Behavioral Analytics for Strategic Mobility: Uncovering Trends in Divvy’s Bike-Share Usage

**Module/Subject Title:** Google Data Analytics

**By:** Hoang Hai Long Do

**Date:** May 2025

***Abstract:***

*This report represents the culmination of a comprehensive data analysis case study, conducted to demons*trate applied proficiency in key tools and methodologies acquired through the Google Data Analytics Professional Certificate on *Coursera. Focusing on Divvy’s bike-share system in Chicago, the project showcases end-to-end analytical workflow, encompassing data cleaning, transformation, and visualization. Using a blend of* ***Spreadsheets, SQL, R (tidyverse), and Power BI****, I systematically uncovered user behavior patterns, operational bottlenecks, and membership conversion insights. Beyond technical execution, the report highlights my capacity for* ***critical thinking, structured problem-solving****, and* ***data-informed storytelling****—positioning data as both a diagnostic lens and a strategic asset for real-world decision-making.*

### **Executive Summary**

This report presents a comprehensive analysis of Divvy’s bike-share usage patterns in March and April 2025, aimed at uncovering strategic levers to boost membership conversion and improve system efficiency. Utilizing real-world data from Divvy’s public trip repository, the study draws on over 669,000 records and applies rigorous preprocessing through SQL and R, followed by visual analytics in Power BI.

Key behavioral trends were identified: members dominate short trips (under 25 minutes), showing routine, commute-driven usage; casual users concentrate activity on weekends and avoid short rides, likely due to fee deterrents. Classic bikes are used predominantly for shorter, localized trips—especially in the city’s center and outer fringes—while electric bikes are used broadly across all geographies and durations.

Geospatial analysis revealed that member users are more widely distributed, with rides extending into residential and peripheral zones, whereas casual usage remains concentrated in central areas. Zones like Lincoln Park and West Town, where both user types overlap, present high-potential targets for conversion campaigns.

Based on these findings, the report recommends three strategic actions: 1) deploy targeted conversion marketing in shared-use hotspots, 2) strengthen classic bike access in the city’s outer corridors, and 3) implement off-peak incentives to normalize casual user behavior toward membership-like engagement. These recommendations aim to align operational planning with user behavior, increase subscription rates, and optimize resource deployment.

### **Business Task Statement:**

To analyze Divvy’s bike-share usage data in order to uncover behavioral patterns, temporal trends, and location-based dynamics that differentiate casual riders from members. The ultimate goal is to identify actionable factors that drive user conversion, enabling Divvy to boost membership rates, optimize operational deployment, and support strategic decision-making across marketing and infrastructure teams.

### **Data Source:**

The dataset utilized in this analysis originates from the Coursera Google Data Analytics Capstone Case Study, titled **“How Does a Bike-Share Navigate Speedy Success?”**, and is publicly hosted via Divvy’s open data portal:<https://divvy-tripdata.s3.amazonaws.com/index.html>. The database is maintained and updated on a monthly cadence, reflecting real-world operational data from Divvy’s bike-sharing system in Chicago.

For this report, two monthly datasets were selected—**March 2025** and **April 2025**—to provide a robust temporal snapshot of rider behavior and system usage. The raw data comprised **298,155 records for March** and **371,341 records for April**, each structured in a tabular format with the following key variables:

* ride\_id: Unique identifier for each ride
* rideable\_type: Type of bike used (e.g., classic, docked, electric)
* started\_at, ended\_at: Timestamp of ride initiation and termination
* start\_station\_name, end\_station\_name: Names of the stations where the ride began and ended
* start\_station\_id, end\_station\_id: Station identifiers
* start\_lat, start\_lng, end\_lat, end\_lng: GPS coordinates for trip origin and destination
* member\_casual: Rider classification (casual or subscribed member)

This foundational dataset enabled a multi-dimensional analysis of rider behavior, spatial demand, and usage segmentation across different user types.

### **Data manipulation documentation:**

To ensure analytical integrity, both raw datasets—March and April 2025—underwent a preprocessing workflow using **SQL for database-level filtering** and **R (Tidyverse + Lubridate)** for refined data transformation.

