



Listas de contenidos disponibles en [ScienceDirect](https://www.sciencedirect.com)

## Datos en breve

revista página de inicio: [www.elsevier.com/locate/dib](http://www.elsevier.com/locate/dib)



### Artículo de datos

# Conjuntos de datos sobre la demanda de reservas hotel CrossMark

Nuno Antonio <sup>a,a,b,n</sup>, Ana de Almeida <sup>a,c,d</sup>, Luis Nunes <sup>b,d</sup>

<sup>a</sup> Instituto Universitário de Lisboa (ISCTE-IUL), Lisboa, Portugal

<sup>b</sup> Instituto de Telecomunicações, Lisboa, Portugal

<sup>c</sup> CISUC, Coimbra, Portugal

<sup>d</sup> ISTAR-IUL, Lisboa, Portugal

### A R T Í C U L O E N F

#### Historia del artículo:

Recibido el 5 de octubre de 2018

Aceptado el 26 de noviembre de 2018

Disponible en línea el 29 de noviembre de 2018

### A B S T R A C

Este artículo describe dos conjuntos de datos con datos de demanda de hoteles. Uno de los hoteles (H1) es un hotel turístico y el otro es un hotel urbano (H2). Ambos conjuntos de datos comparten la misma estructura, con 31 variables que describen las 40.060 observaciones de H1 y las 79.330 observaciones de H2. Cada observación representa una reserva de hotel. Ambos conjuntos de datos comprenden las reservas que debían llegar entre el 1 de julio de 2015 y el 31 de agosto de 2017, incluidas las reservas que llegaron efectivamente y las que se cancelaron. Dado que se trata de datos reales de hoteles, se eliminaron todos los elementos de datos relativos a la identificación del hotel o del cliente. Debido a la escasez de datos comerciales reales para fines científicos y educativos, estos conjuntos de datos pueden desempeñar un papel importante para la investigación y la educación en la gestión de ingresos, el aprendizaje automático o la minería de datos, así como en otros campos.

© 2018 Los autores. Publicado por Elsevier Inc. Se trata de un artículo de acceso bajo licencia CC BY

(<http://creativecommons.org/licenses/by/4.0/>).

### Tabla de especificaciones

Área temática	Hospitality Management
Área temática más específica	Revenue Management Tipo de
datos	Archivos de texto y objetos
R	
Cómo se obtuvieron los datos	Extracción de datos SQL del sistema de gestión de la propiedad (PMS) de los hoteles bases de datos

<sup>n</sup> Autor correspondiente.

Dirección de correo electrónico: [nuno\\_miguel\\_antonio@iscte-iul.pt](mailto:nuno_miguel_antonio@iscte-iul.pt) (N. Antonio).

Formato de los datos	Mixtos (brutos y preprocesados)
Factores experimentales	Algunas de las variables se crearon a partir de otras variables de diferentes tablas de la base de datos. La hora del punto de datos para cada observación se definió como el día anterior a la llegada de cada reserva
Características experimentales	Los datos se extrajeron mediante consultas TSQL ejecutadas directamente en las bases de datos PMS de los hoteles y se empleó R para realizar el análisis de datos
Ubicación de la fuente de datos	Ambos hoteles están situados en Portugal: el H1 en la región turística del Algarve y el H2 en la ciudad de Lisboa.
Accesibilidad de los datos	Los datos se suministran con el documento

Valor de los datos

- El análisis descriptivo puede emplearse para comprender mejor las pautas, tendencias y anomalías de los datos;
- Se utiliza para investigar diferentes problemas como: predicción de cancelación de reservas, clientes segmentación, saciedad del cliente, estacionalidad, entre otros;
- Los investigadores pueden utilizar los conjuntos de datos para comparar los modelos de cancelación de predicciones de reservas con los resultados ya conocidos (por ejemplo, [1]);
- Los investigadores en aprendizaje automático pueden utilizar los conjuntos de datos para evaluar el rendimiento de distintos algoritmos a la hora de resolver el mismo tipo de problema (clasificación, segmentación u otros);
- Los educadores pueden utilizar los conjuntos de datos para problemas de clasificación o segmentación de aprendizaje automático;
- Los educadores pueden utilizar los conjuntos de datos para formarse en estadística o minería de datos.

