INFO - I 590

Data Visualization

Divvy Bikes Data Analysis

Spring 2025

Submitted By

Sri Sai Lahari Gandrapu
Sai Sudha Moola
Yuva Prathima Sola

Abstract

This project explores the Divvy Bikes dataset to uncover usage patterns, understand rider behavior, and present insights that can support decisions related to urban mobility. The analysis highlights key aspects such as how ride activity varies by time of day, the differences between casual and member riders, and which stations experience the most traffic in Chicago. The dataset spans a range of attributes, like start and end locations, trip durations, and types of bikes, collected over time. By applying data cleaning and exploratory data analysis (EDA) methods, the study reveals trends and irregularities in bike usage based on time, day of the week, and geographic location. Visualization tools like heat maps, bar charts, and line graphs present the findings, emphasizing a better understanding of peak usage times and the most frequent stations. Ultimately, this analysis provides actionable recommendations to improve bike-sharing systems, promote sustainable transportation, and support future urban infrastructure planning.

I. Introduction

In recent years, cities worldwide have been exploring innovative strategies to tackle challenges like air pollution, traffic congestion, and sustainability. Chicago stands out with its expanding Divvy bike-sharing program, demonstrating how cycling can contribute to a cleaner and more efficient urban environment. Choosing bikes over short car trips helps ease traffic and reduce carbon emissions, supporting the city's environmental initiatives. This study uses data visualization to analyze Divvy bike usage, aiming to identify ridership trends, peak activity times, and areas where infrastructure improvements could further encourage cycling. Analyzing this data helps demonstrate how Divvy contributes to reducing traffic congestion and promotes the use of eco-friendly transportation options. Additionally, the insights gained can foster community engagement by highlighting the benefits of cycling, ultimately encouraging a shift toward greener travel habits and more active lifestyles.

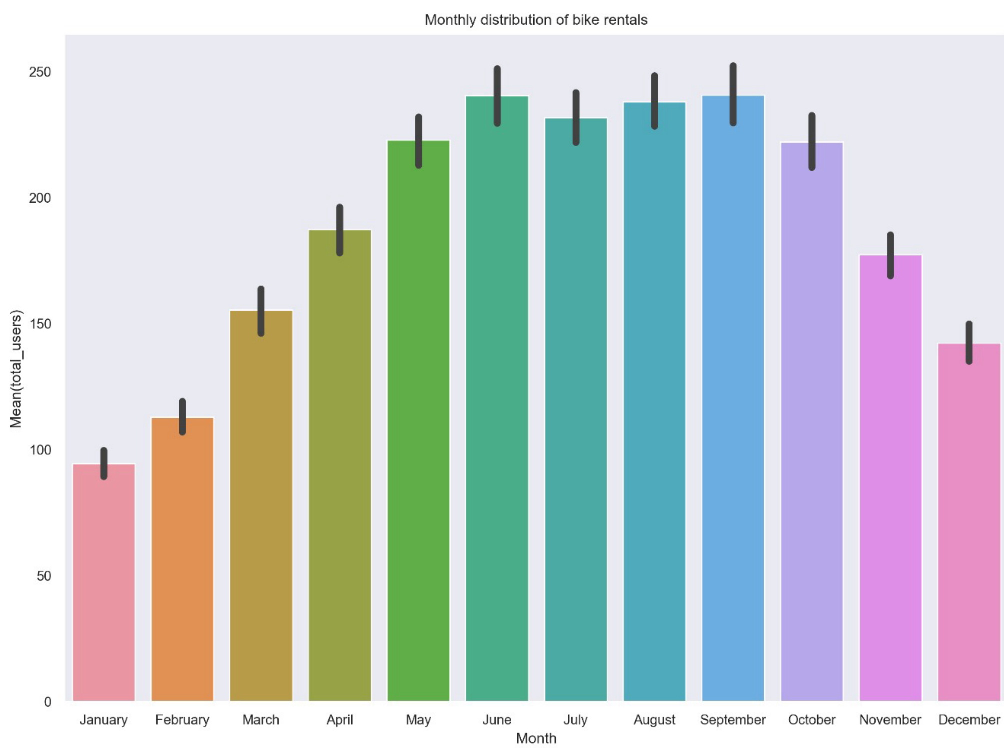
II. Motivation

This project is driven by the opportunity to leverage data visualization in promoting sustainable transportation across Chicago. As the city faces increasing challenges from traffic congestion and environmental concerns, encouraging cycling offers a meaningful solution. Analyzing Divvy bike usage helps reveal patterns that highlight both the environmental benefits of reduced carbon emissions and the need for improved accessibility and infrastructure. By showing how bike-sharing services like Divvy reduce reliance on short car trips and make cycling more practical and appealing for residents, the analysis can foster greater community engagement. Ultimately, the project aims to contribute to Chicago's sustainability goals while promoting a shared commitment to eco-friendly urban travel.

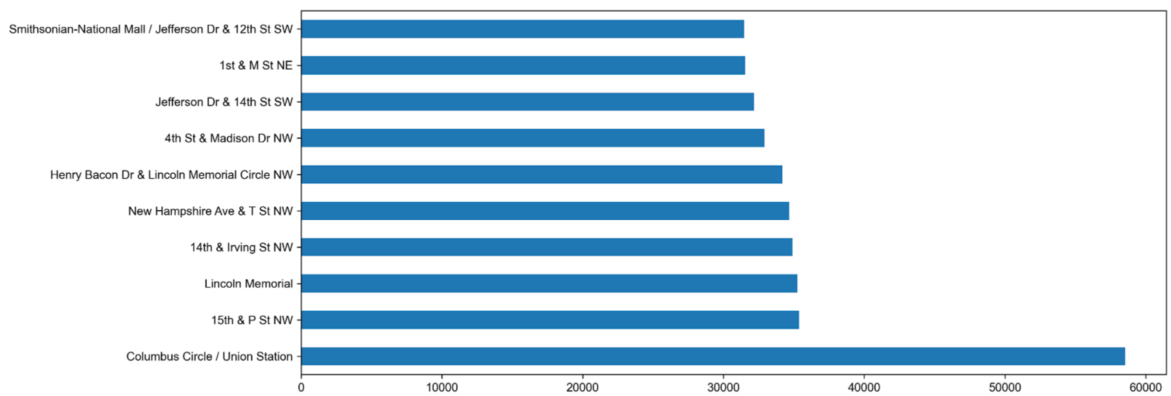
III. Background

1. Capital Bikeshare (Washington, D.C.)

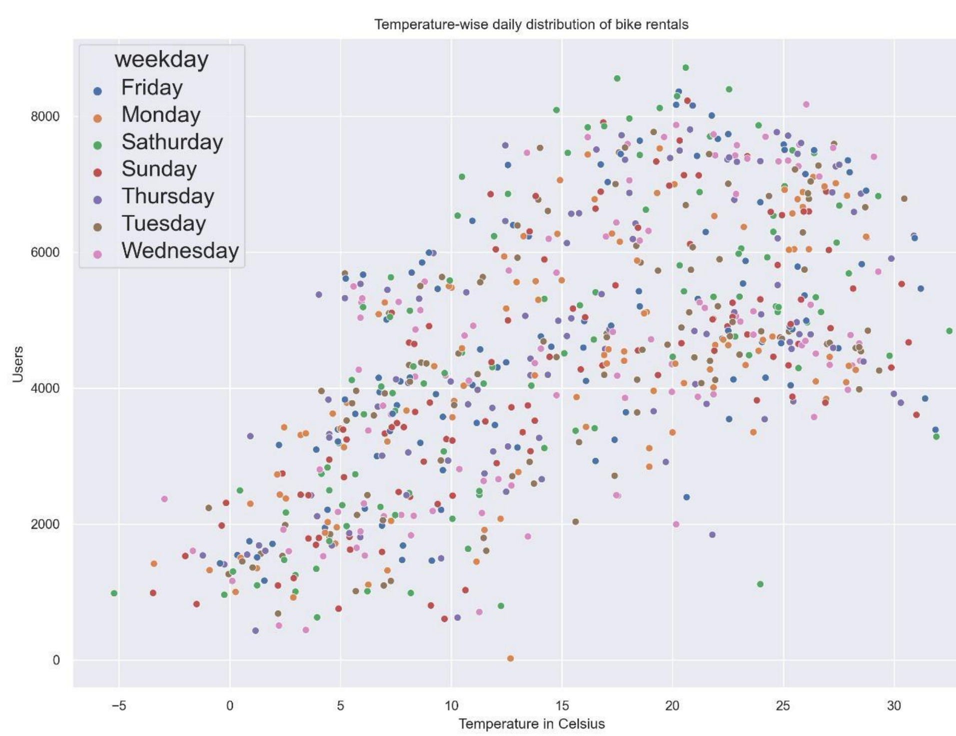
The Capital Bikeshare system has been instrumental in providing comprehensive data on bike-sharing trends, offering insights into temporal and spatial usage patterns. The system effectively captures key behaviors such as weekday commuting spikes and weekend leisure activities, highlighting its dual role in transportation and recreation. Visualizations like bar charts, line charts, and scatter plots have been used to reveal trends that can be leveraged for operational optimization and resource allocation.



Strengths: The monthly distribution bar chart provides a clear view of seasonal usage, with rentals peaking in summer and declining in winter. This seasonal trend enables stakeholders to plan resource allocation effectively during peak and off-peak months. Similarly, the bar chart showcasing station-level usage identifies high-demand stations such as "Columbus Circle/Union Station," providing actionable insights for improving bike distribution and station placement. These visualizations are valuable for guiding marketing strategies and operational decisions, as they pinpoint usage hotspots and seasonal demand fluctuations. The scatter plot showing temperature-wise bike rentals further highlights the positive correlation between warmer weather and increased bike usage, which reflects user behavior influenced by environmental factors.



Weaknesses: Despite its strengths, the scatter plot depicting temperature-wise daily distribution has notable limitations. The widespread data points makes it challenging to draw precise conclusions about the relationship between temperature and bike rentals. The absence of clear differentiation between weekdays and weekends reduces its explanatory power in identifying specific trends. Moreover, the lack of user segmentation, such as separating casual riders and members, restricts the ability to derive targeted insights that could guide marketing or operational strategies. Additionally, while the monthly and station-level visualizations are informative, they do not delve into more granular neighborhood-level dynamics or user-specific behaviors.