**Key Processing Steps:**

1. **Critical Null Removal**:  
    All records with missing essential values (ride\_id, started\_at, ended\_at, start\_lat, start\_lng, end\_lat, end\_lng) were excluded to maintain ride completeness and mapping accuracy.
2. **Categorical Validation**:  
    Unique values from key classification fields (rideable\_type, member\_casual) were inspected to detect anomalies or inconsistencies. No invalid entries were identified.
3. **Text Normalization**:  
    Whitespace inconsistencies in station and rider classification fields were eliminated using string trimming functions.
4. **Duplicate Removal**:  
    Duplicated ride\_ids, if any, were resolved by retaining only the first occurrence based on chronological order (started\_at timestamp).
5. **New Feature Engineering**:  
    Two derived variables were created to enrich the dataset:  
   * duration\_minutes: The time difference between started\_at and ended\_at, calculated in minutes.
   * ride\_weekday: The day of the week extracted from started\_at, to facilitate temporal trend analysis.
6. **Anomaly Filtering**:  
    Implausible rides (e.g., duration of 0 minutes with identical start and end coordinates) were removed to prevent analytical distortion.

This cleaning framework ensured both **data validity** and **modeling readiness**, enabling robust and meaningful visual exploration in Power BI. The accompanying code in SQL and R documents these steps explicitly for reproducibility and audit transparency.

### **Analysis Findings:**

#### **Summary**

The data reveals a clear segmentation in user behavior between casual and member riders. Members overwhelmingly dominate short trips under 25 minutes, particularly on weekdays, suggesting structured, utility-driven patterns such as commuting. Casual users, on the other hand, cluster around weekend and leisure-oriented use, with low engagement in very short rides—likely deterred by pricing mechanics like unlock fees.

Rideable type plays a role in duration: classic bikes are favored for shorter urban trips, especially in central and outer nodes, while electric bikes display broader versatility across both time and geography. Spatially, member users are far more dispersed, reaching into suburban zones, while casual users remain centered in downtown areas.

These patterns point to targeted opportunities: converting casual users in shared-use hotspots, expanding infrastructure to meet routine member needs, and redesigning short-trip incentives to habituate casual riders toward member-like behavior.

