Data Visualization & Communication

Professors: Dimitris Karlis, Nikos Platis

Analyzing Bicycle Usage Data from Washington Capital Bikeshare

Vasileios Ilias Drouzas

Melina Moniaki

Dimitrios Stathopoulos

Giannis Trantalidis

Athens, June 2024

Contents

The primary objective of this study is to comprehensively explore and analyze bicycle usage data from the Washington Capital Bikeshare (CaBi) system. Specifically, we focus on data spanning from January to April in the years 2020, 2021, 2022, and 2023. This period was chosen to capture seasonal variations and longitudinal trends in bike usage across multiple years.

Aim of the Study

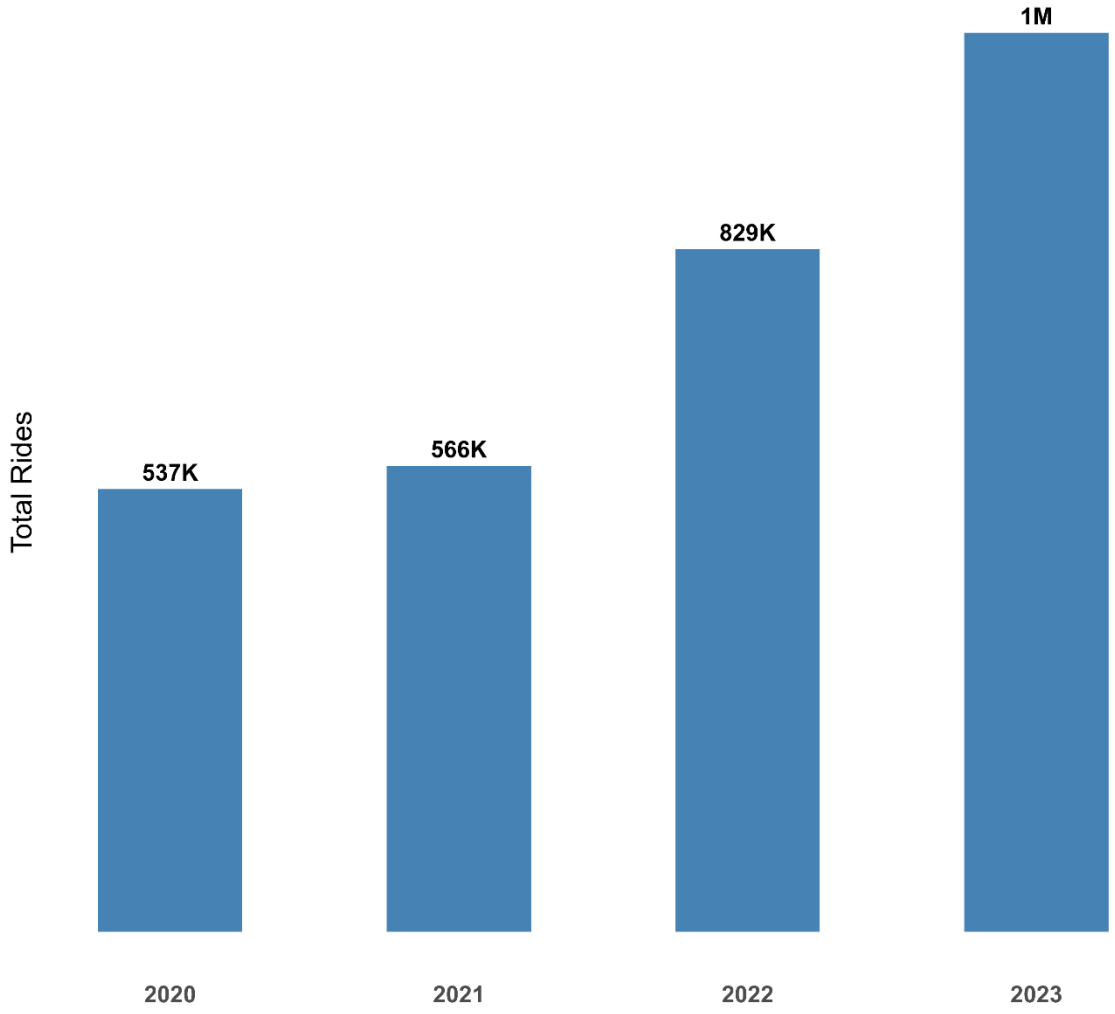
The study aims to delve into the bicycle usage data from the Washington Capital Bikeshare system. Specifically, it will analyze data spanning from January to April over the years 2020, 2021, 2022, and 2023. Notably, this period includes the extension of the stay-at-home order by Governor Inslee on April 2, which lasted through at least May 4. This context provides a unique opportunity to assess how the pandemic and associated restrictions impacted bikeshare usage patterns.

Approach

To achieve the study's objectives, the following approach will be employed:

- **Analyze Usage Patterns**: Examine how bikeshare usage fluctuates over different time periods, considering daily, weekly, and monthly trends. This includes investigating peak usage times and any notable changes in patterns across the four years.
- **Bike Types and Membership Subscriptions**: Identify the most commonly used bike types (e.g., classic, electric, docked) and categorize usage by different membership types (e.g., casual users, annual members). This will help in understanding user preferences and how they have evolved.
- **Station and Route Analysis**: Determine which stations and routes are most frequently used. This involves mapping out high-traffic areas and popular routes, providing insights into user behavior and potential areas for service improvements.
- **Profitability Evaluation**: Assess the profitability of various stations and bike types. By analyzing revenue generated from different sources, the study aims to pinpoint which stations and bike types contribute most to the system's financial health.