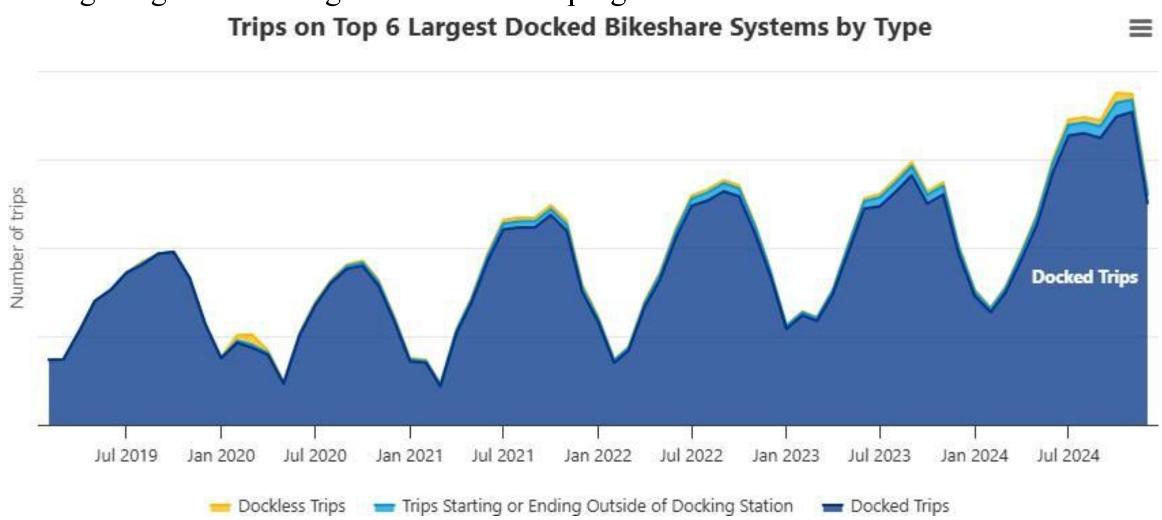


How Our Project Addresses this: By integrating interactive dashboards, our project allows users to filter data by specific stations, timeframes, and user types. The temperature scatter plot, which lacks clarity due to wide variability, is replaced with a heatmap to show the bike rentals. This enables a granular exploration of usage patterns that Capital Bikeshare visualizations cannot offer.

2. Bureau of Transportation Statistics (BTS) Bikeshare Data

The BTS aggregates data from various bike-sharing systems across the U.S., including Divvy, and presents comparative insights. Visualizations include multi-year trend graphs that depict the steady growth of bike-sharing systems, seasonal variation charts showing peak summer ridership, and pie charts breaking down membership versus casual users. This benchmarking approach contextualizes individual systems within the broader landscape of urban mobility.

Strengths: The inclusion of multi-year trends and cross-system comparisons offers a macro-level understanding of bike-sharing adoption. It provides policymakers with a strategic perspective, emphasizing the growth and significance of such programs.



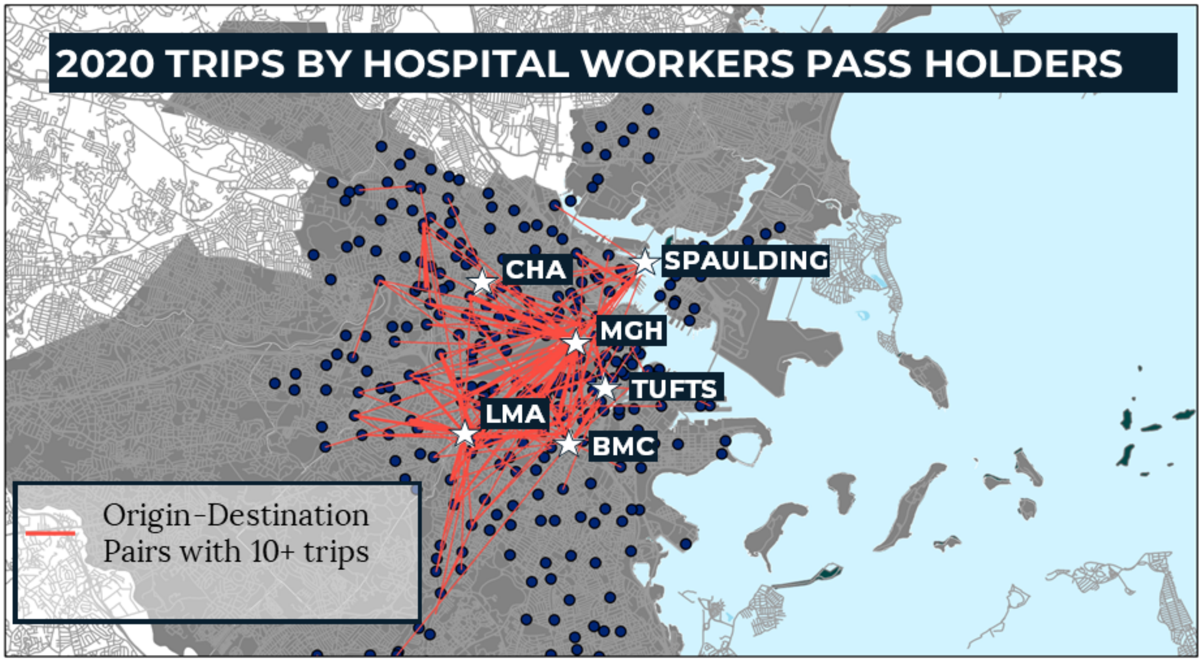
Weaknesses: The dataset's generality limits its applicability for specific cities like Chicago. Geographic insights are sparse, and station-level analysis is absent. The visuals also lack interactivity, reducing their ability to support localized decision-making.

How Our Project Addresses this: We focus specifically on Chicago's Divvy system, leveraging station-level data to create detailed heatmaps and animated maps that capture neighborhood dynamics. By benchmarking against BTS metrics, our project combines micro-level insights with macro-level trends for a comprehensive analysis.

3. Bluebikes (Boston)

The Bluebikes dataset is particularly strong in visualizing trip flows and ride durations. Chord diagrams effectively represent flows between stations, showing the most traveled routes and identifying high-traffic areas. Scatter plots highlight station popularity, while histograms display ride duration distributions for casual and member users. These visuals help stakeholders understand the efficiency and demand of different routes and stations.

Strengths: Bluebikes excels in identifying station-level trends and user behavior through engaging visuals. The chord diagrams are particularly impactful, visually representing the complexity of urban bike flows in an intuitive way.



Weaknesses: While the visualizations are insightful, they lack temporal context, making it difficult to understand how trends evolve over time. Additionally, interactivity is limited, preventing users from customizing views or filtering data by specific criteria like rider type or time of day.

How Our Project Addresses This: Our project introduces temporal dynamics by employing animated maps and interactive heatmaps. These tools not only show station-level trends but also allow users to explore how ridership patterns change over time, bridging a key gap in Bluebikes' analysis.

IV. Objectives

This project aims to deliver meaningful insights that support sustainable urban transportation by analyzing and visualizing data from Chicago's Divvy bike-sharing system. By examining detailed information on rider types, trip lengths, and station activity, the study aims to identify usage patterns, peak times, and areas with high demand. Through the use of interactive dashboards, animated maps, and dynamic heatmaps, the project seeks to equip city planners, policymakers, and stakeholders with a clearer understanding of rider behavior. Beyond highlighting Divvy's role in easing traffic and cutting carbon emissions, the report will offer recommendations for expanding the system, improving user engagement, and enhancing infrastructure, ultimately helping to foster a more sustainable transportation culture in Chicago. The specific objectives of the study are:

- **To Analyze the Divvy Bike Usage Trends in Ridership Patterns from 2021 to 2024:**
    a. Analyze how Divvy bike usage has evolved from 2021 to 2024, focusing on broader trends in ridership over time.
    b. Explore changes in ride volume over the past four years, paying attention to differences in how often members and casual riders use the service.
    c. Investigate whether seasonal changes influence ridership by examining how bike usage fluctuates across different times of the year.
- **To Analyze the Differences in Bike Usage Between Member and Casual Riders:**
    a. Explore how different types of bikes are used by casual and member riders, focusing on their frequency of use and seasonal trends.
    b. Study monthly ridership patterns to gain insights into the habits and level of dependence each group has on the bike-sharing system
- **To Analyze Ride Duration Patterns in the Divvy Bike System**
    a. Study how average ride durations have changed over the last four years, exploring monthly trends and differences between bike types while excluding any outliers that might skew the data.
    b. Compare how long casual riders and members typically use bikes, highlighting behavioral differences and how their ride durations vary month to month.
    c. Look into the average ride times for each type of bike to understand user preferences and the kinds of trips they tend to make.
- **To Analyze Ride Start and End Locations for Divvy Bikes**
  Determine which start and end locations are most commonly used for Divvy bike trips and analyze how their geographic distribution reveals patterns in how people use the service throughout Chicago.

Desired Outcomes for the Analysis of Divvy Bike Usage:

- Well-designed visualizations that highlight trends in bike usage over time, including seasonal shifts and year-to-year variations.
- Observations on how usage patterns differ between casual users and members, along with their impact on system demand.
- A more detailed look into how ride durations, bike preferences, and user behaviors have evolved.
- Evidence-based insights to guide improvements to the Divvy system, such as expanding coverage, adjusting the mix of bike types, and refining service operations.

V. About Dataset

The dataset for this project is taken from the City of Chicago's official Divvy Bike [1] usage data, which provides detailed records of bike trips taken through the Divvy bike-sharing system. While there is data available from the start of usage of these bikes in 2013, for this data analysis, the data spans from January 2021 to October 2024, offering a comprehensive view of bike usage patterns over nearly four years.

