Cyclistic Bike-Share Analysis: Strategies to Convert Casual Riders into Annual Members

1. Introduction

1.1 Project Objective

Cyclistic is a bike-sharing service operating in the city of Chicago, offering both annual subscription plans and pay-per-ride or daily pass options. The primary goal of this project is to analyze how **casual riders** (occasional users) utilize the service compared to **annual members**, in order to define marketing strategies that can convert more casual riders into annual subscribers.

Through this analysis, we aim to:

- Identify **behavioral differences** (average ride duration, usage times and days, bike type preferences, seasonality) between casual riders and members.
- Provide **concrete insights** to develop targeted campaigns that increase annual subscriptions.
- Propose **operational recommendations** and improvement strategies based on data-driven insights.

1.2 Business Question and Stakeholders

The key business question to address is:
“How do usage behaviors differ between casual riders and annual members, and how can these differences be leveraged to convert more casual riders into annual members?”

Primary Stakeholders

🔹 **Marketing Director, Lily Moreno** → Interested in understanding which promotional strategies can be used to increase annual memberships.
🔹 **Cyclistic Marketing Analytics Team** → Responsible for collecting, analyzing, and interpreting data to support strategic marketing decisions.
🔹 **Cyclistic Executive Team** → In charge of approving recommendations and allocating the necessary resources for implementation.

1.3 Report Structure

This report is divided into the following main sections:

- **Data and Methodology**
  Describes the data sources, cleaning procedures, tools used, and key steps in dataset preparation.
- **Exploratory Analysis**
  Explores the results obtained from SQL queries and/or conducted analyses, with a focus on differences between casual riders and members in terms of ride duration, usage times, bike type preferences, and seasonality.
- **Key Insights Summary**
  Summarizes the main findings from the analysis, highlighting the key differences between casual riders and members.
- **Recommendations and Next Steps**
  Proposes actionable strategies to improve the conversion rate of casual riders to annual members, along with suggestions for further analyses and key performance indicators (KPIs) to monitor.
- **Limitations and Future Developments**
  Outlines the limitations encountered (e.g., anonymous data, lack of cost/revenue information) and suggests additional analyses and dataset integrations to refine the strategy.

Note: The **Appendices** contain complete SQL queries and detailed tables, while additional visualizations are compiled in a separate section titled **"Visualizations"**. This ensures that the core report remains focused on insights and recommendations, while technical details and supplementary graphics are placed in their respective support sections.

2. Data and Methodology

2.1 Sources and Licenses

The data used for this analysis comes from Cyclistic's **public dataset**, provided by **Motivate International Inc.** under an open license. The dataset is **anonymous** and does not contain personally identifiable information (PII), covering a **12-month period for the year 2024**.

- **Source:** https://divvy-tripdata.s3.amazonaws.com/index.html
- **Format:** 12 **.csv** files, one for each month (January 2024 - December 2024).
- **License:** Open license allowing data usage for analytical purposes.

The **public availability** and **absence of PII** ensure compliance with **privacy regulations** and **data protection policies**.

2.2 Dataset Description (12 .csv Files)

Each monthly file contains the following fields:

- **ride_id**: Unique identifier for each ride.
- **started_at** and **ended_at**: Timestamps indicating the start and end times of the ride.
- **start_station_name** and **end_station_name**: Names of the departure and arrival stations.
- **start_station_id** and **end_station_id**: Numeric IDs of the stations (sometimes missing).
- **rideable_type**: Type of bike used (classic, electric, etc.).
- **member_casual**: Indicates whether the user is an annual member ("member") or an occasional user ("casual").

Overall, the 12 files contain **several million records**. The **.csv files** were organized into **separate folders** (raw data and cleaned data) and subsequently uploaded into a **Google BigQuery project** to facilitate data management and analysis.

2.3 Data Cleaning and Transformation

2.3.1 Merging and Verifying Monthly Tables

To simplify the management of the **entire year's data**, all monthly tables uploaded to **BigQuery** were merged into a **single final table**. After merging, **record count checks and sample checks** were performed to verify the **absence of duplicates or major anomalies**.

(The detailed SQL query is provided in Appendix A.)

2.3.2 Adding Calculated Columns

Two additional columns were created to **enhance the analysis phase**:

- **ride_length**: Difference between **ended_at** and **started_at**, expressed in **seconds**.
- **day_of_week**: Day of the week (**Monday, Tuesday, etc.**) extracted from **started_at**.

(For details on the SQL queries used, refer to Appendix A.)

2.3.3 Identifying and Removing Anomalies

Records with **null or negative ride_length** were identified, representing **less than 1% of the total dataset**. Given the **small size of this portion**, they were removed to **simplify analysis and prevent distortions**.

This operation was also conducted using **SQL queries in BigQuery**, described in **Appendix A**.

At the end of this process, the **cleaned_data** dataset (without anomalous values) was deemed **ready for analysis**.

2.4 Tools Used

To perform data extraction, cleaning, and analysis, the following tools were utilized:

- **Google Cloud Storage**: Used to upload large files (**over 100 MB**) and make them accessible to **BigQuery**.
- **Google BigQuery**: Managed the **monthly tables**, unified them into a **single dataset**, and executed **cleaning and aggregation queries**.
- **SQL (Query Language)**: Chosen for its **scalability and efficiency** in handling large datasets, as well as its **native integration** with BigQuery.
- **Google Sheets**: Used in a later stage to create **quick visualizations** based on results extracted from **BigQuery**.

With the **cleaned_data** dataset now prepared, we can proceed to the **exploratory analysis**, presented in the next section, to determine **how casual riders differ from annual members** and which strategies can encourage conversion to subscription plans.

3. Exploratory Analysis

Introduction

At this stage, the primary objective is to answer the key question assigned by Marketing Director **Lily Moreno**:
How do annual members and casual riders use Cyclistic bikes differently?

Although the company is also interested in understanding **why casual riders would purchase an annual membership** and **how digital channels can be leveraged to influence them**, this report will primarily focus on **comparing the two user segments (casual vs. members)**. The insights obtained will help shape **marketing strategies** aimed at increasing annual subscriptions and provide **operational recommendations** where applicable.

The following subsections analyze different aspects of bike usage:

- **General overview** (ride volume and total ride duration)
- **Day of the week and hourly analysis**
- **Seasonality (monthly trends)**
- **Bike type preference (classic, electric, scooter)**
- **Ride duration distribution**
- **Median analysis and outliers**

At the end of this section, a summary of **key insights** will be provided, followed by **recommendations** that, while primarily addressing the first business question, may also offer valuable insights for other strategic considerations.

3.1 General Overview

This initial overview focuses on:

- **Total rides by user type**
- **Total ride duration**
- **Average ride duration for casual riders vs. members**

Key Observations

Ride Volume and Total Duration

- **Annual members** take significantly more trips than casual riders (**~3.7 million vs. ~2.1 million**).
- However, **casual riders accumulate a longer total ride duration** (**~901,727 hours vs. ~789,149 hours**), confirming that their trips are, on average, longer.

Frequency vs. Ride Length

- **Members tend to take short and frequent trips**, likely for **commuting purposes**.

- **Casual riders take fewer trips but with significantly longer ride durations**, suggesting a **more recreational or touristic usage**.

Business Implications

- **Extended-duration plans or hybrid packages** could attract casual riders, who tend to use bikes for longer periods.
- **Commuter members**, already engaged with **short and frequent trips**, benefit from cost-effective pricing for daily use.

