It looks like you've uploaded a zip file named `202311-divvy-tripdata.zip`. Let's start by unzipping the file and examining its contents. Then we can proceed with any specific tasks or analysis you have in mind.

I'll begin by unzipping the file and listing its contents.

The zip file contains the following items:

1. `202311-divvy-tripdata.csv`

2. `\_\_MACOSX` (a directory usually created by macOS when zipping files)

We'll focus on the CSV file for analysis. Let's load and inspect the first few rows of the `202311-divvy-tripdata.csv` file to understand its structure and contents.

The CSV file contains data on bike rides with the following columns:

1. `ride\_id`: Unique identifier for each ride.

2. `rideable\_type`: Type of bike (e.g., electric\_bike, classic\_bike).

3. `started\_at`: Start timestamp of the ride.

4. `ended\_at`: End timestamp of the ride.

5. `start\_station\_name`: Name of the starting station.

6. `start\_station\_id`: Identifier for the starting station.

7. `end\_station\_name`: Name of the ending station.

8. `end\_station\_id`: Identifier for the ending station.

9. `start\_lat`: Latitude of the starting point.

10. `start\_lng`: Longitude of the starting point.

11. `end\_lat`: Latitude of the ending point.

12. `end\_lng`: Longitude of the ending point.

13. `member\_casual`: Membership status of the rider (e.g., member, casual).

How would you like to proceed with the analysis? Here are a few options:

- Summary statistics of the rides.

- Analysis of ride duration and distances.

- Visualization of popular stations or routes.

- Membership analysis (member vs casual).

Let me know what specific analysis or insights you're interested in!

Here are the first few rows of the dataset to give you an overview of its structure:

| 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 |

|-----------------|---------------|---------------------|---------------------|-----------------------|------------------|---------------------------|-----------------|-----------|-----------|-----------|-----------|---------------|

| 4EAD8F1AD547356B | electric\_bike | 2023-11-30 21:50:05 | 2023-11-30 22:13:27 | Millennium Park | 13008 | Pine Grove Ave & Waveland Ave | TA1307000150 | 41.881101 | -87.624082 | 41.949473 | -87.646453 | member |

| 6322270563BF5470 | electric\_bike | 2023-11-03 09:44:02 | 2023-11-03 10:17:15 | Broadway & Sheridan Rd | 13323 | Broadway & Sheridan Rd | 13323 | 41.952868 | -87.650035 | 41.952833 | -87.649993 | member |

| B37BDE091ECA38E0 | electric\_bike | 2023-11-30 11:39:44 | 2023-11-30 11:40:08 | State St & Pearson St | TA1307000061 | State St & Pearson St | TA1307000061 | 41.897533 | -87.628694 | 41.897448 | -87.628722 | member |

| CF0CA5DD26E4F90E | classic\_bike | 2023-11-08 10:01:45 | 2023-11-08 10:27:05 | Theater on the Lake | TA1308000001 | Theater on the Lake | TA1308000001 | 41.926277 | -87.630834 | 41.926277 | -87.630834 | member |

| EB8381AA641348DB | classic\_bike | 2023-11-03 16:20:25 | 2023-11-03 16:54:25 | Theater on the Lake | TA1308000001 | Theater on the Lake | TA1308000001 | 41.926277 | -87.630834 | 41.926277 | -87.630834 | member |

This dataset includes detailed information about each ride, including the type of bike, timestamps for the start and end of the ride, station information, geographic coordinates, and the rider's membership status.

Would you like to explore specific aspects or perform a particular analysis on this data?

Let's start with some basic analyses of the cleaned dataset:

1. \*\*Summary Statistics\*\*: Provide a statistical summary of the dataset.

2. \*\*Ride Duration Analysis\*\*: Calculate and analyze the duration of rides.

3. \*\*Trends Over Time\*\*: Analyze trends in bike usage over time.

4. \*\*Popular Stations\*\*: Identify the most popular start and end stations.

5. \*\*Membership Analysis\*\*: Compare usage between members and casual riders.

### 1. Summary Statistics

We'll begin with a statistical summary of the dataset.

