

Project Documentation: Hotel Booking Analysis

By: Sondos Mohamed

1. Project Objective

The objective of this project is to analyze booking data for two hotels (a resort and a city hotel) to build an interactive dashboard in Power BI. The dashboard aims to extract actionable insights regarding booking patterns, identify the primary drivers behind cancellations, and analyze customer behavior to support management in making better strategic decisions to increase revenue and improve operational efficiency.

2. Data Sourcing & Overview

The analysis is based on a single, comprehensive dataset sourced from the file **hotels.csv**. This dataset contains a rich collection of real-world booking information for two types of hotels (a Resort Hotel and a City Hotel) over a period of more than two years.

The dataset is substantial in scale, comprising **approximately 119,000 individual booking records (rows)** and **32 original attributes (columns)**. These attributes provide extensive details for each booking, covering guest information, arrival and departure dates, room specifications, booking channel details, and the final reservation status. The time frame of the data spans from **July 2015 to August 2017**, providing a solid foundation for seasonal and year-over-year performance analysis.

3. Data Cleaning & Transformation (ETL)

All cleaning and transformation steps were executed in the **Power Query Editor** to ensure a clean and reliable data model.

3.1. Handling Nulls & Errors:

- **children column:** 1% null values were detected. These were replaced with **zero (0)**, assuming that a blank field implies no children.
- **adr column:** Null values were replaced with **zero (0)** to ensure the accuracy of financial calculations.
- **country column:** Null values were replaced with the text "Unknown" for clear categorization.

- **agent & company columns:** Nulls were initially handled to signify direct or personal bookings and were definitively managed during the creation of their respective dimension tables.

3.2. Feature Engineering & Keys Creation:

- **BookingID Creation (Fact Table Primary Key):** An **Index Column** starting from 1 was added to create a unique identifier for each booking in the fact table.
- **Date Keys Creation (Numeric):**
 1. **arrival_date_id:** The separate date columns (year, month, day) were merged and then transformed into a numeric key with the format YYYYMMDD.
 2. **reservation_status_date_id:** The date column was directly converted into a numeric key with the same YYYYMMDD format.

3.3. Dimension Tables Creation:

To build a Star Schema model, descriptive columns were extracted to create separate dimension tables. This process was repeated for each dimension:

1. **Duplicate** the main query.
2. **Keep the required column(s)** and remove others.
3. **Remove Duplicates** to get a unique list of values.
4. **Add an Index Column** as a primary key (e.g., CountryID).
5. **Return to the main table** and **Merge** this new key.
6. **Remove the original text column** from the main table

List of created Dimension Tables:

- dim_customer
- dim_date
- dim_hotel
- dim_agent
- dim_company
- dim_country
- dim_meal

- dim_market_segment
 - dim_distribution_channel
 - dim_room_type
 - dim_deposit_type
 - dim_reservation_status
 - dim_customer_type
-

4. Data Modeling

- **Structure:** A **Star Schema** model was built, consisting of one central fact table (Fact_Hotels) surrounded by multiple dimension tables.
 - **Relationships:**
 - Each dimension table was connected to the fact table using a **One-to-Many** relationship.
 - Two relationships were established with dim_date:
 1. An **Active relationship** on Fact_Hotels[arrival_date_id].
 2. An **Inactive relationship** on Fact_Hotels[reservation_status_date_id].
-

5. Analysis using DAX

A comprehensive set of measures was created using DAX to analyze the data:

- **Core KPIs:** [Total Bookings], [Cancellation Rate], [No-Show Rate], [Total Guests].
 - **Financial Measures:** [Total Revenue] (excluding canceled bookings), [ADR], [Forfeited Deposit Revenue].
 - **Time Intelligence Measures:** [Total Revenue LY] and [Revenue YoY %] for year-over-year comparisons.
 - **Advanced Measures:** [Cancellations by Status Date] (using USERELATIONSHIP).
-

6. Visualization & Insights

An interactive 5-page dashboard was designed to tell a comprehensive story of the hotel's performance. Each page answers a specific set of business questions:

1. **Page 1: Overview (Executive Summary)** This high-level summary provides key performance indicators at a glance. It features KPI cards for Total Revenue (\$11.6M), Total Bookings (119K), and Total Guests (217K). It also includes visuals showing the booking volume by hotel type, monthly booking trends, and identifies the top 5 source countries for guests.
2. **Page 2: Customer Analysis** This page provides a deep dive into customer demographics and acquisition channels. It reveals that the Online TA (Travel Agent) is the largest market segment, accounting for 47.4% of bookings, and that Transient customers form the biggest customer type. This is crucial for optimizing marketing spend and channel management.
3. **Page 3: Cancellation Analysis** Focused on understanding the root causes of lost revenue, this page analyzes cancellation patterns. A key insight is that bookings with **"No Deposit"** have a **significantly higher cancellation rate** than other deposit types. It also highlights that the City Hotel experiences a higher volume of cancellations, pointing to an area for further investigation.
4. **Page 4: Booking Patterns** This page analyzes the "when" and "how" of customer booking behavior. It visualizes the seasonality of the Average Daily Rate (ADR), showing a clear peak during the summer months. It also analyzes the average Lead Time, providing insights into the booking window of customers, which is essential for revenue management and forecasting.
5. **Page 5: Room Analysis** This final page evaluates the performance of different room types. The analysis shows that **Room Type 'A' is the most frequently booked**. The page also breaks down the average price (ADR) for each room type, allowing management to identify the most profitable room categories and adjust their inventory and pricing strategies accordingly.
6. **.Advanced Interactive Features**

To elevate the dashboard from a static report to a dynamic and exploratory analytical tool, a suite of advanced Power BI features was incorporated to enhance user experience and enable deep data exploration.

- **Bookmarks & Navigation:**
 - **Description** :A feature that saves a "snapshot" of the current state of a report page (filters, visibility, etc.), enabling the creation of app-like navigation experiences.
 - **Implementation** :A main landing page was designed with custom navigation buttons. Each button is linked to a bookmark that takes the user directly to the relevant analytical page (e.g., "Cancellation Analysis," "Customer Analysis"), streamlining the report exploration.

- **Report Page Tooltips:**

- **Description :** Instead of the default tooltip, custom mini-report pages were designed to appear on hover, providing deeper contextual information.
- **Implementation :** A custom tooltip page) Tooltip_Country (was created. When a user hovers over a country on the map visual, a small line chart appears, showing the monthly booking trend for that specific country, offering an immediate insight into individual market performance.

- **Drillthrough:**

- **Description :** A feature that allows users to navigate from a high-level summary view to a dedicated, automatically filtered detail page based on their selection.
- **Implementation :** A hidden detail page) Agent_Details (was created. Users can now right-click any travel agent in a visual and select "Drillthrough" to be taken to a page that lists every single booking and its financial details for that specific agent.

The dashboard utilized advanced features like **custom icons**, **smart KPI cards** with year-over-year comparisons, and **conditional formatting** to highlight key information.