tiempo, los niveles de ocupación resultado del monitoreo en línea tienden a ser más altos mientras mayor cercanía exista con las fechas de reservación. Una posible explicación para este resultado podría ser que el turista planea su viaje con tal vez uno o dos meses de anticipación.

Destinos internacionales: en este apartado se cubrió información referente a la proyección de demanda para un destino y ocupación media del mes en curso para todo el año 2015, en los destinos internacionales seleccionados: Punta Cana, República Dominicana, y Honolulu, Estados Unidos.

En ambos casos, no fue posible realizar una contabilización en el sistema Datatur de las unidades de alojamiento existentes en cada región porque se encuentran fuera del territorio nacional. No obstante, la prueba piloto cuenta con 124 unidades de alojamiento en Punta Cana y 121 unidades de alojamiento en Honolulu.⁹

Gráfico 3. Proyección de precio promedio por habitación para 2015.
(Monto en dólares)



Fuente: Prueba piloto con información de BeOnPrice.

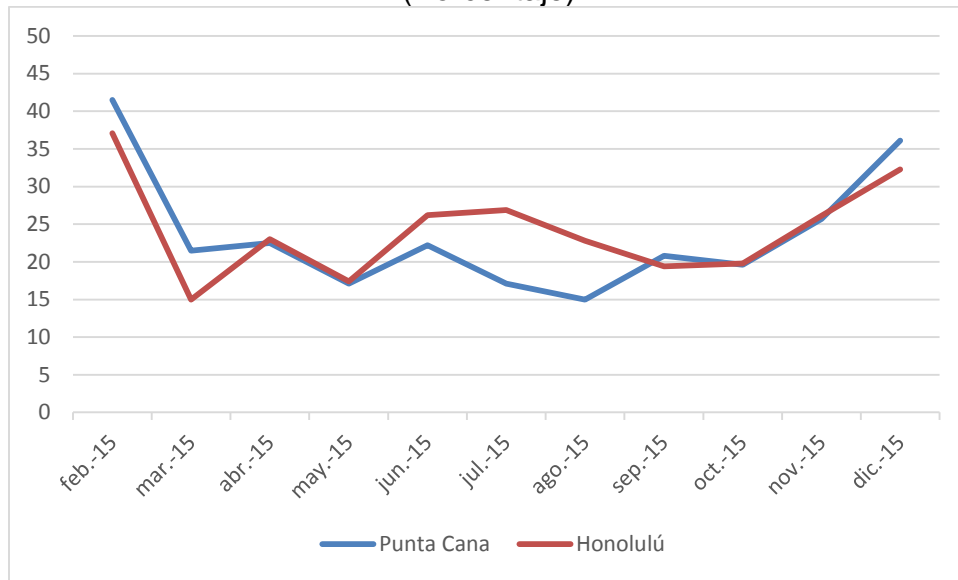
En el Gráfico 3 se puede observar la diferencia en precios en los destinos internacionales: Punta Cana presentará en el primer semestre de 2015 precios promedio por habitación por arriba de los proyectados para Honolulu, en agosto la relación es inversa y en diciembre Punta Cana también presentará precios por arriba de Honolulu. Sin embargo, la diferencia nunca es mayor a 22% entre ambos destinos y se observa estacionalidad en ambos casos.¹⁰

⁹ Fuente: BeOnPrice. La información puede filtrarse atendiendo a diferentes dimensiones como tipo de habitación, tipo de hotel, categoría del hotel y número de habitaciones. Sin embargo, la prueba piloto contempla solamente valores acumulados de las variables como primer acercamiento al monitoreo de actividad hotelera en destinos internacionales.

¹⁰ En los mecanismos de competencia en precios se puede observar que si existe una rigidez en los precios entre competidores y no hay cambios bruscos a la baja/alta entre ellos se puede obtener un indicio de una posible estructura de mercado con poca rivalidad en precios. Como no se observa este comportamiento entre Punta Cana y Honolulu, posiblemente son destinos competidores. Fuente: John Sutton (2007), "Market Structure: Theory and Evidence".