Attributes:

Attribute Name	Data Type	Description
ride Id	string	A unique identifier for each bike ride
rideable_type	Categorical (String)	Type of bike used for the ride. This typically distinguishes between different bike models or types
started_at	Date Time	Timestamp indicating when the ride started.
ended_at	Date Time	Timestamp indicating when the ride ended.
start_station_name	String	Name of the station where the ride began.
start_station_id	String	Unique identifier for the starting station
end_station_name	String	Name of the station where the ride ended.
end_station_id	String	Unique identifier for the ending station
start_lat	Float	Latitude of the starting location of the ride (for geospatial data).

start_lng	Float	Longitude of the starting location of the ride (for geospatial data).
end_lat	Float	Latitude of the ending location of the ride (for geospatial data).
end_lng	Float	Longitude of the ending location of the ride (for geospatial data).
member_casual	Categorical (String)	Type of rider based on membership status. This indicates whether the rider is a casual user or a member of the bike-sharing service

The combination of temporal, spatial, and demographic data will allow for an in-depth analysis of bike usage trends, the identification of peak times and popular routes, and insights into how Divvy's bike-sharing program is supporting sustainable transportation goals in the city.

The data was available as zip files for each month of the year. All the files were merged into a single dataset with a shape of 17,120,977 rows and 13 columns.

```
[ ] bikesData.head()
```

	ride_id	rideable_type	started_at	ended_at	start_station_name	start_station_id	end_station_name	end_station_id	start_lat	start_lng	end_lat	end_lng	member_casual
0	F96D5A74A3E41399	electric_bike	2023-01-21 20:05:42	2023-01-21 20:16:33	Lincoln Ave & Fullerton Ave	TA1309000058	Hampden Ct & Diversey Ave	202480.0	41.924074	-87.646278	41.930000	-87.640000	member
1	13CB7EB698CEDB88	classic_bike	2023-01-10 15:37:36	2023-01-10 15:46:05	Kimbark Ave & 53rd St	TA1309000037	Greenwood Ave & 47th St	TA1308000002	41.799568	-87.594747	41.809835	-87.599383	member
2	BD88A2E670661CE5	electric_bike	2023-01-02 07:51:57	2023-01-02 08:05:11	Western Ave & Lunt Ave	RP-005	Valli Produce - Evanston Plaza	599	42.008571	-87.690483	42.039742	-87.699413	casual
3	C90792D034FED968	classic_bike	2023-01-22 10:52:58	2023-01-22 11:01:44	Kimbark Ave & 53rd St	TA1309000037	Greenwood Ave & 47th St	TA1308000002	41.799568	-87.594747	41.809835	-87.599383	member
4	3397017529188E8A	classic_bike	2023-01-12 13:58:01	2023-01-12 14:13:20	Kimbark Ave & 53rd St	TA1309000037	Greenwood Ave & 47th St	TA1308000002	41.799568	-87.594747	41.809835	-87.599383	member

VI. Workflow

A. Visualization Techniques

This project utilizes a variety of visualization techniques to analyze and present the data on Divvy Bikes in Chicago. Here are the techniques used: Line Charts, Bar Charts, Scatter Plots, Heatmaps, Pie Charts, Box Plots, Rose Charts, and more

The environment and libraries used in the project include:
Google Colab: The team worked collaboratively using Google Colab, which allowed us to share code, results, and progress seamlessly.
Libraries: Pandas, NumPy, Matplotlib, Seaborn, Plotly, Datetime, Plotly Express.

B. Exploratory Data Analysis

Data Preprocessing

As the first step in the analysis, categorical variables were identified along with the unique categories present in each.

```
bikesData['rideable_type'].unique()
```

```
array(['electric_bike', 'classic_bike', 'docked_bike', 'electric_scooter'],
      dtype=object)
```

```
bikesData['member_casual'].unique()
```

```
array(['member', 'casual'], dtype=object)
```

Column names and the values were cleaned and preprocessed to eliminate unnecessary details.

```
bikesData['member_casual'] = bikesData['member_casual'].replace({'member' :'Member', 'casual':'Casual'})
bikesData['rideable_type'] = bikesData['rideable_type'].replace({'electric_bike':'Electric Bike', 'classic_bike':'Classic Bike',
```

C. Feature Engineering

New datetime attributes such as year, month, month name, hour, and ride duration were calculated to facilitate further analysis. These attributes were derived from the original timestamp columns like '*started_at*' and '*ended_at'*, enabling a deeper understanding of trends over time, such as ride patterns throughout the year, month, and day. Additionally, the ride duration attribute was created to help analyze the length of each ride. Based on the hour, the day was divided as shown in the Figure into four periods: morning, afternoon, evening, and night. This division allows for a more detailed analysis of ride patterns throughout the day, helping to identify peak hours for each category of rider and type of bike.

```
def get_time_of_day(hour):
    if 5 <= hour < 12:
        return 'Morning'
    elif 12 <= hour < 17:
        return 'Afternoon'
    elif 17 <= hour < 21:
        return 'Evening'
    else:
        return 'Night'


bikesData['time_of_day'] = bikesData['hour'].apply(get_time_of_day)
```

D. Handling Null Values

Upon examining the dataset for null values, none were found to be significant. The data, sourced directly from the website, was already in a structured and clean format. A few null values were identified in the start and end locations, which could indicate that the bike's system may have malfunctioned or the ride was incomplete before reaching its destination.

VII. Analysis

After addressing the null values, the analysis shifted to examining the categorical variables. One key variable of interest was the distinction between *Member* and *Casual* riders. Their counts were visualized in both normal and logarithmic scales to highlight differences in rider distribution and better handle the varying scales of the data.

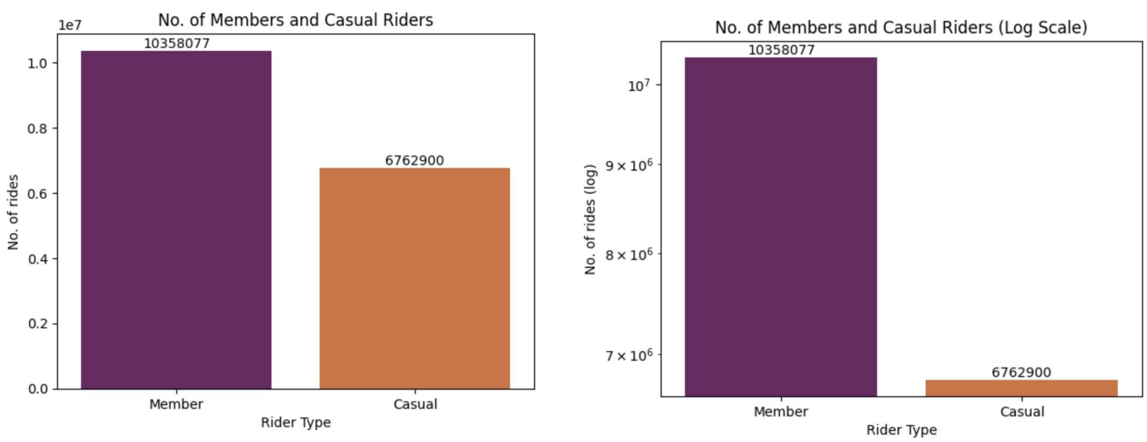


Figure 1. No. of member and causal riders in normal scale and log scale

The plots provide a clear comparison between member and casual riders of Divvy bikes in Chicago, offering valuable insights into rider behavior. Numerically, members accounted for 10,358,077 rides 60.5% of total rides in the past four years, while casual riders logged 6,762,900 rides, resulting in a difference of approximately 3.6 million rides and casual riders contributing 39.5% to the total riders.

The linear scale plot offers a better representation of the data, effectively highlighting the absolute difference between member and casual riders. Given that the counts for both groups are within comparable ranges, the linear scale is more appropriate for this visualization.

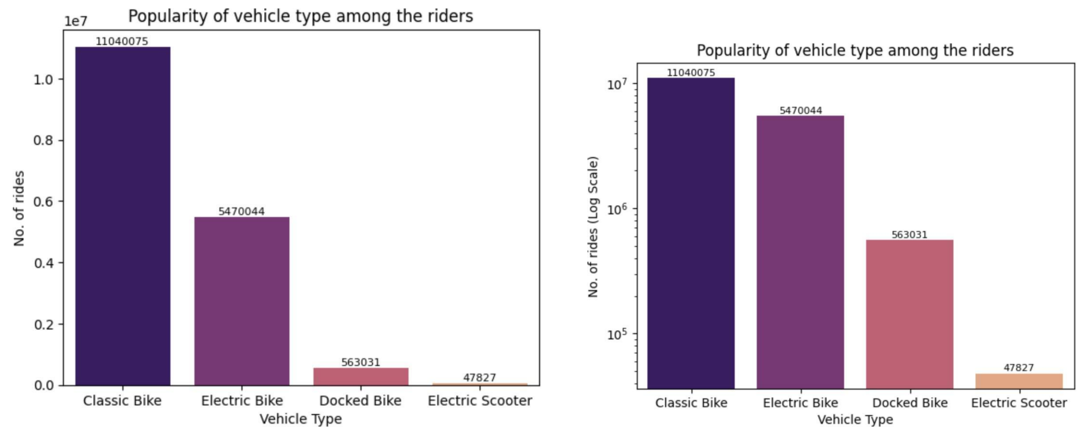


Figure 2. Distribution of rides based on the vehicle type in normal and log scale

The above bar charts show that Classic Bikes are by far the most popular, accounting for 11,040,075 rides, which represents around 64.5% of all rides. Electric Bikes follow with 5,470,044 rides, making up about 32% of the total rides. The usage of Docked Bikes is significantly lower, with only 563,031 rides, and Electric Scooters have the lowest adoption, showing just 47,827 rides, less than 0.3% of the total. The linear scale plot clearly highlights this contrast, especially emphasizing the dominance of Classic Bikes. The raw numerical difference is evident, with the other vehicle types almost appearing negligible in comparison.

The log scale visualization provides better insight into the relationships between the less popular vehicles, showing a clear separation between all four types and making the differences between Docked Bikes and Electric Scooters more apparent. While the linear scale plot is effective for showcasing the dominance of Classic Bikes, the log scale plot helps understand the relationships between less popular options.

