**Universidad Peruana de Ciencias Aplicadas**



**INFORME DE TRABAJO PARCIAL**

Fundamentos de Data Science

Carrera de Ciencias de la Computación

Sección: SC51

|  |  |
| --- | --- |
| Alumnos: | |
| Código | Nombres y apellidos |
| U202216613 | Mancusi Barreda, Giorgio |
| U202214421 | Hualtibamba Valerio, Mathias |
| U202211499 | Osorio Ramos, Fabio André |

2024

**Índice**

[1 Origen de datos 3](#_Toc165595847)

[2 Casos de uso 7](#_Toc165595848)

[2.1 ¿Cuántas reservas se realizan por tipo de hotel? o ¿Qué tipo de hotel prefiere la gente? 8](#_Toc165595849)

[2.2 ¿Está aumentando la demanda con el tiempo? 10](#_Toc165595850)

[2.3 ¿Cuándo se producen las temporadas de reservas: alta, media y baja? 10](#_Toc165595851)

[2.4 ¿Cuándo es menor la demanda de reservas? 12](#_Toc165595852)

[2.5 ¿Cuántas reservas incluyen niños y/o bebes? 13](#_Toc165595853)

[2.6 ¿Es importante contar con espacios de estacionamiento? 15](#_Toc165595854)

[2.7 ¿En qué meses del año se producen más cancelaciones de reservas? 16](#_Toc165595855)

[2.8 ¿Qué tipo de comida los clientes han obtenido dependiendo del hotel? 17](#_Toc165595856)

[2.9 ¿Los usuarios repetidos tienden a consumir / gastar más respecto a los primerizos? 18](#_Toc165595857)

[2.10 ¿En qué mes los turistas gastan menos? 18](#_Toc165595858)

[3 Conclusiones 19](#_Toc165595859)

[4 Bibliografía 21](#_Toc165595860)

# Origen de datos

Dataset: Propiedades y Usos

Este dataset, con los autores Nuno Antonio, Ana de Almedia y Luis Nunes, del 2018 presenta el conjunto de datos de dos hoteles, uno de un hotel resort y el otro de un hotel urbano. Ambos conjuntos de datos comparten la misma estructura, con 31 variables que describen 40.060 observaciones para el hotel resort y 79.330 observaciones para el hotel urbano. Las variables describen factores como las fechas de reserva, la información de los huéspedes, los tipos de habitación y la información de cancelación. Los datos se extrajeron de los sistemas de gestión de propiedades de los hoteles.

