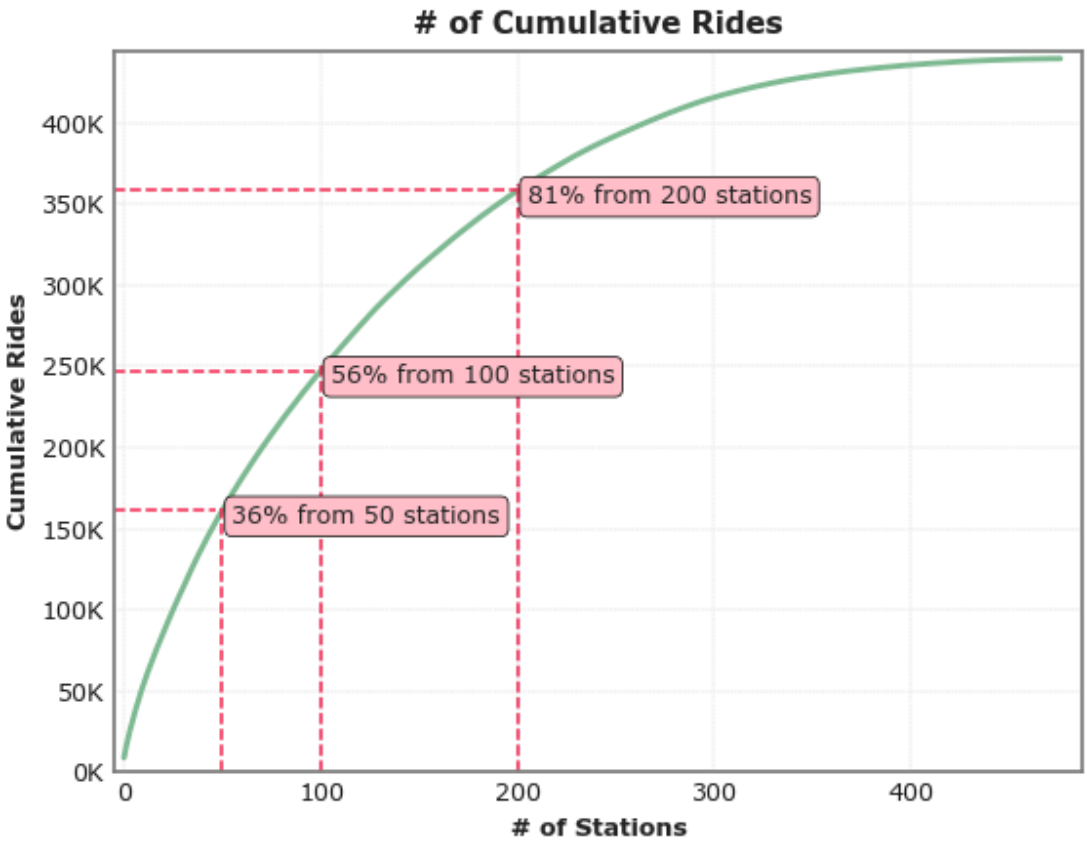
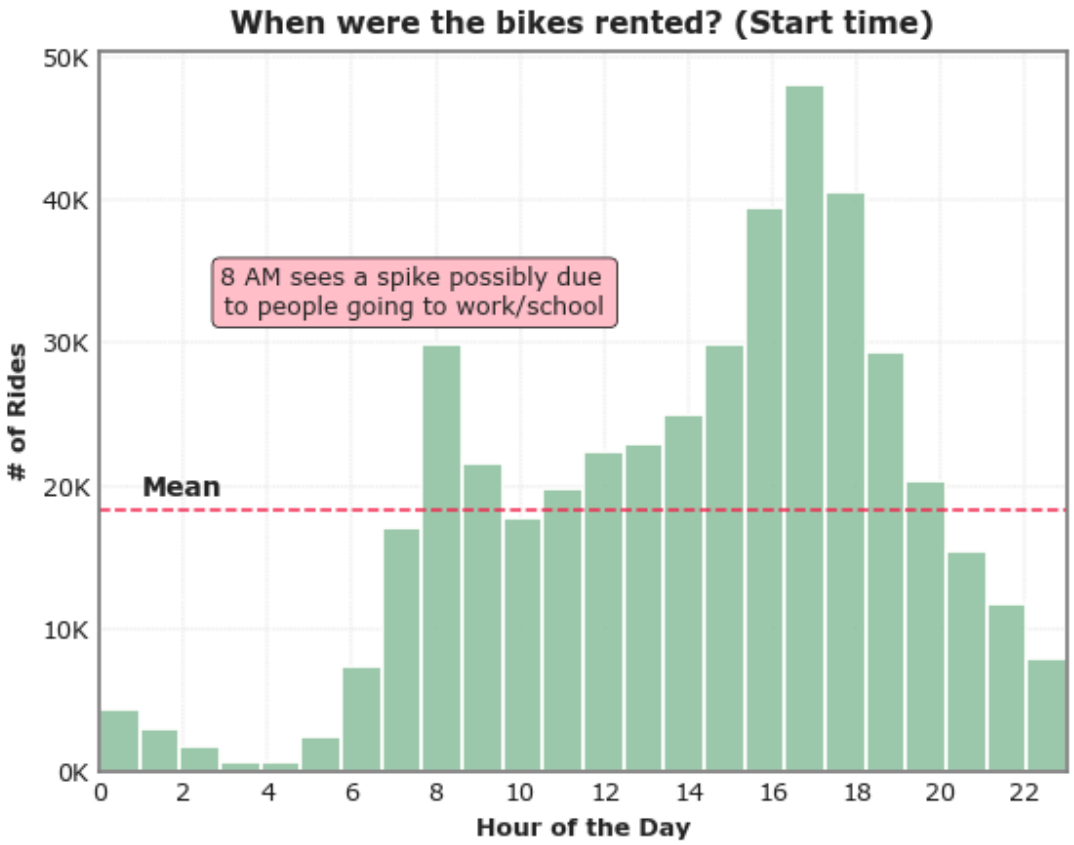
Project 2: BlueBikes Ridership in the Boston Area

1. **# Of Cumulative Rides**
   The Cumulative Rides Plot showcases what proportions of rides occur at similar stations. This generally shows the distribution of blue bike station traffic. As seen below, even though there are over 400 blue bike stations, Only 100 of them account for more than half of all blue bike rides.