1. Datos

En las industrias relacionadas con el turismo y los viajes, la mayoría de las investigaciones sobre problemas de previsión y predicción de la demanda en Revenue Management emplean datos de la industria de la aviación, en el formato conocido como Registro de Nombres de Pasajeros (PNR). Se trata de un formato desarrollado por la industria de la aviación [2]. Sin embargo, los demás sectores del turismo y los viajes, como la hostelería, los cruceros, los parques temáticos, etc., tienen requisitos y particularidades diferentes que no pueden explorarse a fondo sin datos específicos del sector. De ahí que se compartan dos conjuntos de datos de hoteles con datos de demanda para ayudar a superar esta limitación.

Los conjuntos de datos ahora disponibles se recopilaron con el objetivo de desarrollar modelos de predicción para clasificar la probabilidad de cancelación de una reserva hotelera. Sin embargo, debido a las características de las variables incluidas en estos conjuntos de datos, su uso va más allá de este problema de predicción de cancelaciones.

Una de las propiedades más importantes de los datos para los modelos de predicción es no favorecer la fuga de información futura [3]. Para , la marca temporal de la variable objetivo debe ser posterior a la de las variables de entrada. Así, en lugar extraer directamente las variables de la tabla de la base de datos de reservas, cuando estaba disponible, los valores de las variables se extrajeron del registro de cambios de las reservas, con una marca de tiempo relativa al día anterior a la fecha de llegada (para todas las reservas creadas antes de su fecha de llegada).

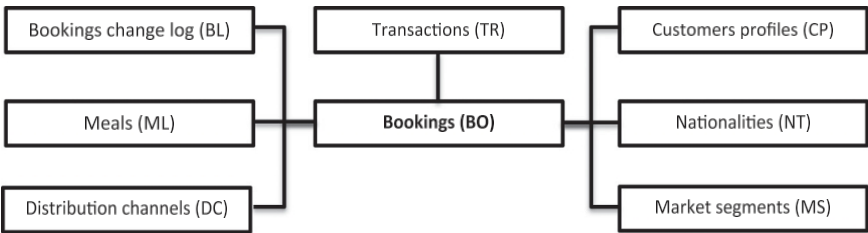


Fig. 1. Diagrama de las tablas de la base de datos PMS de donde se extrajeron las variables.

Cuadro 1

Descripción de las variables.