El comportamiento observado muestra que Punta Cana podría tener precios más bajos en comparación con Honolulu para los meses de enero, febrero, marzo y diciembre; mientras Honolulu podría tener precios más bajos en comparación con Punta Cana para los meses de junio, julio, agosto y septiembre. Alternativamente, los precios en Honolulu tienen una mayor estabilidad, a diferencia de los observados en Punta Cana; lo anterior puede ser reflejo de las condiciones de oferta y demanda en ambos destinos y no necesariamente de una competencia directa en precios entre los destinos internacionales observados.

Gráfico 4. Proyección de ocupación promedio para 2015.
(Porcentaje)



Fuente: Prueba piloto con información de BeOnPrice.

El Gráfico 4 presenta diferencias para la ocupación promedio en 2015 para cada destino, también da cuenta de la estacionalidad que podría mejorarse con una periodicidad semanal: ambos presentan su mayor ocupación para febrero y la brecha entre ocupaciones se observa en junio, julio y agosto; correspondiendo con las mayores diferencias en precio que observamos en el apartado anterior.

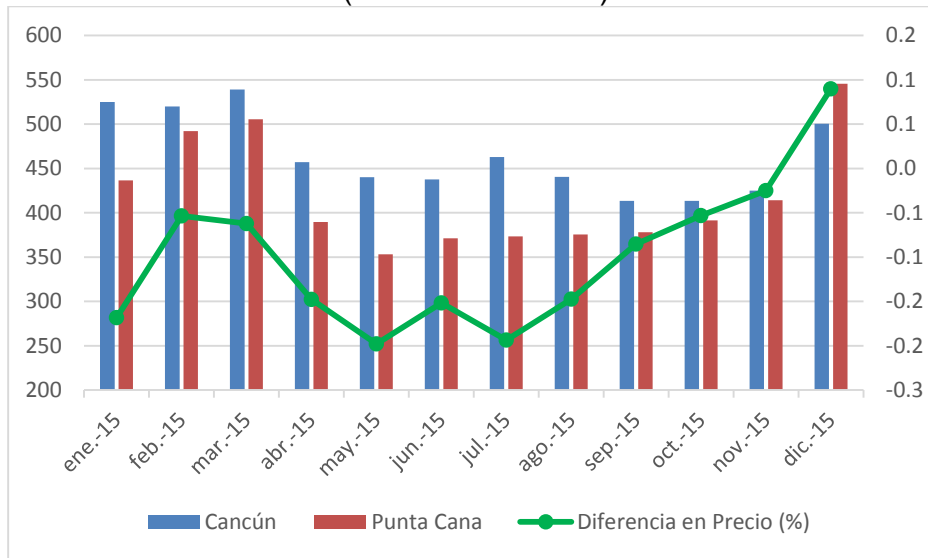
En ambos casos la proyección de la ocupación se encuentra por debajo del 40% en todo el año 2015, considerando que es información de programación de viajes en línea y tiene una alta correlación con la ocupación real. La diferencia en ocupación puede explicarse por diferentes causas, entre las cuáles se encuentran la posible presencia de un tipo de turista, principalmente de Estados Unidos o Canadá que pasa una gran parte del invierno en lugares cálidos, por ejemplo en México y zonas del Caribe; debido a que en Punta Cana se presenta la mayor temperatura mensual promedio que en Honolulu precisamente en los meses que tiene una mayor ocupación.¹¹

¹¹ En el Anexo se presenta un gráfico con la temperatura mensual promedio comparativamente para Cancún, Punta Cana y Honolulu.

Además, la diferencia en ocupación podría explicarse por el tipo de visitante que recibe cada destino, las características de grupo o perfil del turista podrían ser parte de una extensión del presente documento.

Comparativo de destinos internacionales con el destino nacional de Cancún.