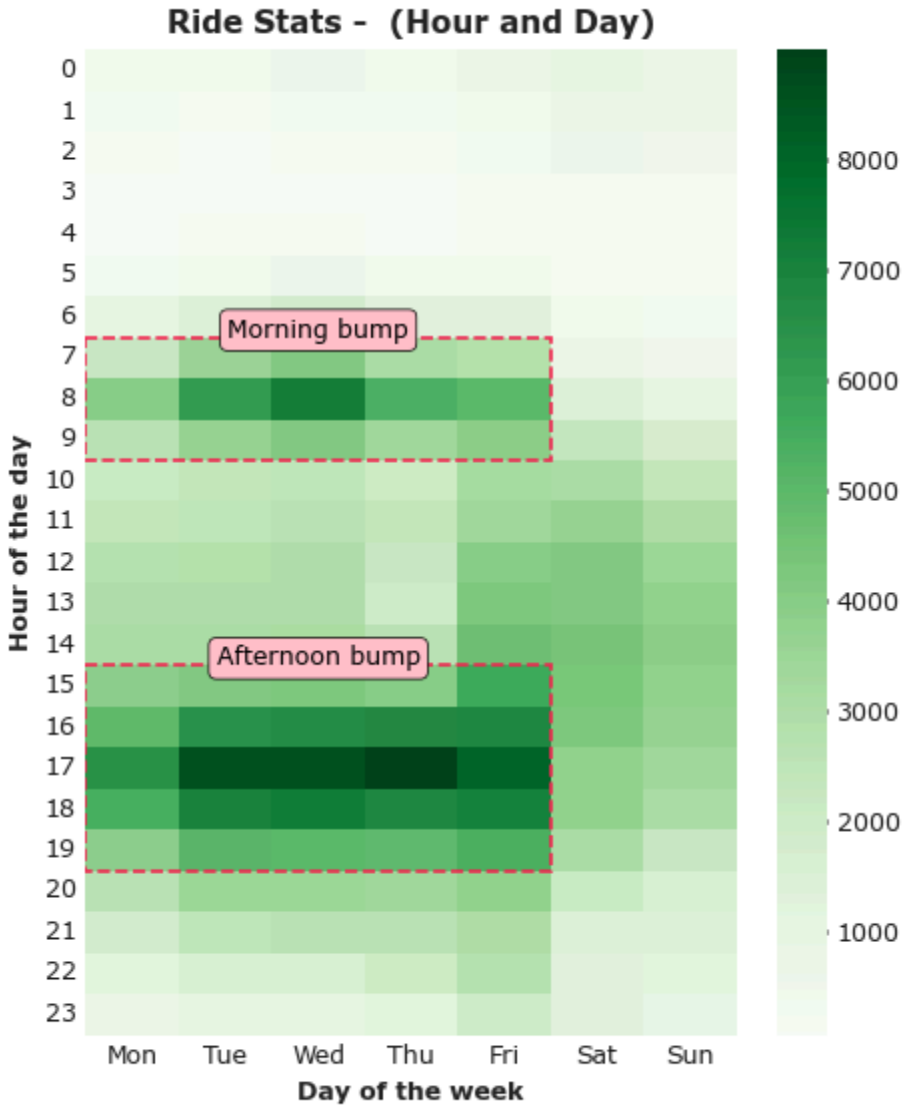
2. **When were bikes rented**

This Histogram compares the number of bike rentals with the time of day rented. This chart shows that there are two spikes during the day for blue bike commuting. The largest occurs between 3:00 pm and 7 pm, and the second peak occurs at 8 am. These peaks most likely correlate to work commutes because they sit close to the start and end times of a 9 to 5 job. The afternoon peak probably also represents leisure rides after work, as there is much more blue bike activity in the afternoon peak compared to the morning peak.