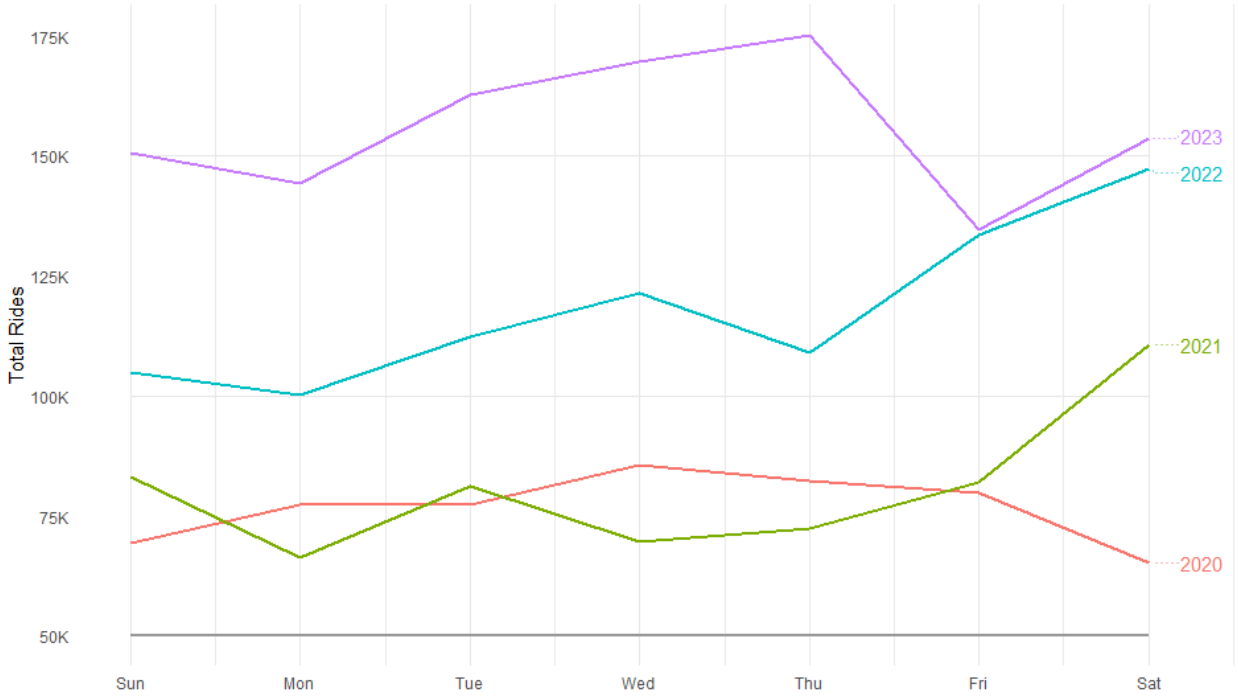
Total rides by year



We see a clear increase in bike rides after 2021, likely influenced by COVID-19. During the quarantine, it is sensible that bicycle usage had been lower and as soon as the isolation measures were aborted, people started to use bicycles as a means of transport more often, possibly to avoid big crowds

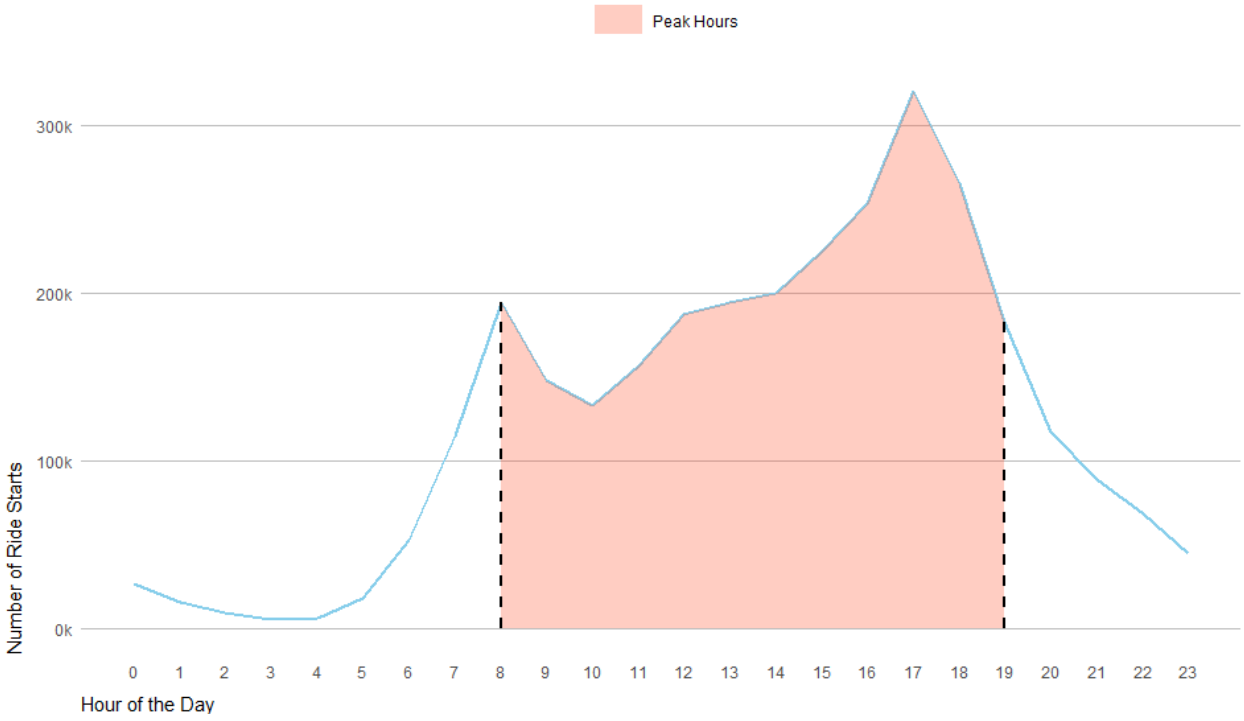
The total number of trips each year went up noticeably, possibly because people turned to biking more during the pandemic, avoiding crowded public transport and seeking outdoor activities.