|  |  |
| --- | --- |
| **Variable** | **Descripción** |
| **ADR** | Tarifa diaria media definida por la suma de todas las transacciones de alojamiento por el número total de noches de estancia |
| **Adults** | Número de adultos |
| **Agent** | Identificación de la agencia de viajes que realizó la reserva |
| **ArrivalDateDayOfMonth** | Día del mes de la fecha de llegada |
| **ArrivalDateMonth** | Fecha de mes de llegada con 12 categorías: "Enero" a "Diciembre" |
| **ArrivalDateWeekNumber** | Número de semana de la fecha de llegada |
| **ArrivalDateYear** | Año de llegada |
| **AssignedRoomType** | Código del tipo de habitación asignada a la reserva. A veces, el tipo de habitación asignado difiere del tipo de habitación reservada debido a razones de operación del hotel (por ejemplo, overbooking) o por solicitud del cliente. El código se presenta en lugar de la designación por razones de anonimato |
| **Babies** | Número de bebés |
| **BookingChanges** | Número de cambios/modificaciones realizados en la reserva desde el momento en que se introdujo la reserva en el PMS hasta el momento del check-in o la cancelación |
| Children | Número de hijos |
| **Company** | Identificación de la empresa/entidad que realizó la reserva o responsable del pago de la misma. Se presenta la identificación en lugar de la designación por razones de anonimato |
| **C**ountry | País de origen. Las categorías están representadas en el formato ISO 3155-3:2013 [[6]](https://www.sciencedirect.com/science/article/pii/S2352340918315191#bib6) |
| **CustomerType** | Tipo de reserva, asumiendo una de las cuatro categorías: |
| Contrato: cuando la reserva tiene un cupo u otro tipo de contrato asociado; |
| Grupo: cuando la reserva está asociada a un grupo; |
| Transitoria: cuando la reserva no forma parte de un grupo o contrato, y no está asociada a otra reserva transitoria; |
| Parte transitoria: cuando la reserva es transitoria, pero está asociada al menos a otra reserva transitoria. |
| **DaysInWaitingList** | Número de días que la reserva estuvo en la lista de espera antes de ser confirmada al cliente |
| **DepositType** | Indicación de si el cliente ha realizado un depósito para garantizar la reserva. Esta variable puede asumir tres categorías: |
| Sin depósito: no se realizó ningún depósito; |
| No reembolso: se realizó un depósito por el valor del costo total de la estadía; |
| Reembolsable: se realizó un depósito con un valor inferior al costo total de la estadía. |
| **DistributionChannel** | Canal de distribución de reservas. El término "TA" significa "Agentes de Viajes" y "TO" significa "Operadores Turísticos" |
| **IsCanceled** | Valor que indica si la reserva fue cancelada (1) o no (0) |
| **IsRepeatedGuest** | Valor que indica si el nombre de la reserva es de un huésped repetido (1) o no (0) |
| **LeadTime** | Número de días transcurridos entre la fecha de entrada de la reserva en el PMS y la fecha de llegada |
| **MarketSegment** | Designación del segmento de mercado. En las categorías, el término "TA" significa "Agentes de Viajes" y "TO" significa "Operadores Turísticos" |
| Meal | Tipo de comida reservada. Las categorías se presentan en paquetes de comidas de hospitalidad estándar: |
| Indefinido/SC – sin paquete de comida; |
| BB – Alojamiento y Desayuno; |
| HB – Media pensión (desayuno y otra comida, generalmente cena); |
| FB – Pensión completa (desayuno, almuerzo y cena) |
| **PreviousBookingsNotCanceled** | Número de reservas anteriores no canceladas por el cliente antes de la reserva actual |
| **PreviousCancellations** | Número de reservas anteriores canceladas por el cliente antes de la reserva actual |
| **RequiredCardParkingSpaces** | Número de plazas de aparcamiento requeridas por el cliente |
| **ReservationStatus** | Último estado de la reserva, asumiendo una de las tres categorías: |
| Cancelado: la reserva fue cancelada por el cliente; |
| Check-Out: el cliente se ha registrado pero ya se ha ido; |
| No-Show: el cliente no se registró e informó al hotel del motivo |
| **ReservationStatusDate** | Fecha en la que se estableció el último estado. Esta variable se puede usar junto con ReservationStatus para comprender cuándo se canceló la reserva o cuándo se retiró el cliente del hotel |
| **ReservedRoomType** | Código del tipo de habitación reservada. El código se presenta en lugar de la designación por razones de anonimato |
| **EstanciasEnWeekendNights** | Número de noches de fin de semana (sábado o domingo) en las que se alojó o reservó el huésped para alojarse en el hotel |
| **StaysInWeekNights** | Número de noches entre semana (de lunes a viernes) en las que se alojó o reservó el huésped para alojarse en el hotel |
| **TotalOfSpecialRequests** | Número de peticiones especiales realizadas por el cliente |

# Casos de uso

Mas allá de los fines expresados por los autores, el uso y manipulación de la información dada nos permite desarrollar un razonamiento lógico mediante la organización de datos y la propuesta de insights. Además, nos permite responder a preguntas como: ¿Quien o quienes se benefician?