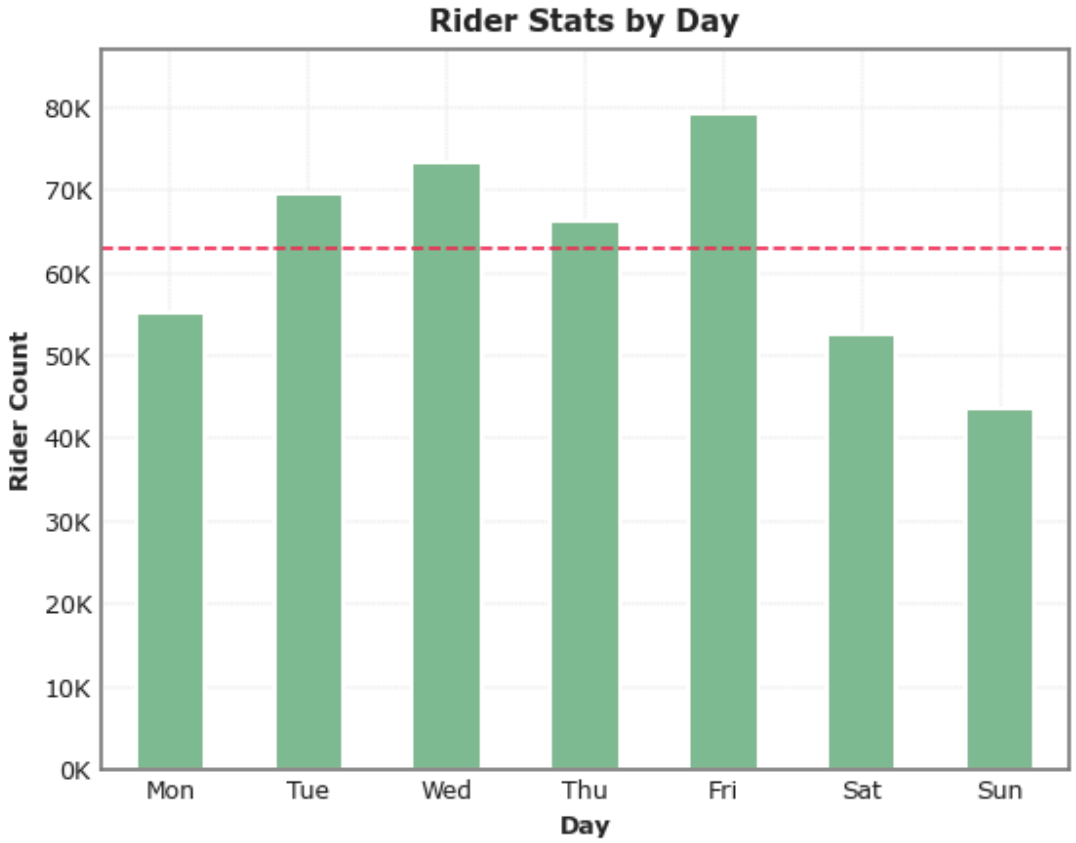
3. **Ride Stats (Hour and Day)**

This heatmap delves deeper into bike ride distribution throughout the week. It shows the morning ride peak and much larger afternoon peak as seen above in Chart 2, but it also shows that blue bikes are ridden much more often on weekdays than on the weekend. This supports the hypothesis that lots of people use blue bikes for a daily commute to work or school.



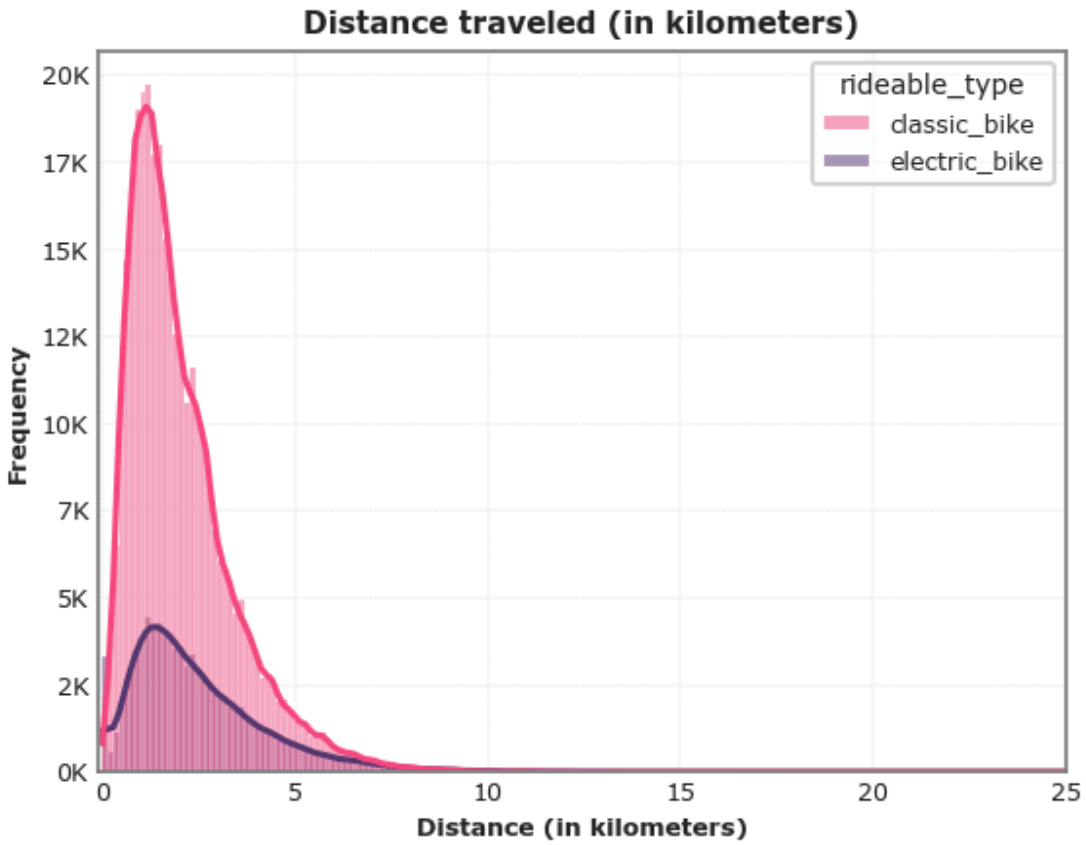
4. **Rider Stats by Day**

This bar graph shows the ride distribution only by day of the week, and is a component of the heatmap in chart 3. This chart similarly shows that blue bikes are used much more on the week days than on the weekends, but also reveals that Friday is the most popular day of the weekfto ride on blue bikes.



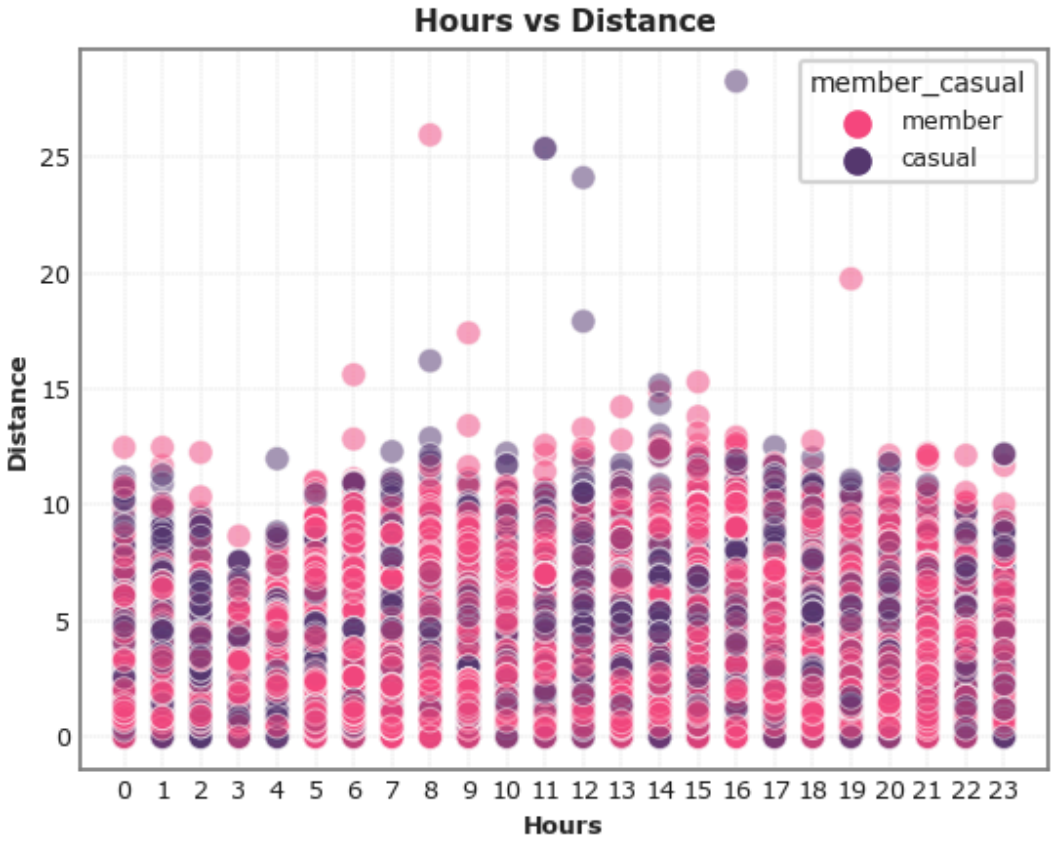
5. **Distance Traveled Using Blue Bikes**

