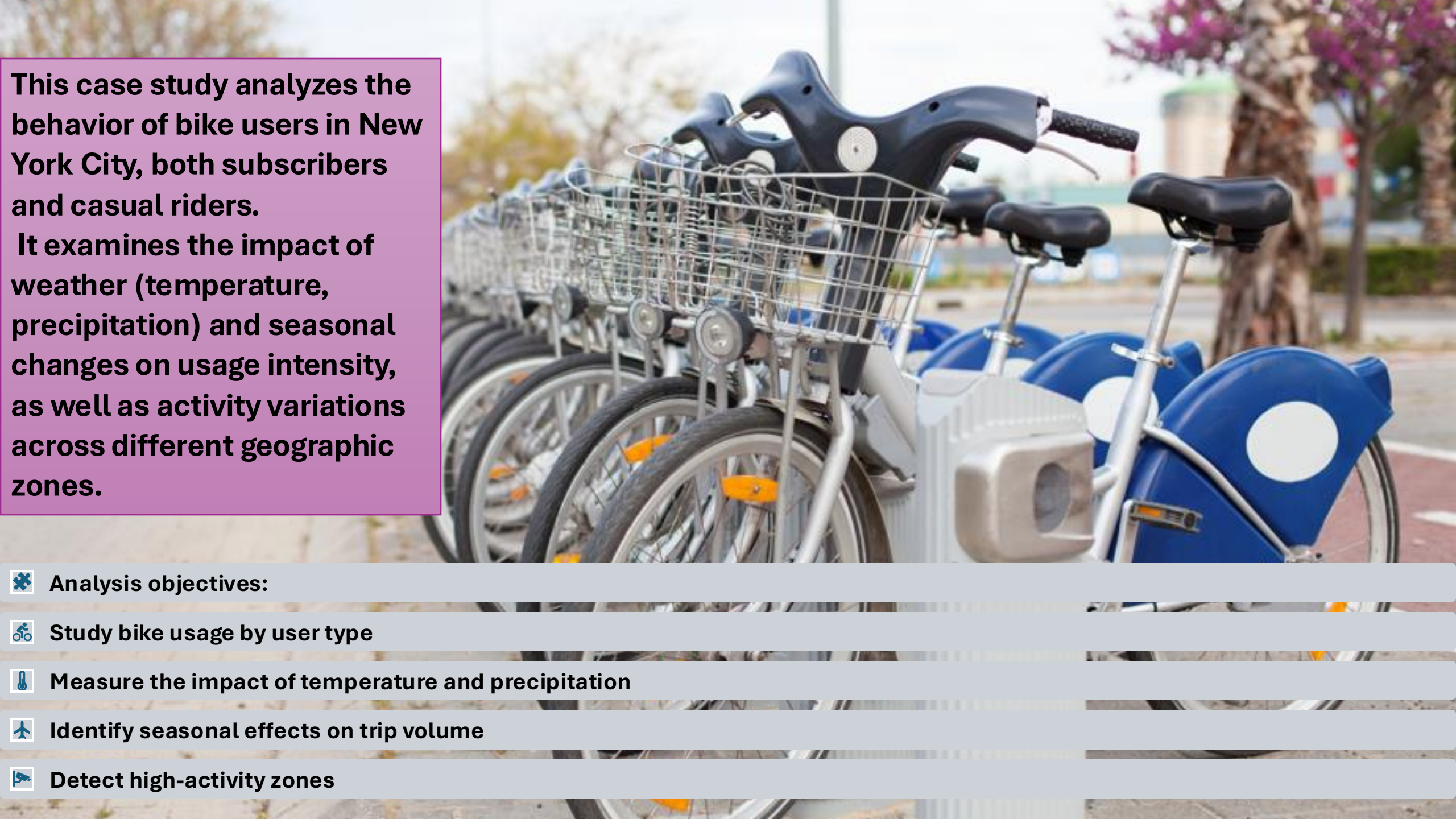
# Understanding Urban Mobility Through BI: The NYC Bike Case

A data-driven exploration of Citibike patterns, regional flow, and weather impact in New York City.