Si bien todo tipo de respuestas estarán orientadas a la industria hotelera y el consumidor (beneficiándose ambos), estas se descompondrán en preguntas específicas:

1. ¿Cuántas reservas se realizan por tipo de hotel? o ¿Qué tipo de hotel prefiere la gente?
2. ¿Está aumentando la demanda con el tiempo?
3. ¿Cuándo se producen las temporadas de reservas: alta, media y baja?
4. ¿Cuándo es menor la demanda de reservas?
5. ¿Cuántas reservas incluyen niños y/o bebes?
6. ¿Es importante contar con espacios de estacionamiento?
7. ¿En qué meses del año se producen más cancelaciones de reservas?
8. ¿Qué tipo de comida los clientes han obtenido dependiendo del hotel?
9. ¿Los usuarios repetidos tienden a consumir / gastar más respecto a los primerizos?
10. ¿En qué mes los turistas gastan menos?

Se hará un análisis a las siguientes preguntas; se propondrá una solución con una visualización adjunta a esta y, finalmente, se darán insights a la industria.

**Desarrollo**

## ¿Cuántas reservas se realizan por tipo de hotel? o ¿Qué tipo de hotel prefiere la gente?

De las 120,000 reservas realizadas:

* Hotel tipo Urbano: 79,330 reservas
* Hotel tipo Resort: 40,060 reservas

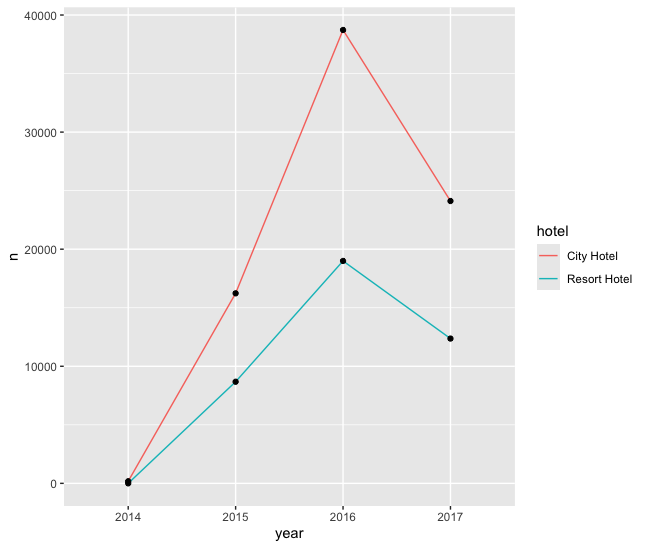
1. A graph of a hotel

   Description automatically generatedMétrica mediante ggplot para el cálculo de la cantidad de usuarios en un hotel

Insight: En caso de construir un nuevo hotel se recomienda que este sea de tipo urbano, ya que es este el tipo de hotel preferido por la mayoría de los clientes.

## ¿Está aumentando la demanda con el tiempo?

Si, se evidencia un aumento de 15,000 clientes a 40,000 clientes entre 2015 a 2016; sin embargo, el tráfico disminuyo a 25,000 en 2017.



1. Imagen de clientes durante el tiempo

Insight: En caso se desee invertir en la industria hotelera, se recomienda estar al tanto del contexto respecto al negocio; esto ya que, si bien hubo un incremento en cuanto al tráfico de clientes, ese se vio considerablemente alterado en el último año.

## ¿Cuándo se producen las temporadas de reservas: alta, media y baja?

La temporada alta es en: Mayo, julio, agosto, octubre  
La temporada media es en: Marzo, abril, junio, septiembre  
La temporada baja es en: Enero, febrero, noviembre, diciembre

