

DESKRIPSI DATASET

Dataset ini berisi data pemesanan sebuah hotel di daerah kota dan sebuah hotel di daerah resor; data yang dikoleksi seperti kapan pemesanan dilakukan, lama tinggal tamu di hotel, banyak tamu dewasa, banyak tamu anak-anak, banyak tamu bayi, jumlah ruang parkir tersedia, dan 25 fitur lainnya. Dataset ini berasal dari sebuah artikel yang diterbitkan oleh ScienceDirect (<https://www.sciencedirect.com/science/article/pii/S2352340918315191>), ditulis oleh Nuno Antonio, Ana Almeida, dan Luis Nunes untuk jurnal *Data in Brief*, Volume 22, Februari 2019. Dari artikel ScienceDirect tersebut, diketahui bahwa kedua hotel berlokasi di Portugal ("H1" berada di daerah resor di Algarve dan "H2" berada di kota Lisbon). Jarak antara dua lokasi ini kurang lebih 280 km dengan mobil dan kedua lokasi berbatasan dengan Atlantik Utara. Tujuan pengumpulan data tersebut adalah untuk memprediksi sebab-sebab kamar dibatalkan.

FITUR

Dataset ini terdiri atas 32 variabel (kolom) dan 119.390 observasi (baris), yang terbagi atas 40.060 observasi "H1" dan 79.330 observasi "H2". Setiap observasi mewakili satu pemesanan hotel. Dataset ini merekam pemesanan hotel yang tiba di hotel antara tanggal 1 Juli 2015 sampai dengan 31 Agustus 2017, termasuk pemesanan yang benar-benar tiba dan pemesanan yang dibatalkan. Berikut adalah deskripsi dan tipe data dari masing-masing variabel dalam dataset:

variable	data type	description
hotel	character	Hotel (H1 = Resort Hotel or H2 = City Hotel)
is_canceled	double	Value indicating if the booking was canceled (1) or not (0)
lead_time	double	Number of days that elapsed between the entering date of the booking into the PMS (Property Management System) and the arrival date
arrival_date_year	double	Year of arrival date
arrival_date_month	character	Month of arrival date
arrival_date_week_number	double	Week number of year for arrival date
arrival_date_day_of_month	double	Day of arrival date
stays_in_weekend_nights	double	Number of weekend nights (Saturday or Sunday) the guest stayed or booked to stay at the hotel
stays_in_week_nights	double	Number of week nights (Monday to Friday) the guest stayed or booked to stay at the hotel
adults	double	Number of adults
children	double	Number of children
babies	double	Number of babies

variable	data type	description
meal	character	Type of meal booked. Categories are presented in standard hospitality meal packages: Undefined/SC – no meal package; BB – Bed & Breakfast; HB – Half board (breakfast and one other meal – usually dinner); FB – Full board (breakfast, lunch and dinner)
country	character	Country of origin. Categories are represented in the ISO 3155–3:2013 format
market_segment	character	Market segment designation. In categories, the term “TA” means “Travel Agents” and “TO” means “Tour Operators”
distribution_channel	character	Booking distribution channel. The term “TA” means “Travel Agents” and “TO” means “Tour Operators”
d	double	Value indicating if the booking name was from a repeated guest (1) or not (0)
previous_cancellations	double	Number of previous bookings that were cancelled by the customer prior to the current booking
previous_bookings_not_canceled	double	Number of previous bookings not cancelled by the customer prior to the current booking
reserved_room_type	character	Code of room type reserved. Code is presented instead of designation for anonymity reasons
assigned_room_type	character	Code for the type of room assigned to the booking. Sometimes the assigned room type differs from the reserved room type due to hotel operation reasons (e.g. overbooking) or by customer request. Code is presented instead of designation for anonymity reasons
booking_changes	double	Number of changes/amendments made to the booking from the moment the

variable	data type	description
		booking was entered on the PMS until the moment of check-in or cancellation
deposit_type	character	Indication on if the customer made a deposit to guarantee the booking. This variable can assume three categories: No Deposit – no deposit was made; Non Refund – a deposit was made in the value of the total stay cost; Refundable – a deposit was made with a value under the total cost of stay.
agent	character	ID of the travel agency that made the booking
company	character	ID of the company/entity that made the booking or responsible for paying the booking. ID is presented instead of designation for anonymity reasons
days_in_waiting_list	double	Number of days the booking was in the waiting list before it was confirmed to the customer
customer_type	character	Type of booking, assuming one of four categories: Contract - when the booking has an allotment or other type of contract associated to it; Group – when the booking is associated to a group; Transient – when the booking is not part of a group or contract, and is not associated to other transient booking; Transient-party – when the booking is transient, but is associated to at least other transient booking
adr	double	Average Daily Rate as defined by dividing the sum of all lodging transactions by the total number of staying nights
required_car_parking_spaces	double	Number of car parking spaces required by the customer
total_of_special_requests	double	Number of special requests made by the customer (e.g. twin bed or high floor)
reservation_status	character	Reservation last status, assuming one of three categories:

variable	data type	description
		Canceled – booking was canceled by the customer; Check-Out – customer has checked in but already departed; No-Show – customer did not check-in and did inform the hotel of the reason why
reservation_status_date	double	Date at which the last status was set. This variable can be used in conjunction with the ReservationStatus to understand when was the booking canceled or when did the customer checked-out of the hotel

VARIABEL OUTPUT (Target yang diinginkan)

is_canceled	double	Value indicating if the booking was canceled (1) or not (0)
-------------	--------	---

SOAL

Dipersilahkan kepada peserta untuk memprediksi output dari file hotel_bookings.csv dan tuliskan hasilnya dalam format Excel, namai form Excel tersebut dengan format

NAMA_TIM_UNIV_KETUA_TIM.xls