El Khlife Messsoud

April 2025



This case study analyzes the behavior of bike users in New York City, both subscribers and casual riders.
It examines the impact of weather (temperature, precipitation) and seasonal changes on usage intensity, as well as activity variations across different geographic zones.

- Analysis objectives:
- Study bike usage by user type
- Measure the impact of temperature and precipitation
- Identify seasonal effects on trip volume
- Detect high-activity zones

# Data Sources Overview

| Dataset | Description | Source |
|---|---|---|
| citibike_trips | Bike trip records including user type, duration, start/stop time, and station coordinates | bigquery-public-data.new_york_citibike |
| zip_codes | Geospatial boundaries used to locate start and end stations within ZIP code areas | bigquery-public-data.geo_us_boundaries |
| gsod20* | Daily weather data: mean temperature, wind speed, and precipitation | bigquery-public-data.noaa_gsod |
| codepostal | Custom table mapping ZIP codes to boroughs and neighborhoods | my-project-trees-in-nyc-53016.cyclistic |

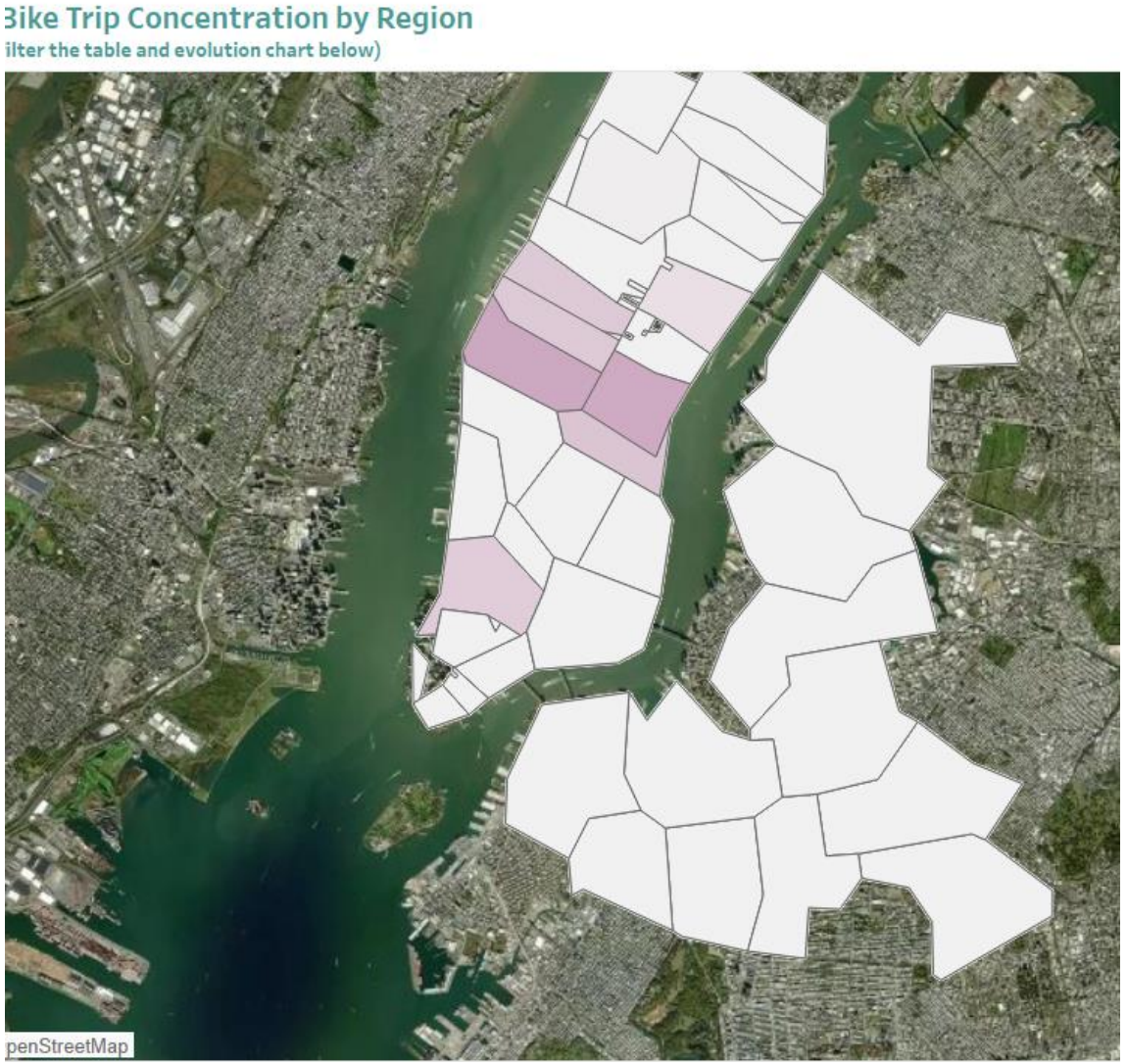
## Data Preparation Summary

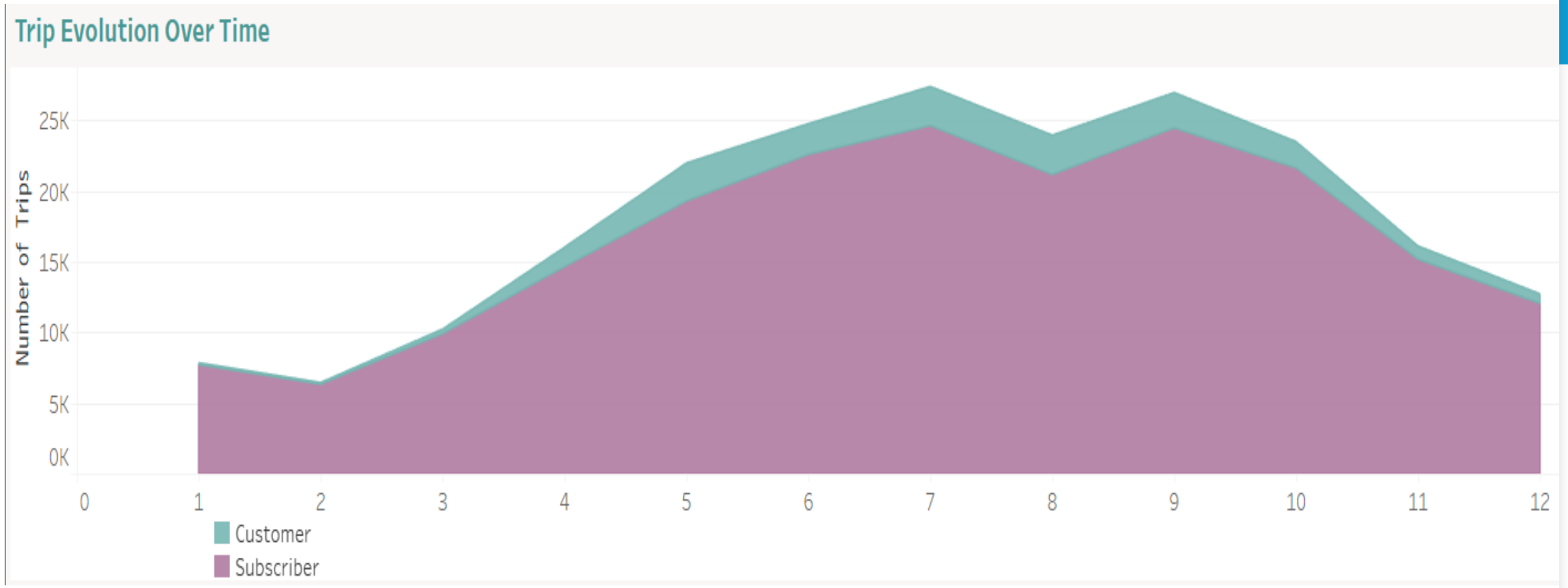
- Geospatial joins between station coordinates and ZIP code boundaries
- Temporal join with weather data by matching trip date to weather date
- Added custom columns:
  - tripminutes: rounded trip duration (10-minute intervals)
  - Number of trips : aggregated number of trips
  - Adjusted start_day and stop_day by +5 years for modern display
- Filtered data:
  - Weather station: New York Central Park (WBAN 94728)

