Ds4200 Project: Shefali Verma, Tessa Joseph, Sohan Ramchandani

Link to Website: <https://tessjoseph.github.io/ds4200finalproject/index.html>

**Paragraphs to explain design Ideas for each plot:**

1. Geospatial: Station Location VS Flow
   1. This visualization was created to showcase both the geographic locations of MBTA stations and their total daily flow. Using Folium, we plotted each station as a colored dot on a street map of Boston—color-coded by line (e.g., Red Line = red, Green Line = green, etc.). The size of each dot corresponds to the station’s average daily flow, with larger dots indicating higher levels of commuter traffic. To account for the wide range in flow between the busiest and least busy stations, we applied a logarithmic scale to the dot sizes, ensuring clear visual differentiation. For added precision, we implemented an interactive tooltip feature, allowing users to hover over a station to view its exact flow value. We also overlaid Boston’s neighborhood (ward) boundaries in a simple, thin black outline to help users understand the spatial context of each station without detracting from the main data. Notably, transfer stations—those serving multiple lines—are represented in black to emphasize their exceptionally high levels of activity.
2. Line plot: Flow by Route and Time of Day
   1. This visualization was designed to compare the total average ridership across the four main MBTA lines—Red, Orange, Blue, and Green—over the course of a day. To achieve this, we created a multi-line plot where each line is color-coded to match the corresponding train line (e.g., the Red Line is shown in red, the Blue Line in blue, etc.). The x-axis represents time intervals during the MBTA's operating hours, while the y-axis shows the total average number of people entering and exiting stations on each line. A legend in the top right corner clearly indicates which color corresponds to which line, and gridlines are included to enhance readability and support easy comparison
3. Heat Map: Flow by Station and Time of Day
   1. This visualization was created to compare the average flow at each individual station across the four MBTA subway lines. To do this, we developed an interactive heatmap, with separate panels for each line—and for each branch of the Green Line. Radio buttons at the top of the visualization allow users to toggle between the different heatmaps. As in the previous plot, the x-axis represents time intervals during the MBTA’s service hours, while the y-axis displays individual stations on the selected line. Color intensity in the heatmap reflects the volume of average flow, with darker, deeper hues indicating higher levels of activity at a given station and time.
4. Altair Bubble Chart: Boardings per Stop by Route
   1. This visualization was created to display the total boardings by station and route for the MBTA system across different times of day. Using Altair, we plotted each station as a bubble—color-coded by transit line (e.g., Green Line = green, Red Line = red, etc.). The size of each bubble corresponds to the total number of boardings onto that station, with larger bubbles representing higher ridership levels. To account for the different times of day, the chart categorizes the data into distinct time periods, such as morning peak, midday, and evening. Interactive tooltips allow users to hover over each bubble to view detailed information, including the station name, route, and exact boarding numbers. This design makes it easy to compare total boardings across both individual stations and entire lines, while also identifying peak ridership periods. By combining interactivity, proportional sizing, and categorical color encoding, the visualization offers a clear and accessible method for analyzing MBTA transit usage.
5. D3 Horizontal Bar Chart: Busiest Stations by Day Type and Year
6. The bar chart was created to highlight the top 20 busiest MBTA stations based on total daily activity, measured by the combined number of boardings and alightings. Using an interactive horizontal bar chart, we represented each station with a colored bar, where the length corresponds to the average number of passengers. Stations are color coded according to their subway line, allowing for immediate recognition of route specific trends. To account for variation in traffic patterns , dropdown menus were added at the top, allowing us to switch between different seasons, and day types. This symanic setup provides a comparative view of how station usage changes over time and across different types of days, offering insights into which stations consistently see high demand and how weekend ridership differs from weekday commuting patterns.