This histogram comares all blue bike rides by their travel distance for both class bikes and E bikes. Both bike types have only one peak and while there are far fewer E-bike rides, both peaks occur around the same distance, which is about 1.5 km, or almost 1 mile. THis shows that most blue bike rides are somewhere around 1 mile, and reveals the intended travel range users have in mind when they choose to use a blue bike.

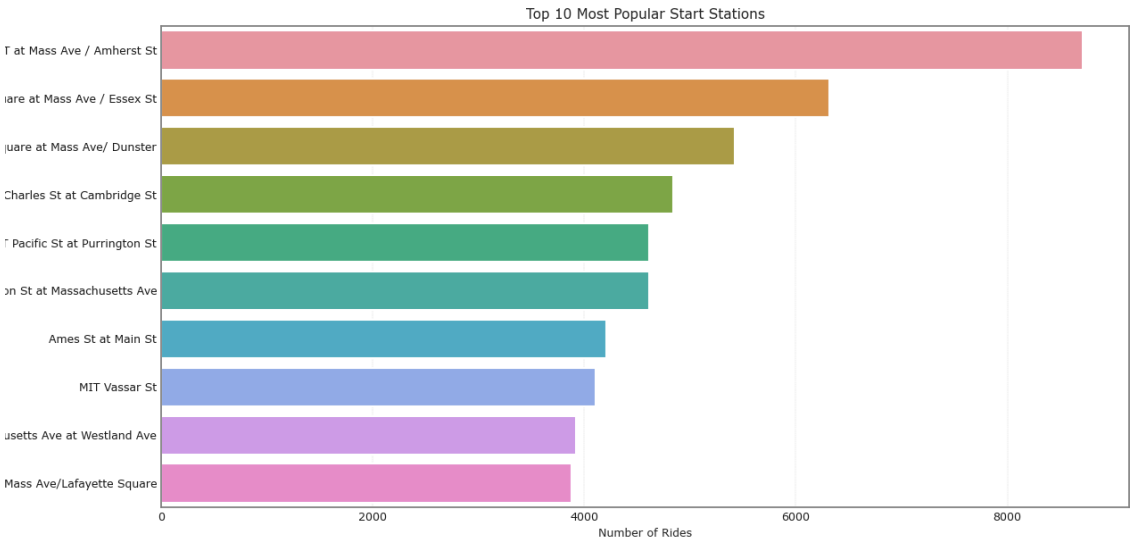


6. **Hours Vs Distance**

This scatterplot shows the distribution of ride distance during different times of the day for casual riders as well as subscription members. Interestingly, the distance distribution seems to be roughly equal for all times at the day, with the exception of a small dip around 3 and 4 am, and a small peak at 2pm. This shows that more often than not, time of day does not really affect the intended ride distance for both casual and subscription members.

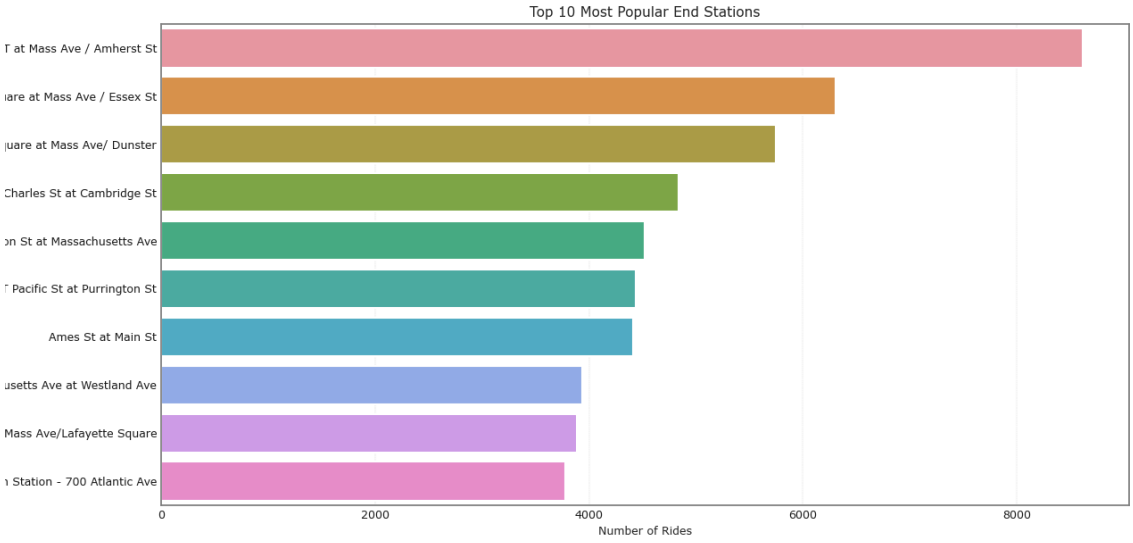


7. **Bar Plot of Most Popular Start Stations**
   The bar plot of the most popular start stations displays the top 10 stations where trips begin most frequently. This visualization helps identify key locations where bike demand is highest, which is crucial for bike redistribution strategies and station placement. By knowing the most popular start stations, we can infer high-traffic areas and potentially predict where bikes are most needed. If certain stations consistently appear at the top, they might be central business districts or residential hubs.