A screenshot of a computer

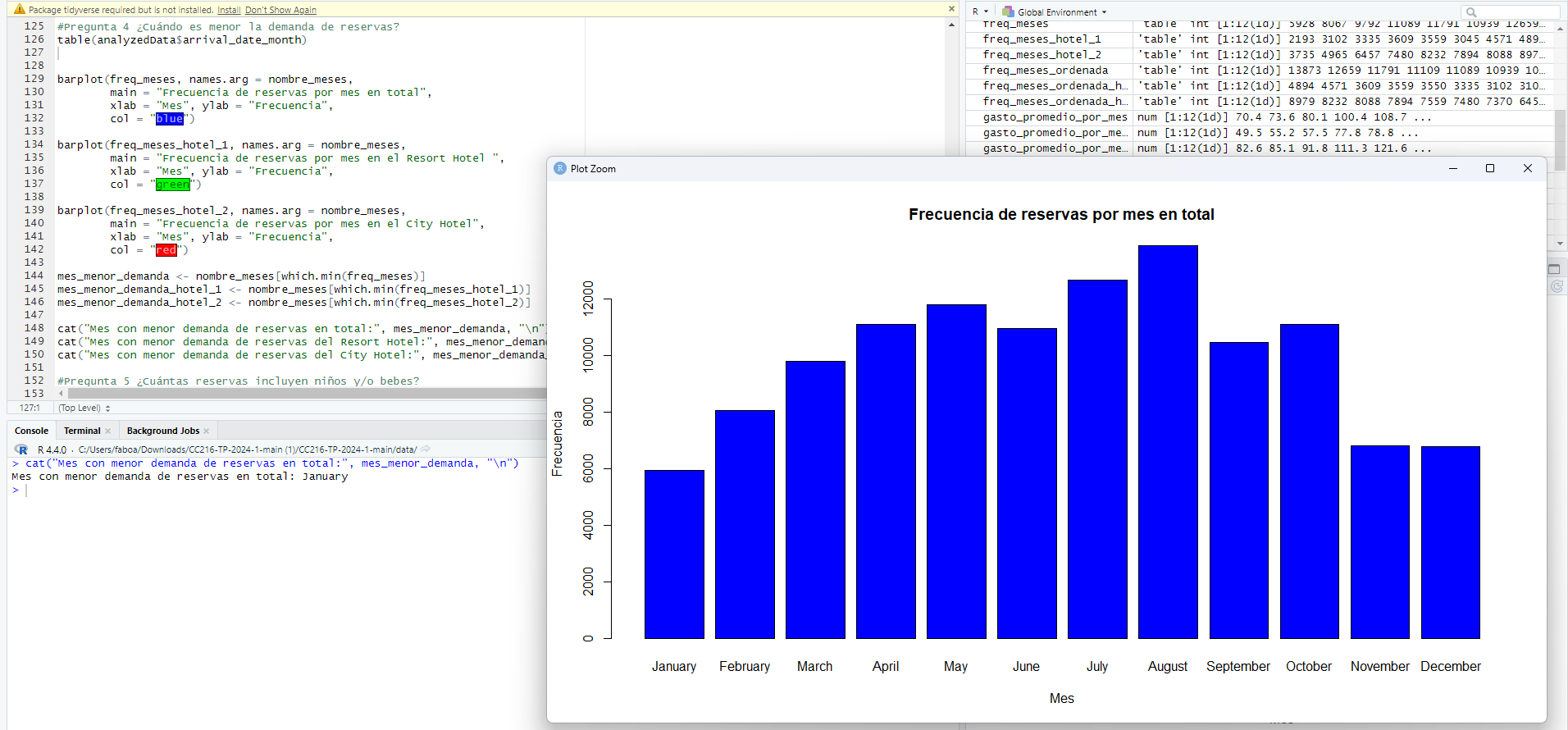
Description automatically generated

1. Figura de frecuencia por meses

Insight: El patrón de temporadas de reservas revela que la demanda turística está influenciada por factores estacionales, como el clima y las vacaciones escolares. La temporada alta coincide con los meses de verano y octubre, mientras que la temporada baja abarca los meses de invierno y fin de año. Este conocimiento es útil para ajustar estrategias de precios y promociones, maximizando los ingresos durante los períodos de alta demanda y atrayendo clientes durante la temporada baja con ofertas especiales.

## ¿Cuándo es menor la demanda de reservas?

La demanda de reserva es menor en enero.



1. Figura de reserva por meses en total

Insight: Probablemente debido a la recuperación financiera después de las festividades de fin de año o el clima desfavorable las reservas en enero son considerablemente bajas. Esta información es esencial para que las empresas turísticas ajusten sus estrategias de marketing y precios durante este período de menor demanda.

## ¿Cuántas reservas incluyen niños y/o bebes?

Se han encontrado 9506 reservas de las cuales incluyen niños y/o bebes.