**Data Selection Note**

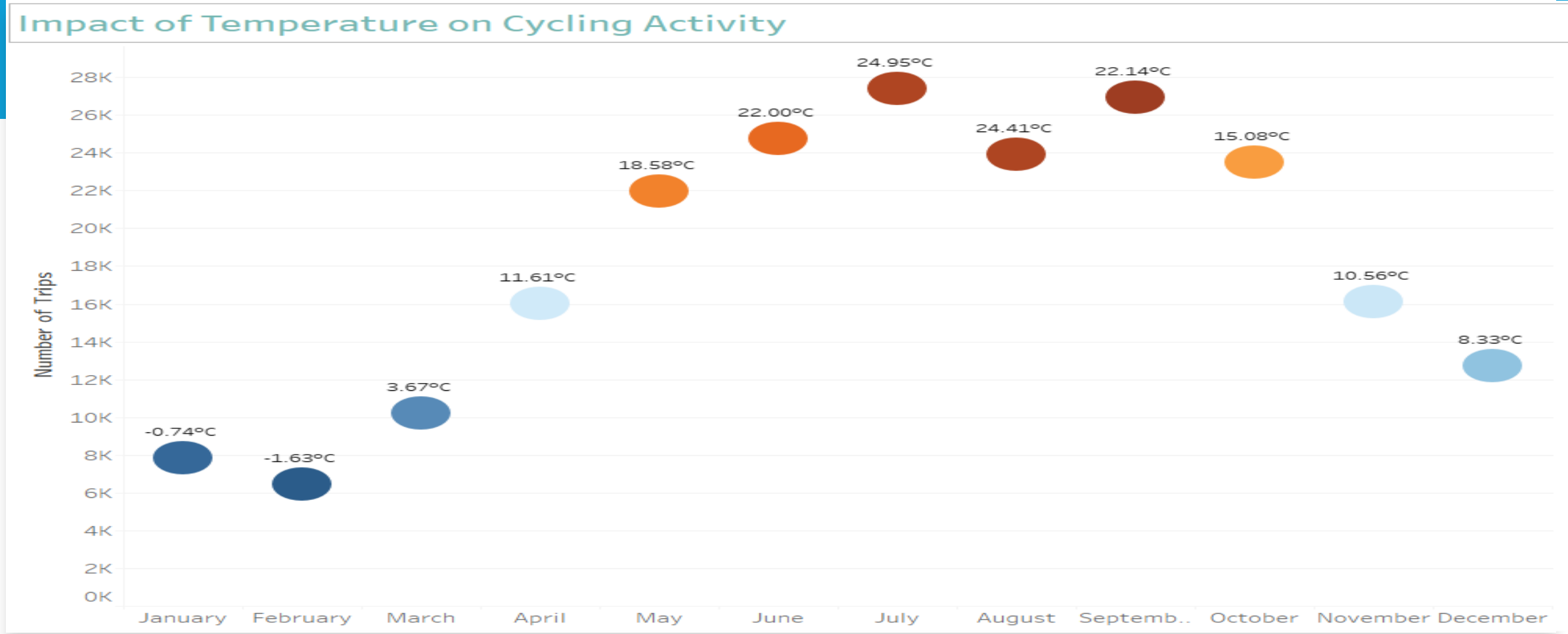
To ensure consistency and completeness in the analysis, only the years **2019 and 2020** were selected.
These two years provide the **most complete datasets** for both bike trips and weather, covering all 12 months.