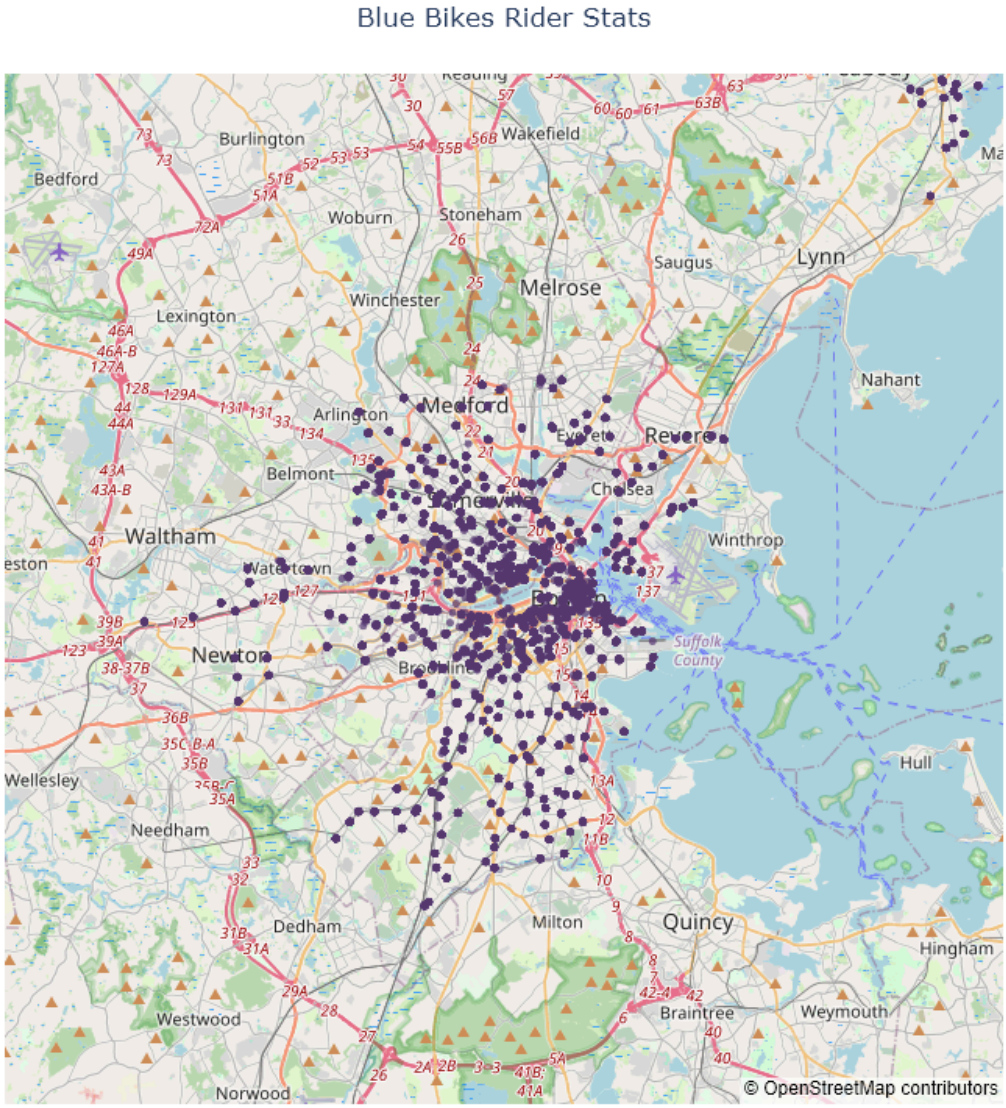


8. **Bar Plot of Most Popular End Stations**

   Similarly, the bar plot of the most popular end stations shows the top 10 stations where trips end most frequently. This visualization complements the start station analysis by highlighting destinations that are commonly reached. Understanding these endpoints can assist in optimizing station locations and ensuring that bikes are available where users are most likely to need them next. If certain end stations align with key transit points or popular destinations, this indicates a pattern in user commuting or travel behavior.

