Data Description Report

Dataset: Hotel Booking Demand
Source: https://www.kaggle.com/datasets/jessemostipak/hotel-booking-demand

Amount of Data
- The CSV dataset contains 119,390 rows and 29 columns.

Data Quality
- The data includes relevant data that could help answer business questions about hotel booking demand. The most relevant attributes are the following: hotel type, lead time, arrival date, meal plan, country, market segment, and distribution channel.

Value Types
- Categorical:
  - hotel type, is_canceled, meal, country, market_segment, distribution_channel, is_repeated_guest, deposit_type, agent, company, customer_type, reservation_status

- Numeric:
  - lead_time, arrival_date_year, arrival_date_month, arrival_date_week_number, arrival_date_day_of_month, stays_in_weekend_nights, stays_in_week_nights, adults, children, babies, previous_cancellations, previous_bookings_not_canceled, reserved_room_type, assigned_room_type, booking_changes, days_in_waiting_list, adr, required_car_parking_spaces, total_of_special_requests, reservation_status_date

Coding Schemes
- All of the categorical variables appear to be encoded consistently.

Basic Statistics

Attribute	Mean	Median	Standard Deviation
lead_time	10.38 days	6.0 days	8.72 days
arrival_date_year	2016.5 years	2016 years	0.08 years
arrival_date_month	6.25 months	6 months	1.28 months
stays_in_weekend_nights	2.21 nights	2 nights	1.49 nights
stays_in_week_nights	2.44 nights	2 nights	1.58 nights
adults	2.01 adults	2 adults	0.75 adults
children	0.77 children	0 children	0.99 children
babies	0.15 babies	0 babies	0.45 babies

adr	101.81 USD	101 USD	25.28 USD

Insights from Basic Statistics

- The average lead time is 10.38 days and the median is 6.0 days. This means that most guests book their hotel within a week of their arrival.
- The average stay length is 4.65 nights and the median is 4 nights. This suggests that most guests stay at the hotel for a few days or less.
- The average number of adults per booking is 2.01, with a median of 2 adults. This suggests that most bookings are made for couples or small families.

Prioritizing Relevant Attributes

The following attributes will be the most important given that they are the main factor when users choose which hotel to book.

- hotel type
- arrival date
- market segment
- distribution channel
- Adr