Variable	Tipo	Descripción	Fuente/Ingeniería
<i>ADR</i>	Numérico	Tarifa media diaria definida por [5]	BO, BL y TR / Calculada dividiendo la suma de todas las transacciones de alojamiento por el número total de noches de estancia.
<i>Adultos</i>	Entero	Número de adultos	BO y BL
<i>Agente</i>	ID categórico	de la agencia de viajes que realizó la reserva(s)	BO y BL
<i>ArrivalDateDayOfMonth</i>	Integer	Día del mes de la fecha de llegada	BO y BL
<i>ArrivalDateMonth</i>		Mes categórico de la fecha de llegada con 12 categorías: "Enero" a "Diciembre"	BO y BL
<i>ArrivalDateWeekNumber</i>	Integer	Número de semana de la fecha de llegada	BO y BL
<i>ArrivalDateYear</i>	Integer	Año de la fecha de llegada	BO y BL
<i>AssignedRoomType</i>	Categorical	Código del tipo de habitación asignada al BO y BL reserva. A veces el tipo de habitación asignado difiere del tipo de habitación reservado debido a razones operativas del hotel (por ejemplo, overbooking) o a petición del cliente. El código se presenta en lugar de la designación por razones de anonimato	
<i>Bebés</i>	Entero	Número de bebés	BO y BL
<i>BookingChanges</i>	Entero	Número de cambios/enmiendas realizados a la reserva desde el momento en que ésta se introdujo en el PMS hasta el momento de la facturación o cancelación	BO y BL/Calculado sumando el número de iteraciones únicas que cambian algunos de los atributos de la reserva, a saber: personas, fecha de llegada, noches, tipo de habitación reservada o comida.
<i>Hijos</i>	Entero	Número de hijos	BO y BL/Suma de ambos pagaderos y hijos no abonados
<i>Empresa</i>		Identificación categórica de la empresa/entidad que realizó la reserva de BO y BL. o responsable del pago del reserva. Se presenta el DNI en lugar de la por razones de anonimato.	
<i>País</i>	Categórico	País de origen. Las categorías se representan en el formato ISO 3155-3:2013 [6].	BO, BL y NT
<i>TipoCliente</i>	Categórico	Tipo de reserva, asumiendo uno de cuatro categorías: Contrato: cuando la reserva tiene asociada una adjudicación u otro tipo de contrato; Grupo: cuando la reserva está a un grupo; Transitoria: cuando la reserva no parte de un grupo o contrato, y no está asociada a otra reserva transitoria; Transitoria-parte: cuando la reserva es transitoria, pero está asociada al menos a otra reserva transitoria.	BO y BL
<i>DaysInWaitingList</i>	Entero	Número de días que la reserva estuvo en la lista de espera antes de que se confirmara al cliente	BO/Calculado restando la fecha de confirmación de la reserva al cliente de la fecha de entrada de la reserva en el PMS.
<i>DepositType</i>	Categórica	Indicación de si el cliente realizó un depósito para garantizar la reserva. Esta variable puede asumir tres categorías Sin depósito: no se ha realizado ningún depósito;  No reembolso - se realizó un depósito por el valor del coste total de la estancia;  Reembolsable: se ha realizado un depósito por un valor inferior al coste total de la estancia.	BO y TR/Valor calculado en función de los pagos identificados para la reserva en la tabla de transacciones (TR) antes de la fecha de llegada o cancelación de la reserva. En caso de que no se haya encontrado ningún pago, el valor es "Sin depósito". Si el pago fue igual o superior al coste total de la estancia, el valor se establece como "No reembolsado". En caso contrario, el valor se establece como "Reembolsable"

Cuadro 1 (continuación)

Variable	Tipo	Descripción	Fuente/Ingeniería
<i>Canal de distribución</i>	Catagórico	Canal de distribución de reservas. El término "AT" significa "Agencias de Viajes" y "TO" significa "turoperadores"	BO, BL y DC
<i>IsCanceled</i>	Valor categórico	que indica si la reserva se ha cancelado (1) o no (0)	BO
<i>IsRepeatedGuest</i>	Valor categórico	que indica si el nombre de la reserva era de un invitado repetido (1) o no (0)	BO, BL y C/ Variable creada por verificando si un perfil estaba asociado con el cliente de la reserva. En caso afirmativo, y si la fecha de creación del perfil del cliente era anterior a la fecha de creación de la reserva en la base de datos del PMS, se asumía que la reserva procedía de un cliente repetido.
<i>Plazo de entrega</i>	Entero	Número de días transcurridos entre la fecha de entrada de la reserva en la fecha del PMS de la fecha de llegada y la fecha de llegada	BO y BL/ Resta de la entrada
<i>Segmento de mercado</i>	Designación	categoría del segmento de mercado. En categorías, el término "AT" significa "Agencias de Viajes" y "TO" significa "Tour Operadores".	BO, BL y MS
<i>Comida</i>	Catagórica	Tipo de comida reservada. Las categorías se presentan en paquetes estándar de comidas de hospitalidad: No definido/SC - sin régimen de comidas; BB - alojamiento y desayuno; Media pensión (desayuno y otra comida, normalmente la cena); FB - Pensión completa (desayuno, almuerzo y cena)	BO, BL y ML
<i>ReservasAnterioresNoCanceladas</i>	Entero	Número de reservas anteriores no cancelada por el cliente antes de la reserva en curso	BO y BL / En caso de que no hubiera ningún perfil de cliente asociado a la reserva, el valor se establece en 0. caso contrario, el valor es el número de reservas con el mismo perfil de cliente creadas antes de la reserva actual y no canceladas.
<i>Anulaciones anteriores</i>	Entero	Número de reservas anteriores que fueron cancelada por el cliente antes de la reserva en curso	BO y BL/ En caso de que no hubiera perfil de cliente asociado a la reserva, el valor se establece en 0. caso contrario, el valor es el número de reservas con el mismo perfil de cliente creadas antes de la reserva actual y canceladas.
<i>RequiredCardParkingSpaces requeridas</i>	Entero	Número de plazas de aparcamiento por el cliente	BO y BL
<i>ReservationStatus</i>	Catagórico	Último estado la reserva, suponiendo uno de los siguientes tres categorías: Cancelado: el cliente ha cancelado la reserva; Check-Out - el cliente se ha registrado pero ya se ha marchado; No presentación: el cliente no se ha presentado y ha informado al hotel del motivo.	BO