Questions Answered

1. What are the trends in Divvy bike usage over different time periods?

Analysis of Data:

Our visual analysis of ridership trends over time and the distribution of rides throughout a month utilizes a combination of line and bar charts. These visualizations help uncover patterns and trends in ridership, offering insights into both long-term fluctuations and daily or weekly patterns.

1.1 How has the Divvy Bikes ridership changed over the past four years?

Visualization Methods:

- **Line Chart:** The ridership over the past years is illustrated using a line chart, providing a clear view of how ridership has evolved over the past four years.
- Figure 3 presents total ride counts from 2021 through October 2024, with the y-axis anchored at zero. This follows standard data visualization guidelines intended to avoid exaggerating minor fluctuations. However, the relatively modest year-to-year changes are difficult to observe in this format. As discussed in class, when the goal is to highlight small variations, it can be appropriate to adjust the y-axis. Following this guidance, Figure 3.1 focuses on the period from 2021 to 2023 and does not start the y-axis at zero. This makes the fluctuations more visible, though it comes with the trade-off of potentially overstating the magnitude of change. To ensure both clarity and transparency, we have included both versions for comparison.

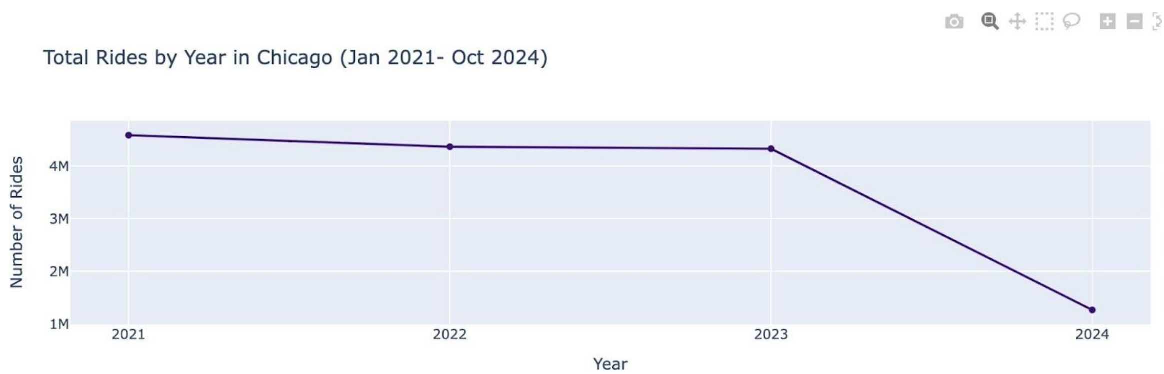


Figure 3. Variation in ride distribution from 2021 to 2024

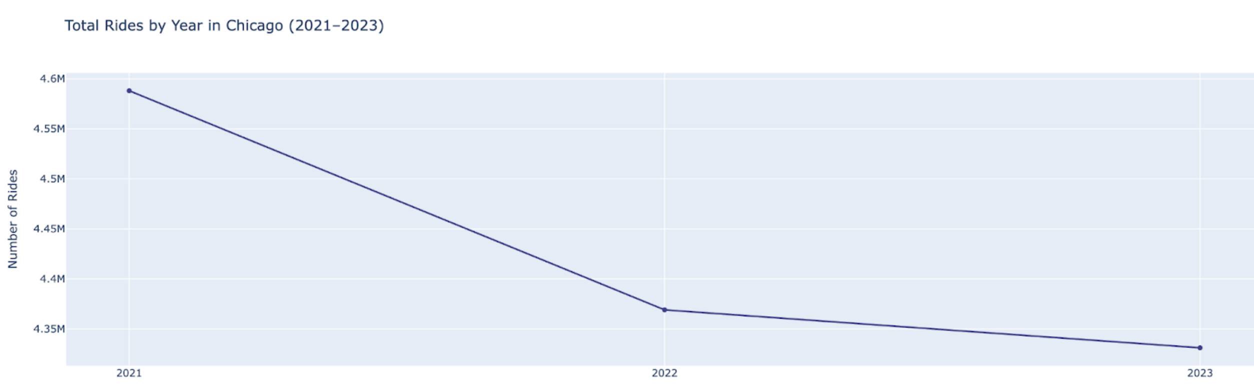


Figure 3.1 Variation in ride distribution from 2021 to 2023

1.2 Are there any seasonal and monthly variations in the ridership patterns?

- **Bubble Chart:** This chart shows the distribution of rides over the course of a month, offering a snapshot of how ridership varies from month to month. The size of each bubble corresponds to the total number of rides for each specific month, revealing peaks and valleys that are indicative of user behavior, demand fluctuations, and seasonal variations.

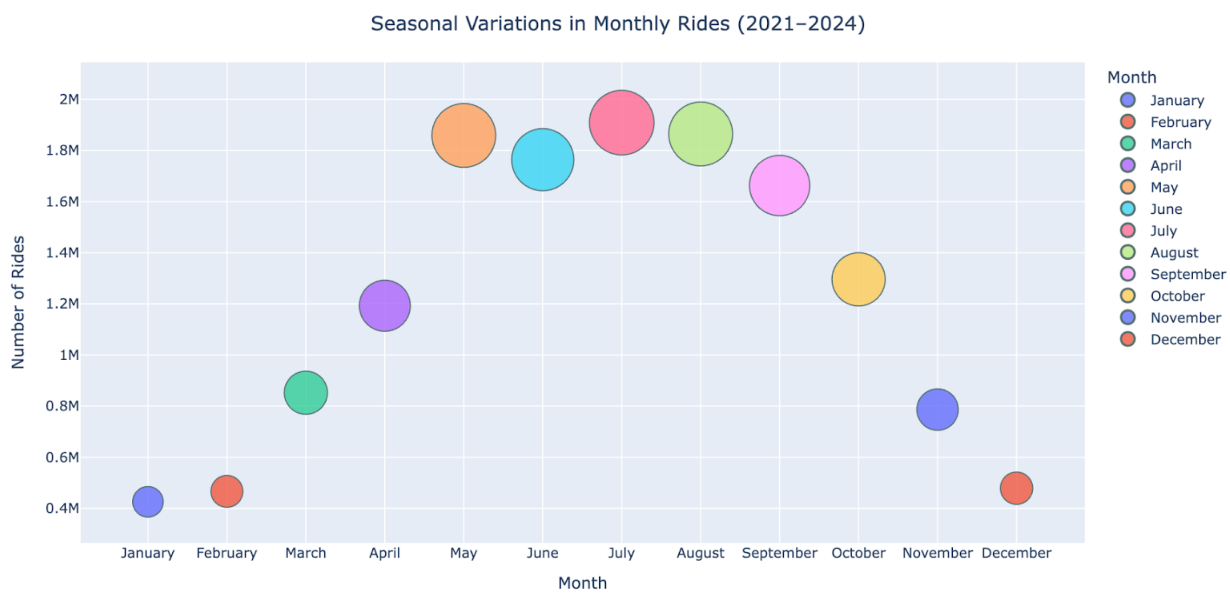


Figure 4. Monthly variations in Divvy bikes usage

Rose Chart: This Rose chart illustrates the seasonal distribution of rides over a year. Each "petal" represents one season, and its radial length, and therefore its area, is scaled to the total number of rides recorded during that period. Summer's petal extends furthest from the center, reflecting the highest demand in warm months; spring and autumn petals are of moderate length, indicating steady but lower ridership; and winter's petal is markedly shorter, revealing a pronounced decline in user activity when weather conditions are less favorable. By arranging these values around a full circle, the chart not only highlights the relative magnitudes of ridership by season but also underscores the cyclical nature of demand, with clear peaks in summer and troughs in winter that correspond to predictable shifts in user behavior and climate.

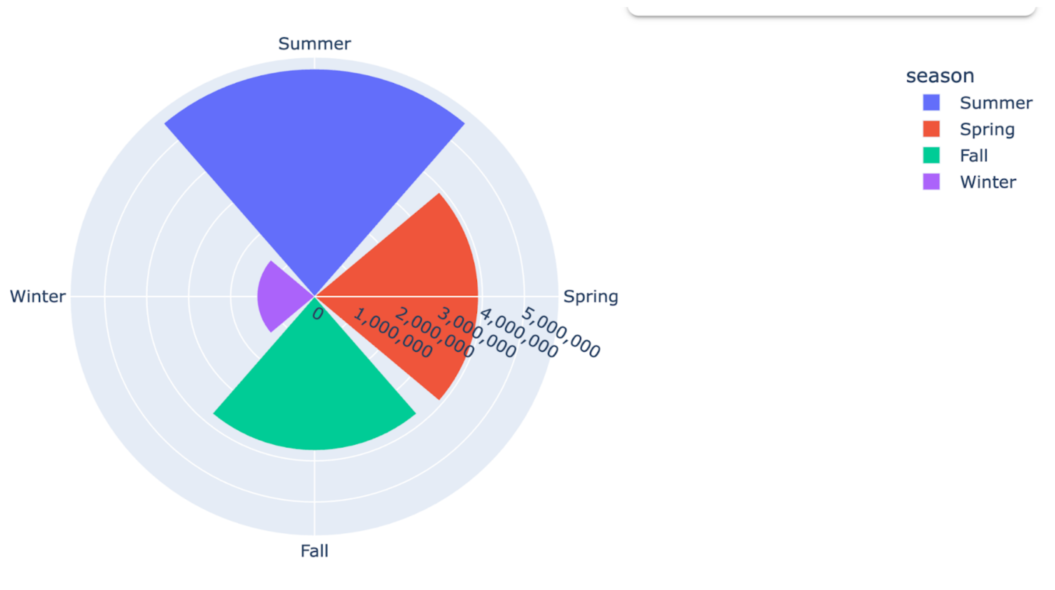


Figure 5. Seasonal variations in the Divvy bikes

1.3 How does the ride's traffic vary during different times of the day and what is the time during which there is the highest demand for divvy bikes?