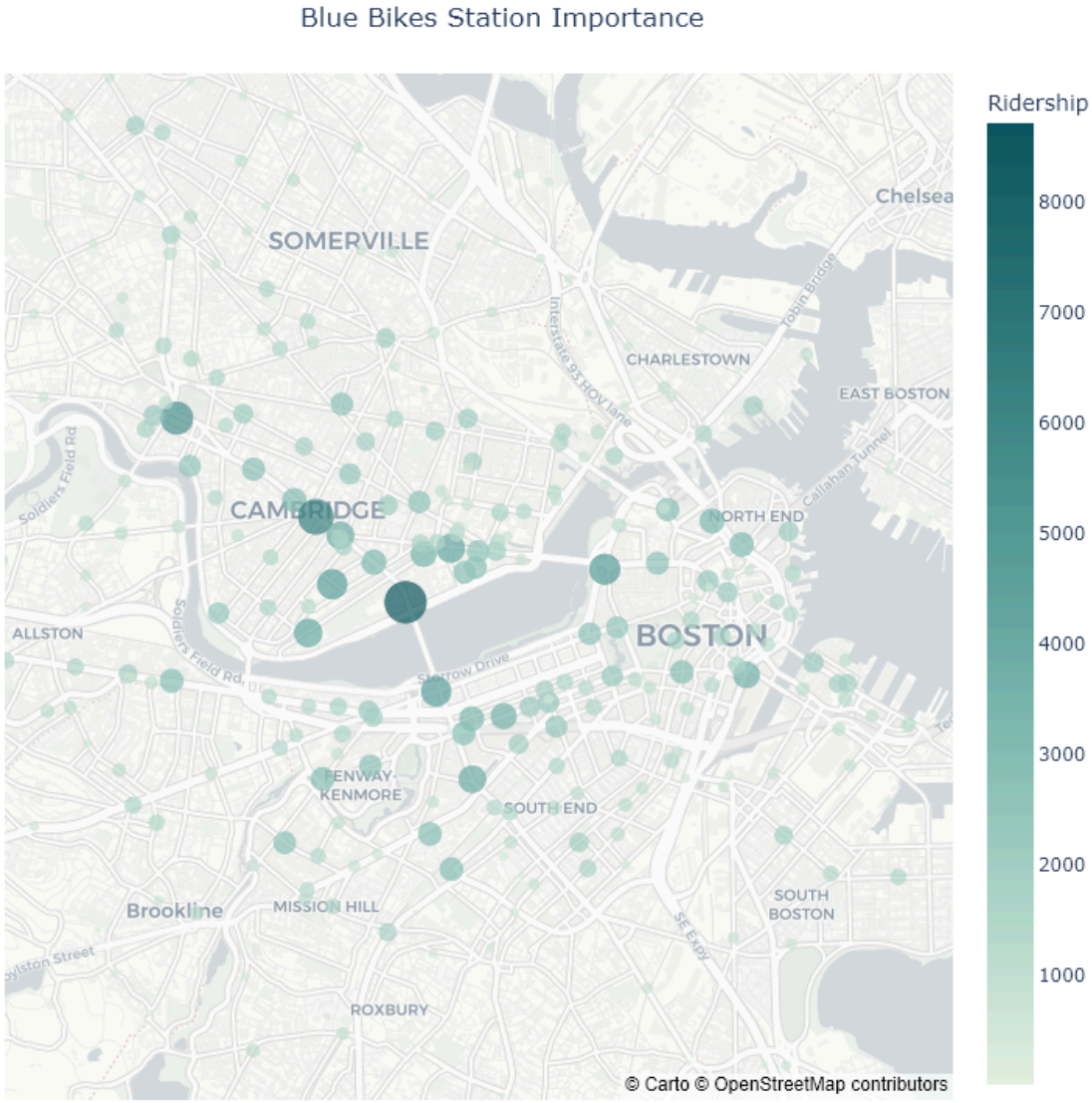


9. **Geomapping Train Stations**
   Plotting all the Blue Bikes stations over the map of the Greater Boston region shows a highly concentrated number of stations in the downtown area. There are bike stations around almost all MBTA bus/train stations.



10. Geomapping Train Station Importance

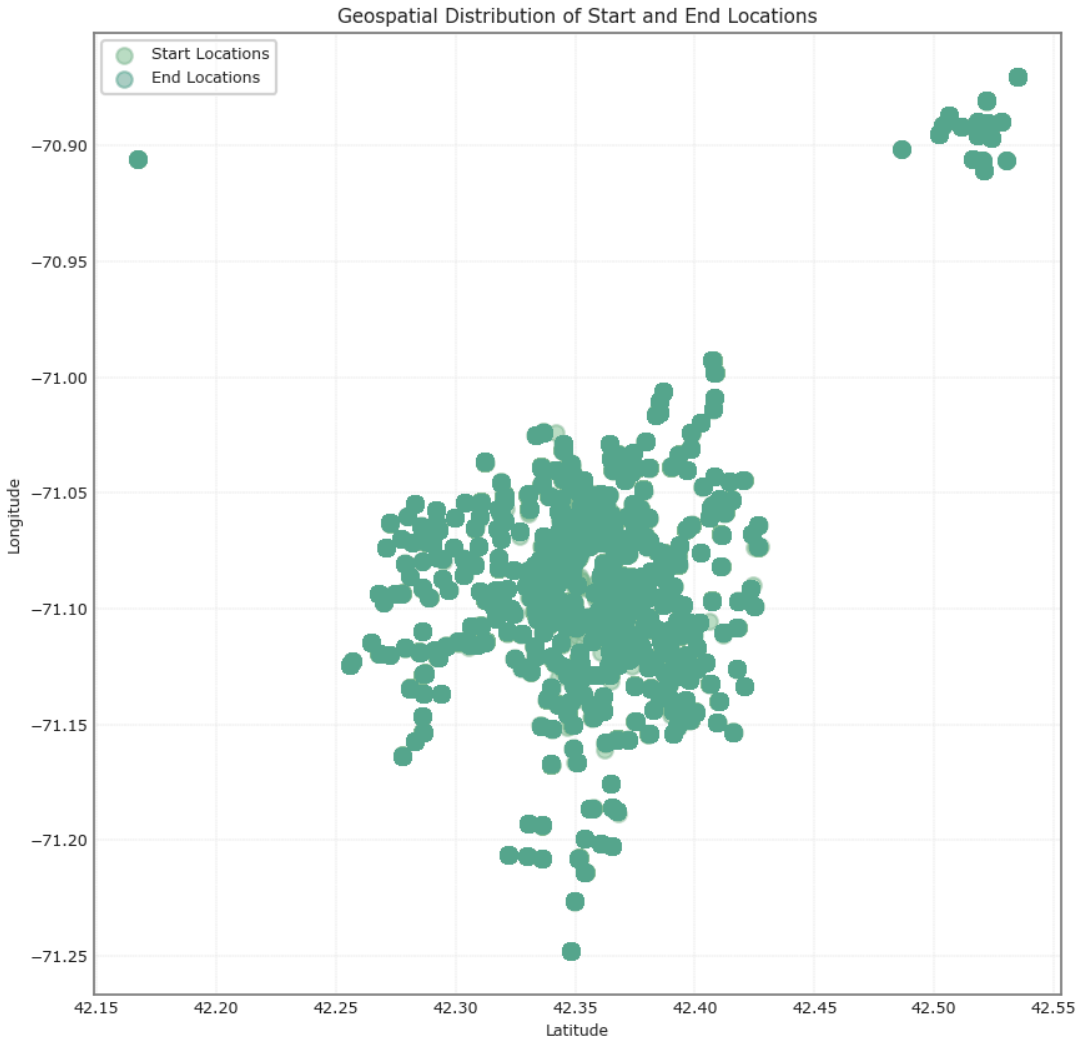
The area around MIT is a blue bike hotspot based on the high ridership. The neighborhoods near the Charles river attract many riders probably due to the scenic drive along the river.