### Summary Statistics

#### Numerical Columns:

| Statistic | start\_lat | start\_lng | end\_lat | end\_lng |

|-----------|-----------|-----------|---------|---------|

| Count | 274,798 | 274,798 | 274,798 | 274,798 |

| Mean | 41.8968 | -87.6444 | 41.8972 | -87.6446|

| Std Dev | 0.0456 | 0.0252 | 0.0457 | 0.0253 |

| Min | 41.6486 | -87.8441 | 41.6485 | -87.8440|

| 25% | 41.8779 | -87.6581 | 41.8782 | -87.6584|

| Median | 41.8938 | -87.6418 | 41.8940 | -87.6418|

| 75% | 41.9242 | -87.6287 | 41.9253 | -87.6292|

| Max | 42.0649 | -87.5282 | 42.0649 | -87.5282|

#### Categorical Columns:

| Column | Unique Values | Most Frequent Value | Frequency |

|----------------------|---------------|----------------------------------------|-----------|

| ride\_id | 274,798 | Unique for each ride | N/A |

| rideable\_type | 2 | classic\_bike | 175,677 |

| start\_station\_name | 1,060 | Clinton St & Washington Blvd | 2,598 |

| start\_station\_id | 1,049 | WL-012 | 2,598 |

| end\_station\_name | 1,086 | Clinton St & Washington Blvd | 2,706 |

| end\_station\_id | 1,072 | WL-012 | 2,706 |

| member\_casual | 2 | member | 202,701 |

### 2. Ride Duration Analysis

Next, we'll calculate and analyze the duration of rides. The duration will be calculated as the difference between `ended\_at` and `started\_at`.

### Ride Duration Analysis

#### Summary Statistics for Ride Duration (in seconds):

| Statistic | Value |

|-----------|------------|

| Count | 274,798 |

| Mean | 770.39 |

| Std Dev | 1,846.15 |

| Min | -3,274.00 |

| 25% | 298.00 |

| Median | 498.00 |

| 75% | 851.00 |

| Max | 89,850.00 |

### Notes:

- There are some negative durations, which indicate data issues or incorrect timestamps. We should filter out these anomalies.

### 3. Trends Over Time

Let's analyze the trends in bike usage over time by examining the number of rides per day.

### Trends Over Time

The plot above shows the number of rides per day, indicating the daily usage trends for the Divvy bike-sharing service. We can see fluctuations in the number of rides, which might be influenced by various factors such as weather, weekdays vs. weekends, and events.

### 4. Popular Stations

Next, let's identify the most popular start and end stations. We'll look at the top 10 stations for both starting and ending rides.

#### Most Popular Start Stations

#### Most Popular Start Stations

| Station Name | Number of Rides |

|--------------------------------|-----------------|

| Clinton St & Washington Blvd | 2,598 |

| University Ave & 57th St | 2,445 |

| Kingsbury St & Kinzie St | 2,400 |

| Ellis Ave & 60th St | 2,329 |

| Streeter Dr & Grand Ave | 2,172 |

| Clinton St & Madison St | 2,147 |

| Clark St & Elm St | 2,060 |

| Ellis Ave & 55th St | 1,845 |

| Morgan St & Polk St | 1,785 |

| Wells St & Elm St | 1,763 |

#### Most Popular End Stations

| Station Name | Number of Rides |

|--------------------------------|-----------------|

| Clinton St & Washington Blvd | 2,706 |

| University Ave & 57th St | 2,399 |

| Kingsbury St & Kinzie St | 2,328 |

| Ellis Ave & 60th St | 2,296 |

| Clinton St & Madison St | 2,250 |

| Streeter Dr & Grand Ave | 2,156 |

| Clark St & Elm St | 2,048 |

| Ellis Ave & 55th St | 1,873 |

| Canal St & Adams St | 1,824 |

| Wells St & Concord Ln | 1,805 |

### 5. Membership Analysis

Finally, let's compare the usage between members and casual riders. We'll look at the number of rides and average ride duration for each group.

### Membership Analysis

| Membership Type | Number of Rides | Average Ride Duration (seconds) |

|-----------------|-----------------|---------------------------------|

| Casual | 72,077 | 1,073.15 |

| Member | 202,689 | 663.26 |

### Insights:

- \*\*Number of Rides\*\*: Members take significantly more rides compared to casual riders.

- \*\*Average Ride Duration\*\*: Casual riders have a longer average ride duration compared to members.

Would you like to delve deeper into any of these analyses or explore other aspects of the data?

The visualizations have been saved successfully. You can download them using the following links:

- [Download Visuals Part 1](sandbox:/mnt/data/divvy\_tripdata\_visuals\_1.png)

- [Download Visuals Part 2](sandbox:/mnt/data/divvy\_tripdata\_visuals\_2.png)

If you need any further analysis or additional visualizations, feel free to let me know!