Data Dictionary

Project Name: Hotel Booking

Description: This project involves using historical sales data to predict future sales trends and optimize inventory management.

Columns - 28

Rows - 119391

1]

Data Element: hotel

Description:Type of hotel (either "City Hotel" or "Resort Hotel")

Data Type:Categorical

Business Rule: Can only be one of two values: "City Hotel" or "Resort Hotel."

2]

Data Element: is_canceled

Description: Indicates whether the booking was canceled (1) or not (0).

Data Type: Integer

Business Rule: Can only take values of 0 (not canceled) or 1 (canceled).

3]Data Element: lead_time

Description: Number of days between booking date and arrival date.

Data Type: Integer

Business Rule: Must be a positive integer, typically within 0 to 365 days.

4]

Data Element: arrival_date_year

Description: The year of arrival.

Data Type: Integer

Business Rule: Must be a valid year, generally between 2015 and 2017 in the dataset.

5]

Data Element: arrival_date_month

Description: The month of arrival.

Data Type: Categorical

Business Rule: Limited to 12 months: "January," "February," etc

6]

Data Element: arrival_date_day_of_month

Description: The day of the month for arrival.

Data Type: Integer

Business Rule: Must be an integer between 1 and 31, based on the month.

7]

Data Element:stays_in_weekend_nights

Description: Number of weekend nights (Saturday and Sunday) the guest stayed.

Data Type: Integer

Business Rule: Must be a non-negative integer, usually 0 to 7.

8]

Data Element: stays_in_week_nights

Description: Number of weekday nights (Monday to Friday) the guest stayed.

Data Type: Integer

Business Rule: Must be a non-negative integer, usually 0 to 14.

9]

Data Element: adults

Description: Number of adults in the booking.

Data Type: Integer

Business Rule: Must be a non-negative integer, usually 1 to 30.

10]

Data Element: children

Description: Number of children in the booking.

Data Type: Integer

Business Rule: Must be a non-negative integer, usually 0 to 10.**Business Rule:** Must be a positive integer, unique within the dataset

---

11]

Data Element: babies

Description: Number of babies in the booking.

Data Type: Integer

Business Rule: Must be a non-negative integer, usually 0 to 10

. --

12]

Data Element: meal

Description: Type of meal plan selected (e.g., Bed & Breakfast, Full Board, etc.).

Data Type: Categorical

Business Rule: Must be one of the specified meal types.

13]

Data Element: country

Description: Country of origin of the guest.

Data Type: Categorical

Business Rule: Must be a valid country code (e.g., PRT, GBR, ESP, etc.).

14]

Data Element: market_segment

Description: Market segment designation (e.g., Online Travel Agent, Direct, etc.).

Data Type: Categorical

Business Rule: Must be one of the predefined market segments.

15]

Data Element: distribution_channel

Description: Channel used for booking (e.g., Direct, Travel Agency, etc.).

Data Type: Categorical

Business Rule: Must be one of the predefined distribution channels.

16]

Data Element: is_repeated_guest

Description: Indicates whether the guest has made a repeat booking (1) or not (0).

Data Type: Integer

Business Rule: Must be either 0 (not a repeated guest) or 1 (repeated guest).

Business Rule: Must be a positive integer, unique within the dataset

17]

Data Element: previous_cancellations

Description: Number of previous bookings that were canceled.

Data Type: Integer

Business Rule: Must be a non-negative integer, usually 0 to 10.

18]

Data Element: previous_bookings_not_canceled

Description: Number of previous bookings that were not canceled.

Data Type: Integer

Business Rule: Must be a non-negative integer, usually 0 to 10.

19]

Data Element: reserved_room_type

Description: Room type reserved for the booking.

Data Type: Categorical

Business Rule: Must be one of the specified room types.

20]

Data Element: assigned_room_type

Description: Actual room type assigned to the booking.

Data Type: Categorical

Business Rule: Must be one of the specified room types.

21]

Data Element: booking_changes

Description: Number of changes made to the booking.

Data Type: Integer

Business Rule: Must be a non-negative integer, usually 0 to 10.

22]

Data Element: deposit_type

Description: Type of deposit made for the booking (e.g., No Deposit, Non-Refund, etc.).

Data Type: Categorical

Business Rule: Must be one of the specified deposit types.

23]

Data Element: days_in_waiting_list

Description: Number of days the booking was on the waiting list.

Data Type: Integer

Business Rule: Must be a non-negative integer, usually 0 to 20.

24]

Data Element: customer_type

Description: Type of customer (e.g., Contract, Transient, etc.).

Data Type: Categorical

Business Rule: Must be one of the specified customer types.

25]

Data Element: adr (Average Daily Rate)

Description: Average daily rate for the booking.

Data Type: Float

Business Rule: Must be a positive number, typically within a reasonable range (e.g., 0 to 500).

26]

Data Element: required_car_parking_spaces

Description: Number of car parking spaces required.

Data Type: Integer

Business Rule: Must be a non-negative integer, usually 0 to 5.

27]

Data Element: total_of_special_requests

Description: Total number of special requests made by the guest.

Data Type: Integer

Business Rule: Must be a non-negative integer, usually 0 to 10.

28]

Data Element: reservation_status_date

Description: The date when the reservation status was last updated.

Data Type: Date (or String, depending on how it is stored in the dataset)

Business Rule: Must be a valid date format (MM-DD-YYYY) and should not be a future date.