*(See details and queries in **Appendix B.1 – "General Overview"**)*

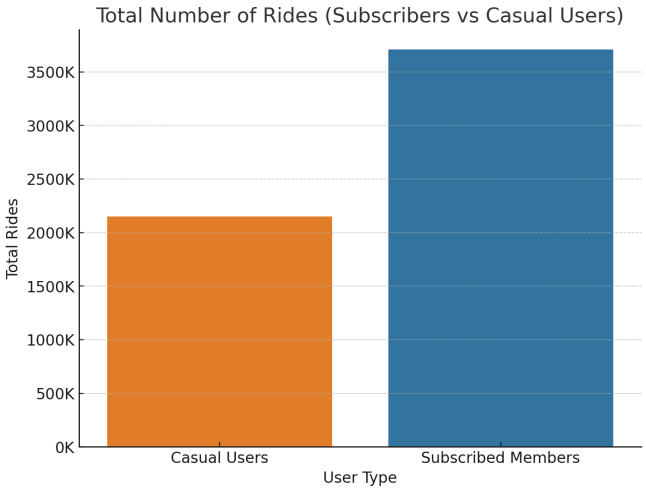


Figure 1 – Total Number of Rides (Members vs. Casual Riders)

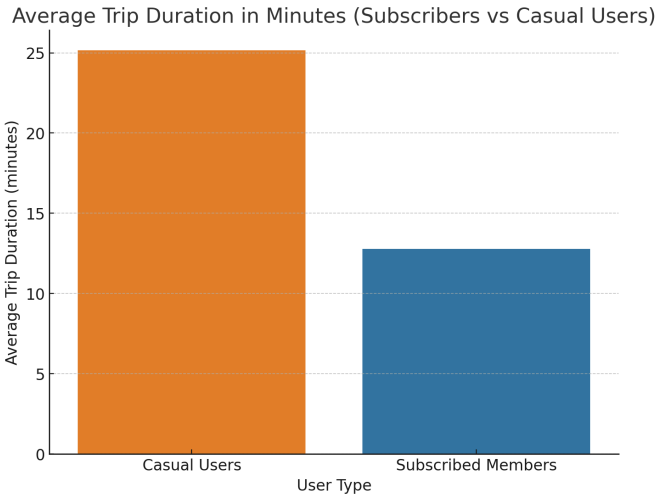


Figure 2 – Average Ride Duration (Members vs. Casual Riders)

3.2 Daily and Hourly Analysis

This section examines **service usage based on the day of the week and hourly distribution**, distinguishing between the two user segments.

3.2.1 Weekly Usage Distribution

Members

- Peak usage occurs on **weekdays (Tuesday, Wednesday, Thursday)**, aligning with **work-related or commuting purposes**.
- **Shorter average ride duration (~12–14 minutes).**

Casual Riders

- Prefer **weekends (Saturday and Sunday)**, with **significantly longer rides (~28–30 minutes).**
- Usage is primarily **recreational or tourism-driven**.

Implications

- **Weekend plans or special discounts** could help convert casual riders who are more active on weekends.
- **Increase bike availability in tourist areas** during weekends to accommodate higher demand.

3.2.2 Hourly Usage Distribution

Members

- **Two distinct peaks:**
  - **Morning (7–9 AM)** and **late afternoon (4–6 PM)**, consistent with **commuting patterns**.
- **Stable average ride duration (~11–13 minutes).**

Casual Riders

- **Gradual increase in usage throughout the morning, peaking between 2 and 5 PM.**
- **Average ride duration is often more than double that of members.**

Implications

- **Afternoon promotions:** Membership discounts between **2–5 PM** to encourage casual riders to subscribe.
- **Fleet optimization:** Allocate additional bikes to **central and tourist areas during peak afternoon hours**.

*(See details and queries in **Appendix B.2 – "Daily and Hourly Analysis"**)*

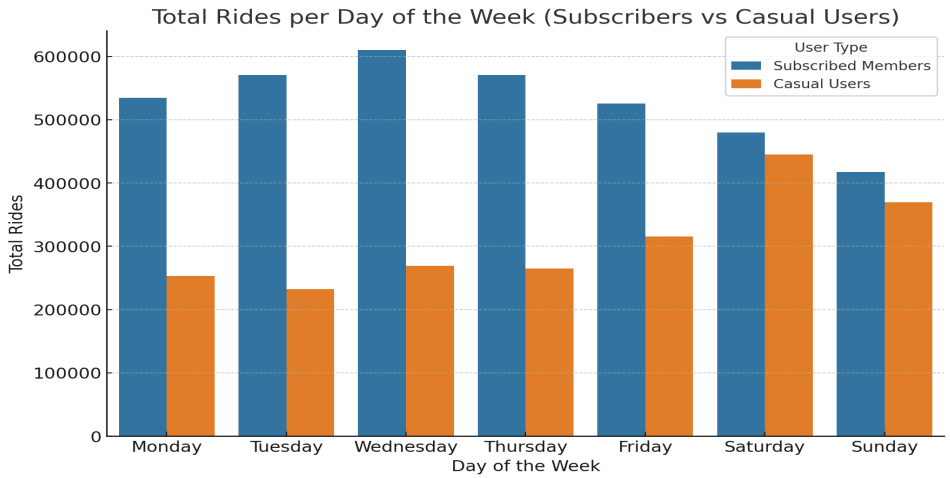


Figure 3 - Total Number of Rides by Day of the Week (Members vs. Casual Riders)

3.3 Seasonality (Monthly Analysis)

This section examines monthly service usage to identify potential seasonal variations.

Seasonal Peaks for Casual Riders

- Rides increase significantly from May through September, reaching a high of about 346k in September.
- Their average ride duration remains consistently higher than that of members, pointing to more recreational or tourist-focused usage during warmer months.

Usage Trend Among Members

- Members maintain higher ride counts overall, with a peak in September (~474k rides) and a moderate decline in colder months.
- Although their usage also rises in summer, the variation is less extreme than that of casual riders.

Implications

- **Summer promotions**: Offer seasonal memberships or extended passes to attract casual riders when they are most likely to use the service.
- **Winter incentives**: Introduce loyalty programs or discounts in colder months to retain casual riders' interest and encourage members to continue their subscriptions.

*(See details and queries in **Appendix B.3 – "Monthly Analysis"**)*

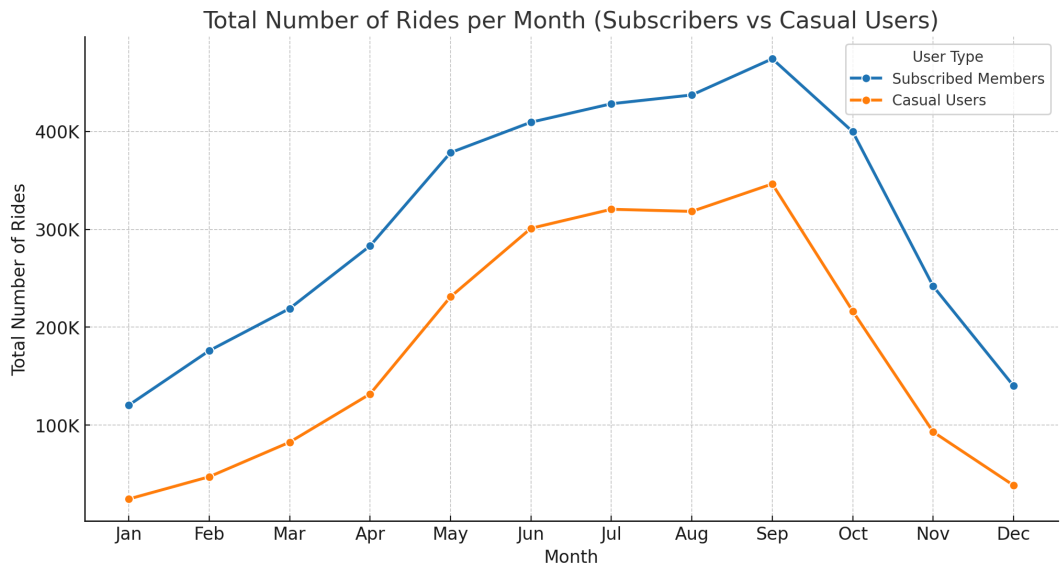


Figure 4 - Total Number of Rides by Month (Members vs. Casual Riders)

3.4 Bike Type (Classic vs. Electric vs. Scooter)

This section examines the bike type preferences of members and casual riders across classic bikes (**classic_bike**), electric bikes (**electric_bike**), and scooters (**electric_scooter**).