Cuadro 1 (continuación)

Variable	Tipo	Descripción	Fuente/Ingeniería
<i>ReservationStatusDate</i>	Fecha	Fecha en la que se estableció el último estado. Esta variable se puede utilizar junto con <i>ReservationStatus</i> para saber cuándo se canceló la reserva o cuándo abandonó el cliente el hotel.	BO
<i>ReservedRoomType</i> código es	Código categórico del tipo de habitación reservada. El	presentado en lugar de la designación por razones de anonimato	BO y BL
<i>EstanciasEnNochesDeFinDeSemana</i>	Entero	Número de noches de fin de semana (sábado o BO y BL/ Calculado contando las domingos) el huésped se alojó o reservó hasta el número de noches de fin de semana de la estancia en el hotel	
<i>EstanciasEnNochesSemana</i>	Entero	Número de noches entre semana (de lunes a viernes- BO y BL/Calculado contando las días) el huésped se alojó o reservó para alojarse número de noches de la semana del total en el hotel	
<i>TotalOfSpecialRequests</i>	Entero	Número de solicitudes especiales realizadas por el cliente BO y BL/Suma de todas las solicitudes especiales (por ejemplo, cama doble o piso alto)	

<sup>a</sup> Se presenta el DNI en lugar de la designación por razones de anonimato.

No todas las variables de estos conjuntos de datos proceden de las tablas de la base de datos de reservas o de registro de cambios. Algunas proceden de otras tablas y otras se han creado a partir de distintas variables de diferentes tablas. En la Fig. 1 se presenta un diagrama con las tablas de la base de datos PMS de las que se extrajeron las variables. En la sección siguiente se ofrece una descripción detallada de cada variable.

## 2. Diseño experimental, materiales y métodos

datos se obtuvieron directamente de los servidores de las bases de datos PMS de los hoteles ejecutando una consulta TSQL en SQL Server Studio Manager, la herramienta de entorno integrado para gestionar bases de datos Microsoft SQL [4]. Esta consulta recogía en primer lugar el valor o ID (en el caso de claves externas) de cada variable de la tabla BO. A continuación, se comprobaba si la tabla BL había sufrido alguna alteración con respecto al día anterior a la llegada. Si se encontraba una alteración, se utilizaba el valor presente en la tabla BL. Para todas las variables con valores en tablas relacionadas (como comidas, canales de distribución, nacionalidades o segmentos de mercado), se recuperaron sus valores relacionados. En el Cuadro 1 figura una descripción detallada de las extraídas, su origen y los procedimientos de ingeniería empleados en su creación.

La PMS asegura que no existen datos ausentes en las tablas de su base de datos. Sin embargo, en algunas variables categóricas como Agente o Empresa, "NULL" se presenta como una de las categorías. Esto no debe considerarse como un valor que falta, sino como "no aplicable". Por ejemplo, si una reserva "Agente" se define como "NULL" significa que la reserva no procedía de una agencia de viajes.

En los cuadros 2–7 se presentan estadísticas resumidas de los conjuntos de datos de ambos hoteles. Estas estadísticas se obtuvieron utilizando el paquete R "skimr" [7].

