

Assignment: Hotel Booking Analysis

Weight: 10%

Using Power Query and/or Advanced Excel

Steps to follow:

1. Data Import and Initial Setup

- **Import Data:** Start by importing the dataset into Power Query. Ensure all columns are loaded correctly, with the appropriate data types.
- **Remove Unnecessary Columns:** Remove sensitive or unnecessary columns like name, email, phone_number, and credit_card since they don't add value to the analysis and may contain artificial or PII (Personal Identifiable Information) i.e. sensitive information.

2. Data Cleaning

- **Handle Null or Missing Values:**
 - Identify columns with missing values, such as children or agent, and decide how to handle them (e.g., replace with 0 if missing in children or use "Unknown" for categorical values like agent).
- **Correct Data Types:** Verify and change data types where necessary:
 - Convert dates (like reservation_status_date) to date format.
 - Ensure numeric columns like adr, total_of_special_requests, and lead_time are set to number format.
- **Remove Duplicates:** Check for and remove any duplicate rows to ensure data integrity.

3. Data Transformation and Feature Engineering

- **Create New Columns:**
 - **Stay Duration:** Add a new column for total_stay by summing stays_in_weekend_nights and stays_in_week_nights.
 - **Is Family Booking:** Create a column to identify family bookings by checking if children + babies > 0.
 - **Booking Lead Time Range:** Classify lead_time into categories (e.g., short, medium, long).
 - **Cancellation Rate by Country:** Aggregate the dataset to calculate the cancellation rate for each country, giving insights into regions with high cancellation tendencies.

4. Data Aggregation and Summarization

- **Analyze Key Metrics:** Create aggregations for key business metrics:
 - **Average Daily Rate (ADR) per Market Segment:** Use Group By to find the average adr by market_segment.
 - **Cancellation Rates:** Calculate the cancellation rate for each market_segment, distribution_channel, and customer_type.
 - **Special Requests:** Count the average total_of_special_requests by hotel type (City Hotel vs. Resort Hotel).