**Gráfico 5. Proyección de precio promedio por habitación para 2015.
(Monto en dólares)**

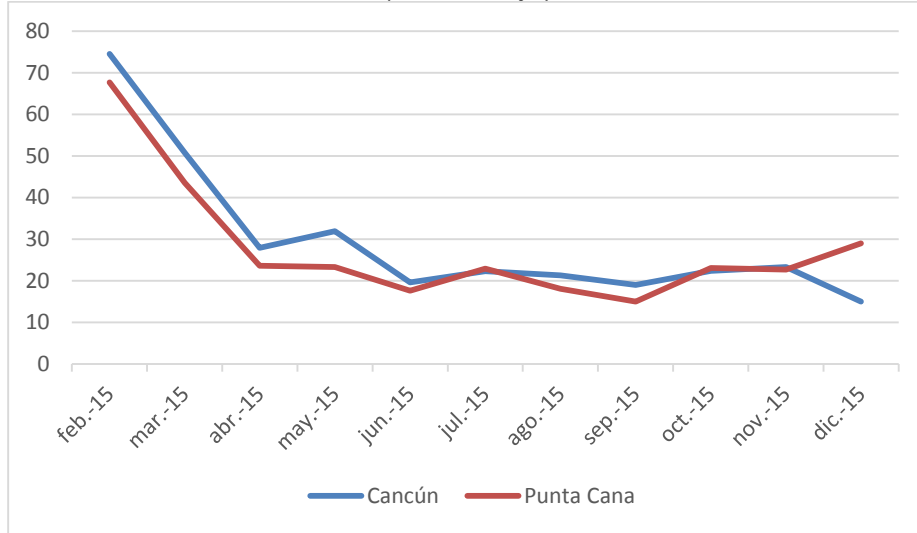


Fuente: Prueba piloto con información de BeOnPrice.

En el Gráfico 5 se puede observar que Cancún presentará en 2015 precios promedio por habitación por arriba de los proyectados para Punta Cana, con excepción del mes de diciembre. Sin embargo, la diferencia nunca es mayor al 20%. Adicionalmente, la brecha entre ambos precios no se mantiene en todo el periodo, en ocasiones es mayor como en mayo, junio y julio; y la brecha se incrementa en cada mes del segundo semestre; por lo que se puede observar de manera mensual ambos presentan estacionalidad.¹²

¹² En los mecanismos de competencia en precios se puede observar que si existe una rigidez en los precios entre competidores y no hay cambios bruscos a la baja/alta entre ellos se puede obtener un indicio de una posible estructura de mercado con poca rivalidad en precios. Como no se observa este comportamiento entre Cancún y Punta Cana, posiblemente son destinos competidores. Fuente: John Sutton (2007), “*Market Structure: Theory and Evidence*”.

Gráfico 6. Proyección de ocupación promedio para 2015.
(Porcentaje)



Fuente: Prueba piloto con información de BeOnPrice.

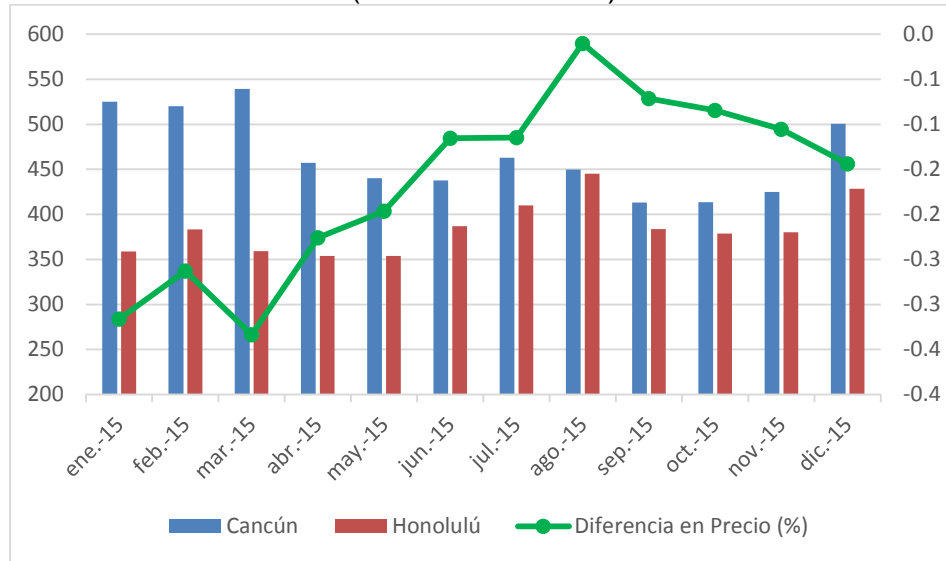
El Gráfico 6 señala que Cancún presentará sus más altos niveles de ocupación promedio en los meses de febrero y marzo, por su parte, Punta Cana presentará los niveles más altos de ocupación promedio en las mismas fechas y con un repunte en diciembre; lo que podría explicarse porque las reservaciones en línea son poco anticipadas.