Los que no estén familiarizados con las operaciones hoteleras deben tener cuidado. En el sector hotelero, es bastante habitual que los clientes cambien los atributos de su reserva, como el número de personas, la duración de la estancia o el tipo de habitación que prefieren, ya sea en el momento de registrarse o durante su estancia. También es habitual que los hoteles no conozcan la nacionalidad correcta del cliente hasta el momento del check-in. Por lo tanto, aunque la captura de datos se haya realizado considerando un intervalo de tiempo anterior a la fecha de llegada, es comprensible que la distribución de algunas variables difiera entre las reservas no canceladas y las canceladas. En consecuencia, el uso de estos conjuntos de datos puede requerir que se tenga en cuenta esta diferencia en la distribución. Esta diferencia puede observarse en los gráficos de tabla de las Fig. 2 y Fig. 3. Los gráficos de tabla son un potente método de visualización y se produjeron con el paquete *tabplot* R [8] que permite la exploración y el análisis de grandes conjuntos de datos multivariantes. En los gráficos de tabla, cada columna representa una variable y cada fila, un contenedor con un número predefinido de observaciones. En estas dos, cada casilla

Cuadro 2  
Estadísticas resumidas del conjunto de datos H1 - Variables de fecha.

Variable	Min	Max	Mediana	Único
<i>ReservationStatusDate</i>	2014-11-18	2017-09-14	2016-07-31	913

Cuadro 3  
Estadísticas resumidas del conjunto de datos H1 - Variables categóricas.

Variable	Único	Recuentos máximos
<i>Agente</i>	186	240: 13 095, NULO: 8 209, 250: 2 869, 241: 1 721
<i>FechaLlegadaMes</i>	12	Ago: 4 894, Jul: 4 573, abr: 3 609, may: 3 559
<i>AssignedRoomType</i>	11	A: 17 046, D: 10 339, E: 5 638, C: 2 214
<i>Empresa</i>	236	NULO: 36 952, 223: 784, 281: 138, 154: 133
<i>País</i>	125	PRT: 17 630, GBR: 6 814, ESP: 3 957, IRL: 2 166
<i>TipoCliente</i>	4	Tra.: 30 209, Tra.-Partido: 7 791, Con.: 1 776, Gro.:284
<i>Tipo de depósito</i>	3	No Dep: 38 199, No Dep.: 1 719, Ref.: 142
<i>Canal de distribución</i>	4	TA/TO: 28 295, Dir.: 7 865, Cor: 3 269, Und.: 1
<i>IsCanceled</i>	2	0: 28 938, 1: 11 122
<i>IsRepeatedGuest</i>	2	0: 38 282, 1: 1 778
<i>Segmento de mercado</i>	6	OnL.: 17 729, Off.: 7472, Dir.: 6 513, Gro.: 5 836
<i>Comida</i>	5	BB: 30 005, HB: 8 046, Und.: 1 169, FB: 754
<i>ReservationStatus</i>	3	C.Out: 28 938, Can.: 10 831, No-Show: 291
<i>ReservedRoomType</i>	10	A: 23 399, D: 7 433, E: 4 892, G: 1610

Cuadro 4  
Estadísticas resumidas del conjunto de datos H1 - Variables enteras y numéricas.

Variable	Media	SD	P0	P25	Mediana	P75	P100
<i>ADR</i>	94.95	61.44	-6.38	50	75	125	508
<i>Adultos</i>	1.87	0.7	0	2	2	2	55
<i>ArrivalDateOfMonth</i>	15.82	8.88	1	8	16	24	31
<i>FechaLlegadaNúmeroSemana</i>	27.14	14.01	1	16	28	38	53
<i>LlegadaFechaAño</i>	2016.12	0.72	2015	2016	2016	2017	2017
<i>Bebés</i>	0.014	0.12	0	0	0	0	2
<i>BookingChanges</i>	0.29	0.73	0	0	0	0	17
<i>Niños</i>	0.13	0.45	0	0	0	0	10
<i>DaysInWaitingList</i>	0.53	7.43	0	0	0	0	185
<i>LeadTime</i>	92.68	97.29	0	10	57	155	737
<i>AnteriorBookingsNotCanceled</i>	0.15	1	0	0	0	0	30
<i>AnteriorCancelaciones</i>	0.1	1.34	0	0	0	0	26
<i>Plazas de aparcamiento obligatorias</i>	0.14	0.35	0	0	0	0	8
<i>EstanciasEnNochesDeFinDeSemana</i>	1.19	1.15	0	0	1	2	19
<i>EstanciasEnNochesDeSemana</i>	3.13	2.46	0	1	3	5	50
<i>TotalDePeticionesEspeciales</i>	0.62	0.81	0	0	0	1	5