Hotel tipo urbano: 5473 de estas reservas pertenecen a este tipo de hotel.

Hotel tipo resort: 4031 de estas reservas pertenecen a este tipo de hotel.

A graph of a graph showing a number of hotels

Description automatically generated with medium confidence

1. Imagen de datos con mínimo un bebe o un niño

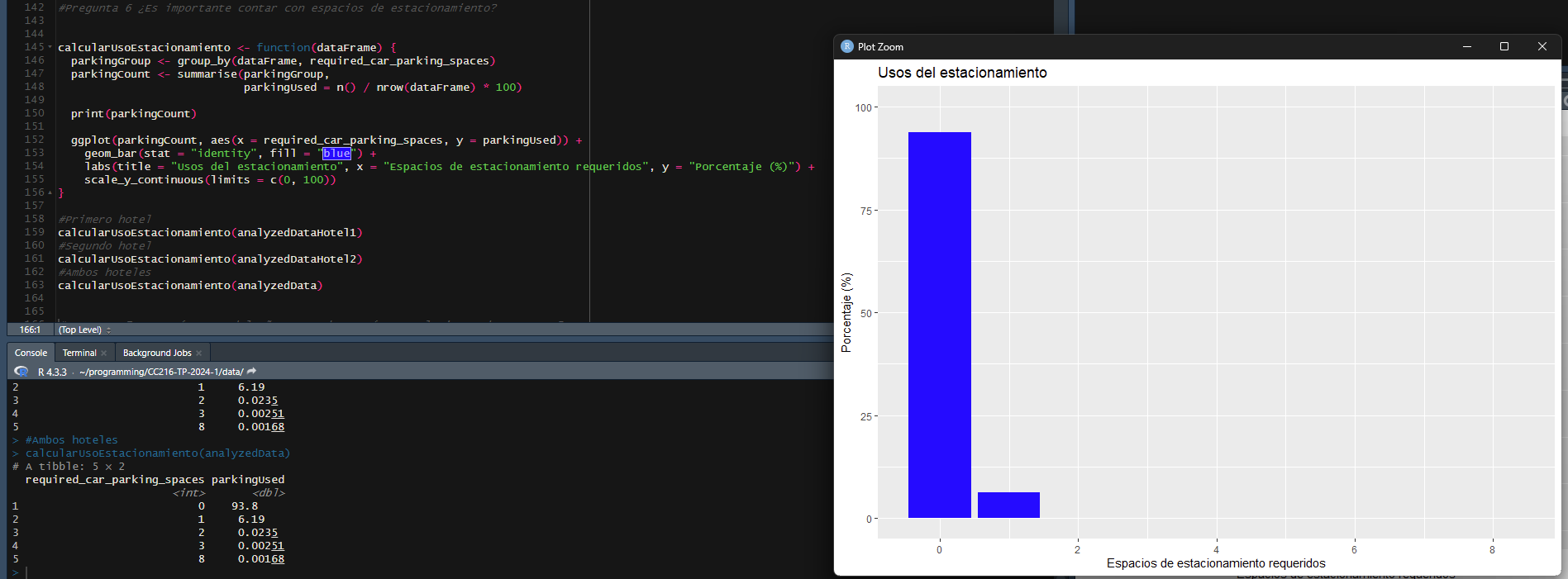
A graph of a bar chart

Description automatically generated

1. Imagen de datos con mínimo un bebe y un niño

Insight: Considerando que, aproximadamente, el 10% de los clientes va a tener un niño o un bebe, se recomienda ofrecer paquetes personalizados para este tipo de clientes.

## ¿Es importante contar con espacios de estacionamiento?

No, la disponibilidad de estacionamiento no marcaria una diferencia mínimamente notable en el tráfico de clientes. El análisis de datos indica que un 93.8% no ocupa estacionamiento, mientras que apenas un 6.19% ocupa un espacio de estacionamiento.