En ambos casos se observa una ocupación por debajo del 40% en todo el año 2015, esto podría reflejar que Cancún tiene mayores niveles de reservaciones en línea anticipada que Punta Cana. De nuevo, la diferencia en ocupación podría explicarse por la presencia de turistas principalmente de Estados Unidos o Canadá que pasa una gran parte del invierno en lugares cálidos, por ejemplo en México y zonas del Caribe; debido a que en noviembre y diciembre son los únicos meses en donde la temperatura mensual promedio de Punta Cana es mayor que en Cancún y precisamente son los meses en donde la ocupación promedio tiene un repunte en Punta Cana.¹³

Además, la diferencia en ocupación podría explicarse por el tipo de visitante que recibe cada destino, las características de grupo o perfil del turista podrían ser parte de una extensión del presente documento.

¹³ En el Anexo se presenta un gráfico con la temperatura mensual promedio comparativamente para Cancún, Punta Cana y Honolulu.

Gráfico 7. Proyección de precio promedio por habitación para 2015.
(Monto en dólares)



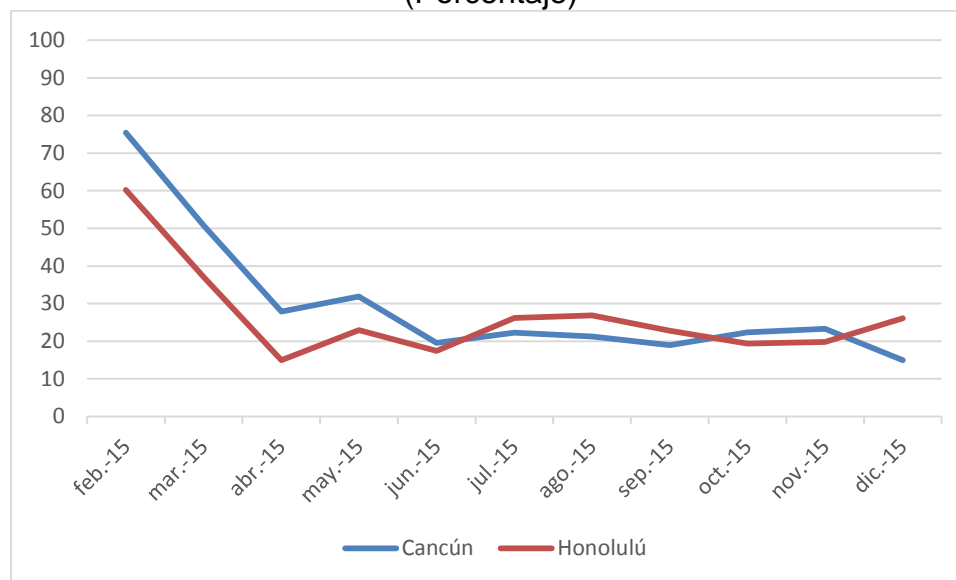
Fuente: Prueba piloto con información de BeOnPrice.

En el Gráfico 7 se puede observar la diferencia en precios de un destino nacional (Cancún) con un posible destino competidor (Honolulu): Cancún presentará en todo el año 2015 precios promedio por habitación por arriba de los proyectados para Honolulu, además, la mayor diferencia se encuentra en enero con aproximadamente 30% y en ambos casos se puede observar una posible estacionalidad en precios.¹⁴

En este caso, la brecha entre ambos precios no se mantiene en todo el periodo, en ocasiones es mayor como en agosto; y la brecha se incrementa en cada mes también hasta agosto y después disminuye en los últimos cuatro meses del año.

¹⁴ En los mecanismos de competencia en precios se puede observar que si existe una rigidez en los precios entre competidores y no hay cambios bruscos a la baja/alta entre ellos se puede obtener un indicio de una posible estructura de mercado con poca rivalidad en precios. Como no se observa este comportamiento entre Cancún y Honolulu, posiblemente son destinos competidores. Fuente: John Sutton (2007), "Market Structure: Theory and Evidence".

Gráfico 8. Proyección de ocupación promedio para 2015.
(Porcentaje)



Fuente: Prueba piloto con información de BeOnPrice.