Total rides by day of the week



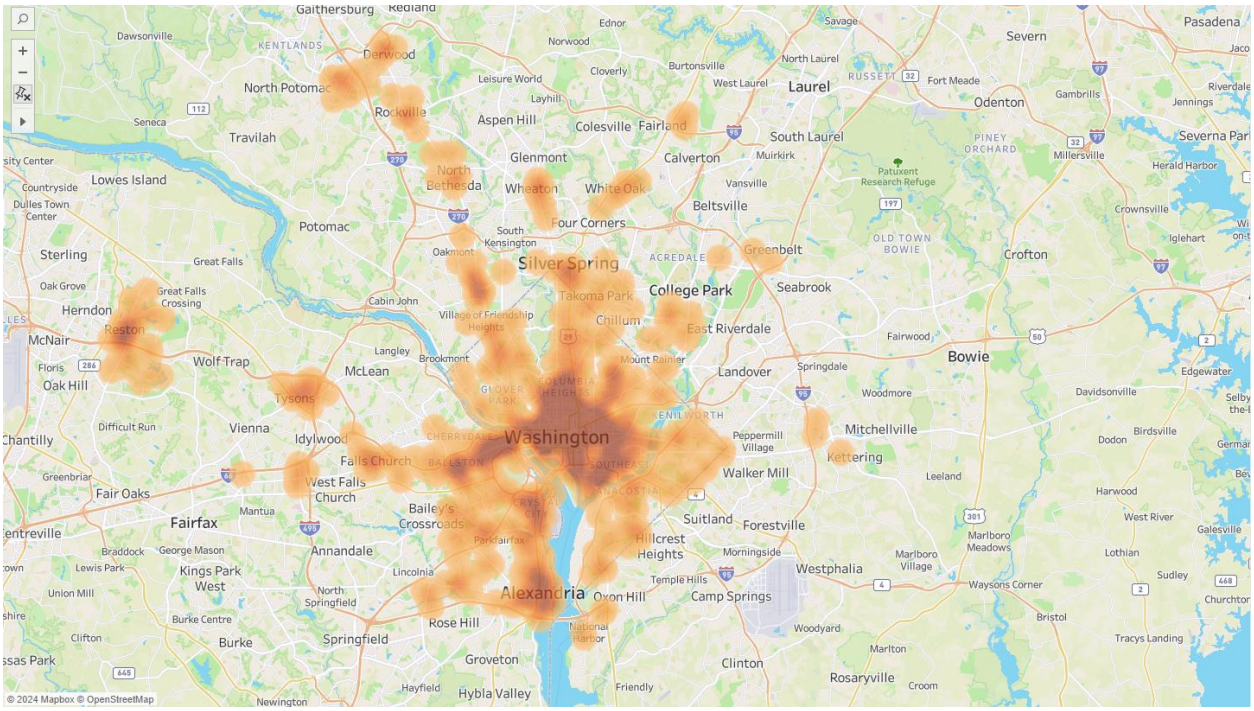
We notice that March and April are the months with greater bike engagement (except 2020). We consider this sensible, since in the winter months the cold and unpredictable weather discourages people from using bicycles. In 2020, with COVID measures at their peak, it is interesting to notice that January is on average higher than the other months; this can be explained by the fact that COVID-19 started to rise more after February, and soon after quarantine measures were taken.

Ride Start Time Distribution



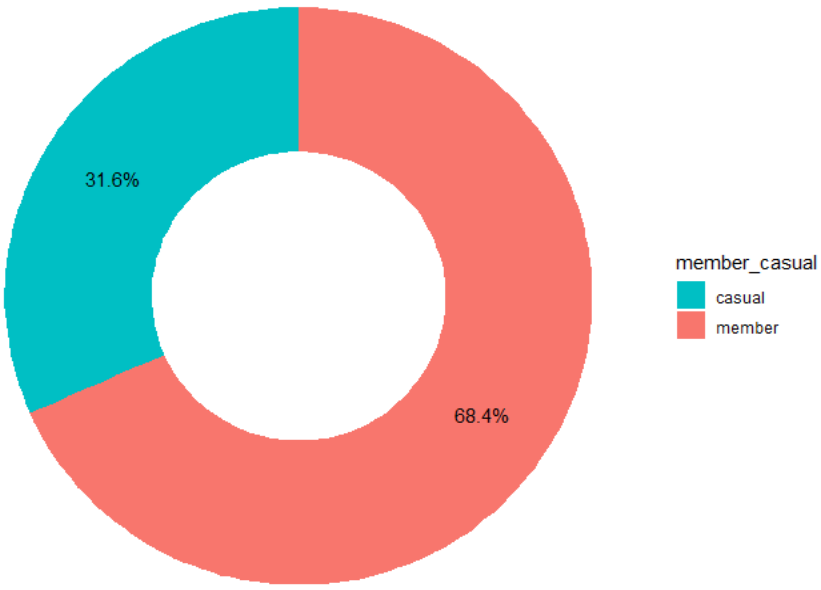
In this plot, we are examining the hourly distribution of bike starts. Active seems to be starting from 5 or 6 in the morning and keep going until late at midnight. It is also observable that the peak hours are roughly from 8a.m. through to 7p.m., this is expected to regarding the fact that in general the active hours of humans are mostly on those hours. We can also observe a drop from 8 o'clock until 12 o'clock in the morning suggesting that potential bike users might be at their workplaces, homes etc.