#### **Duration distribution by User Type and Rideable Type**

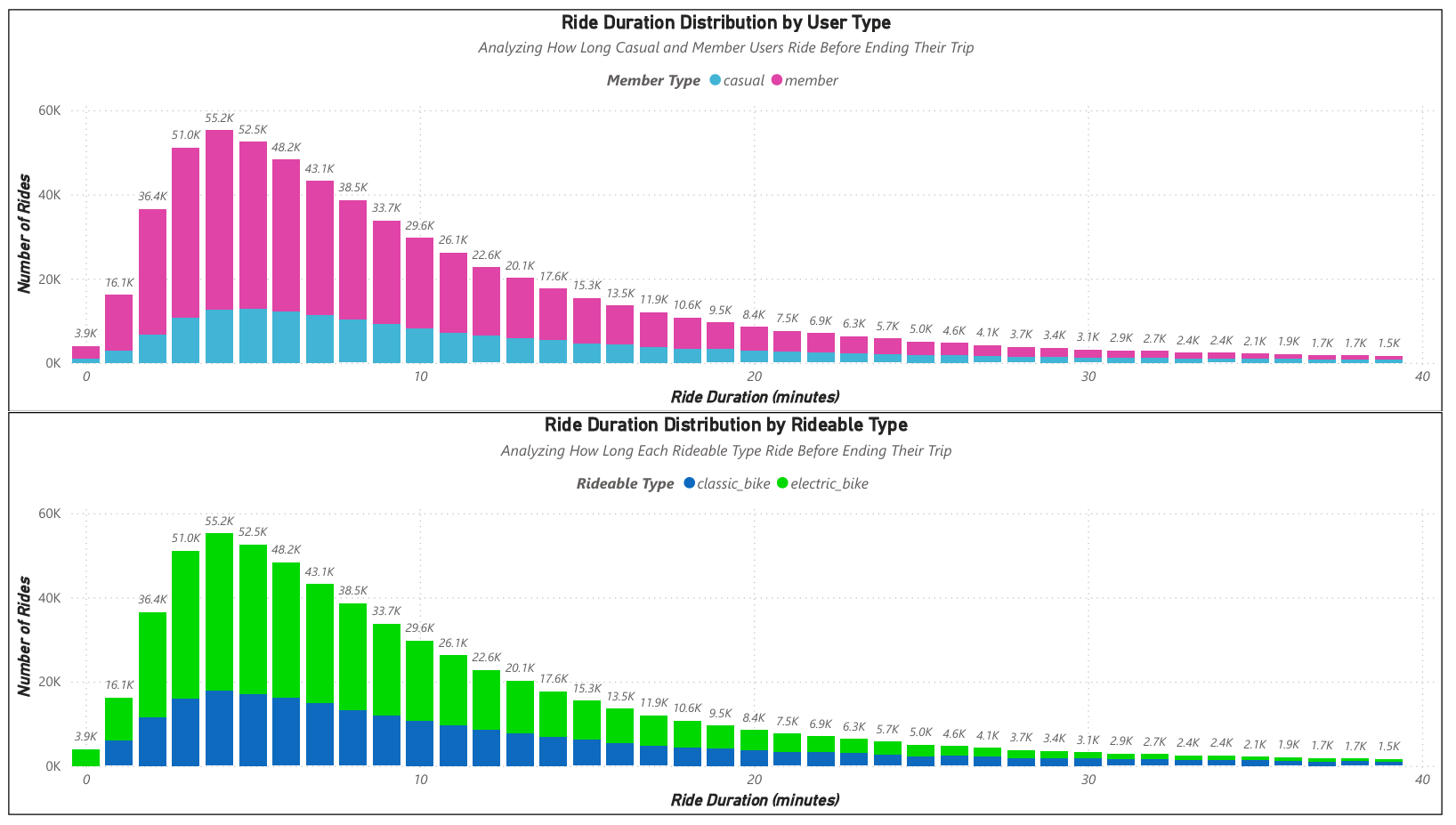


Figure 1. Duration by User Type and Rideable Type

* **Short Trips Are Dominated by Members -**Rides under 25 minutes are largely made by members, showing consistent and frequent usage, likely for commuting or short errands.
* **Casual Users Rarely Take Very Short Rides -**Casual usage below 3 minutes is low—likely due to unlock fees making brief rides uneconomical.
* **Classic Bikes Are Used More for Short Rides -**Classic bikes dominate rides under 20 minutes, showing that electric bikes are used more broadly, even in longer durations.
* **Ride Behavior Converges Over Time -**After 25 minutes, the difference between member and casual users, and between bike types, evens out—suggesting shared behavior in longer, possibly leisure rides.