Electric Bikes

- Most popular choice for both user segments.
- Members take frequent but shorter trips (**~11 min average**).
- Casual riders have slightly longer rides (**~14–15 min**), especially on weekends and in the afternoon.

Classic Bikes

- Casual riders take significantly longer rides (**up to ~38 min**), mainly for leisure/tourism, peaking on weekends and midday hours.
- Members maintain steady usage (**14–15 min**), mostly for weekday commuting (**peaks at 7–9 AM and 4–6 PM**).

Scooters

- Least used vehicle type, especially among members.
- Casual riders use them slightly more (**~11–12 min**), primarily on weekends and during leisure hours.

Business Implications

- **Special offers on e-bikes**: A dedicated membership plan for e-bike users, particularly casual riders on weekends, could be beneficial.
- **Tourist bundles with classic bikes**: Casual riders who take longer rides (~38 min) could be attracted by partnerships with hotels or travel agencies.
- **Scooter fleet assessment**: Low usage among members suggests evaluating fleet size or launching targeted promotions (e.g., free trials or discounts).

*(See details and queries in **Appendix B.4 – "Bike Type Analysis"**)*

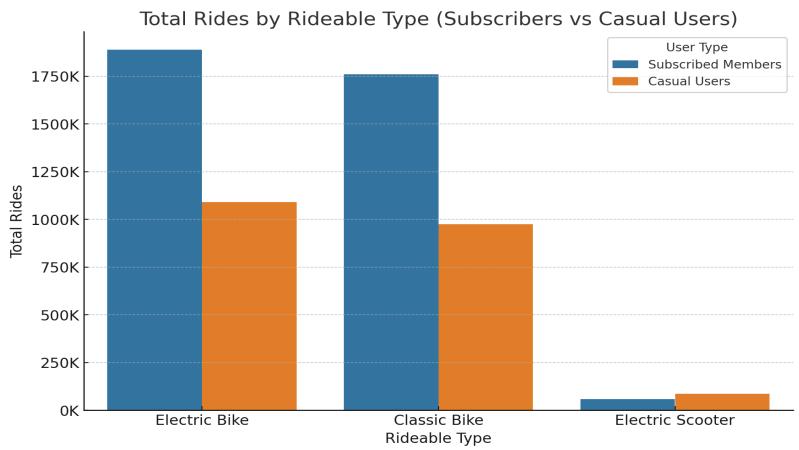


Figure 5 - Total Number of Rides by Vehicle Type (Members vs. Casual Riders)

3.5 Ride Duration Distribution (Time Range Analysis)

This section categorizes trips into predefined duration ranges (**<5 min, 5–15 min, 15–30 min, 30–60 min, >60 min**) to highlight how members and casual riders use the service in terms of travel time.

Key Findings

Members

- Over **75% of rides last less than 15 minutes**, confirming a strong preference for short commuting trips.
- Only a small fraction (**>60 min trips, ~25,969 rides, less than 1% of total trips**) consists of very long rides, likely due to exceptional circumstances or occasional usage.

Casual Riders

- More than **37% of trips exceed 15 minutes**, and nearly **130,000 rides last over an hour**.
- This distribution suggests a **recreational or tourism-oriented usage**, with fewer time constraints compared to commuters.

Business Implications

- **Extended ride-time memberships**
    - Subscription plans covering **rides up to 45–60 minutes** without extra charges could appeal to casual riders who typically take longer trips.
- **Progressive pricing model**
    - Implementing **higher fees beyond a set threshold (e.g., 30 or 45 min) for non-members** could encourage more casual riders to subscribe, given their frequent ride durations exceeding 30 minutes.

*(See details and queries in **Appendix B.5 – "Ride Duration Distribution"**)*

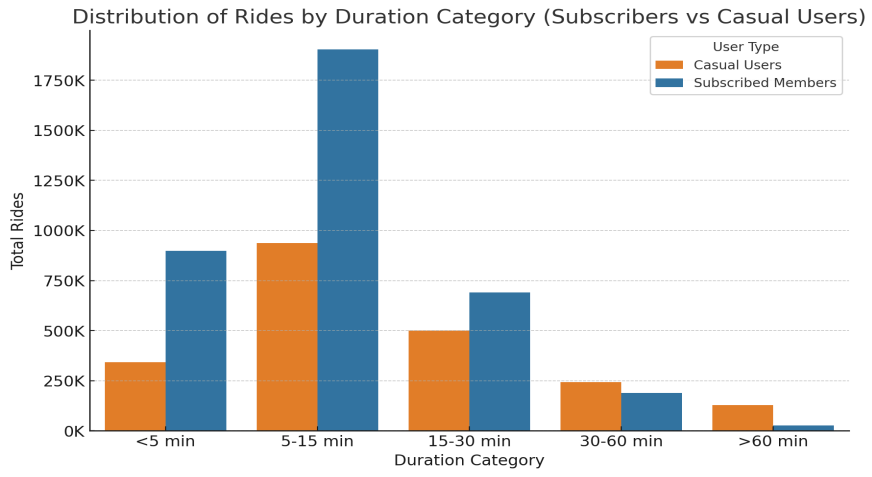


Figure 6 - Trip Distribution by Duration Range (Members vs. Casual Riders)

3.6 In-Depth Analysis: Median, Percentiles, and Outliers

To complete the ride duration analysis, the impact of potential outliers was assessed, particularly **exceptionally short (<1 min) or long (>18 hours) trips**. These values can distort the mean, making it less representative.

- **Trips >18 hours:** Represent **~0.17% of total rides**. While some may be due to technical errors, others could stem from legitimate multi-hour day passes. Given their minimal impact, they are noted but not excluded.
- **Trips <1 minute:** Account for **~2.2% of total rides**. These may include accidental unlocks or immediate ride cancellations. Since no clear policy identifies them as errors, they are retained in the dataset, but their presence is highlighted.

3.6.1 Calculation of Median and Percentiles

In addition to the mean, the **median (50th percentile)** and key quantiles (**25th, 75th, 90th, and 95th percentiles**) were computed using the **APPROX_QUANTILES()** function. The full SQL query is provided in **Appendix B.6 – "Median and Percentiles Queries"**.

Median (50th Percentile)

- **Members: 521 seconds (~8.7 min)**
- **Casual Riders: 723 seconds (~12.0 min)**
  → Casual riders take longer trips even at the **midpoint of the distribution**.

25th and 75th Percentiles

- **Members: 305 sec (~5 min) and 887 sec (~14.8 min)**
- **Casual Riders: 401 sec (~6.7 min) and 1364 sec (~22.7 min)**
  → **75% of casual riders' trips last up to ~23 min**, compared to ~15 min for members.

Higher Percentiles (90th, 95th)

- **Members:** Up to **~1900 sec (~31.7 min at the 95th percentile)**
- **Casual Riders: ~3960 sec (~66 min at the 95th percentile)**
  → Confirms that **a significant share of long-duration rides belongs to casual riders**.