1. Figura del uso del estacionamiento por cliente

Insight: En caso una nueva marca hotelera desee distribuir sus inversiones, y esta sea limitada; entre las necesidades no debe encontrarse la inversión en un estacionamiento, ya que apenas afectara al tráfico de clientes y, por ende, los ingresos.

## ¿En qué meses del año se producen más cancelaciones de reservas?

La mayor cantidad de cancelaciones de reservas en el año se generan en enero; este comportamiento se ve más pronunciado en los hoteles tipos urbano.

A graph of a graph of a graph

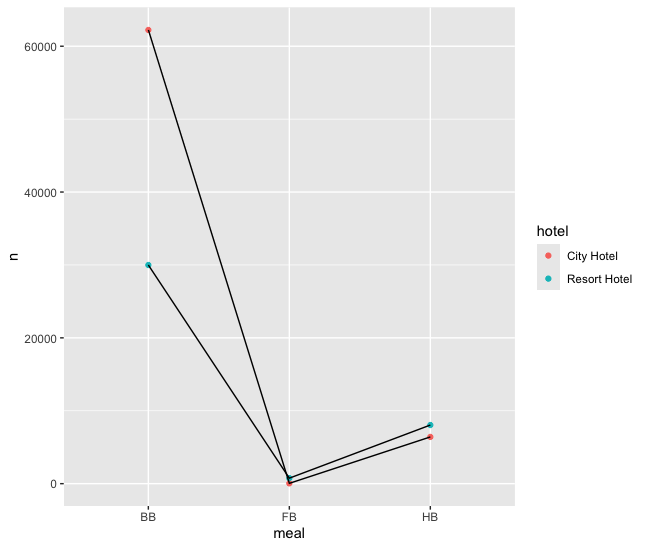
Description automatically generated with medium confidence

1. Figura de las cancelaciones ordenado por meses

Insight: Los factores económicos o climáticos pueden afectar el flujo de cancelaciones de reservas. Se recomienda ofrecer otro tipo de promociones para no perder a estos clientes.

## ¿Qué tipo de comida los clientes han obtenido dependiendo del hotel?

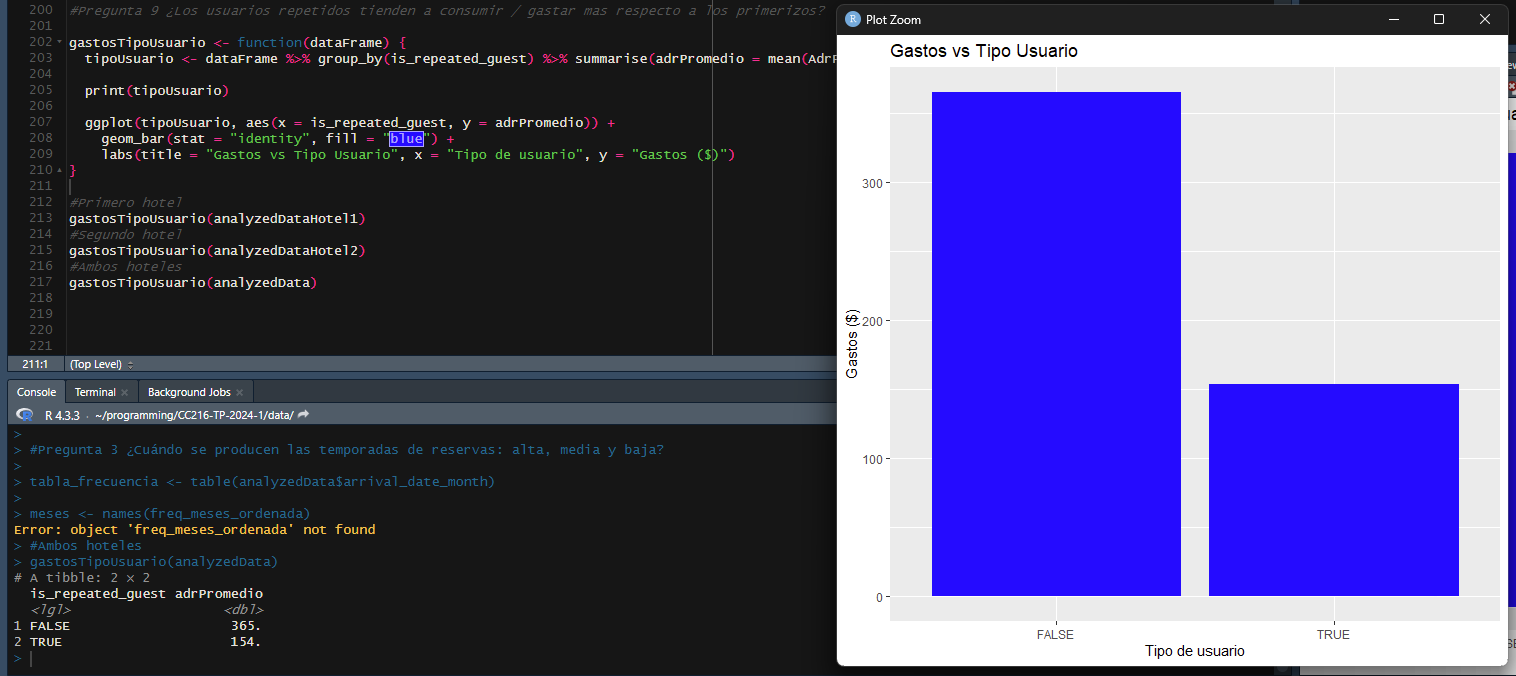
Se ha encontrado que hubo una gran cantidad de clientes que han solicitado el tipo de comida en la cama, *BB* o *Bed & Breakfast, HB o Half Board* y *FB* o *Full Board*. Dicha información aplica tanto para los hoteles tipo resort o urbano.