#### **Weekday Distribution by User Type and Rideable Type**

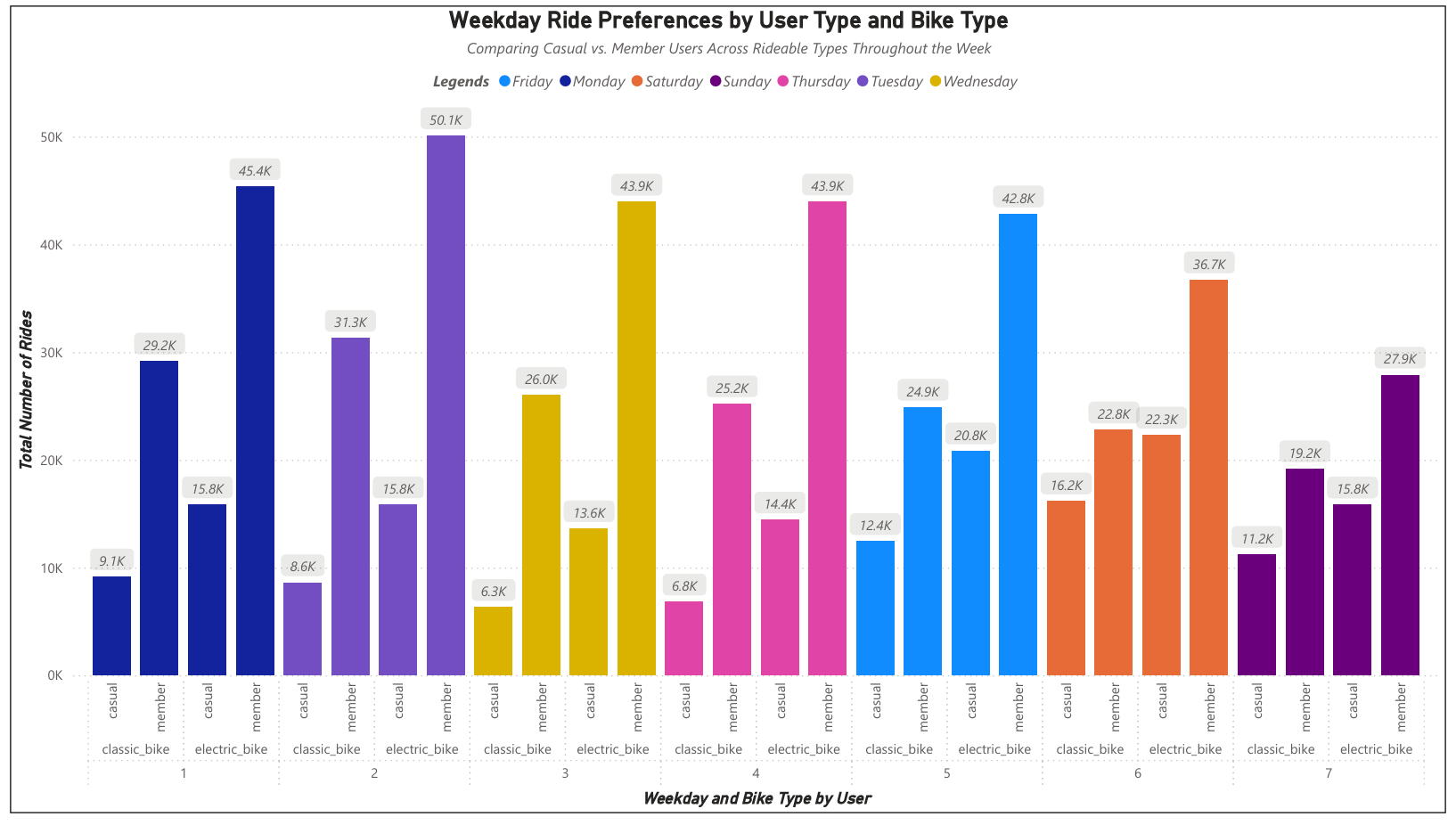


Figure 2. Weekday Distribution by User Type and Rideable Type

* **Members Ride More on Weekdays**Member rides peak on Thursday and Tuesday. Usage is steady across the workweek, showing routine behavior.
* **Casual Use Peaks on Weekends**Casual riders spike on Saturday and Sunday. This shows leisure-based intent, not daily utility.
* **Electric Bikes Are Preferred by Members**Across all days, members use electric bikes more than casual users—implying higher willingness to pay or efficiency needs.