El Gráfico 8 presenta diferencias para la ocupación promedio en 2015 para cada destino, también da cuenta de la estacionalidad que podría mejorarse con una periodicidad semanal: ambos presentan su mayor ocupación en los primeros tres meses del año. La ocupación se encuentra por debajo del 40% en todo el año 2015.

Por lo tanto, se puede argumentar que tanto Punta Cana y Honolulu, considerando los resultados de la presente prueba piloto podrían presentar un comportamiento con cambios notables en comparación de precios y con cambios marginales en comparación de ocupación promedio con respecto a Cancún.

4.2 Consulta de opiniones de los consumidores (*sentiment analysis*).

En cuanto a la generación e intercambio de opiniones, los usuarios desean expresar su versión y conocimiento sobre los productos, servicios, experiencias que disfrutan en un destino turístico. La forma más común de intercambio de opinión está representada por los comentarios vertidos en una plataforma de comentarios (blog). Toda ésta información, por lo general carece de la estructura por lo que su extracción se dificulta, incluso en los sistemas actuales, debido a las miles de fuentes que se encuentran disponibles en línea.

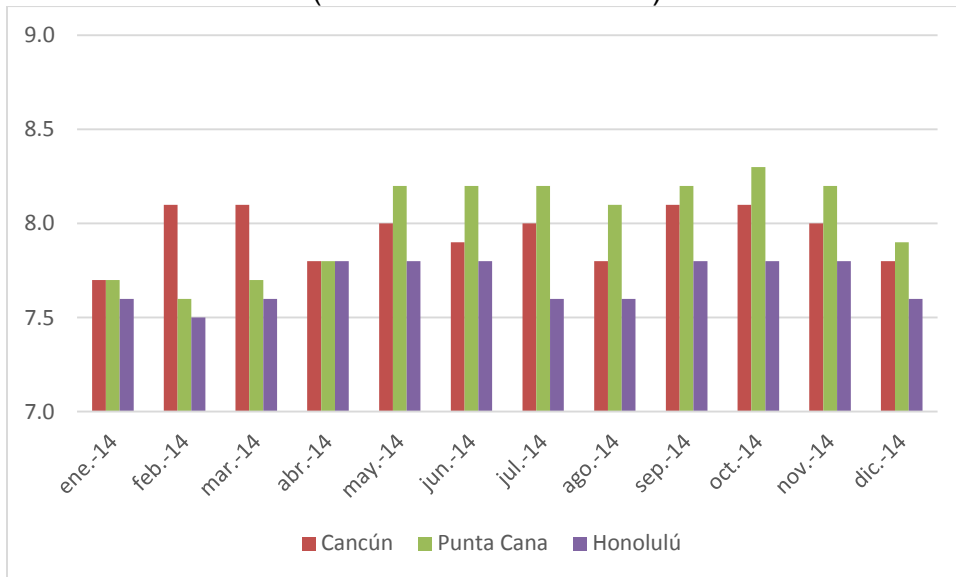
Sin herramientas de extracción de datos, los analistas del sector público tendrían que leer varios cientos de informes de texto, artículos de periódicos y los foros que se encuentran en las redes sociales para obtener un análisis del mercado estudiado. Además, manualmente descargar diversas aplicaciones para allegarse de la información e incluso tener la claridad sobre los hechos, descartando la información subjetiva en cada una de las fuentes de información. Debido a lo anterior, es importante considerar el uso de información con base en extractores de información.

Existen estudios empíricos como el de Lisi (2014) donde es posible especificar los servicios disponibles y los servicios proporcionados en un destino turístico. Existe una evaluación del usuario que permite clasificar los alojamientos dentro de un rango que tiene como extremos las definiciones “Excelentes” a “Terribles”. En el documento antes citado, las propiedades del extractor de datos permiten hacer referencia a los sitios por su nombre y lugares por dirección, código postal, ciudad y país.

4.2.1 Prueba piloto sobre opinión en destinos nacionales e internacionales seleccionados.

A continuación se presenta la información sobre el nivel de satisfacción de los turistas hospedados en Cancún, Punta Cana y Honolulu para el año 2014 (el nivel de satisfacción se mide en un rango de 0 a 10). En este apartado se cubrió información para realizar una comparación entre Cancún y los destinos internacionales seleccionados. A través de la solución en línea se consultó el módulo de satisfacción vía los comentarios que arroja el acopio de datos de los algoritmos de BeOnPrice para todas las categorías. Las unidades de alojamiento calificadas por turistas en la prueba piloto presentan los siguientes números: Cancún (165 unidades); Punta Cana (64 unidades); y Honolulu (96 unidades).