Line Chart: This chart shows the distribution of rides over the course of a day, offering a snapshot of how ridership varies by hour. The height of each point corresponds to the total number of rides for that specific hour, revealing peaks during the early morning and late afternoon, which align with typical commuting times. These patterns highlight user behavior and demand fluctuations throughout the day.

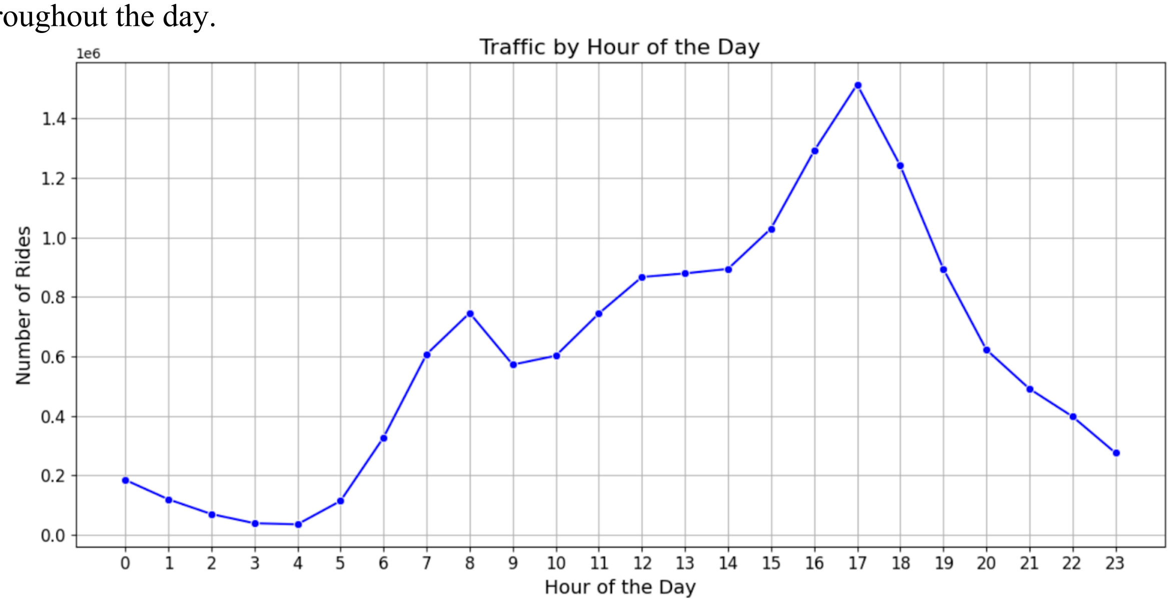


Figure 6. Hourly distribution of Divvy bikes' usage

Results:

- **Annual Ride Distribution**: A steady decline in total rides is observed from 2021 to 2024, reflecting changes in user preferences or external factors influencing ridership.
- **Seasonal Patterns**: Summer emerges as the busiest season, with over 5.5 million rides, while winter sees a significant drop, with just over 1.3 million rides.
- **Monthly Variations**: Usage peaks during the summer months (June to August) and hits its lowest in January and December, indicating strong seasonality in ridership.
- **Hourly Traffic Trends**: Divvy bike usage shows distinct peaks during weekday mornings (8 AM) and evenings (5 PM), corresponding to commute hours

Insights:

- **Seasonal Dependency**: Divvy bike usage is highly influenced by weather conditions, with warmer months driving higher ridership.
- **Commuter-Focused Usage**: The hourly traffic patterns highlight Divvy's role as a key commuting solution for urban professionals during weekdays.
- **Operational Optimization**: The clear seasonality and time-of-day preferences offer opportunities to align bike availability, maintenance schedules, and promotional efforts with user demand

2. How does bike usage differ between member riders and casual riders in terms of bike types, frequency of use, and seasonal patterns?

Data Analysis:

Our analysis compares the usage patterns of Divvy bikes by members and casual riders using visualizations like bar charts and heatmaps. These tools help uncover differences in bike type preferences, ride frequencies, and seasonal trends. The data highlights distinct behaviors in how these two groups engage with the system. Members exhibit consistent and frequent usage, while casual riders show sporadic and leisure-oriented behavior. Preferences for bike types also vary between the two groups, reflecting their differing needs and priorities.

2.1 How Does Bike Usage Frequency Differ Between Members and Casual Riders?

Donut Chart: This chart represents the number of rides categorized by rider type, casual and member. The height of each bar corresponds to the total rides for each rider type, offering a clear comparison between the two groups. The chart reveals that members significantly outnumber casual riders in overall usage, indicating stronger engagement and consistent reliance on the service by subscribed members.

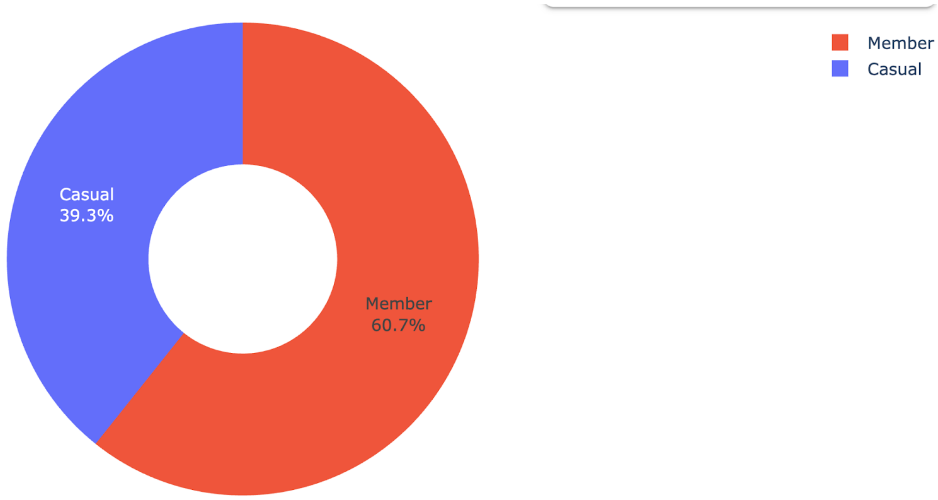


Figure 12. Donut chart representing the number of rides based on the rider type

Grouped Bar Chart: This chart illustrates the distribution of rides throughout the week, comparing the usage patterns of casual and member riders on each day. The height of each bar reflects the total number of rides for a specific day, highlighting distinct patterns. Members dominate the ridership on weekdays, indicating regular commuting behavior, while casual riders show increased activity on weekends, suggesting leisure-oriented usage. This division emphasizes the varying preferences and schedules of the two user groups.

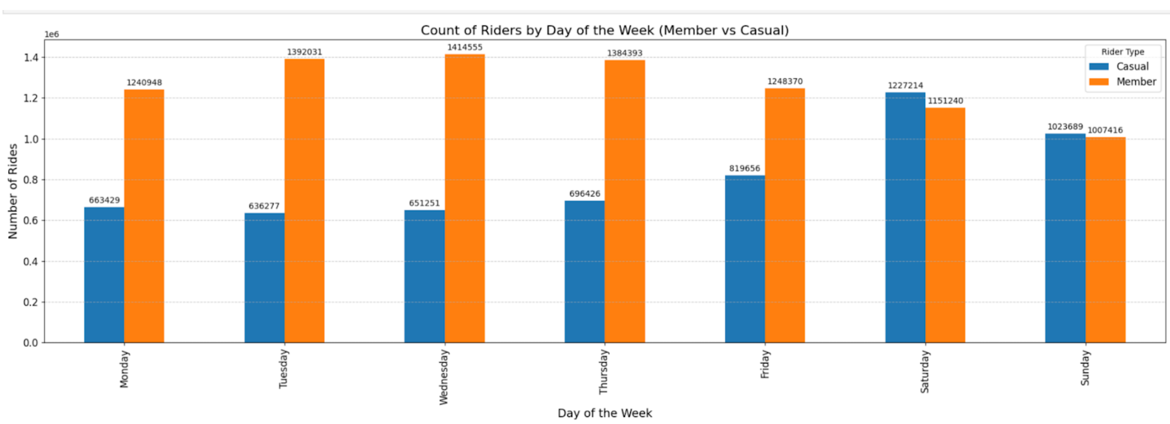


Figure 13. Ride patterns among the riders using the Divvy Bikes during different days of the week

Grouped Bar Chart: This chart illustrates monthly bike usage trends for Divvy members and casual riders. Members consistently outpace casual riders in total rides across all months, showcasing habitual usage. Casual riders exhibit strong seasonality, with a steep increase from spring (March) to summer peaks (July), followed by a sharp decline in fall and winter.

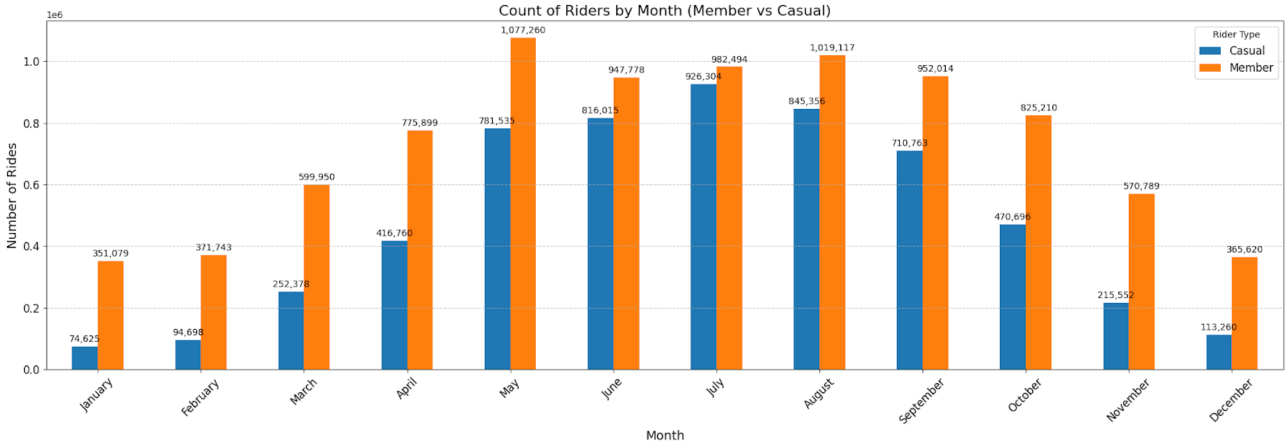


Figure 14. Ride patterns among the riders using the Divvy Bikes on different months of the year

Results

- Members significantly outnumber casual riders in total bike usage, demonstrating a stronger engagement and reliance on the system (Figure 12).
- Members consistently show higher ridership compared to casual riders on all days. However, casual riders display a relatively stronger presence on weekends, aligning with leisure-oriented usage (Figure 13).
- Members display consistent reliance on Divvy, likely for commuting or routine travel.