#### **Geographical Distribution by Rideable Type**

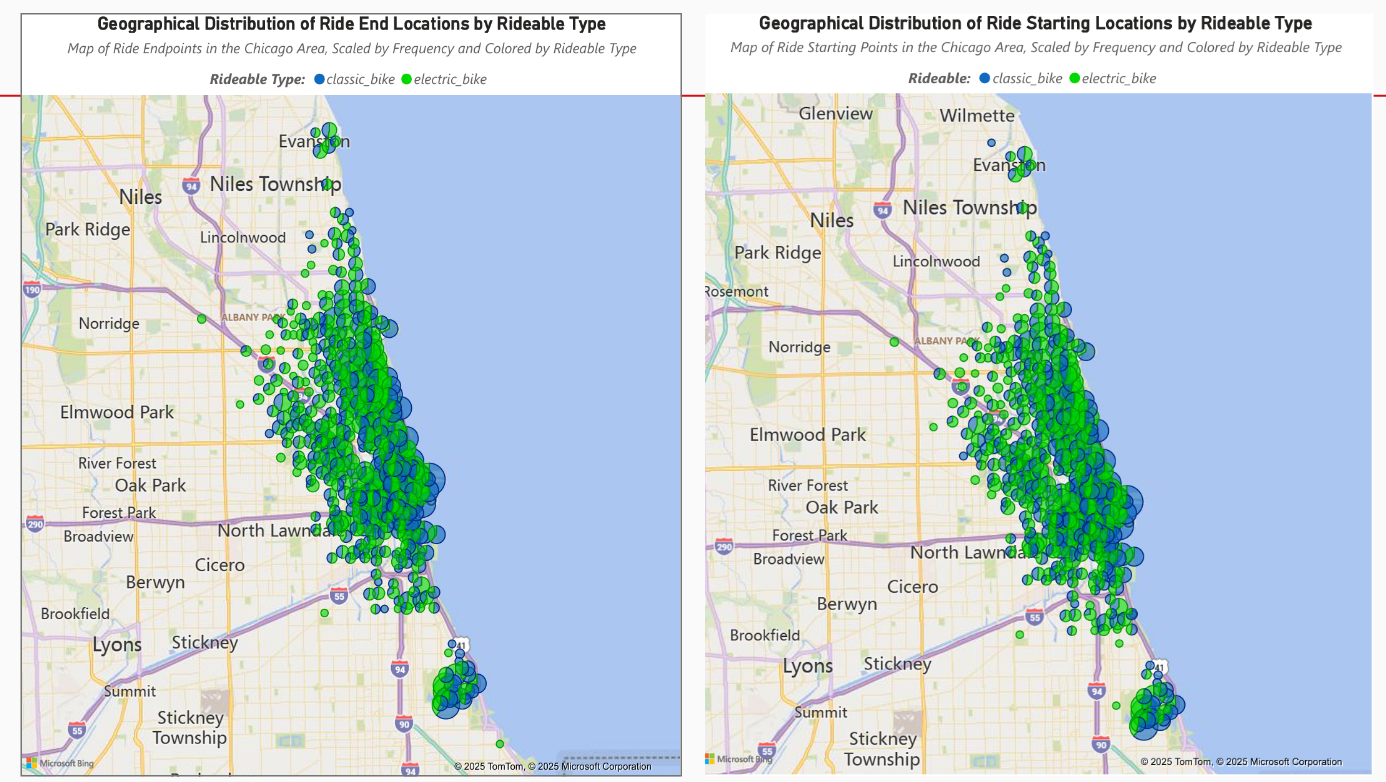


Figure 3. Geographical Distribution by Rideable Type, scaled by Frequency

* **Classic Bike Usage in Outer Areas**Classic bike usage increases in the outer upper and lower regions of the map, likely reflecting designated station areas or transit nodes.
* **Classic Bike Cluster Downtown**There is dense clustering in the city center, suggesting classic bikes are favored where infrastructure is concentrated.
* **Electric Bikes Dominate Everywhere**From core to outskirts, electric bikes see broad use—suggesting universal appeal across user types and regions.