1. Figura de tipos de comida seleccionado por tipo de hotel

Insight: Se recomienda potenciarla promoción de servicios tipo BB en los hoteles, ya que esto ha obtenido una gran cantidad de usuarios a comparación de las otras dos opciones.

## ¿Los usuarios repetidos tienden a consumir / gastar más respecto a los primerizos?

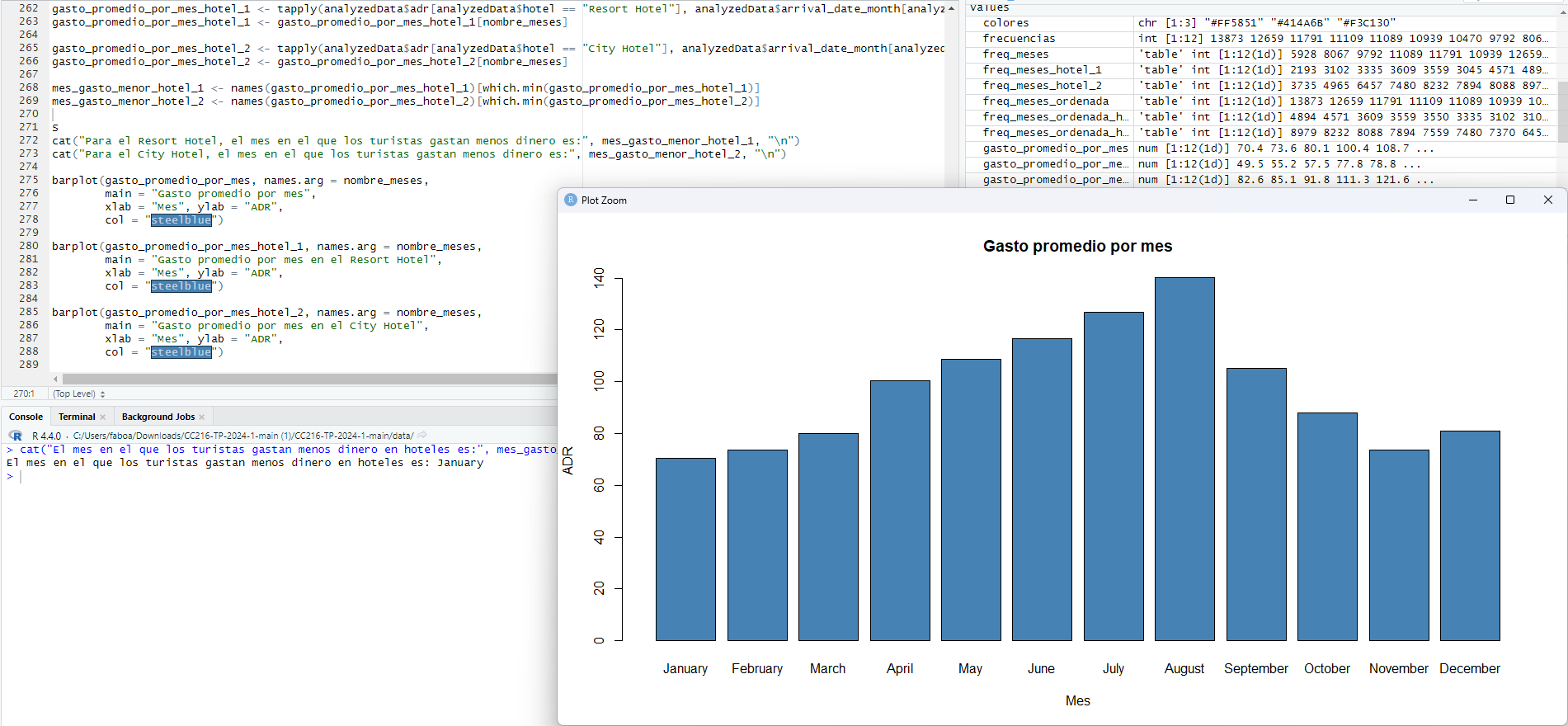
Considerando como un usuario repetido aquel que viene, como mínimo, por segunda vez: No; de hecho, se evidencia un gasto significativamente menor a un cliente que visita por primera vez el hotel. Los clientes nuevos gastan en promedio 365$ durante su estadía, mientras que aquellos que no es su primera vez gastan en promedio 154$. Esto puede deberse principalmente a que aquel que viene por segunda vez no experimenta tanto como un cliente nuevo; es decir, este ya sabe cuánto y donde gastar el dinero.

1. Gastos por tipo de usuario, si este ha estado anteriormente en el hotel o no

Insight: En caso se detecte un cliente como “no nuevo”, promocionarle los nuevos servicios ofrecidos por el hotel y catalogarlas como una experiencia nueva, así como evidenciar las diferencias con las experiencias ofrecidas en el pasado.

## ¿En qué mes los turistas gastan menos?

Los turistas gastan menos dinero en enero.



1. Figura de gasto por turista

