Advanced Hotel Booking Demand Analysis for Strategic Decision Making

Overview:

In this mini project, you will analyze a real-world dataset containing hotel booking demand data. The goal is to apply data-driven decision-making techniques, focusing on descriptive analytics, linear regression, time series analysis, and business interpretation. You will explore different customer segments and trends related to booking cancellations, leading to actionable insights for hotel management. You will submit a short report (not exceeding 3000 words) and present your findings in a 15-minute presentation.

Dataset:

1. The dataset can be found as a CSV here hotel_bookings.csv
2. The dataset description is provided in this article Hotel booking demand datasets - ScienceDirect
3. Example business analytics reports:

   CBRE: example_CBRE_report.pdf

   Savill: example_Savill_report.pdf

The dataset contains 31 variables, including booking details, customer demographics, and booking outcomes (e.g., cancellations). Some key variables include:

- **Average Daily Rate (ADR)**
- **Lead Time**
- **Market Segment**
- **IsCancelled** (whether a booking was cancelled)

Project Requirements:

1. **Descriptive Analytics for Segments (20 points)**
   - **Objective**: Summarize key patterns and trends in the data for different market segments.
   - **Task**: Calculate and present summary statistics (mean, median, mode, standard deviation) for variables like ADR, lead time, and cancellations across different market segments such as:
     - Market Segment (e.g., Transient, Group, Corporate)
     - Distribution Channel (e.g., Travel Agents, Direct)
     - Other segments of your choice

- **Deliverables**:
    - Visualizations (bar charts, histograms) showing key patterns within segments.
    - Discuss how different segments behave differently (e.g., cancellation rates by segment, ADR trends).
    - Identify any outliers or unique behaviors in certain segments.

2. **Segment-Based Linear Regression Models (30 points)**
    - **Objective**: Build and compare linear regression models for different market segments.
    - **Task**:
        - Perform separate linear regressions where **ADR** is the dependent variable, and **Lead Time** is the independent variable for different market segments (e.g., Transient, Group).
        - Interpret the coefficients and the model fit for each segment. How does the relationship between ADR and lead time differ between segments?
    - **Deliverables**:
        - Scatter plots for each segment with their respective regression lines.
        - Discussion of how pricing strategies can vary across segments based on lead time.
        - Compare R-squared values for different models and evaluate which segment is most predictable.
3. **Trendline and Time Series Analysis (25 points)**
    - **Objective**: Identify and interpret trends in booking volumes and cancellations over time.
    - **Task**:
        - Create a time series plot showing booking volumes and/or cancellation rates over time (e.g., monthly or yearly trends).
        - Add a trendline to the time series and interpret whether the trends are increasing, decreasing, or stable. Focus on potential seasonality in bookings or cancellations.
        - Any other trends that you observe? Are there differences in the trendlines across segments?
    - **Deliverables**:
        - Time series plot with a trendline.
        - Discussion of seasonal patterns (e.g., high cancellation rates in specific months) and long-term trends.
        - Explain how the trends might impact revenue management decisions or promotional strategies.
4. **Advanced Business Insights (25 points)**
    - **Objective**: Provide strategic recommendations based on your analyses.
    - **Task**:

    - Based on your findings from the descriptive analytics, linear regressions, and time series analysis, provide actionable business recommendations.
    - Discuss strategies to optimize pricing based on lead time and market segments, and suggest ways to reduce cancellations (e.g., targeting specific segments with higher cancellation rates).
    - Provide insights into potential promotional strategies or changes in booking policies to mitigate risks such as high cancellation rates during certain periods.
- **Deliverables**:
    - A summary of your business insights and recommendations tailored to different market segments and booking trends.
    - Justify your recommendations with data, citing specific trends and relationships observed in your analysis.

Report Structure:

Your report should be structured as follows:

- **Introduction**: Briefly introduce the purpose of the project and the dataset.
- **Descriptive Analytics**: Summarize key statistics and patterns for different segments.
- **Segment-Based Linear Regression Models**: Present the regression models for each segment, findings, and interpretation.
- **Trendline and Time Series Analysis**: Describe the time series trends and their implications.
- **Advanced Business Insights**: Provide actionable recommendations for hotel management based on your analyses.

Presentation Guidelines:

- **Time**: 15 minutes (10 minutes for presentation + 5 minutes for Q&A).
- **Content**:
    - Provide a concise summary of your analysis, focusing on the **most critical and interesting** findings. Remember that there are 4 groups working on this data set, so focus on what you think are unique insights that you discover yourself
    - Use visual aids (charts, graphs, tables) to explain your insights clearly.
    - Conclude with a discussion of your business recommendations, segment-based insights, and strategies to manage booking and cancellation patterns.
- **Format**: You may use slides (e.g., PowerPoint) or a live demo of your analysis in a tool like Jupyter Notebook.

Submission:

- **Report**: use your Jupyter notebook as your report. Make sure you still do all the proper formatting to make the notebook presentable.
- **Presentation**: Be prepared to present your findings in class.

Evaluation Criteria:

- **Depth of Analysis** (40%)
- **Interpretation of Results and Business Recommendations** (30%)
- **Presentation Clarity and Professionalism** (20%)
- **Creativity and Advanced Insights** (10%)