contiene 100 observaciones. Las barras de cada variable muestran el valor medio para las variables numéricas o la frecuencia de cada nivel para las variables categóricas. Analizando estas figuras se puede comprobar que, para ambos hoteles, la distribución de variables como *Adultos*, *Niños*, *EstanciasEnNochesDeFinDeSemana*, *EstanciasEnNochesDeSemana*, *Comida*, *País* y *TipoDeHabitaciónAsignada* es claramente diferente entre reservas no canceladas y canceladas.

Cuadro 5

Estadísticas resumidas del conjunto de datos H2 - Variables de fecha.

Variable	Min	Max	Mediana	Único
<i>ReservationStatusDate</i>	2014-10-17	2017-09-07	2016-08-10	864

Cuadro 6

Estadísticas resumidas del conjunto de datos H2 - Variables categóricas.

Variable	Único	Recuentos máximos
<i>Agente</i>	224	9: 31 955, NULO: 8 131, 1: 7 137, 14: 3 640
<i>FechaLlegadaMes</i>	12	Ago: 8 983, May: 8 232, Jul: 8 088, Jun: 7 894
<i>AssignedRoomType</i>	9	A: 57 007, D: 14 983, E: 2 168, F: 2 018
<i>Empresa</i>	208	NULO: 75 641, 40: 924, 67: 267, 45: 250
<i>País</i>	166	PRT: 30 960, FRA: 8 804, DEU: 6 084, GBR: 5315
<i>TipoCliente</i>	4	Tra.:59 404, Tra.-P.: 17 333, Con.: 2 300, Gro.:293
<i>Tipo de depósito</i>	3	No Dep.: 66 442, No Dev.: 12 868, Ref.: 20
<i>Canal de distribución</i>	5	TA/TO: 68 945, Dir.: 6 780, Cor: 3 408, GDS: 193
<i>IsCanceled</i>	2	0: 46 228, 1: 33 102
<i>IsRepeatedGuest</i>	2	0: 77 298, 1: 2 032
<i>Segmento de mercado</i>	8	Onl.: 38 748, Off.: 16 747, Gro.: 13 975, Dir.: 6 093
<i>Comida</i>	4	BB: 62 305, SC: 10 564, HB: 6 417, FB: 44
<i>ReservationStatus</i>	3	C.Out: 46 228, Can.: 32 186, No-Show: 916
<i>ReservedRoomType</i>	8	A: 62 595, D: 11768, F: 1 791, E: 1 553

Cuadro 7

Estadísticas de resumen del conjunto de datos H2 - Variables enteras y numéricas.

Variable	Media	SD	P0	P25	Mediana	P75	P100
<i>ADR</i>	105.3	43.6	0	79.2	99.9	126	5400
<i>Adultos</i>	1.85	0.51	0	2	2	2	4
<i>ArrivalDateOfMonth</i>	15.79	8.73	1	8	16	23	31
<i>FechaLlegadaNúmeroSemana</i>	27.18	13.4	1	17	27	38	53
<i>LlegadaFechaAño</i>	2016.17	0.7	2015	2016	2016	2017	2017
<i>Bebés</i>	0.0049	0.084	0	0	0	0	10
<i>BookingChanges</i>	0.19	0.61	0	0	0	0	21
<i>Niños</i>	0.091	0.37	0	0	0	0	3
<i>DaysInWaitingList</i>	3.23	20.87	0	0	0	0	391
<i>LeadTime</i>	109.74	110.95	0	23	74	163	629
<i>AnteriorBookingsNotCanceled</i>	0.13	1.69	0	0	0	0	72
<i>AnteriorCancelaciones</i>	0.08	0.42	0	0	0	0	32
<i>Plazas de aparcamiento obligatorias</i>	0.024	0.15	0	0	0	0	3
<i>EstanciasEnNochesDeFinDeSemana</i>	0.8	0.89	0	0	1	2	16
<i>EstanciasEnNochesDeSemana</i>	2.18	1.46	0	1	2	3	41
<i>TotalDePeticionesEspeciales</i>	0.55	0.78	0	0	0	1	5

