Flight Delay Analysis Using the 'nycflights13' Dataset

Procedure

The analysis was conducted using R with packages 'nycflights13', 'dplyr', 'ggplot2', and 'lubridate'. Key aspects investigated include:

- 1. **Average Delay by Destination**: Determining the average delay per destination and visualizing the spatial distribution.
- 2. **Plane Age and Delays**: Assessing if there's a correlation between the age of a plane and the average delay.
- 3. **Frequency of Flights and Delays**: Filtering flights to those with planes that have flown at least 100 flights.
- 4. **Airtime and Duration Comparison**: Comparing the airtime of flights with the actual duration between departure and arrival.
- 5. **Day of the Week and Delays**: Identifying which day of the week typically experiences the least delay.

Results and Discussion

1. Average Delay by Destination:

- The analysis reveals the average delay experienced by flights to various destinations.
- A spatial plot was generated to visualize these delays across different geographic locations.

2. Plane Age and Delays:

- The relationship between the age of planes and the average delay was visualized.
- A scatter plot with a linear regression line showed the trend between plane age and delay times.

3. Frequency of Flights and Delays:

- Flights were filtered to include only those with planes that have completed at least 100 flights.
- This subset could provide insights into the reliability and delay patterns of frequently used aircraft.

4. Airtime and Duration Comparison:

- A comparison between the scheduled airtime and the actual duration of flights was performed.
- The plot indicates discrepancies and aligns them with potential causes such as airport locations and time zone changes.

5. Day of the Week and Delays:

- Analysis of delays across different days of the week was conducted to determine the optimal day for travel to minimize delays.
 - The resulting visualization highlights the days with the lowest average delays.