Color Density Map Plot



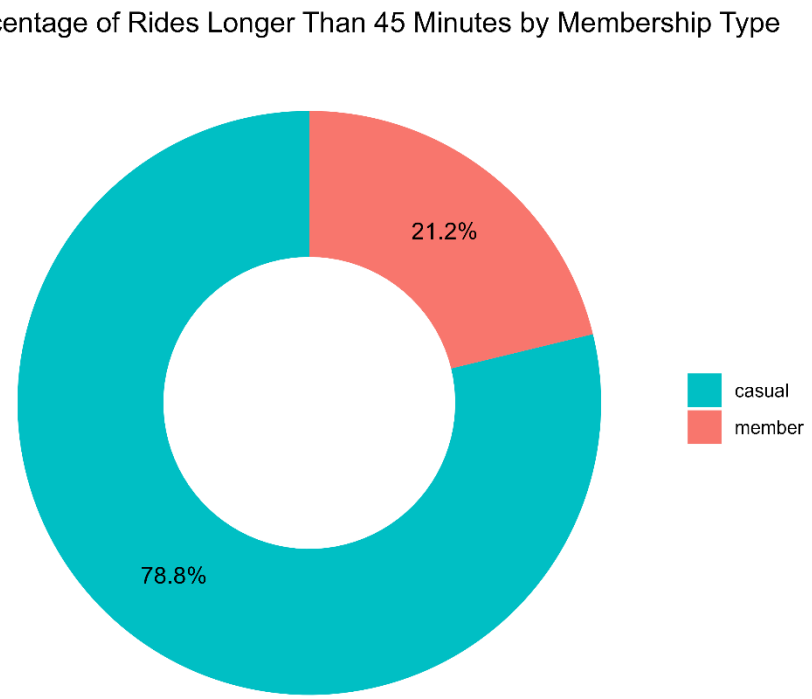
This is a color intensity map, each of the stations are colored with the orange palette. Darker spots indicate high activity of bike rides starting and more opaque spots indicate low activity(bike ride starts). We can observe that locations like Washington and Alexandria show high activity and locations close i.e. to College Park have low activity.

Membership Ratio

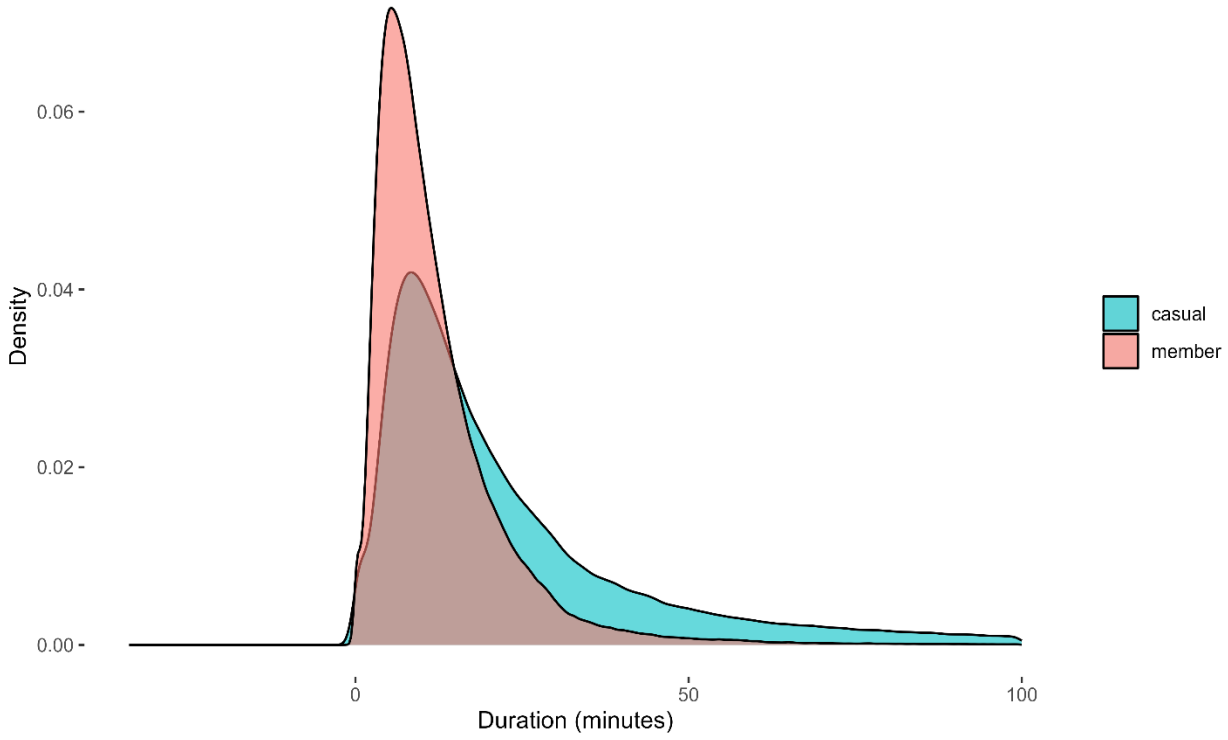


Ride usage data indicates that member users account for a significantly higher proportion of rides (approximately 70%) compared to casual riders (approximately 30%).

