Bookin Information for a Multi-service Business

Data Cleaning and Exploratory Data Analysis (EDA) Report

Project Overview

The dataset provided contains booking information for a multi-service business. It includes data for class bookings, subscriptions, facility rentals, and birthday party reservations. The dataset is a real-world export and may contain inconsistencies, missing values, and other imperfections that require cleaning and preprocessing.

Dataset Columns

The dataset initially contained the following columns:

- **Booking ID**
- **Customer ID**
- **Customer Name**
- **Booking Type**
- **Booking Date**
- **Status**
- **Class Type**
- **Instructor**
- **Time Slot**
- **Duration (mins)**
- **Price**
- **Facility**
- **Theme**
- **Subscription Type**
- **Service Name**
- **Service Type**
- **Customer Email**
- **Customer Phone**

Data Cleaning Process

1. Handling Missing Values

- Identified missing values in various columns.

- The column **'Subscription Type'** had no values, so it was dropped.

- The **'Time Slot'** column had missing values, which were filled using:

    1. **Mode Imputation**: Filled missing values with the most frequent time slot.

    2. **Forward Fill (ffill)**: Filled missing values with the previous row's value.

    3. **Backward Fill (bfill)**: Filled missing values with the next row's value.

- The **'Facility'** and **'Service Name'** columns contained the same values, so missing values in one column were filled using the other.

- After merging, the **'Service Name'** column was dropped.

- The **'Class Type'** column was also dropped after confirming it duplicated values present in other columns.

2. Handling Duplicate Rows

- Checked for duplicate rows in the dataset.

- Printed the number of duplicate rows found (if any).

3. Handling Categorical Data

- Replaced missing values in categorical columns with placeholders:

    o **'Instructor'** → Replaced missing values with 'No Instructor'.

    o **'Theme'** → Replaced missing values with 'Not Specified'.

4. Handling Numerical Data

- The **'Duration (mins)'** column had missing values, which were replaced with the median value of the column.

5. Handling Contact Details

- Missing values in contact details were replaced with placeholders:

    o **'Customer Email'** → 'No Email Provided'.

    o **'Customer Phone'** → 'No Phone Provided'.

6. Saving Cleaned Data

- The cleaned dataset was saved to a new Excel file:
  File : Cleaned_Dataset_for_dataAnlysis.xlsx

Exploratory Data Analysis (EDA)

After cleaning, we performed EDA to identify trends and insights:

1. **Distribution of Booking Types**

   o Identified the most popular booking types (e.g., Class, Facility,Birthday).

2. **Peak Booking Times**

   o Analyzed booking trends based on time slots.

3. **Revenue Analysis**

   o Calculated revenue trends by price, duration, and service type.

4. **Instructor-Based Analysis**

   o Determined which instructors had the most bookings.

5. **Customer Engagement**

   o Identified frequent customers based on booking history.

This cleaned dataset is now ready for **Power BI visualization**.

Power BI Dashboard Components

The dashboard contains multiple visualizations that provide insights into booking trends, revenue generation, and customer preferences.

1. Bar Chart: Total Bookings by Facility Category

- Displays the number of bookings for different facility categories.

- Helps in identifying which facility is the most popular among customers.

2. Line Chart: Monthly Revenue Generation by Booking Status

- Shows revenue trends over months based on booking status (e.g., confirmed, pending).

- Helps businesses track financial performance and seasonal trends.

3. Bar Chart: Total Revenue by Facility

- Represents total revenue generated by each facility.

- Helps in determining which facility contributes the most to revenue.

4. Line Chart: Booking Trends by Time Slot

- Shows peak hours for bookings throughout the day.

- Helps in optimizing facility usage and resource allocation.

5. Vertical Bar Chart: Total Revenue by Booking Type

- Displays revenue generation for different booking types (e.g., class bookings, Facility, birthday parties).

- Helps in understanding the most profitable service type.

6. **Pie Chart: Distribution of Booking Types**

   - Represents the proportion of different service types in total bookings.

   - Helps in visualizing customer preferences and demand for services.

Business Insights and Decision-Making

- **Identifying High-Demand Facilities** → Focus marketing efforts on the most popular facilities.

- **Revenue Optimization** → Track monthly revenue trends and adjust pricing strategies.

- **Customer Behavior Analysis** → Optimize staffing and resource allocation based on peak booking times.

- **Booking Type Analysis** → Invest in service types that generate higher revenue.