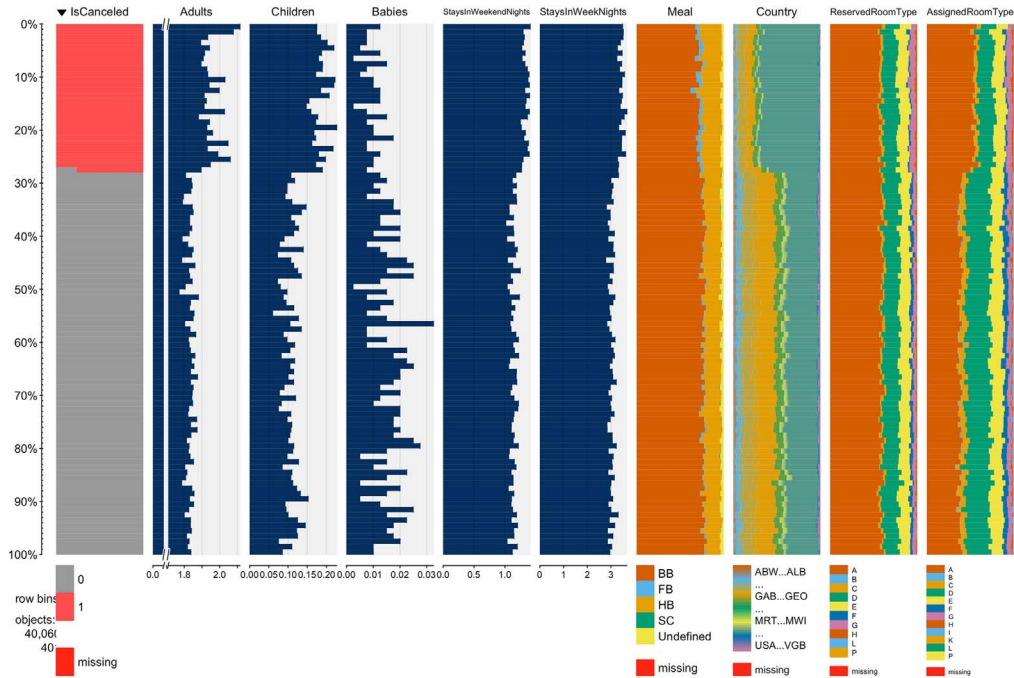


Fig. 2. Visualización parcial del conjunto de datos H1 de todas las observaciones.