In the year 2020, there is a noticeable decline in the number of rides taken by casual riders.



Distribution of Ride Duration by Membership Type



In the second diagram (bar plot), we break the previous percentages annually, yet confirming again that in all the years 2020-2024, members tend to make more rides than the casual users.

Finally, we find out an interesting fact: Only 21.2% of the users that exceed the duration of 45 minutes are members (remind that members have the first 45 minutes free). The majority of users that do lengthy rides are casuals.

Cyclistic members have a 45-minute free ride period as part of their membership plan.

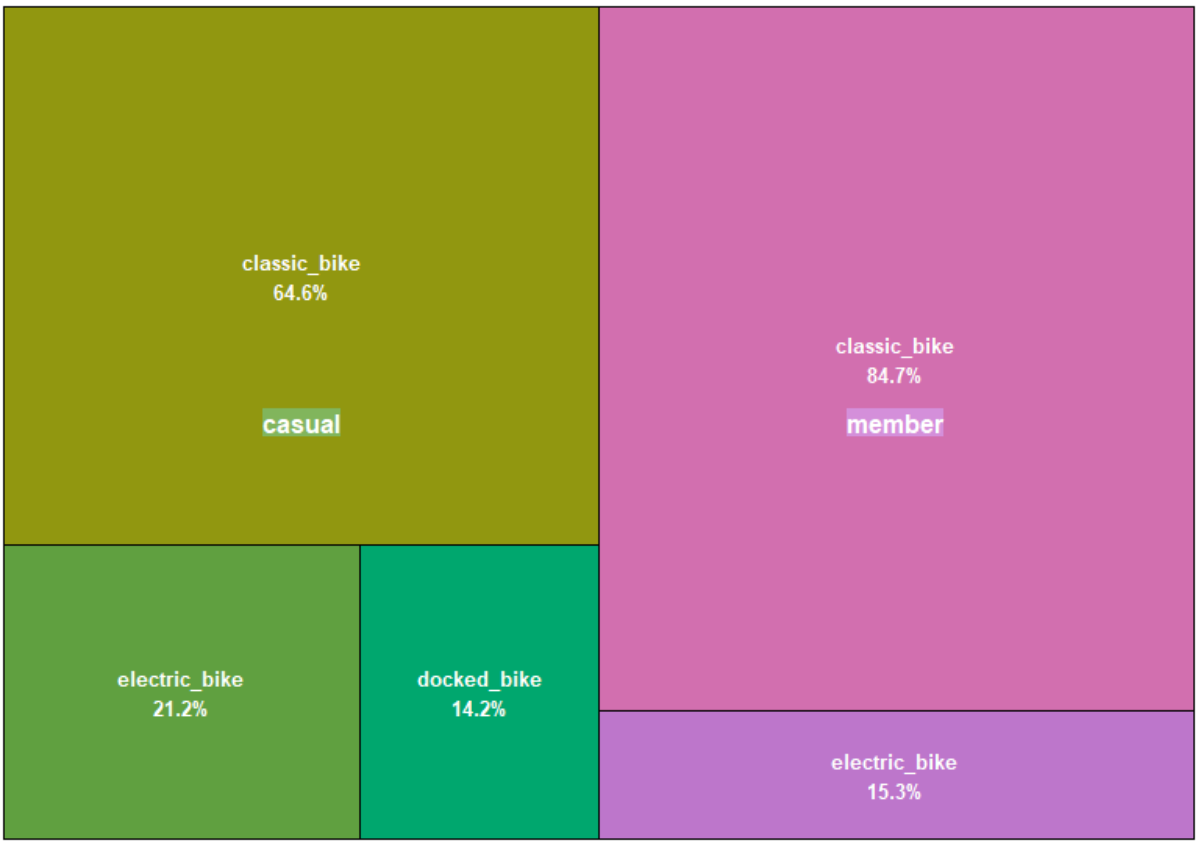
We observe that member riders perform shorted rides.

This could lead to two possible scenarios: members may favor shorter rides to stay within the free ride window, or individuals who anticipate primarily taking shorter rides may choose to become members to take advantage of the free ride period.