Gráfico 9. Reputación en línea para 2014.
(destinos seleccionados)



Fuente: Prueba piloto con información de BeOnPrice.

En el Gráfico 9 se observa que la satisfacción de los turistas en general supera los 7.5 puntos para los tres destinos en todo el año 2014. Este tipo de resultados puede ser útil para observar los periodos en donde se observa una reducción y los repuntes para poder mejorar la reputación en un destino específico. Los tres destinos presentan un comportamiento positivo en el nivel de satisfacción, alcanzando niveles de entre 7.5 y 9 puntos. Por lo tanto, se puede argumentar que los tres destinos correspondientes tuvieron una percepción positiva en sus visitantes y tienen diferencias para cada uno de los meses, con un incremento de Punta Cana para el segundo semestre de 2014.

Es importante considerar que Honolulu obtiene las opiniones con un menor nivel de satisfacción en todos los meses de 2014, principalmente cuando se hace referencia al mes de febrero (7.6) con 1,118 comentarios calificados. En segundo término, Punta Cana obtiene un buen nivel de opiniones en el segundo semestre del año 2014, observando su mejor indicador en el mes de octubre (8.3) con 252 comentarios calificados.

Por último, Cancún obtiene un buen nivel de opiniones en el primer semestre del año 2014, considerando que su mejor indicador (8.1) se mantiene para cuatro meses: febrero, marzo, septiembre y octubre; con 1,217; 1,539; 1,128; y 1,175 comentarios calificados, respectivamente.

5. Conclusiones.

En la prueba piloto expuesta en el presente documento se reunió información sobre proyecciones en tiempo real para el año 2015 de los precios promedio por habitación y ocupación promedio con reservaciones en línea para dos destinos a nivel nacional: Ciudad de México y Cancún; así como de dos destinos a nivel internacional: Punta Cana y Honolulu. Luego, se observó la calificación que obtuvieron en la satisfacción de los turistas para el año 2014.

De ésta manera, considerando los resultados de la extracción realizada el día 16 de enero de 2015 como prueba piloto, existe la posibilidad de que la ocupación promedio presenta una estacionalidad similar en la comparación de Cancún con respecto a cada uno de los destinos internacionales, considerando una periodicidad mensual. Además, los tres destinos seleccionados presentan un comportamiento positivo en el nivel de satisfacción de los turistas, alcanzando niveles de entre 7.5 y 9 puntos, considerando todo el año 2014.

Este estudio es el primero en investigar el uso productivo de información prospectiva y retrospectiva a partir de Big Data para realizar comparaciones tanto a nivel nacional como internacional en destinos turísticos. Sin embargo, la prueba piloto considera variables a nivel acumulado, en otras palabras, no se contempló la diferencia en precios promedio por tipo de habitación ni tampoco la ocupación promedio por categoría de hotel. En el caso de realizar un análisis con mayor detalle sobre las variables anteriores y también con una periodicidad mayor (para disminuir el efecto estacional) podrían mejorarse sustancialmente los resultados obtenidos.

En términos generales, la prueba piloto indica que el dinamismo de la actividad hotelera puede ser estudiado con mayor detalle considerando la información que surge en tiempo real sobre los destinos turísticos. En particular, son estudiados los precios promedio para una habitación y su comparación a nivel destino turístico; la ocupación promedio en el destino y su comparación durante todo un año; y la percepción de los turistas por destino como complemento para entender la rivalidad en precios que pueden presentar dos destinos a nivel internacional con uno de los destinos turísticos de mayor reconocimiento nacional: Cancún, Quintana Roo.

6. Anexo.

Cuadro 1. Interacción electrónica en el lugar de mercado, con ejemplos de turismo.