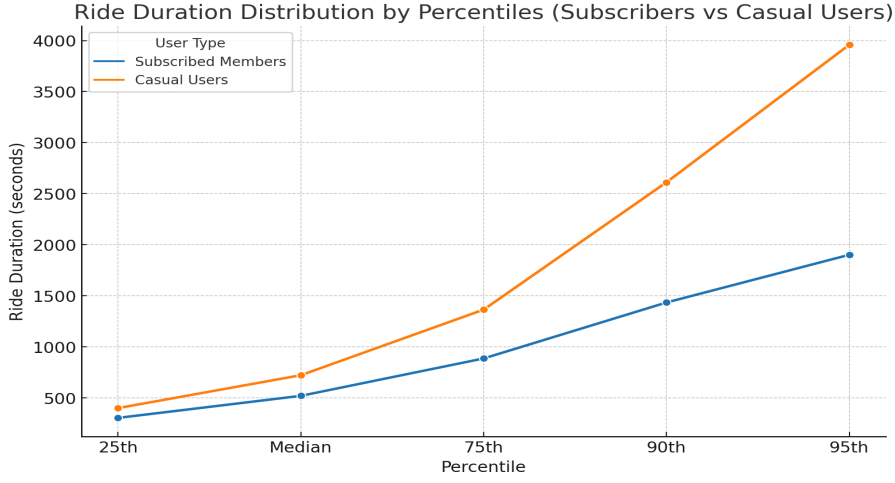


Figure 7 - Ride Duration Distribution by Percentiles (Members vs. Casual Riders)

3.6.2 Sensitivity Analysis and Outlier Impact

To assess the effect of extreme ride durations, a **sensitivity analysis** was conducted:

- **Trips >18 hours:** Filtering out **~0.17% of records**, the **casual riders' average ride duration drops from ~1509 sec to ~1225 sec (~25 min to ~20 min)**, while the **members' average remains largely unchanged**.
- **Trips <1 min:** Filtering out **2.2% of records**, the **mean increases slightly**, but does **not change the overall conclusion** that **casual riders take longer trips than members**.

Conclusion:

Removing extreme outliers **slightly shifts the average values**, but **does not alter the core finding**: **casual riders have significantly longer ride durations and greater variance in trip length compared to members**.

4. Summary of Key Insights

Ride Duration
Casual riders take **significantly longer rides** across all time frames, with **more than 37% of trips exceeding 15 minutes** and around **5–6% lasting over 60 minutes**.
Members prefer **short trips**, with **over 75% of rides lasting less than 15 minutes**, reflecting **frequent and commuter-oriented usage**.

Days and Hours
Members show **peak usage on weekdays (Tuesday, Wednesday, Thursday)**, mainly **during commuting hours (7–9 AM, 4–6 PM)**.
Casual riders prefer **weekends (Saturday and Sunday)**, taking **longer trips** with **peak usage between 2–5 PM**.

Seasonality
Both segments see a **peak in total rides during September** (~474k for members and ~346k for casual riders), indicating a broadly similar monthly pattern. However, **average ride duration differs significantly**: members maintain relatively stable times (around 8–14 minutes), while **casual riders exceed 20 minutes in the summer months**, highlighting a more recreational or tourist-driven use linked to warmer weather.

Bike Type Preference
E-bikes are **the most widely used option** for both segments, but members use them in a **more functional** way (short, frequent trips).
Classic bikes attract **much longer rides among casual riders** (up to **38 min on average**), especially on **weekends**.
Scooters remain **less relevant for members**, while casual riders use them **slightly more**, but with **shorter trip durations than classic bikes**.

These findings, supported by the **quantile analysis (median and percentile)**, highlight **two distinct user segments**:

- **Casual riders** with a **strong leisure/tourist focus** (weekends, summer, longer rides).
- **Annual members** with a **regular, commuter-driven pattern** (weekdays, shorter rides).

These differences provide **key insights for designing marketing strategies, pricing models, and promotional campaigns** to increase **annual memberships**.

5. Recommendations and Next Steps

Implement Weekend Plans and Extended Passes

- **Reason**: Casual riders primarily use the service on **weekends** for **longer trips**.
- **Action**: Introduce a **"Weekend Plus" membership**, allowing trips up to **60 minutes without extra fees**, making an annual plan more attractive than single-ride purchases.

Targeted Promotions During Afternoon Hours

- **Reason**: Casual riders peak between **2–5 PM**, with **longer rides**.
- **Action**: Offer **discounts on membership upgrades** for casual users exceeding **30 minutes per ride** in this time slot. Implement **automated email or push notifications** highlighting **potential savings** with an annual membership.

Tourism-Focused Classic Bike Bundles

- **Reason**: Casual riders prefer **classic bikes** for **long rides**, especially in **summer months**.
- **Action**: Establish **hotel partnerships** and introduce **"City Tour" packages**, integrating seasonal or annual memberships to encourage **tourists to opt for longer-term plans**.

E-Bike and Commuter-Focused Plans

- **Reason**: Both **casual riders and members** frequently use **e-bikes**, but members primarily for **short commutes**.
- **Action**: Develop a **"Commuter Membership"**, offering **priority access during peak hours** and **discounted short-ride rates**.

Progressive Pricing for Longer Rides

- **Reason**: Casual riders often take **longer trips exceeding 15–30 minutes**.
- **Action**: Implement **incremental pricing** for non-members on rides **beyond 30 minutes**, encouraging **annual subscription adoption** to avoid higher per-ride costs.

Next Steps

- **A/B Testing** of pricing strategies (**weekend pass, summer pass, commuter plan**).
- **Monitor Conversion Rates**: Track **casual-to-member upgrade rates** over **3–6 months**.
- **Customer Feedback Analysis**: Evaluate **user reviews, surveys, and suggestions** to refine membership offerings.

6. Limitations and Future Developments

While this analysis provides a comprehensive overview of **casual riders vs. members**, several **limitations** must be considered:

Key Limitations

- **Lack of Cost/Revenue Data**:
    - No financial data is available to assess the **profitability of each plan**. Recommendations are based on ride volume and duration rather than margins.
- **Anonymous Dataset**:
    - Without **unique user IDs**, we **cannot track customer retention** or identify frequent casual riders who may be strong membership candidates.
- **Outlier Ride Durations**:
    - Rides exceeding **18 hours** or lasting **less than 1 minute** were partly included. If they are **data recording errors**, they may slightly affect average calculations.
- **No Geospatial Analysis**:
    - Some **start/end station data is missing**, preventing an in-depth look at **touristic vs. residential traffic patterns** for bike allocation optimization.

Future Developments

- **Integration with Location & Weather Data**
    - Understanding **usage patterns under different weather conditions** and **urban mobility trends**.
- **Churn Rate Analysis**
    - If user IDs become available, we could track **membership cancellations** and analyze **repeat casual users' conversion potential**.
- **Demographic Insights**
    - Incorporating **age, gender, and income data** could enhance segmentation, allowing for **more targeted marketing campaigns**.

Appendix

Appendix A – Data Preparation & Cleaning

This section outlines the steps taken to integrate, clean, and prepare the dataset for analysis.

A.1 Creation of the Unified Table

```
CREATE OR REPLACE TABLE
`fluted-bot-433911-f3.cyclistic_data_2024.all_trips` AS
SELECT * FROM
`fluted-bot-433911-f3.cyclistic_data_2024.january_2024`
UNION ALL
SELECT * FROM
`fluted-bot-433911-f3.cyclistic_data_2024.february_2024`
UNION ALL
...
UNION ALL
SELECT * FROM
`fluted-bot-433911-f3.cyclistic_data_2024.december_2024`;
```

A.2 Verification of Records in the Unified Table

```
SELECT COUNT(*) AS total_rows
FROM `fluted-bot-433911-f3.cyclistic_data_2024.all_trips`;

SELECT *
FROM `fluted-bot-433911-f3.cyclistic_data_2024.all_trips`
LIMIT 10;
```

A.3 Creation of ride_length and day_of_week

```
CREATE OR REPLACE TABLE
`fluted-bot-433911-f3.cyclistic_data_2024.all_trips_cleaned` AS
SELECT *,
       TIMESTAMP_DIFF(ended_at, started_at, SECOND) AS ride_length,
       FORMAT_TIMESTAMP('%A', started_at) AS day_of_week
FROM `fluted-bot-433911-f3.cyclistic_data_2024.all_trips`;
```