Monthly total rides by year

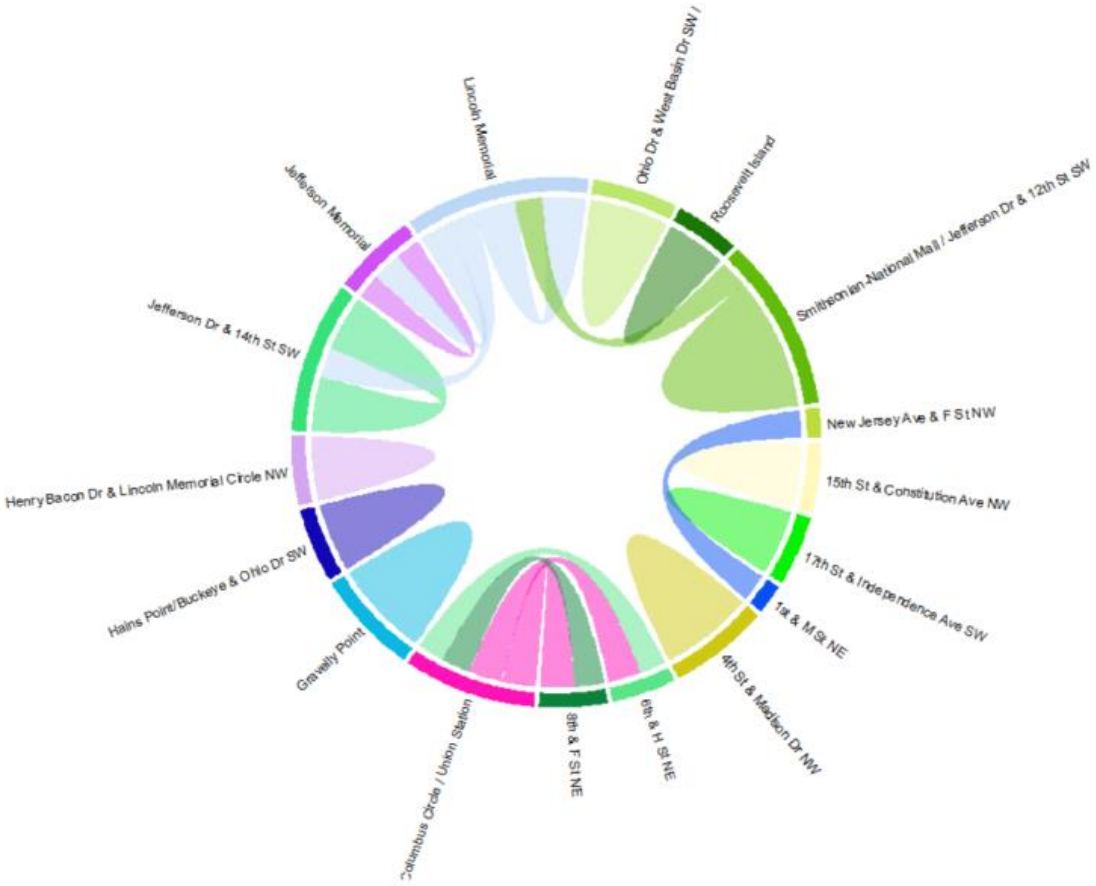
For the years 2021, 2022, and 2023, the data shows that the total number of rides taken by both members and casual riders starts lower in January and exhibits an upward trend until April. This pattern could potentially be attributed to improving weather conditions during spring months.

Bike Usage Percentage by Member Type and Bike Type



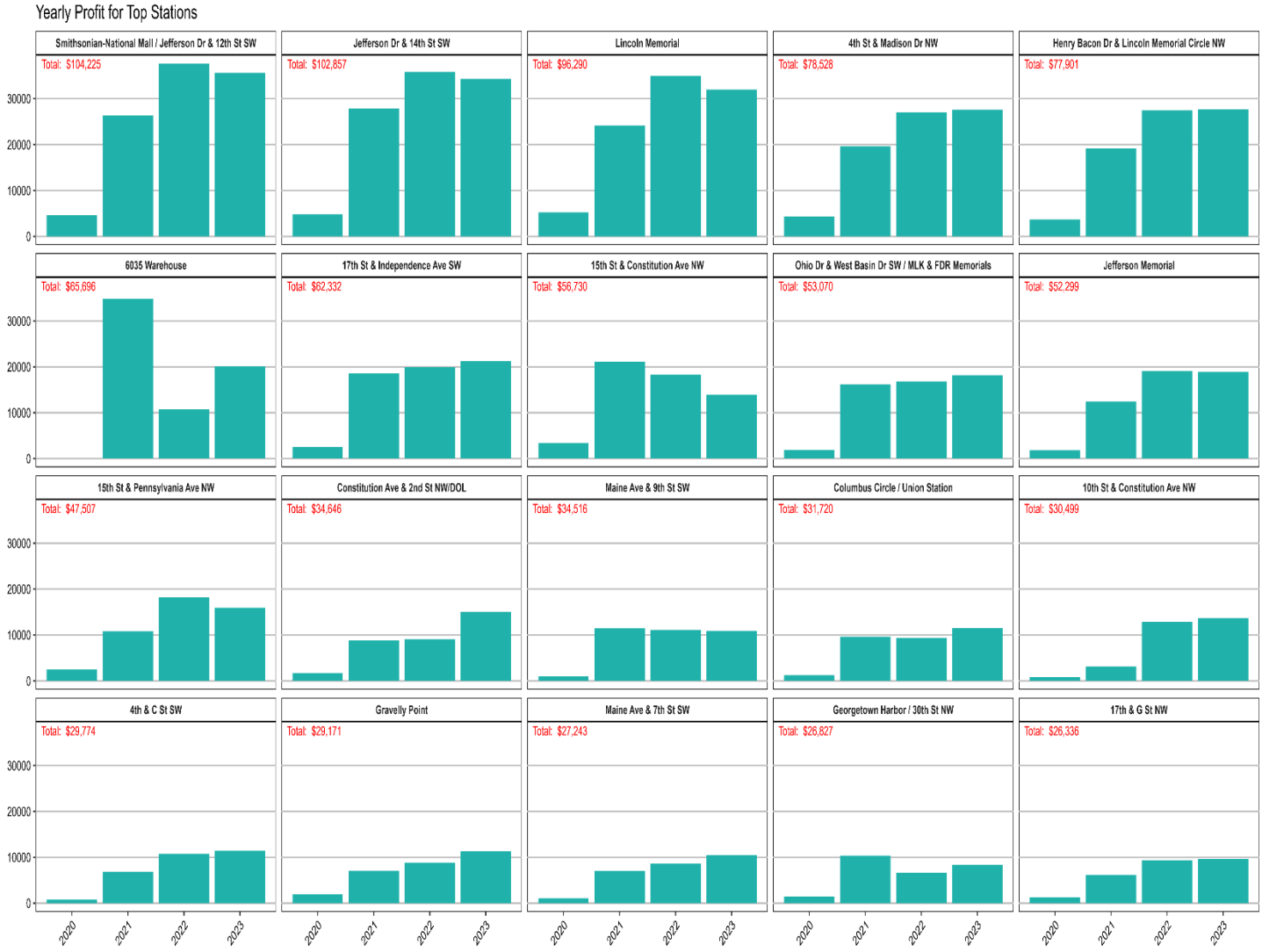
We notice that the majority of users, whether they are casual users or members, tend to use the classic bike more than the others. Classic bikes are preferred for their convenience and ease of use, appealing to a broad user base seeking reliable transportation options. Some casual users opt for the electric bike, while the docked bike is a minority and chosen only by casual users. Note that electric bikes attract casual users drawn to their assisted pedaling capabilities, offering a comfortable and efficient alternative for longer or challenging routes.

