

# CS7DS4 / CSU44065 Data Visualization

## Assignment 3: Addressing Complexity

Tanmay Samdani  
Student No: 24348137

### Declaration

"I have read and I understand the plagiarism provisions in the General Regulations of the University Calendar for the current year, found at <http://www.tcd.ie/calendar>. I have also completed the Online Tutorial on avoiding plagiarism 'Ready Steady Write', located at <http://tcd-ie.libguides.com/plagiarism/ready-steady-write>."

## 1 Tools and Technologies

The following tools and libraries were used to create the visualization dashboard:

- **Python**: For data processing and dashboard development.
- **Pandas**: For data preprocessing, cleaning, and feature engineering.
- **Plotly**: For creating interactive visualizations (box plots, scatter plots, heatmaps, etc.).
- **Dash**: For building the interactive web application.
- **Dash Bootstrap Components**: For styling and layout enhancements.
- **Chardet**: For encoding detection when reading CSV files.

## 2 Dataset Description

The dataset used consists of Airbnb booking data with the following attributes:

- **\*\*Original Columns\*\***: Start Date, End Date, Number of Adults, Children, Infants, Earnings, Location, Unit Type, Status.
- **\*\*Derived Attributes\*\***:
  - **Stay Duration**: Calculated as the difference between the end and start dates.
  - **Total Guests**: Sum of adults, children, and infants.
  - **Revenue per Guest**: Earnings divided by the total number of guests.

- **Booking Month, Weekday, Seasonal Period:** Extracted for time-based analysis.

The dataset is complex due to its variety of numerical, categorical, and temporal data, which necessitates multiple visualization idioms to uncover insights effectively.

### 3 Visualization Tasks

The following tasks are supported by the dashboard:

1. **\*\*Identify High-Performing Locations\*\*:** **Chart:** Box Plot of Earnings by Location.
2. **\*\*Analyze Guest Composition\*\*:** **Chart:** Stacked Bar Chart of Guest Composition by Location.
3. **\*\*Track Monthly Earnings Trends\*\*:** **Chart:** Line Chart of Monthly Earnings.
4. **\*\*Identify Seasonal Patterns\*\*:** **Chart:** Bar and Line Combination Chart for Seasonal Performance.
5. **\*\*Explore Correlations Between Metrics\*\*:** **Chart:** Correlation Heatmap of Key Metrics.

### 4 Encoding Channels and Idioms

- **\*\*Box Plot\*\*:** Position and length of boxes encode earnings distribution and outliers.
- **\*\*Stacked Bar Chart\*\*:** Color segments represent guest categories (adults, children, infants).
- **\*\*Line Chart\*\*:** Lines and markers encode earnings trends over time.
- **\*\*Heatmap\*\*:** Color intensity shows earnings patterns by month and weekday.
- **\*\*Scatter Plot\*\*:** Position, size, and color encode stay duration, earnings, and total guests.
- **\*\*Pie Chart\*\*:** Segment size shows earnings distribution by unit type.

### 5 Novelty and Complexity

**Novelty:**

- The dashboard combines multiple visualization idioms (box plot, line chart, heatmap, scatter plot) into a single cohesive interface.
- Interactive filtering allows users to explore different subsets of the data dynamically.

**Complexity:**

- **\*\*Data Variety\*\*:** The dataset includes numerical, categorical, and temporal data.

- **\*\*Interactive Elements\*\***: Dropdowns and sliders enable real-time data manipulation.
- **\*\*Multiple Facets\*\***: Visualizations are organized into rows and columns to compare different insights side-by-side.

## 6 Critical Analysis

### Strengths:

- Interactive and intuitive dashboard design.
- Supports multiple analytical tasks with varied visualization idioms.
- Professional styling enhances readability and user experience.

### Weaknesses:

- Performance may degrade with larger datasets.
- Complexity may require user guidance for full utilization.

# Complete Visualization Dashboard

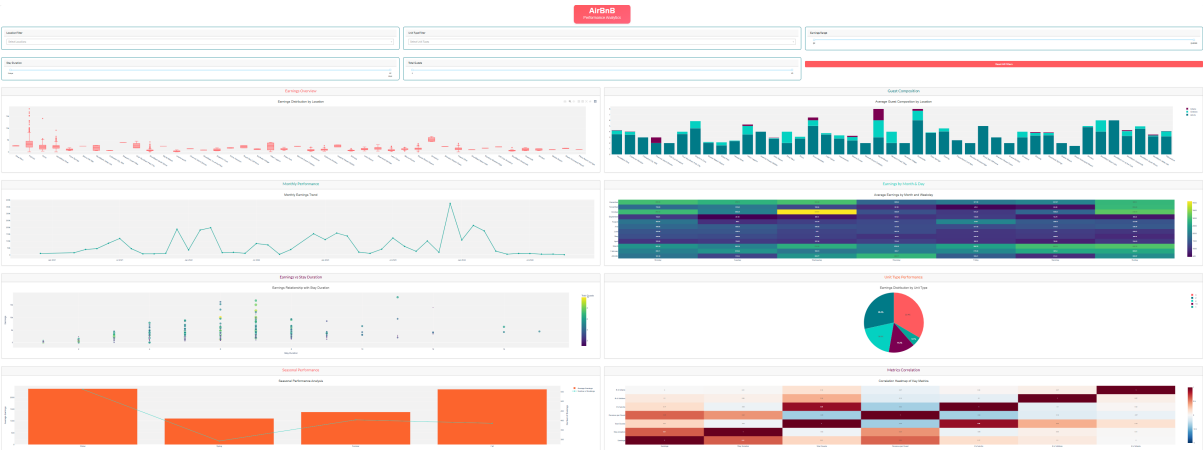


Figure 1: Complete Airbnb Performance Analytics Dashboard