A.4 Detection of Null or Negative Values

```
SELECT COUNT(*) AS null_or_negative_count
FROM `fluted-bot-433911-f3.cyclistic_data_2024.all_trips_cleaned`
WHERE ride_length <= 0 OR ride_length IS NULL;
```

A.5 Creation of the Final Cleaned Table

```
CREATE OR REPLACE TABLE
`fluted-bot-433911-f3.cyclistic_data_2024.cleaned_data` AS
SELECT *
FROM `fluted-bot-433911-f3.cyclistic_data_2024.all_trips_cleaned`
WHERE ride_length > 0;
```

Appendix B

Appendix B.1 – Query for Total Rides and Total Ride Duration

B.1.1 - Query: Basic statistics (total rides and average ride duration by user category).

```
SELECT
    member_casual,
    COUNT(*) AS total_rides,
    ROUND(AVG(ride_length), 2) AS avg_ride_length_seconds,
    ROUND(MAX(ride_length), 2) AS max_ride_length_seconds,
    ROUND(MIN(ride_length), 2) AS min_ride_length_seconds
FROM `fluted-bot-433911-f3.cyclistic_data_2024.cleaned_data`
GROUP BY member_casual;
```

B.1.2 Total ride duration by user category.

```
SELECT
    member_casual,
    COUNT(*) AS total_rides,
    SUM(ride_length) AS total_ride_duration_seconds,
    ROUND(SUM(ride_length) / 3600, 2) AS total_ride_duration_hours
FROM `fluted-bot-433911-f3.cyclistic_data_2024.cleaned_data`
GROUP BY member_casual;
```

Appendix B.2 – Daily and Hourly Analysis

This appendix contains the SQL queries used to analyze the distribution of rides by day of the week and by hour of the day, distinguishing between casual riders and annual members.

B.2.1 – Distribution by Day of the Week

```
SELECT
    member_casual,
    day_of_week,
    COUNT(*) AS total_rides,
    ROUND(AVG(ride_length), 2) AS avg_ride_length_seconds
FROM `fluted-bot-433911-f3.cyclistic_data_2024.cleaned_data`
GROUP BY member_casual, day_of_week
ORDER BY member_casual, day_of_week;
```

B.2.2 – Hourly Distribution (Total Rides)

```
SELECT
  EXTRACT(HOUR FROM started_at) AS hour_of_day,
  member_casual,
  COUNT(*) AS total_rides
FROM `fluted-bot-433911-f3.cyclistic_data_2024.cleaned_data`
GROUP BY hour_of_day, member_casual
ORDER BY hour_of_day, member_casual;
```

B.2.3 – Hourly Distribution (Average Ride Duration)

```
SELECT
  EXTRACT(HOUR FROM started_at) AS hour_of_day,
  member_casual,
  ROUND(AVG(ride_length), 2) AS avg_ride_length_seconds
FROM `fluted-bot-433911-f3.cyclistic_data_2024.cleaned_data`
GROUP BY hour_of_day, member_casual
ORDER BY hour_of_day, member_casual;
```

B.2.4 – Weekday vs Weekend

```
SELECT
  CASE
    WHEN day_of_week IN ('Saturday', 'Sunday') THEN 'Weekend'
    ELSE 'Weekday'
  END AS day_type,
  EXTRACT(HOUR FROM started_at) AS hour_of_day,
  member_casual,
```

```
  COUNT(*) AS total_rides,
  ROUND(AVG(ride_length), 2) AS avg_ride_length_seconds
FROM `fluted-bot-433911-f3.cyclistic_data_2024.cleaned_data`
GROUP BY day_type, hour_of_day, member_casual
ORDER BY day_type, hour_of_day, member_casual;
```

Appendix B.3 – Monthly Analysis

This appendix includes the SQL query used to analyze how the number of rides and the average duration vary by month for the two user segments.

```
SELECT
  EXTRACT(MONTH FROM started_at) AS month,
  member_casual,
  COUNT(*) AS total_rides,
  ROUND(AVG(ride_length), 2) AS avg_ride_length_seconds
FROM
  `fluted-bot-433911-f3.cyclistic_data_2024.cleaned_data`
GROUP BY
  month, member_casual
ORDER BY
  month, member_casual;
```

Appendix B.4 – Bike Type Analysis

B.4.1 – Usage of Different Bike Types

```
SELECT
    rideable_type,
    member_casual,
    COUNT(*) AS total_rides,
    ROUND(AVG(ride_length), 2) AS avg_ride_length_seconds
FROM `fluted-bot-433911-f3.cyclistic_data_2024.cleaned_data`
GROUP BY rideable_type, member_casual
ORDER BY total_rides DESC;
```

B.4.2 – Hourly and Daily Differences by Bike Type

```
SELECT
  rideable_type,
  member_casual,
  CASE
    WHEN day_of_week IN ('Saturday','Sunday') THEN 'Weekend'
```

```
    ELSE 'Weekday'
  END AS day_type,
  EXTRACT(HOUR FROM started_at) AS hour_of_day,
  COUNT(*) AS total_rides,
  ROUND(AVG(ride_length), 2) AS avg_ride_length_seconds
FROM `fluted-bot-433911-f3.cyclistic_data_2024.cleaned_data`
GROUP BY rideable_type, member_casual, day_type, hour_of_day
ORDER BY rideable_type, member_casual, day_type, hour_of_day;
```

Appendix B.5 – Ride Duration Distribution

```
SELECT
  member_casual,
  CASE
    WHEN ride_length < 300 THEN '<5 min'
    WHEN ride_length BETWEEN 300 AND 900 THEN '5-15 min'
    WHEN ride_length BETWEEN 900 AND 1800 THEN '15-30 min'
    WHEN ride_length BETWEEN 1800 AND 3600 THEN '30-60 min'
    ELSE '>60 min'
  END AS duration_category,
  COUNT(*) AS total_rides
FROM `fluted-bot-433911-f3.cyclistic_data_2024.cleaned_data`
GROUP BY member_casual, duration_category
ORDER BY member_casual, duration_category;
```