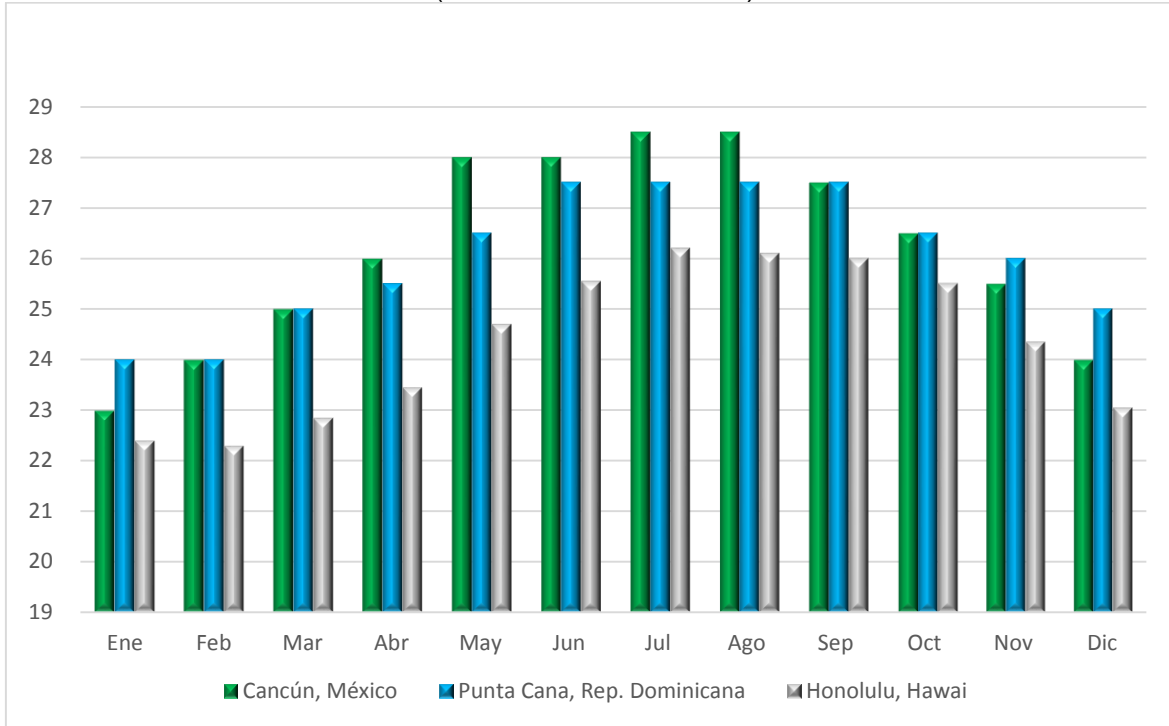
e-Negocio	Negocio	Cliente	Gobierno
Negocio	B2B Extranet entre los hoteleros y los tour operadores.	B2C Aplicaciones de comercio dónde el cliente compra boletos de avión.	B2G Interacción de negocios con departamentos gubernamentales, p. ej. Desarrollo hotelero requiere permiso de obras.
Cliente	C2B Clientes registrando sus preferencias sobre la lealtad a un hotel o aerolínea/ clubs exclusivos.	C2C Clientes informando a otros clientes sobre un bien o una mala práctica en el vicio o el producto.	C2G Clientes aplicando para visas, solicitando mapas y localizando información del destino.
Gobierno	G2B Gobierno informando a los hoteles sobre la legislación acerca de la seguridad en los alimentos o los impuestos.	G2C Gobierno informando a los clientes sobre regulaciones, visa o requerimientos de vacunación.	G2G Gobiernos interactuando en materia de políticas de turismo o solicitando asistencia técnica a través de organizaciones como la Asociación Mundial del Turismo.

Fuente: Conde (2011), p. 5.

Modelos de Negocios en el Comercio Electrónico. Adaptación de Conde (2011), p. 6.

- **Modelo Fachada** (aparador electrónico).- recolección de información que proveen los clientes y se muestra en una plataforma comercial.
- **Modelo de Subasta** (deRemate.com).- los vendedores colocan el precio al cual están dispuestos a vender su producto o servicio e invitan a los compradores a hacer su oferta para seleccionar a un único comprador.
- **Modelo del Portal** (conjunto de varios negocios).- utiliza las tecnologías de la información para reunir a compradores y vendedores a través de medios electrónicos, para crear un mercado virtual.
- **Modelo de Precios Dinámicos** (depreciación temporal o de acuerdo con la oferta y la demanda).- los vendedores ofrecen sus productos a los compradores y estos a su vez compran el producto a quien lo vende al precio más bajo.
- **Modelo de Trueque** (intercambio).- los usuarios intercambian productos en la red.
- **Modelo de Rebaja** (descuentos).- los vendedores ofrecen sus productos a los compradores y estos a su vez compran el producto con descuentos por volumen.
- **Modelo de Productos y Servicios Gratis** (promociones y clubes de fidelidad).
- **Modelo de intercambios B2B** (compras al mayoreo).- los participantes son negocios u otro tipo de organizaciones con gran volumen de compra.