#### **Geographical Distribution by Member Type**

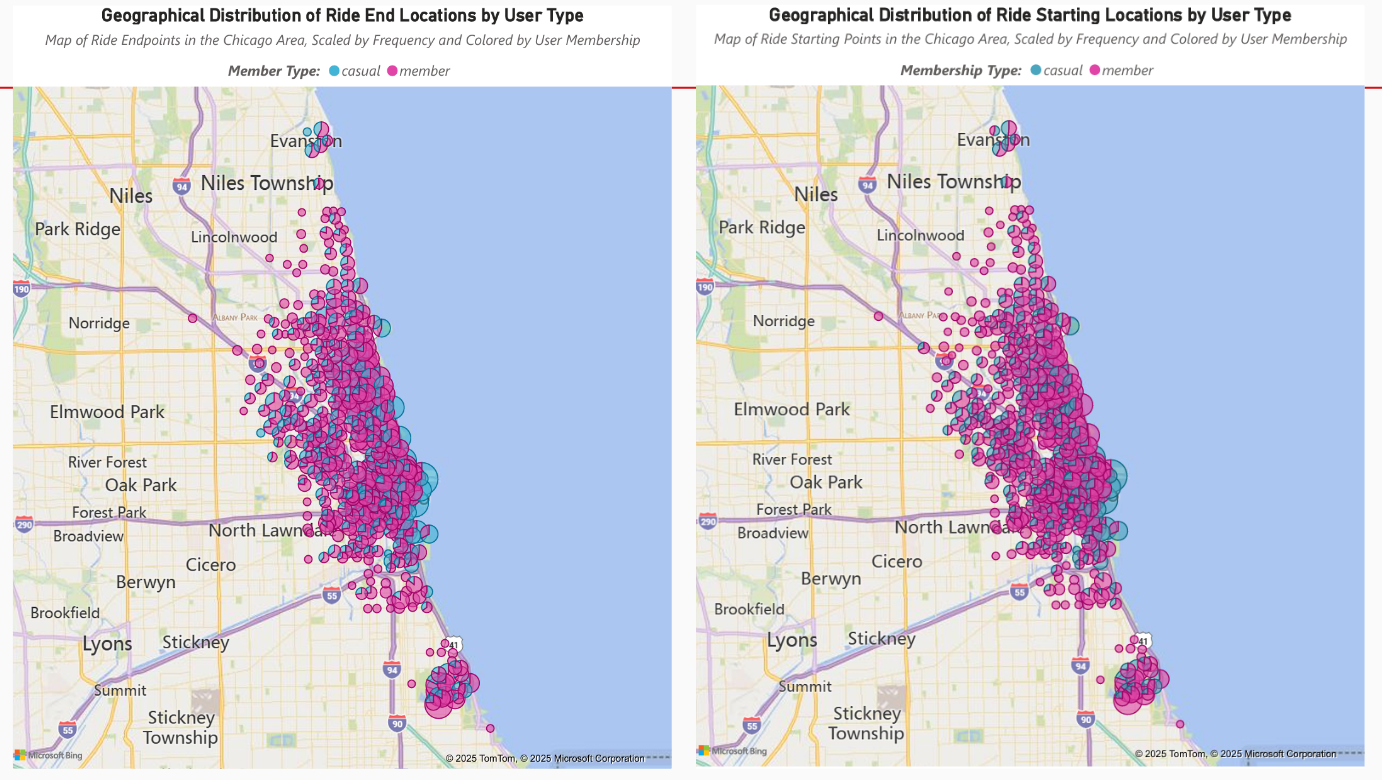


Figure 4. Geographical Distribution by Member Type, scaled by Frequency

* **Lack of Casual Penetration in Outskirts**Casual users are rarely observed beyond central districts. This may reflect a lack of awareness, infrastructure, or perceived value for single-ride users in those areas.
* **Potential Conversion Opportunity**Shared hotspots—especially in mid-density zones like West Town and Lincoln Park—show overlapping use by both groups. These could serve as prime locations for marketing aimed at casual-to-member conversion.
* **Peripheral Reach – Member Users**Member rides extend into the far north and south zones, especially near residential belts. This implies members use Divvy for commuting or daily errands beyond the city’s commercial heart.

### **Recommendations:**

1. **Deploy Conversion Campaigns in Overlapping Zones**Target areas like Lincoln Park and West Town—where both user types overlap—for promotional offers (e.g., first-month free membership, QR code prompts at stations) to convert casual users riding frequently in high-traffic areas.
2. **Expand Classic Bike Infrastructure in Outer Zones**Given the consistent use of classic bikes in the city’s periphery, expand or maintain stations there to support routine member commutes and build predictability in access.
3. **Incentivize Off-Peak Short Rides for Casuals**Introduce off-peak discounts or ride bundles to reduce the impact of unlock fees on short trips. This nudges casual users into member-like behavior and builds habitual use patterns.

### **Conclusion:**

The analysis exposes distinct patterns in how casual users and members engage with Divvy’s system. Members ride more, ride shorter, and ride further—indicating utility. Casuals ride less, on weekends, and stay central—showing leisure. These insights present direct levers: target overlapping zones for conversion, strengthen classic bike access in outer areas, and use pricing to shift casual behavior. With focused action, Divvy can drive higher engagement, operational efficiency, and sustained growth.