11. Scatter Plot of Start and End Locations

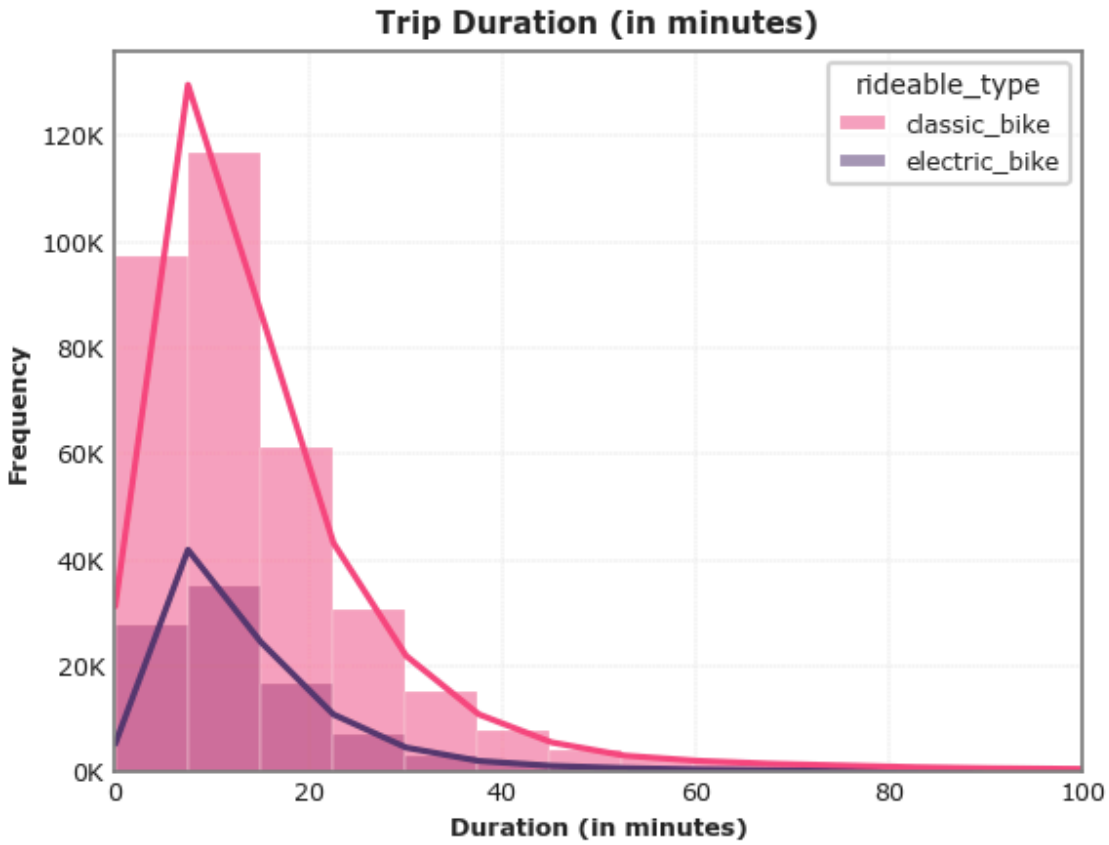
The scatter plot of start and end locations visualizes the geographic distribution of where trips begin and end. This geospatial analysis was chosen to understand the spatial dynamics of the bike-sharing system. The plot reveals areas with high densities of bike usage and helps identify clusters of activity. If we see that start and end points are

concentrated in specific areas, it might suggest high activity zones that require more attention for bike availability and maintenance.



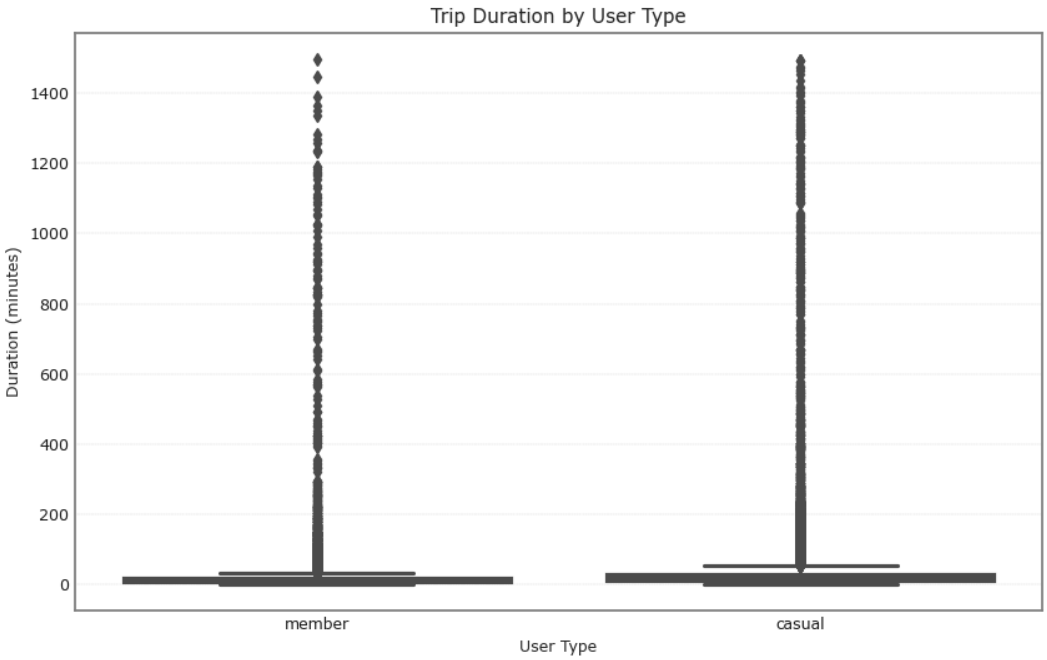
12. Histogram of Trip Durations (Classic + E Bike)

The histogram of trip durations shows the distribution of how long the bike trips last. This visualization was chosen because it provides a clear picture of the most common trip durations. From this histogram, we can identify if most trips are short or long, and detect any anomalies or outliers in trip durations. For instance, if we see a peak at around 15-20 minutes, it suggests that most users use the bikes for relatively short trips, likely for commuting or quick errands.