Temperatura Mensual Promedio en °C
(destinos seleccionados)



Fuentes:

<http://blog.felicesvacaciones.es/70-tiempo-clima-y-temperaturas-en-punta-cana>

<http://www.cancunadventure.com.mx/informacion/clima.asp>

<http://es.climate-data.org/location/109/>

Bibliografía.

- [1] Conde, J.; Schmidt, E.; y Ochoa, I. (2011), “*El turismo electrónico, una necesidad para las empresas del sector*”, Revista de Investigación en turismo y desarrollo local. Vol 4, No. 9. Febrero.
- [2] Crandall, David; Lars Backstrom, Daniel Huttenlocher and Jon Kleinberg (2009), “*Mapping the World’s Photos*”, Department of Computer Science, Cornell University.
- [3] Economic Commission for Latin America and the Caribbean (CLAC) (2014), “*Big data and open data as sustainability tools*”, A working paper prepared by the Economic Commission for Latin America and the Caribbean, United Nations, October 2014.
- [4] Executive Office of the President (2014), “*Big Data: seizing opportunities, preserving values*”, May 2014.
- [5] Ferrara, Emilio; Pasquale De Meo; Giacomo Fiumara; Robert Baumgartner (2014), “*Web Data Extraction, Applications and Techniques: A Survey*”, Jun.
- [6] Fiumara, Giacomo (2007), “*Automated Information Extraction from Web Sources: a Survey*”, Italy.
- [7] Hal R. Varian (2013), “*Big Data: New Tricks for Econometrics*”, April.
- Haller, M.; B Proll; W. Retschitzegger; A M. Tjoa & R.R. Wagner (2000), “*Integrating Heterogeneous Tourism Information in TIScover-The MIRO-Web Approach*”, Information and Communication Technologies in Tourism, Barcelona.
- [8] Heerchap, Nico; Shirley Ortega, Alex Priem & May Offermans (2014), “*Innovation of tourism statistics through the use of new big data sources*”
- [9] Herget, Jan & Jan Fait (2014), “*Possibilities and limitations of data.mining in the Czech Tourism research*, Czech Tourism Insittute.
- [10] John Sutton (2017), “*Market Structure: Theory and Evidence*”, Handbook of Industrial Organization, Volume 3, Elsevier.
- Lisi, Francesca A. & Floriana Esposito (2014), “*Semantic Web Services for Integrated Tourism in the Apulia Region*”, Italy.
- [11] Sarawagi, Sunita (2008), “*Information Extraction*”, Indian Institute of Technology.
- [12] Saggion, Horacio & Adam Funk (2004), “*Extracting Opinions and Facts for Business Intelligence*”, EU-funded MUSING project.
- [13] Organización Mundial del Turismo (2014), Turismo y nuevas tecnologías – Seminario técnico, San José (Costa Rica), 14 y 15 de mayo de 2013, OMT, Madrid
- [14] Walchhofer, Norbert; Milan Hronský; Michael Poettler; Robert Baumgartner & Karl A. Froeschl (2010), “*Semantic Online Tourism Market Monitoring*”, Semantic Systems and Services, Austria.
- Páginas web.
- [1] Schaal, D. (2012). How Booking.com turned the other OTAs into converts. Recuperado 20 de enero de 2015 de: <http://skift.com/2012/06/25/how-booking-com-conquered-world/>

La serie de Documentos de Investigación Estadística y Económica presenta resultados preliminares de investigación realizados en la Secretaría de Turismo con el propósito de generar intercambio y debate de ideas para el desarrollo del sector turismo. El contenido de los Documentos de Investigación Estadística y Económica, así como los argumentos vertidos, son responsabilidad exclusiva de los autores y no reflejan necesariamente los de la Secretaría de Turismo.