Top 20 Most Common Bike Routes



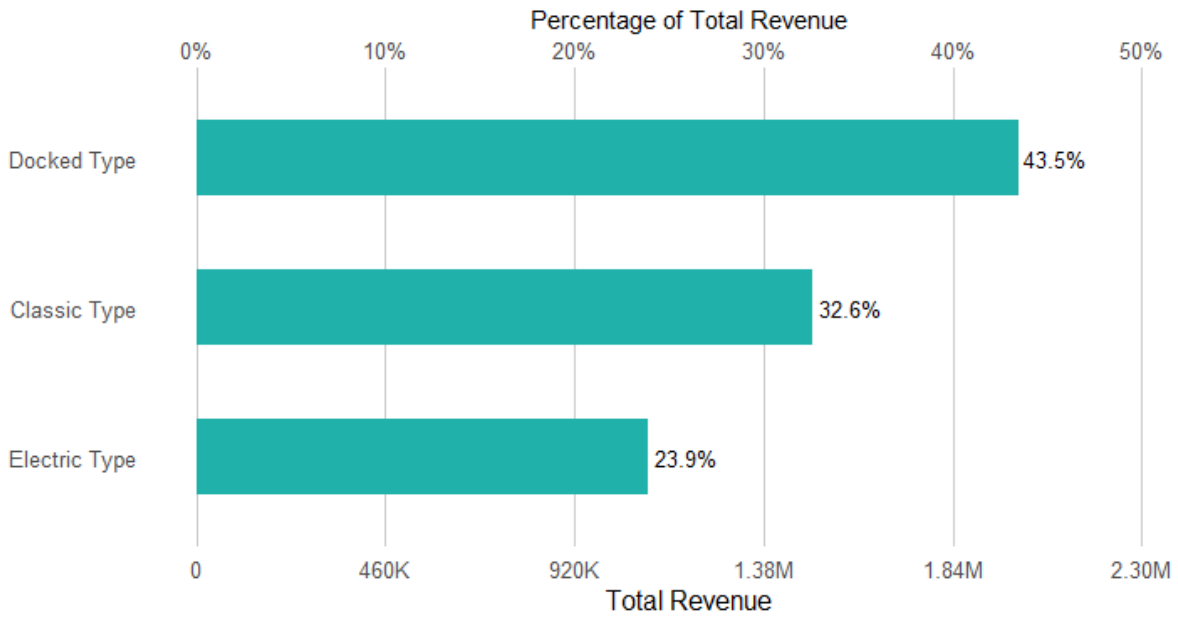
This plot illustrates the most frequently used bike routes. It is evident that the most common routes are concentrated around stations near the capital of Washington. In this chord plot, each starting station is represented by a specific color, and its chords indicate connections to various end stations. The thickness of each chord represents the frequency of the route—the thicker the chord, the more commonly used the route. Additionally, the length of the arc for each station indicates its overall activity level; longer arcs denote busier stations.

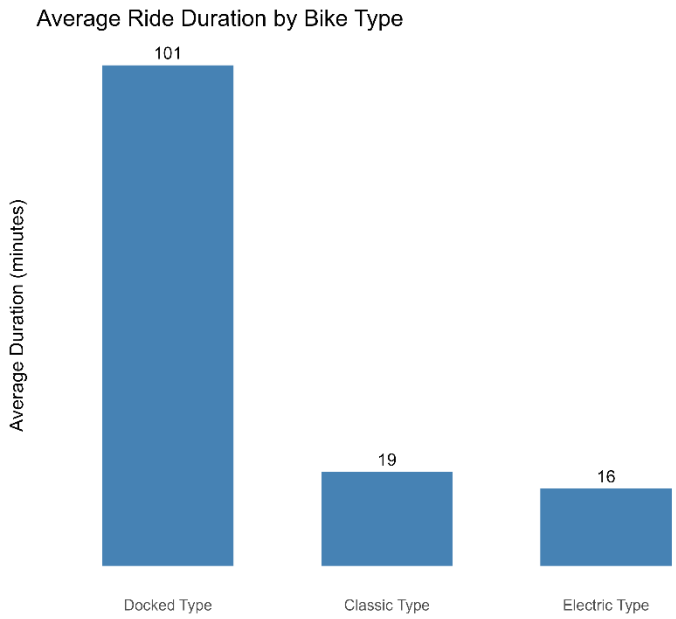
Total Profit over time for top Stations



As we suspected, the most profits arise after 2020 for all the stations, when quarantine measures were relaxed and people tended to the use of bicycles to avoid crowded public means of transport. For most stations, the most profitable years have been 2022 and 2023.