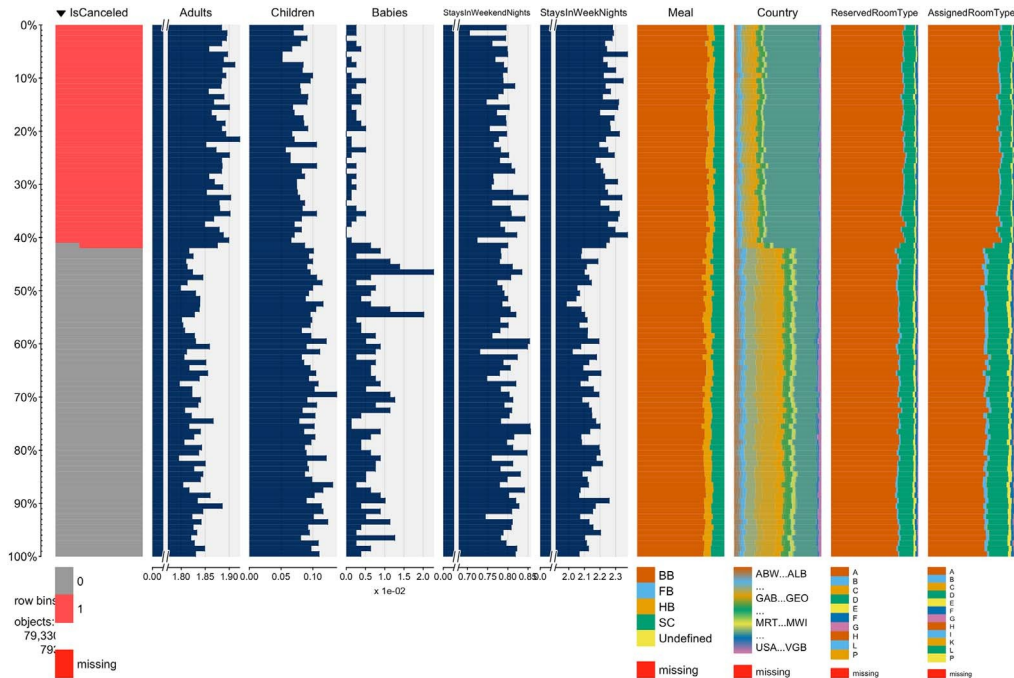


Fig. 3. Visualización parcial de todas las observaciones del conjunto de datos H2.



## Agradecimientos

Los autores desean dar las gracias a la administración de los hoteles por permitir que sus datos se compartan públicamente.

## Documento de transparencia. Material complementario

El documento de transparencia asociado a este artículo puede consultarse en la versión en línea en <https://doi.org/10.1016/j.dib.2018.11.126>.

## Apéndice A. Información complementaria

Los datos suplementarios asociados a este artículo pueden consultarse en la versión en línea en <https://doi.org/10.1016/j.dib.2018.11.126>.

## Referencias

- [1] N. Antonio, A. Almeida, L. Nunes, Predicting hotel bookings cancellation with a machine learning classification model, in: Proceedings of the 16th IEEE International Conference Machine Learning Application, IEEE, Cancún, Mexicopp. 1049–1054. doi:[10.1109/ICMLA.2017.00-11](https://doi.org/10.1109/ICMLA.2017.00-11), 2017.
- [2] Organización de Aviación Civil Internacional, Guidelines on Passenger Name Record (PNR) data, (2010). <[https://www.iata.org/iata/passenger-data-toolkit/assets/doc\\_library/04-pnr/New%20Doc%209944%201st%20Edition%20PNR.pdf](https://www.iata.org/iata/passenger-data-toolkit/assets/doc_library/04-pnr/New%20Doc%209944%201st%20Edition%20PNR.pdf)> (consultado 17 Febrero 2016).
- [3] D. Abbott, Applied Predictive Analytics: Principles and Techniques for the Professional Data Analyst, Wiley, Indianapolis, IN, USA, 2014.
- [4] Microsoft, SQL Server Management Studio (SSMS), (2017). <<https://docs.microsoft.com/en-us/sql/ssms/sql-server-management-studio-ssms>> (consultado el 24 de marzo de 2018).
- [5] American Hotel & Lodging Association, Uniform System of Accounts for the Lodging Industry, 11.ª edición revisada., Educational Institute, Nueva York, 2014.
- [6] Organización Internacional de Normalización, Códigos de país ISO 3166-3:2013, <<https://www.iso.org/obp/ui/#iso:std:iso:3166:-3:ed-2:v1:en,fr>> (consultado el 24 de marzo de 2018), 2013.
- [7] A. McNamara, E.A. de la Rubia, H. Zhu, S. Ellis, M. Quinn, skimr: Resúmenes compactos y flexibles de datos. Versión del paquete R 1.0.1 <<https://CRAN.R-project.org/package%20skimr>> , 2018.
- [8] M. Tennekes, E. de Jonge, tabplot: Tableplot, una visualización de grandes conjuntos de datos <<https://CRAN.R-project.org/package%20tabplot>> , 2017.