Insights

- The higher number of rides by members suggests that they are the core users of the bike-sharing system, with more consistent and frequent usage (Figure 12).
- Members exhibit a structured usage pattern centered around commuting times, while casual riders are more inclined towards leisure or sporadic activities, as indicated by their weekend and summer months peaks (Figures 13 and 14).

2.2 What Are the Bike Type Preferences of Members and Casual Riders?

Heatmap: This chart shows the relationship between different bike types (Classic Bike, Electric Bike, and Docked Bike) and rider types (Casual and Member), providing a clear overview of ride distribution. The intensity of the color corresponds to the total number of rides within each category, revealing patterns of preference. Lighter shades represent higher ride counts, highlighting the dominance of Classic Bikes and Electric Bikes among both Casual and Member riders, while Docked Bikes show minimal usage.

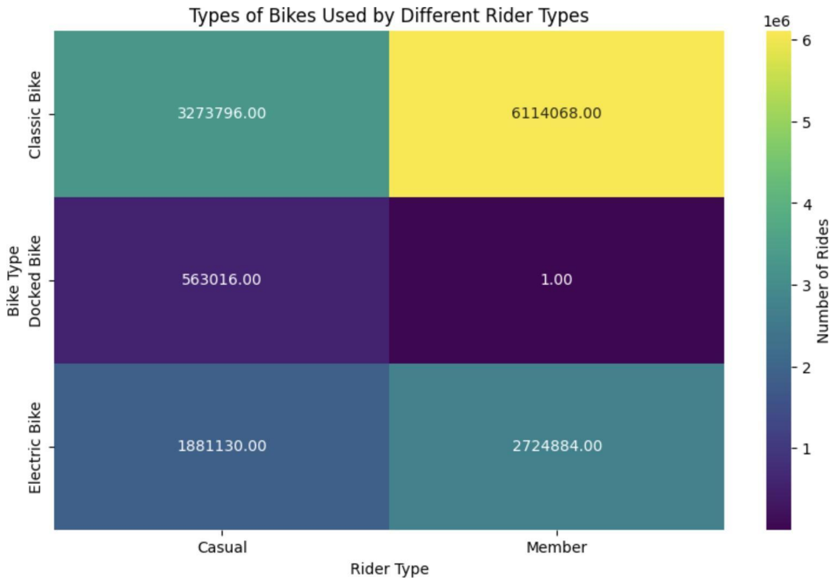


Figure 15. Heat map showing the relation between bike type and rider type

Results:

- Classic Bikes are the most widely used by both Members and Casual riders, with Members showing significantly higher usage.
- Electric Bikes are the second most preferred type, with Members utilizing them more frequently than Casual riders.
- Docked Bikes have the lowest usage, especially among Members, where utilization is almost negligible.

Insights:

- A larger fleet of Classic and Electric Bikes should be prioritized to meet user demand effectively.
- The minimal usage of Docked Bikes indicates they could be phased out or repurposed for better resource optimization.
- The higher preference for Electric Bikes among Members suggests a growing demand for faster and more convenient options, which could guide future investments in fleet upgrades.

3. Ride duration analysis

3.1 How has the average ride duration for Divvy bikes varied over the past four years, and what does this trend reveal about changes in rider behavior or system usage?

Explanation

The table shows the average ride duration for Divvy bikes from 2021 to 2024. A clear decline in average ride duration is observed over the years, with a significant drop between 2021 (21.81 minutes) and 2022 (17.09 minutes). The durations stabilize in 2023 and 2024, averaging around 15.9 minutes, suggesting a shift toward shorter trips, possibly due to changes in user behavior or operational factors like bike availability.

	year	ride_duration
0	2021.0	21.811925
1	2022.0	17.096774
2	2023.0	15.954268
3	2024.0	15.874831

Figure 21. Change in ride duration (in minutes) from 2021 to 2024.

Results

- 2021 recorded the highest average ride duration at 21.81 minutes.
- A sharp decrease is observed in 2022, with ride durations dropping to 17.09 minutes.
- Ride durations stabilize around 15.9 minutes in 2023 and 2024.
- The steady trend in recent years suggests a shift in rider behavior or optimized trip patterns.

Insights

- The decline in average ride duration may indicate more frequent short-distance trips, aligning with increased urban commuting needs.
- Operational improvements, such as better station distribution, may have contributed to shorter trips.
- Casual riders, who tend to take longer trips, may have reduced engagement relative to members.
- Further investigation into ridership patterns or external factors (e.g., weather, pricing) could explain the decline.

3.2 How do ride durations for Divvy bikes vary across different months of the year, and what patterns can be observed in monthly usage trends after removing outliers?

Explanation

The box plot displays the monthly variation in ride durations, with outliers removed to focus on typical patterns. The median ride duration remains relatively stable throughout the year, with minor variations. Summer months (July to September) show slightly higher median durations and greater variability, reflecting increased leisure and recreational trips during warmer weather.

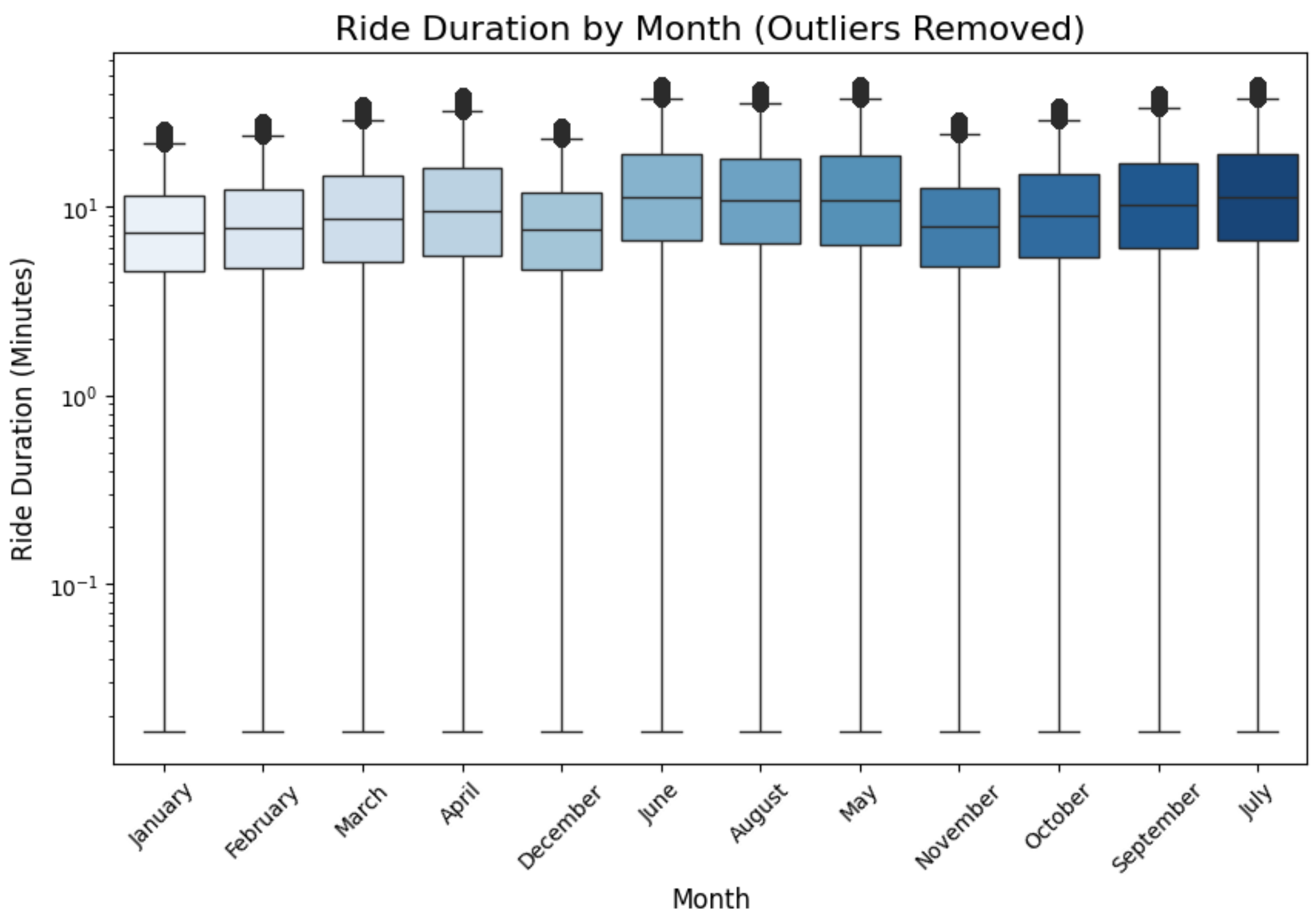


Figure 22. Variation in ride durations based on the month.

Results

- Median ride durations are consistent across most months, ranging between 10 to 15 minutes.
- Summer months (July, August, September) show slightly higher medians and wider distributions, indicating more varied trip purposes.

- Winter months (December to February) exhibit more uniform and shorter ride durations, likely due to limited outdoor activity.
- Outliers are more frequent in summer months, likely reflecting occasional long leisure trips.

Insights

- Summer months see greater variability in ride durations, driven by recreational and leisure activities.
- Winter months emphasize shorter, utilitarian trips, likely dominated by members commuting.
- Consistency in median durations suggests a stable core user base with routine travel needs.
- Seasonal trends highlight opportunities to optimize services for both short commutes and longer leisure trips during peak seasons.