Total Revenue by Bike Type



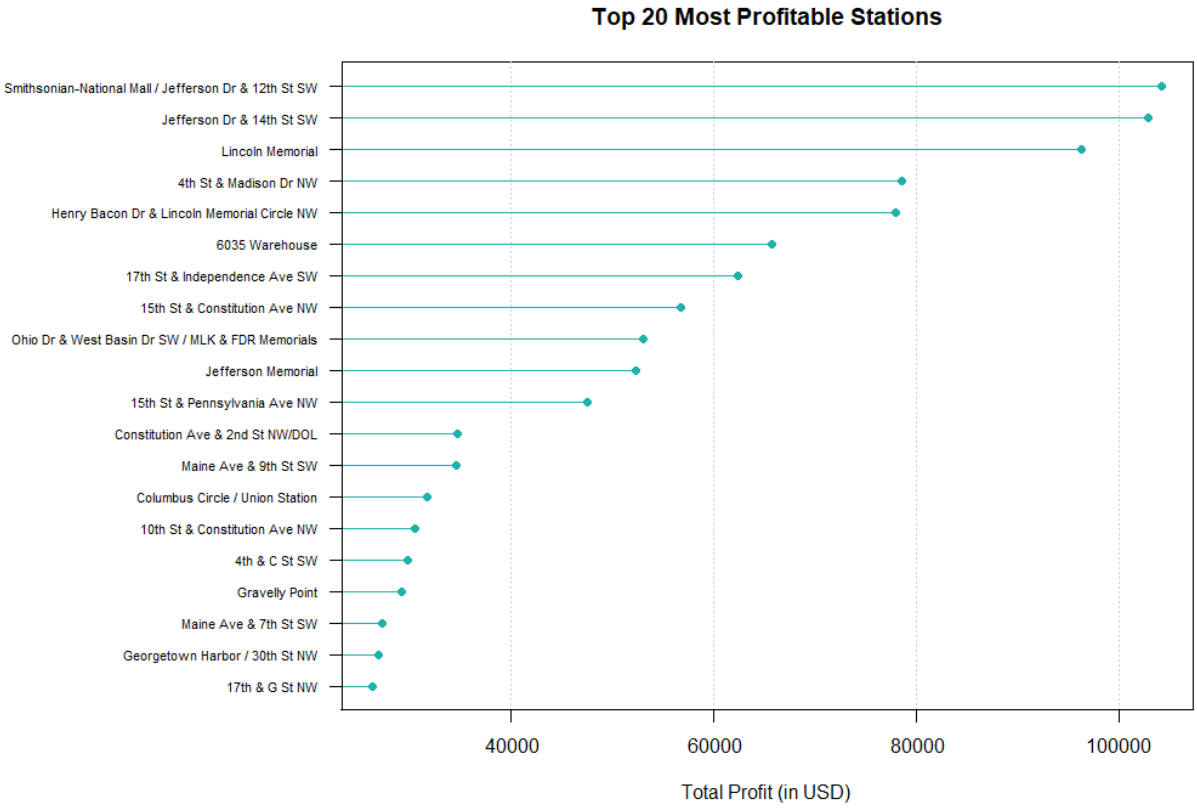


The most profitable bike type is the docked bike, gaining profits of about 2 million dollars (43.5% of the total revenue), followed by the classic bike overcoming 1.5

million (32.6% of the total revenue). Finally, the electric type brings the least profit (less than 1 million and 23.9% of the total revenue).

Even though the docked bike is the least type of bicycle that is used, as we found out earlier, its users tend to use it for a big amount of time, thus making a bigger profit for the company.

Most profitable stations over the years



The first three stations (Smithsonian- National Mall/ Jefferson Dr & 12th/14th St SW and Lincoln Memorial) have been the most profitable ones, gaining over

100.000 dollars (the first two) and approximately 95.000 (Lincoln Memorial). These three stations are all central ones and very close to each other, so it is not surprise to see them top the standings. The majority of the stations achieve a profit below 60.000 dollars.

Conclusions

Despite being rarely used, docked bikes generate the greatest profit for the company due to their lengthy rides. The most profitable stations are centrally located, benefiting from high user traffic. The pandemic and quarantine measures in 2020 led to a decline in bicycle usage and the company's profits, although there is a yearly growing trend in bike rides and usage. Ride starts are concentrated around Washington, Alexandria, and nearby areas, with peak usage occurring between 8:00 a.m. and 7:00 p.m. The most common bike routes are centered around stations near the capital, indicating high activity in these regions.

Contributions

Drouzas Vasilis handled the following diagrams:

1) Bike Usage percentage by member type and bike type
2) Number of trips by member type
3) Monthly total rides by year
4) Yearly total profit per membership type
5) Most profitable stations by rides and revenue (interactive #1)

Moniaki Melina handled the following diagrams:

1)Top 20 Most Common Bike Routes (Chord Plot)

2)Most profitable Stations over the years (Lollipop Plot)

3)Total Profit over time for top Stations (Trellis Plot)

4) Washington Capital Bikeshare Stations (Interactive #2 with tableau)

Dimitris Stathopoulos handled the following diagrams:

1) Total rides by day of the week (Line Plot)

2) Ride Start Time Distribution (Line Plot)

3) Total Revenue by Bike Type (Horizontal Bar Plot)

4) Color Density Map Plot (Color Intensity Map Plot)

Yannis Trantalidis handled the following diagrams**:**

1) Total rides by year
2) All plots in Membership Ratio
3) Distribution of Ride Duration by membership type
4) Monthly Total Rides by year
5) Average Ride Duration by Bike type