- The heat map reveals a particularly high concentration of cycling activity in the heart of Manhattan. The intensity of the colors clearly indicates that this dense urban area constitutes the epicenter of cycling in New York. The visualization allows users to click on specific regions to filter data in the associated tables and charts, providing a dynamic way to explore geographic patterns.

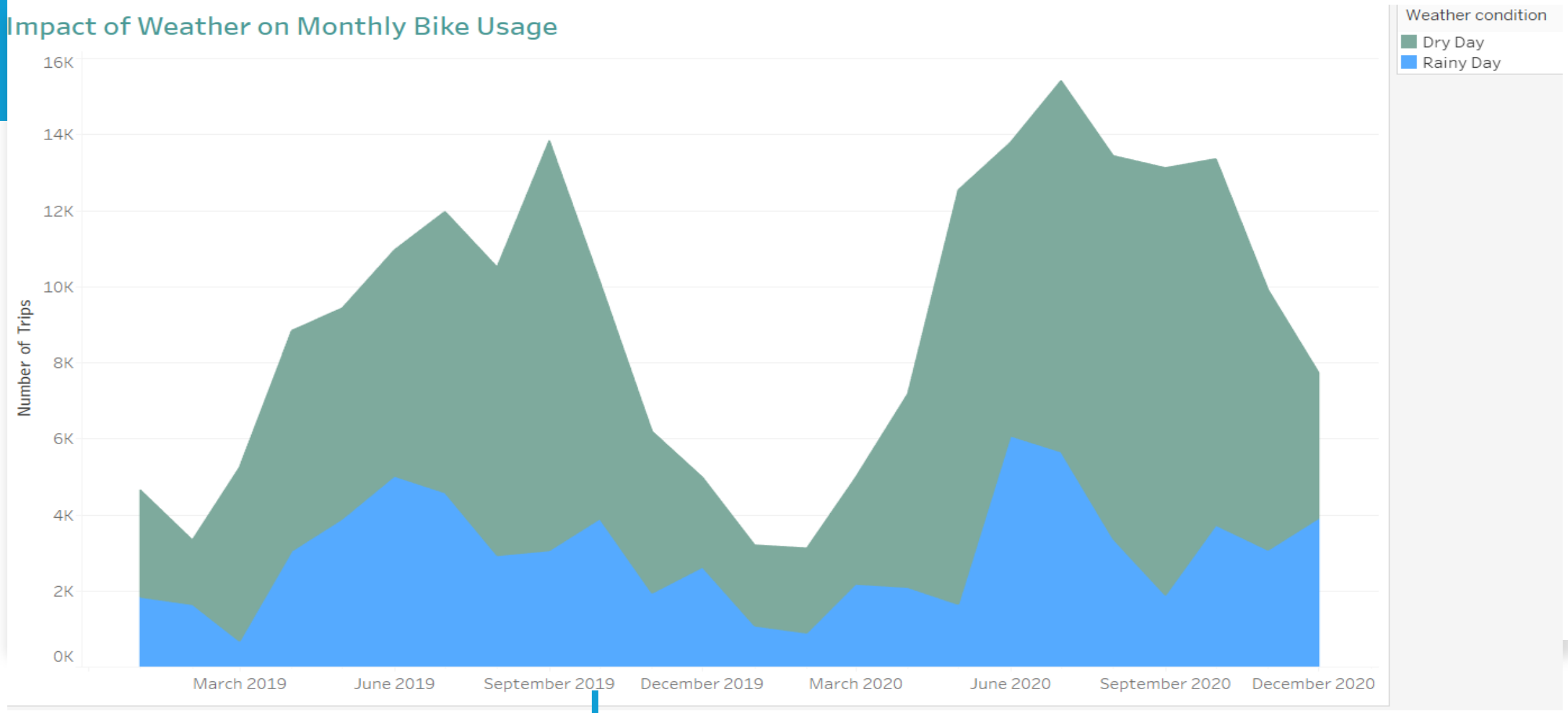





- This area chart compares monthly bike usage between Subscribers and Customers.
  Subscribers consistently take more trips, peaking in summer (June–September).
  Customers follow the same seasonal pattern, but with fewer rides.