3.3 How does the average ride duration differ between Divvy casual riders and members, and what does this reveal about their usage behavior?

Explanation

The bar chart compares the average ride duration between casual riders and members. Casual riders have a significantly higher average ride duration (over 25 minutes) compared to members, whose average ride duration is approximately 12 minutes. This suggests that casual riders use Divvy bikes for leisure or recreational purposes, while members use them for shorter, utilitarian trips such as commutes

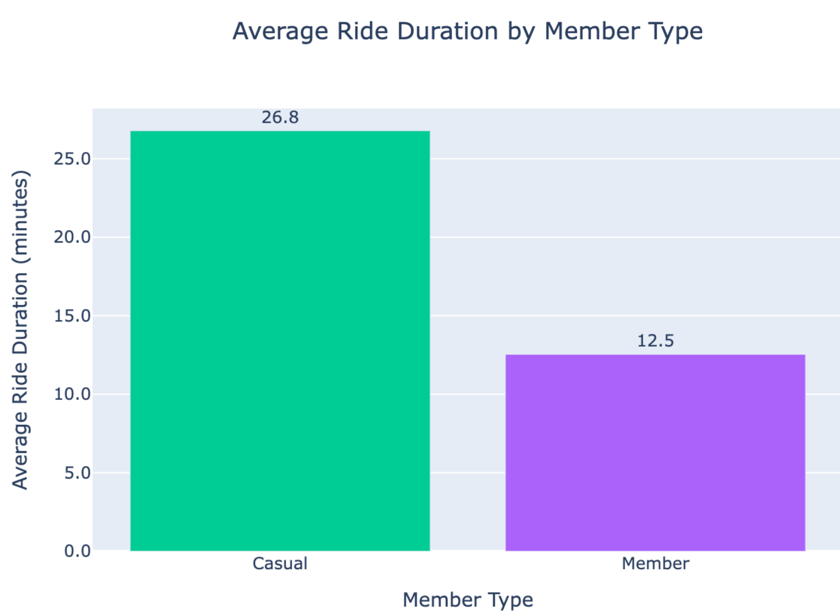


Figure 23. Average ride duration based on the rider type

Results

- Casual riders have an average ride duration of over 25 minutes.
- Members' average ride duration is significantly lower, at around 12 minutes.
- The difference highlights contrasting trip purposes between the two groups.
- Members focus on shorter, practical trips, while casual riders prioritize leisure or exploration

Insights

- Casual riders' longer durations indicate potential for promoting extended rental plans.
- Members' shorter durations emphasize Divvy's role as a commuting tool.
- Tailored strategies can target casual riders with packages for recreational trips.
- The data underscores the need to balance fleet availability for varying trip lengths

3.4 How do average ride durations for Divvy members and casual riders vary across different months, and what trends can be observed in their usage patterns throughout the year?

Explanation

The heatmap shows the average ride duration for members and casual riders by month. Casual riders consistently have longer ride durations than members, with peaks during spring and summer months (March to July). Members maintain relatively consistent average durations year-round, with minor increases during summer, suggesting routine usage unaffected by seasonal trends.

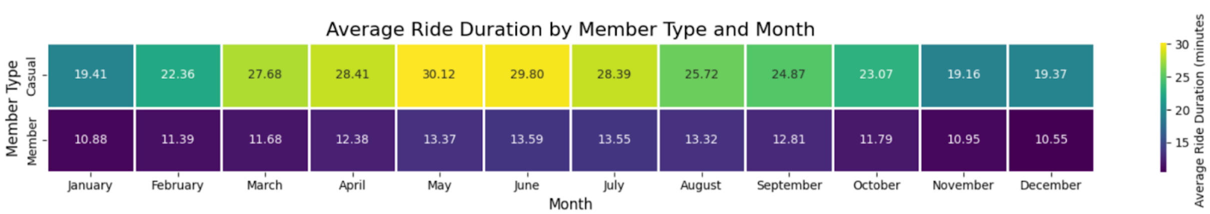


Figure 24. Variations in ride duration (in minutes) based on the rider type and the month

Results

- Casual riders’ ride durations peak in March (27.68 minutes) and June (30.12 minutes), reflecting increased leisure activity.
- Members’ ride durations remain stable, ranging between 10.5 and 13.5 minutes, with slight increases in summer months.
- Casual riders show a clear seasonal pattern, with shorter durations in winter months (around 19 minutes).
- Members exhibit steady year-round patterns, emphasizing their utilitarian use of Divvy bikes.

Insights

- Casual riders' ride durations are influenced by weather and recreational opportunities, peaking during warmer months.
- Members prioritize shorter, consistent trips, likely for commuting or errands.
- Marketing efforts can focus on promoting seasonal plans for casual riders.
- The data highlights the need for operational planning to accommodate seasonal demand fluctuations.

3.5 How does the average ride duration vary across different bike types in the Divvy system, and what does this reveal about user preferences and trip purposes?

The heatmap compares the average ride duration for classic, docked, and electric bikes. Docked bikes have the longest average ride duration (66.33 minutes), suggesting occasional or unique usage scenarios. Classic bikes have an average duration of 17.61 minutes, indicating their suitability for regular, short trips. Electric bikes have the shortest average duration (13.31 minutes), reflecting their efficiency and preference for quick commutes.

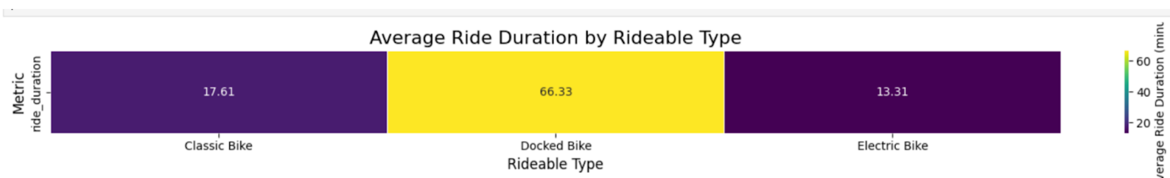


Figure 25. Average ride duration based on the vehicle type

Results

- Docked bikes are used for the longest average trips (66.33 minutes), indicating occasional or special-purpose usage.
- Classic bikes average 17.61 minutes, highlighting their dominance for regular, mid-length trips.
- Electric bikes are used for the shortest trips, averaging 13.31 minutes, emphasizing their efficiency for quick travel.

Insights

- Docked bikes likely cater to users with less frequent but longer rides, requiring further analysis of their deployment.
- Classic bikes remain the primary choice for regular users with moderate trip durations.
- The short duration of electric bike trips underscores their value for fast, efficient commutes, aligning with urban mobility trends.
- Optimizing the distribution of electric bikes in high-demand areas could enhance system efficiency and user satisfaction

4. Ride analysis by location

4.1 Which part of Chicago is widely used by different riders?

Density Map: This map visualizes the concentration of start locations for member riders, providing a clear indication of areas with high bike usage. The intensity of color corresponds to the density of start locations, with red regions representing the highest activity. The map highlights key hotspots, likely reflecting popular residential, commercial, or transit areas that are central to member riders' commuting or activity patterns.

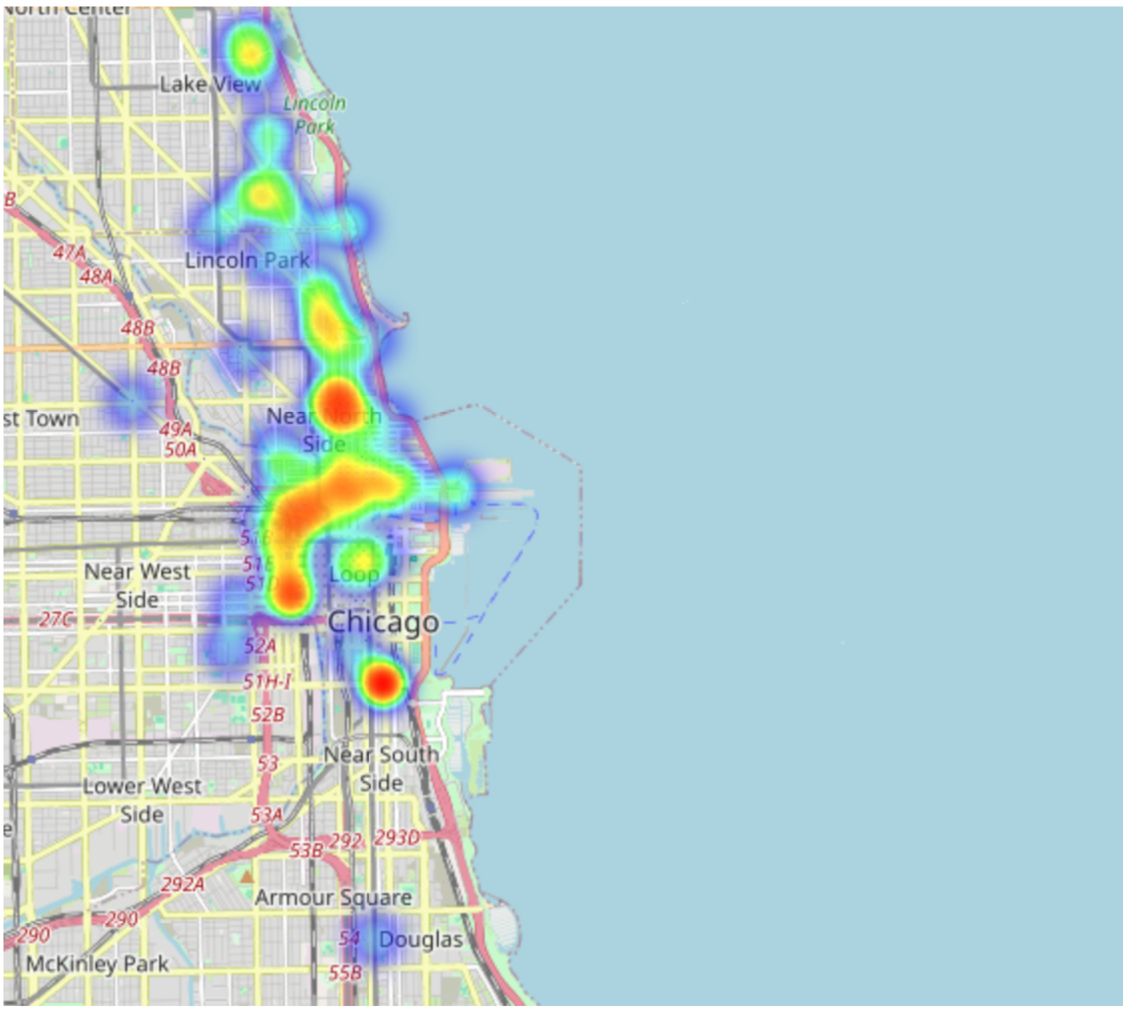


Figure 26. Density map showing the start locations of member riders.