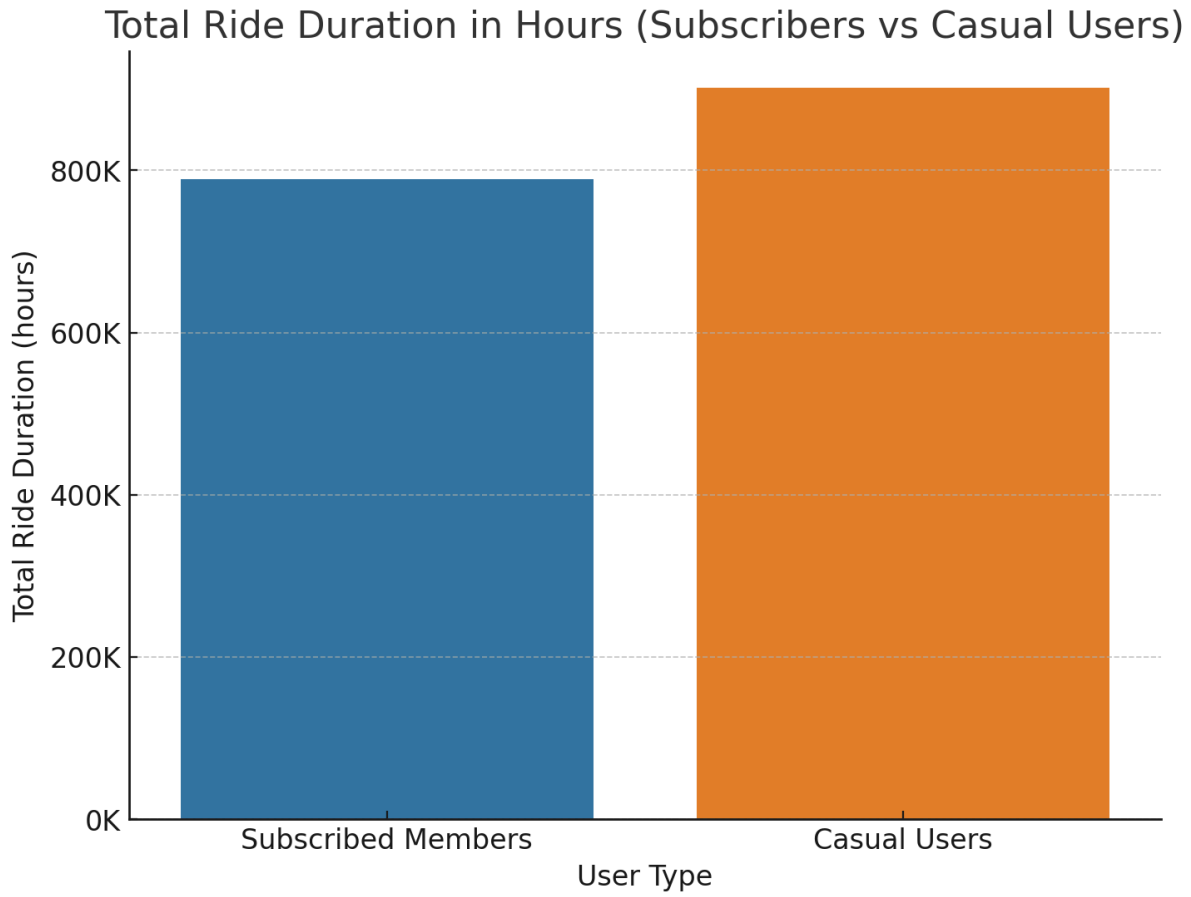
Appendix B.6 – Key SQL Queries for Median and Percentiles

```
SELECT
  member_casual,
  APPROX_QUANTILES(ride_length, 100)[OFFSET(50)] AS
median_ride_length,
  APPROX_QUANTILES(ride_length, 100)[OFFSET(25)] AS percentile_25,
  APPROX_QUANTILES(ride_length, 100)[OFFSET(75)] AS percentile_75,
  APPROX_QUANTILES(ride_length, 100)[OFFSET(90)] AS percentile_90,
  APPROX_QUANTILES(ride_length, 100)[OFFSET(95)] AS percentile_95
FROM `fluted-bot-433911-f3.cyclistic_data_2024.cleaned_data`
GROUP BY member_casual;
```

Additional Visualizations

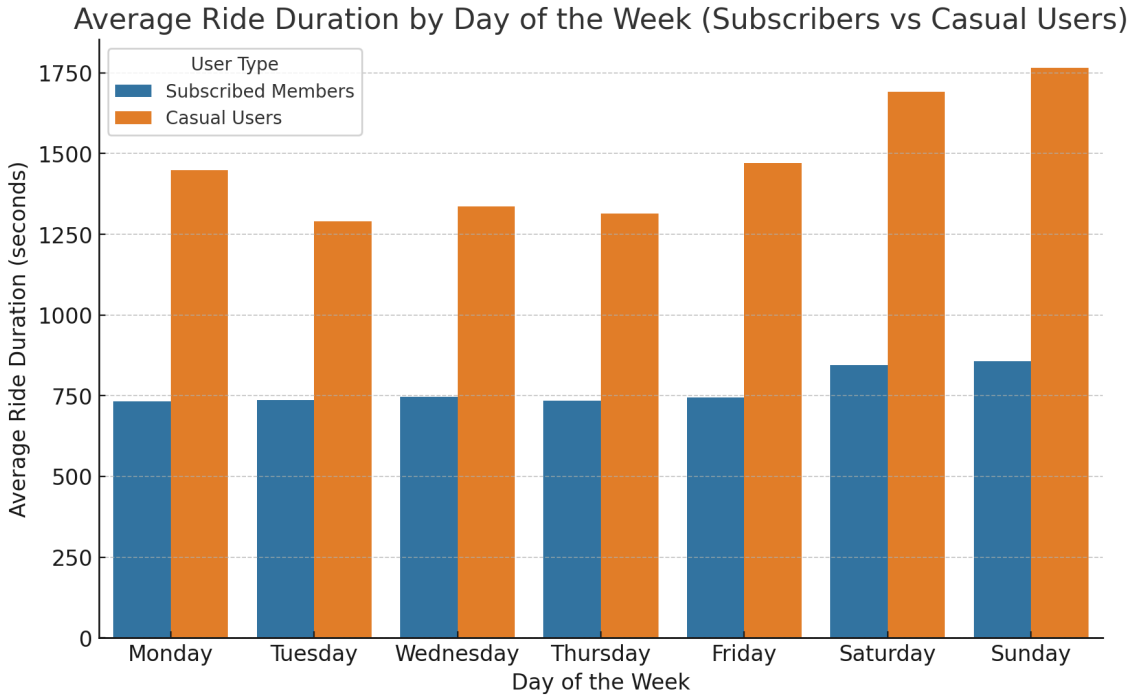
Section B1 – Total Rides and Ride Duration

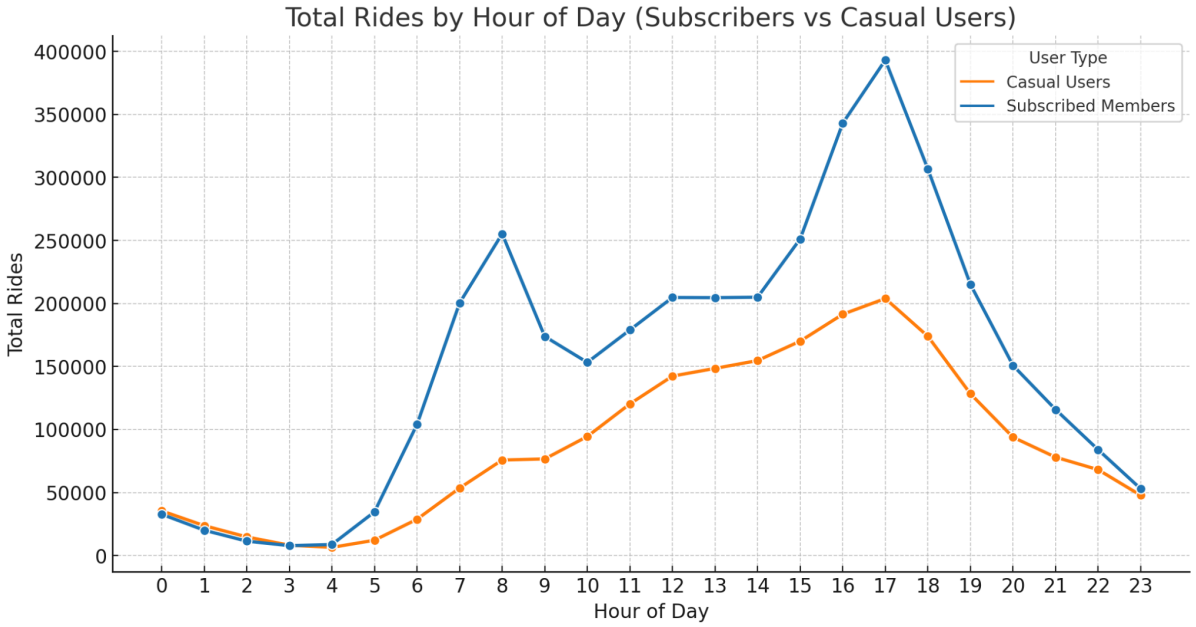
The chart displays the total ride duration for members and casual riders. Although members take more trips, casual riders accumulate a higher total ride time, indicating a more recreational/touristic usage.

