**Booking Data Cleaning and Analysis Report by Sudhanshu Singh**

**Step 1: Cleaning and Preparing the Data**

**1.1 Converting Excel to CSV**

i) The dataset is in ".xlsx" format and is converted to ".csv" using Excel.

ii) Spaces in column names are replaced with underscores to reduce error chances in SQL.

**1.2 Setting Up the MySQL Database**

i) A database named "assessment" is created to store the booking data.

ii) A table called "booking\_data" is created with the required columns and correct data types.

**1.3 Importing Data into MySQL**

i) The CSV file is copied to MySQL’s "Uploads" folder because MySQL only allows access to this location.

ii) The data is loaded into MySQL using "LOAD DATA INFILE", and missing values in important columns are handled properly.

**Step 2: Transforming the Data**

**2.1 Fixing Date Format Issues**

i) A new column, "Date\_Of\_Booking", is added to store the date in the proper format.

ii) The "Booking\_Date" column is converted from text to date format and stored in "Date\_Of\_Booking".

iii) After verifying the conversion, the "Booking\_Date" column is dropped to keep the data clean.

**2.2 Handling Missing and Duplicate Data**

i) The "Subscription\_Type" column is deleted since it has no values.

ii) Duplicate "Booking\_ID" entries are checked.

iii) Missing values in "Class\_Type", "Instructor", "Facility", and "Theme" are replaced with default values like "Not Specified" or "Unknown" to maintain consistency.

**Step 3: Analysing the Data**

**3.1 Finding Key Metrics**

i) The average duration of each service type is calculated.

ii) Yearly revenue is analysed to track business growth.

iii) The number of bookings for each "Booking\_Type" is counted.

iv) The top five most booked services and their total revenue are identified.

**3.2 Booking Trends**

i) Monthly trends are analysed to find peak booking periods.

ii) The top 10 highest-paying customers are identified based on total spending.

**3.3 Checking Service Performance and Popularity**

i) The number of pending and confirmed bookings is analysed to track service reliability.

ii) The most booked time slots are identified.

iii) The most used facilities are found, along with their share of total bookings.

**3.4 Exporting Data for Further Use**

i) The cleaned and processed data is retrieved and exported for further use in reporting and business analysis.

**Step 4: Creating a Power BI Dashboard**

**4.1 Loading Data into Power BI**

i) The cleaned CSV file is loaded into Power BI using "Get Data" > "Text/CSV", and an interactive dashboard is created.

**Conclusion**

i) The booking data is successfully cleaned and imported into MySQL.

ii) Issues like missing values and incorrect formats are fixed.

iii) Key insights like revenue trends, customer spending, and booking trends are identified.

iv) The data is visualized in "Power BI", making it easier to explore booking trends and overall business performance.