Density Map: This map visualizes the concentration of start locations for casual riders, highlighting areas with high bike usage. The intensity of color corresponds to the density of start locations, with red regions representing the highest activity. The map emphasizes hotspots, likely associated with popular tourist attractions, recreational areas, or leisure destinations frequented by casual riders.

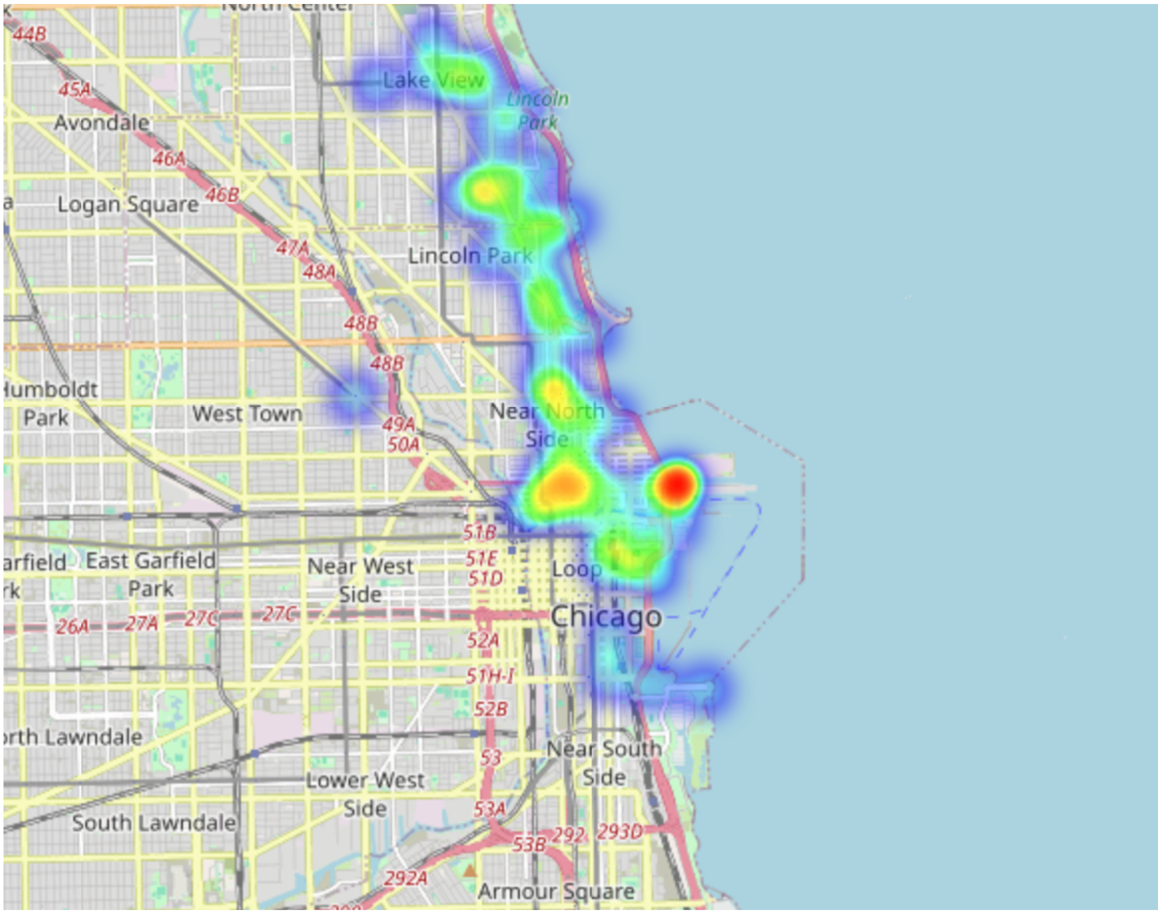


Figure 27. Density map showing the start locations of casual riders.

Results:

- The density map for member riders (Figure 26) indicates that their bike usage is concentrated in central and densely populated areas of the city, including downtown and major commuting hubs.
- Casual riders (Figure 27) exhibit a higher concentration of start locations near tourist attractions, recreational zones, and waterfront areas, suggesting a preference for leisure-oriented trips.

Insights:

- Member riders' concentration in downtown and central areas reflects their primary usage for commuting or routine activities, making these areas crucial for maintaining bike availability and station capacity (Figure 26).
- Casual riders are more inclined towards leisure and recreational activities, as highlighted by their start locations near popular tourist and waterfront areas. This suggests that bike-sharing systems could focus on enhancing user experience in these zones, such as providing clear wayfinding and adequate bike availability during weekends and peak tourist seasons (Figure 27).
- Understanding these differing usage patterns can help operators optimize bike distribution and service offerings to meet the distinct needs of members and casual riders.

4.2 What are the most popular start locations and end locations?

Heatmap: This heatmap illustrates the number of rides between the top 10 start and end station pairs, offering insight into the most frequently traveled routes. The color intensity reflects the volume of rides, with darker shades indicating higher counts. Notably, routes such as Streeter Dr & Grand Ave to DuSable Lake Shore Dr & Monroe St, and Michigan Ave & Oak St to Millennium Park, show particularly high usage, suggesting these corridors serve as major connectors for riders. These high-traffic routes likely align with key commercial, tourist, or recreational zones, emphasizing their role in the overall bike share network.

Figure 28. Ride counts of top 10 locations in Chicago

Results:

- The heatmap of station-to-station ride counts shows that the busiest connection is between Streeter Dr & Grand Ave and Millennium Park, with 37,821 rides, indicating a high-demand corridor likely influenced by its proximity to tourist attractions and central activity zones.
- Other high-volume links include DuSable Lake Shore Dr & Monroe St to Clark St & Elm St, and

Michigan Ave & Oak St to Millennium Park, each showing over 15,000 rides.
- In contrast, the least busy link among the top 10 stations is from Kingsbury St & Kinzie St to Theater on the Lake, with only 1 recorded ride, highlighting limited direct traffic between these specific points.

Insights:

- The high volume of rides between key downtown and lakefront stations suggests strong usage for both commuting and tourism-related activities. These connections likely reflect well-traveled routes between popular destinations and key access points to the city's core.
- Streeter Dr & Grand Ave consistently appears as a major hub, indicating its strategic location for both member and casual riders. Enhancing infrastructure and bike availability around this station could improve overall system performance.
- The wide range in ride counts across station pairs suggests varying levels of connectivity and rider preference. Low-volume links, while part of the top 10 stations, may benefit from targeted outreach or infrastructure improvements to boost utilization.
- Understanding which station pairs drive the most traffic can inform operational decisions such as rebalancing, maintenance prioritization, and future station siting.

Conclusion

The Divvy bike-sharing system serves as a vital component of urban mobility in Chicago, catering to both commuting and recreational needs. Members predominantly use Divvy bikes for short, consistent trips, showcasing its role as a practical transportation option for daily commutes.

In contrast, casual riders engage with the system for longer, more leisure-oriented trips, particularly during the summer months. Seasonal trends play a significant role, with summer seeing peak ridership and winter experiencing sharp declines, especially among casual riders, who are more influenced by weather conditions.

Classic bikes are the most widely used across both rider types, reflecting their reliability and versatility for regular trips. Electric bikes, while used less overall, show promise for quick, efficient commutes, especially among casual riders. Docked bikes cater to niche scenarios with longer average trip durations. Spatial analysis highlights downtown Chicago and the lakefront as key hubs for both trip starts and ends, driven by commuting and tourist activities. However, peripheral areas show lower usage, indicating opportunities for system expansion and increased accessibility.

To improve operational efficiency and user satisfaction, Divvy should focus on seasonal planning, ensuring adequate bike availability during peak demand periods and introducing promotions or incentives during off-peak seasons. Expanding coverage to low-density areas can attract new users and balance system usage. Tailored offerings, such as leisure packages for casual riders and optimized placement of electric bikes, can address diverse user needs. By implementing these strategies, Divvy can enhance its role in promoting sustainable transportation while maximizing system utilization and user engagement.

Future Work

Future work could begin by enriching Divvy trip logs with weather conditions, special-event schedules, and transit data to explain sudden ridership swings and feed ARIMA or LSTM models for precise station-level demand forecasting and smarter bike rebalancing. Clustering riders by trip duration, distance, and time-of-day preferences would support targeted incentives, such as off-peak discounts or gamified challenges, to boost usage when it's needed most. Applying graph-analysis techniques like community detection and betweenness centrality can uncover critical corridors and underserved neighborhoods, guiding where to add or expand stations. Overlaying these insights with demographic and socioeconomic layers will reveal equity gaps and inform outreach or subsidized membership programs. Finally, packaging everything into an interactive dashboard would turn our visualizations into a dynamic decision-support system, empowering planners and operators to optimize the bike-share network in real time.

References

[1] https://divvy-tripdata.s3.amazonaws.com/index.html
[2] https://github.com/davidellavalle/Capital-Bikeshare
[3] https://data.bts.gov/stories/s/Bikeshare-and-E-scooter-Table-of-Contents/fseh-ipec/
[4] https://bluebikes.com/system-data
[5] OpenAI, "ChatGPT," o4-mini, May 3, 2025. [Online]. Available: https://chat.openai.com/