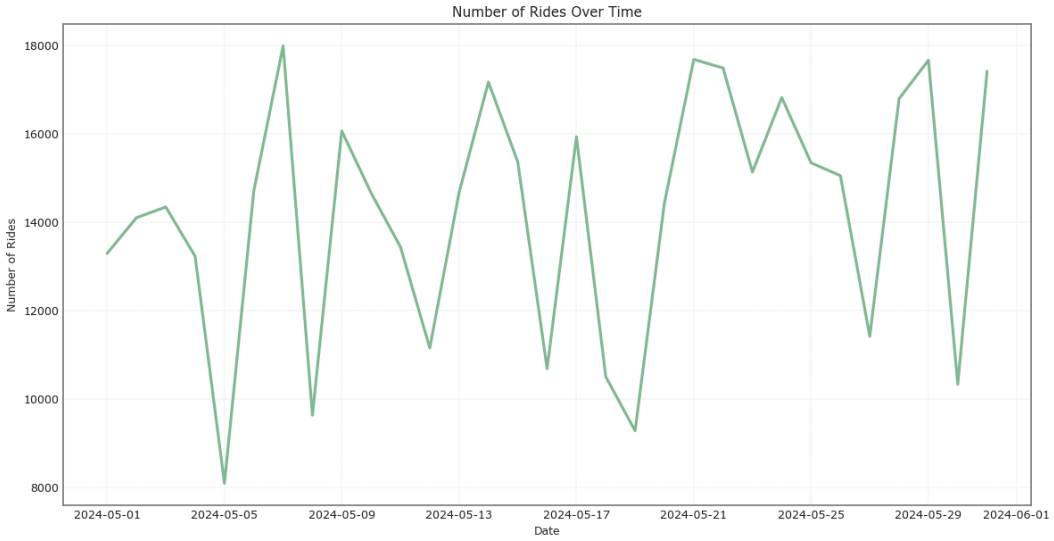
13. Box Plot of Trip Durations for Members vs Casual Users

The box plot compares the distribution of trip durations between members and casual users. This visualization was chosen to highlight differences in usage patterns between these two groups. The box plot shows the median trip duration, the interquartile range, and any potential outliers. Insights from this plot might reveal that one group tends to have longer or shorter trips on average, indicating different usage behaviors. For example, if casual users have a wider range of trip durations, it might suggest they use the service for more varied purposes compared to members.

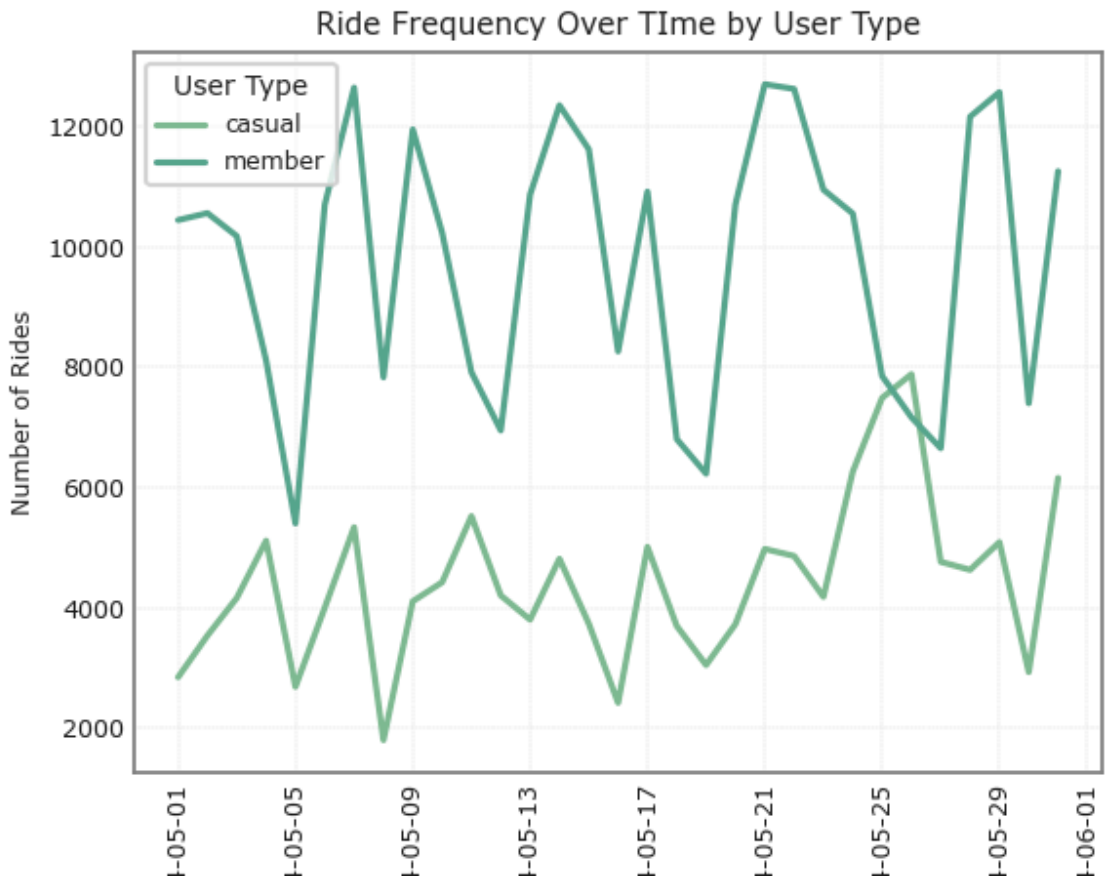


14. Line Plot of Number of Rides Over Time

The line plot showing the number of rides over time illustrates the temporal trends in bike usage. This visualization was chosen to analyze how bike usage varies over days, weeks, or months. It helps identify patterns such as peak usage times, trends in increasing or decreasing usage, and seasonal effects. For instance, we might observe higher bike usage during weekdays, indicating commuter patterns, or spikes in usage during certain months, suggesting seasonal influences.



15. **Line Plot Comparing Number of Rides for Members vs Casual Users Over Time**
    This line plot compares the number of rides over time between members and casual users. It was chosen to highlight temporal differences in usage between these two groups. The plot can reveal trends such as whether membership growth is steady, if casual user rides peak during certain times (e.g., weekends or holidays), and how external factors might differently impact each group. For example, a sudden increase in casual user rides during a specific period might indicate a promotional event or favorable weather conditions.



Findings and Insights

From these visualizations, several key insights can be gathered. We can determine the typical duration of trips and how they differ between user types, identify the most popular start and end stations, understand the spatial distribution of bike trips, and analyze temporal trends in bike usage. These insights are crucial for optimizing the bike-sharing system, including improving bike distribution strategies, enhancing user experience, and planning future expansions. The differentiation between member and casual user behavior can guide targeted marketing and membership retention efforts. Overall, these visualizations provide a comprehensive understanding of the usage patterns and operational needs of the bike-sharing service.