Insight: Debido a la menor demanda, enero podría ser un momento ideal para planificar vacaciones más económicas. Esto podría atraer a aquellos que prefieren evitar las multitudes y que están dispuestos a sacrificar ciertas comodidades por precios más bajos. Las empresas turísticas podrían capitalizar la tendencia ofreciendo paquetes y promociones diseñados para atraer a este segmento de mercado durante enero.

# Conclusiones

En síntesis, podemos afirmar:

* En la industria hotelera, son los tipos resort aquellos que reciben una mayor cantidad de clientes e ingresos.
* Comprender las temporadas es fundamental para ajustar las estrategias de precios, promociones y gestión de recursos de acuerdo con la demanda esperada en cada período.
* Los clientes nuevos tienden a gastar más que los clientes recurrentes. Esta conclusión resalta la importancia de ofrecer experiencias novedosas y promocionar constantemente nuevos servicios y ofertas, incluso para los huéspedes habituales, con el fin de mantener su interés y niveles de gasto.
* Se requiere incluir servicio de comida a la habitación, este se ha encontrado una gran cantidad de usuarios que lo solicitan sea de resort o tipo urbano

# Bibliografía

Antonio, N., de Almeida, A., & Nunes, L. (2019). Hotel booking demand datasets. *Data in Brief*, *22*, 41–49. <https://doi.org/10.1016/J.DIB.2018.11.126>