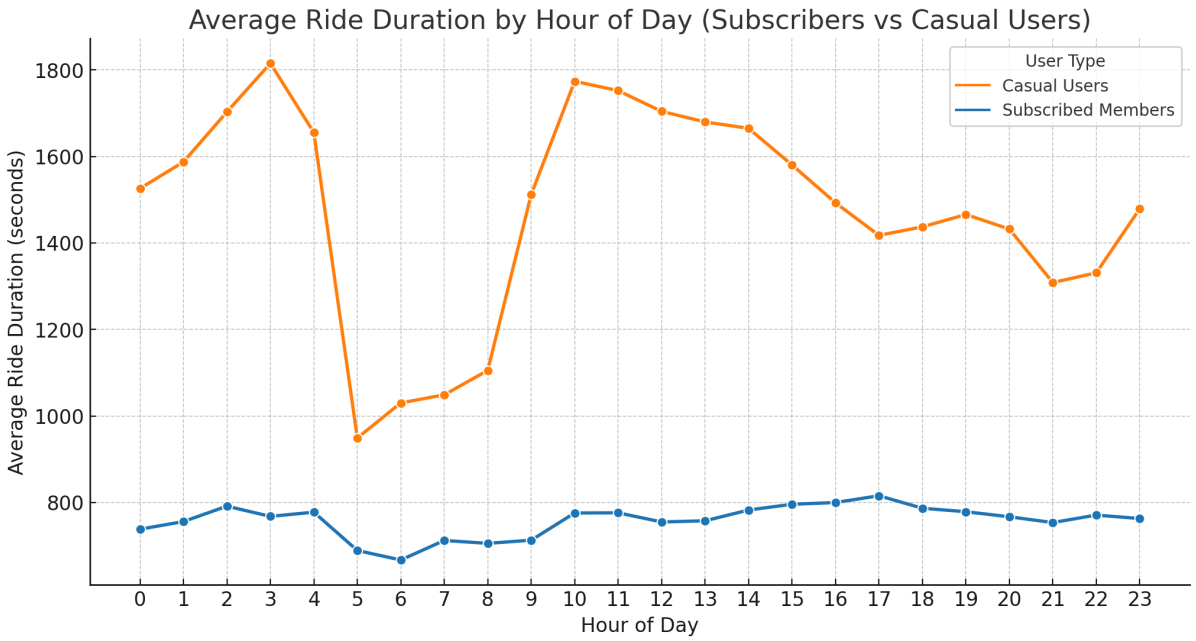


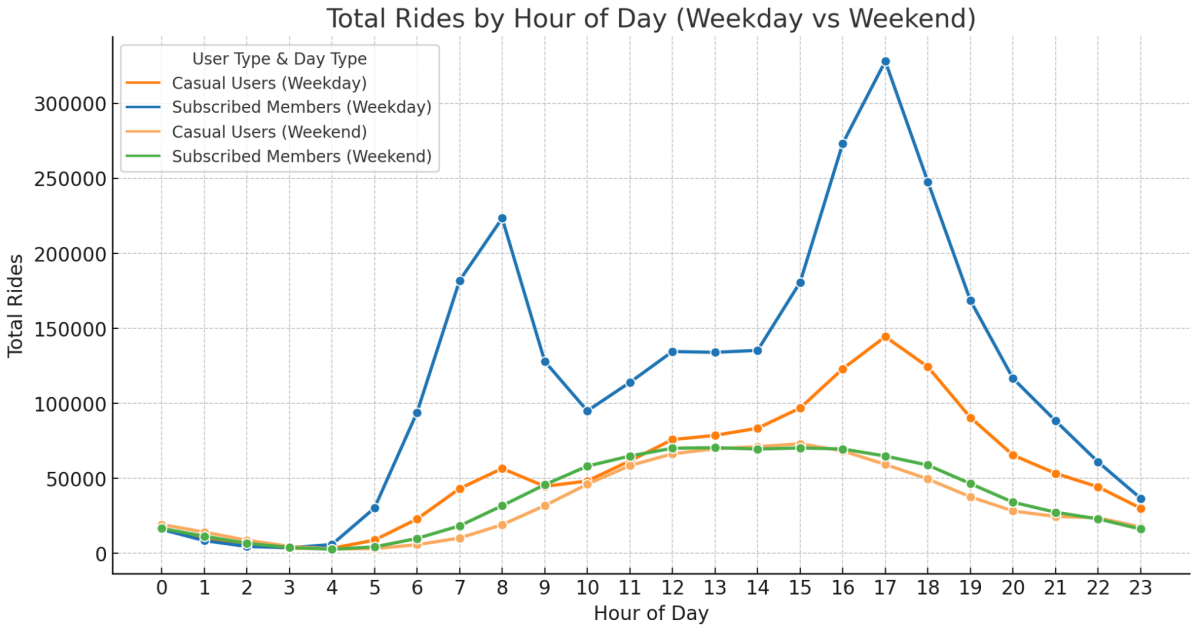
Section B2 – Daily and Hourly Analysis

This section presents a series of visualizations that highlight differences in service usage between members and casual riders based on days of the week and hourly distribution.



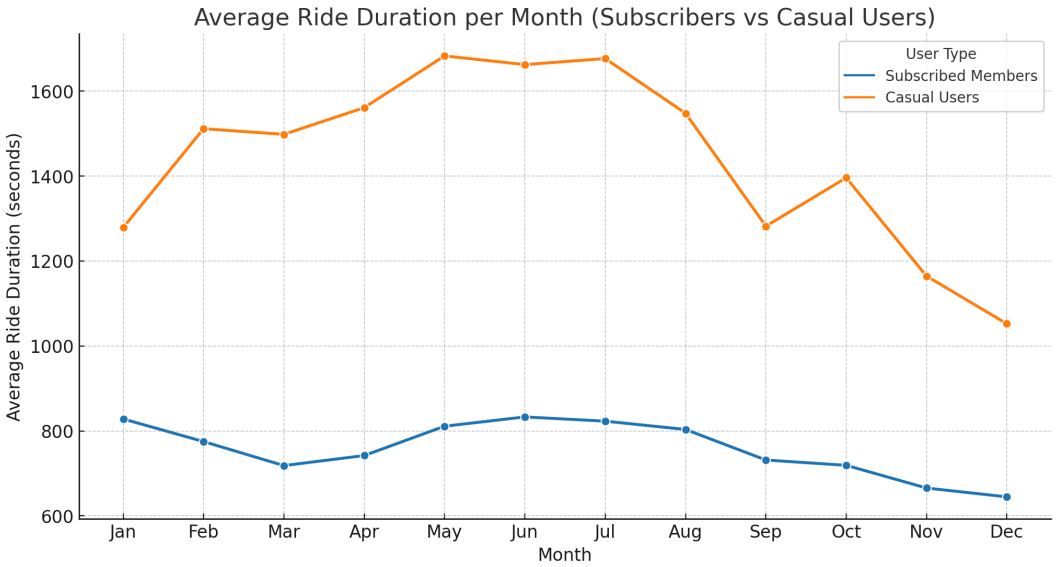



Average Ride Duration by Hour of Day (Subscribers vs Casual Users)
User Type
Casual Users
Subscribed Members
Average Ride Duration (seconds)
1800
1600
1400
1200
1000
800
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23
Hour of Day


Total Rides by Hour of Day (Weekday vs Weekend)
User Type & Day Type
Casual Users (Weekday)
Subscribed Members (Weekday)
Casual Users (Weekend)
Subscribed Members (Weekend)
Total Rides
300000
250000
200000
150000
100000
50000
0
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23
Hour of Day

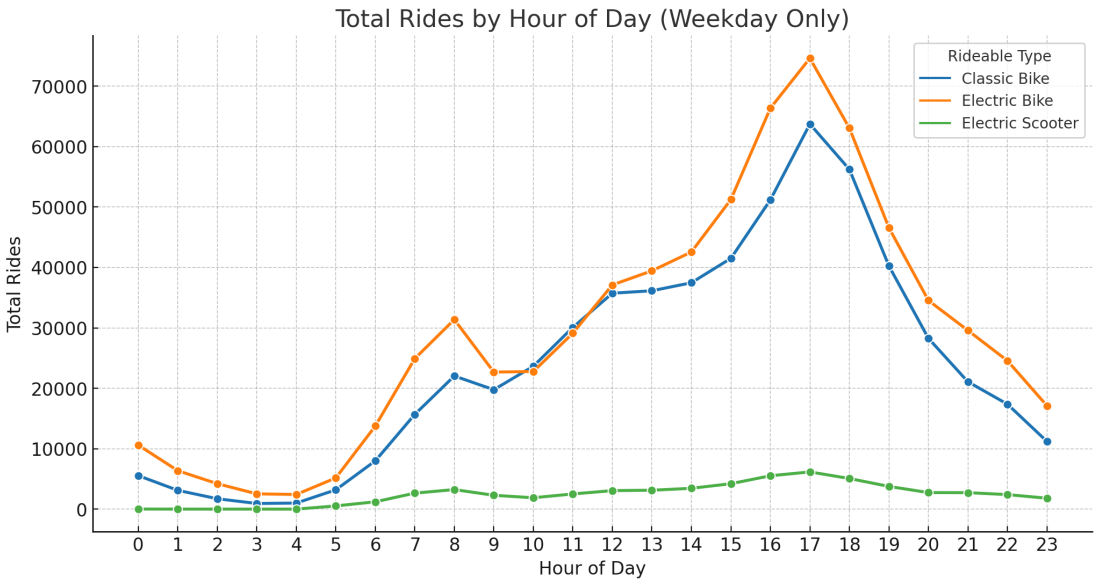
Section B3 – Monthly Analysis (Seasonality)

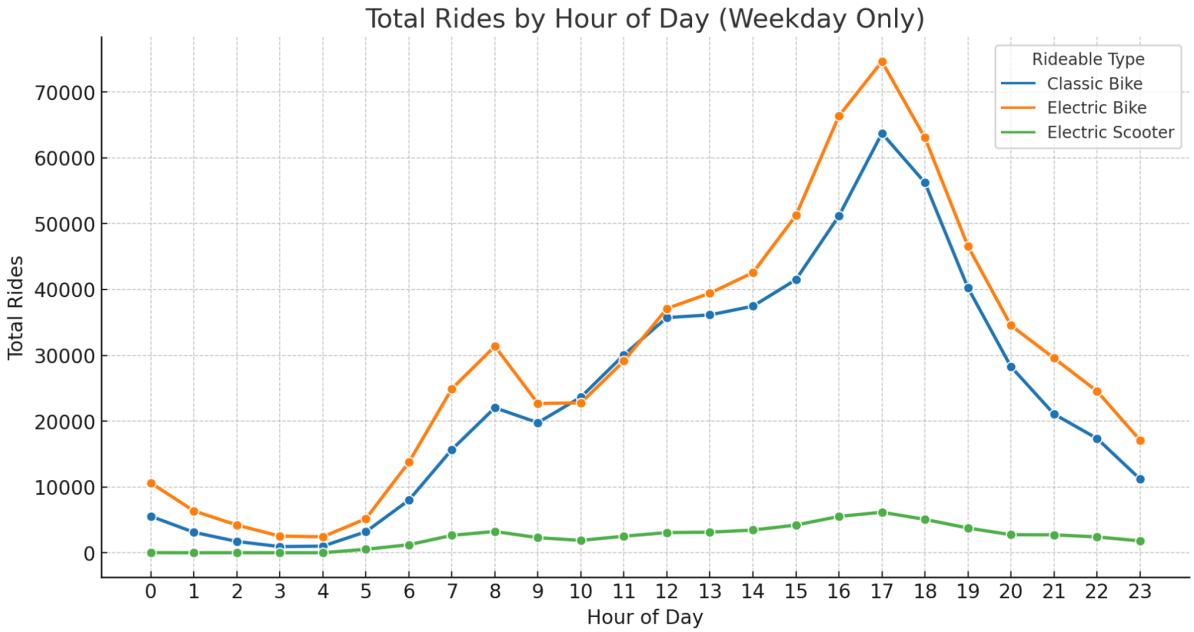
This visualization shows the **average ride duration per month**, revealing seasonal trends. Casual riders take longer trips in warmer months, peaking between May and September, while members maintain a stable ride duration year-round.

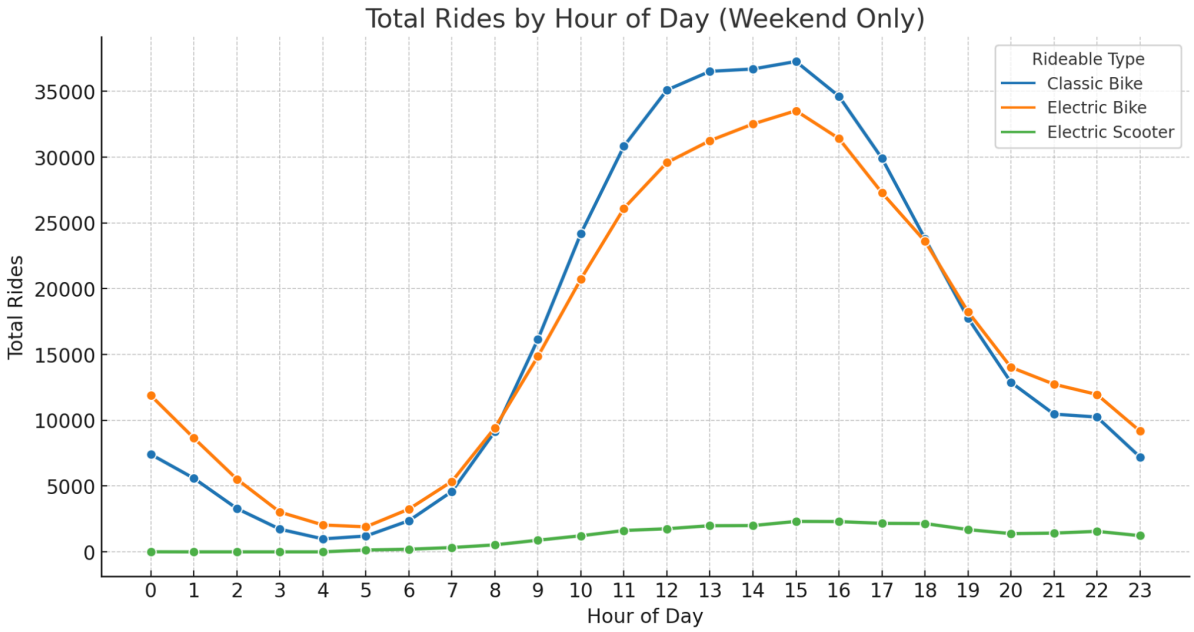


Section B4 – Bike Type Analysis

This section includes visualizations on the usage of different bike types (classic, electric, scooter) by members and casual riders, distinguishing between weekdays and weekends. It also highlights differences in average ride duration for each bike type.




Total Rides by Hour of Day (Weekday Only)
Rideable Type
Classic Bike
Electric Bike
Electric Scooter
Total Rides
0
10000
20000
30000
40000
50000
60000
70000
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23
Hour of Day


Total Rides by Hour of Day (Weekend Only)
Rideable Type
Classic Bike
Electric Bike
Electric Scooter
Total Rides
0
5000
10000
15000
20000
25000
30000
35000
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23
Hour of Day

Section B5 – Ride Duration Distribution

Visualization of ride duration by time range. Members mostly take short trips (5–15 min), while casual riders have longer rides, with many exceeding 30 minutes.