  Trip volume declines for both groups in the last months of the year.



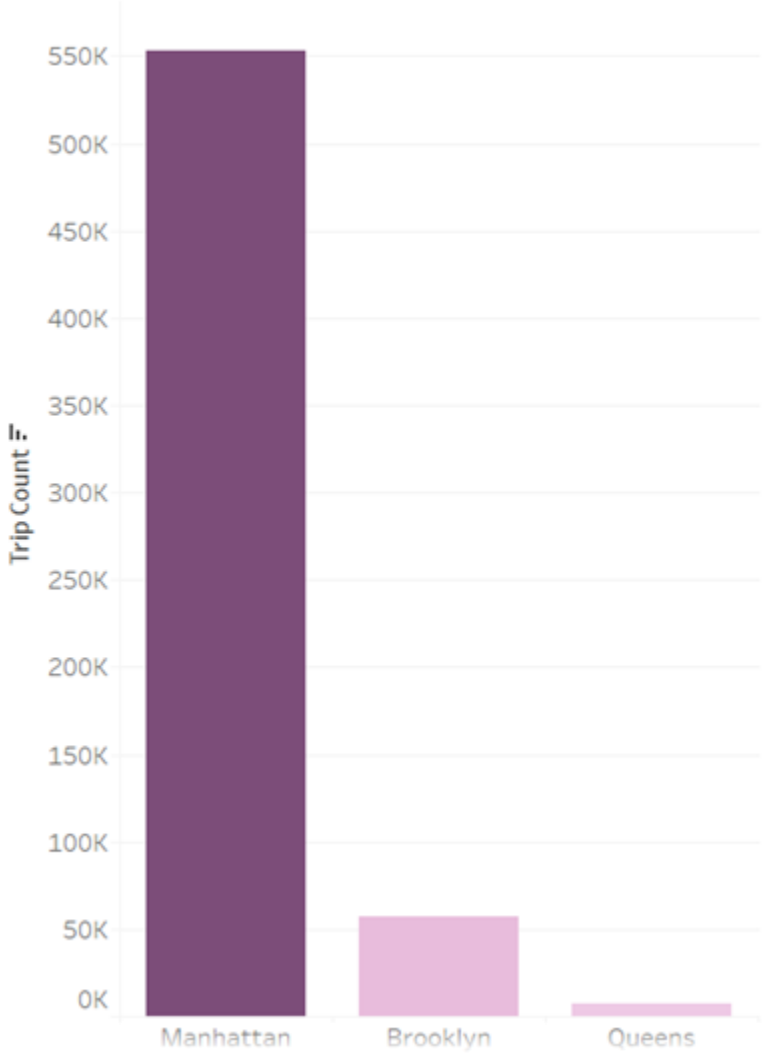

- Bike trips increase as temperature rises, starting in May.
  The peak is in July (24.95°C) with around 28000 trips.
  From October, trips drop steadily, reaching a low in February (-1.63°C, ~6000 trips).
  The bubble chart shows this seasonal pattern using size and color to reflect temperature.




- The charts show how rain affects bike usage.

There's a clear negative relationship: more rain means fewer trips.
The data compares dry-weather rides to rainy-day rides.
Trips drop sharply when it rains, especially from May to September.
Even in good seasons, rain discourages biking

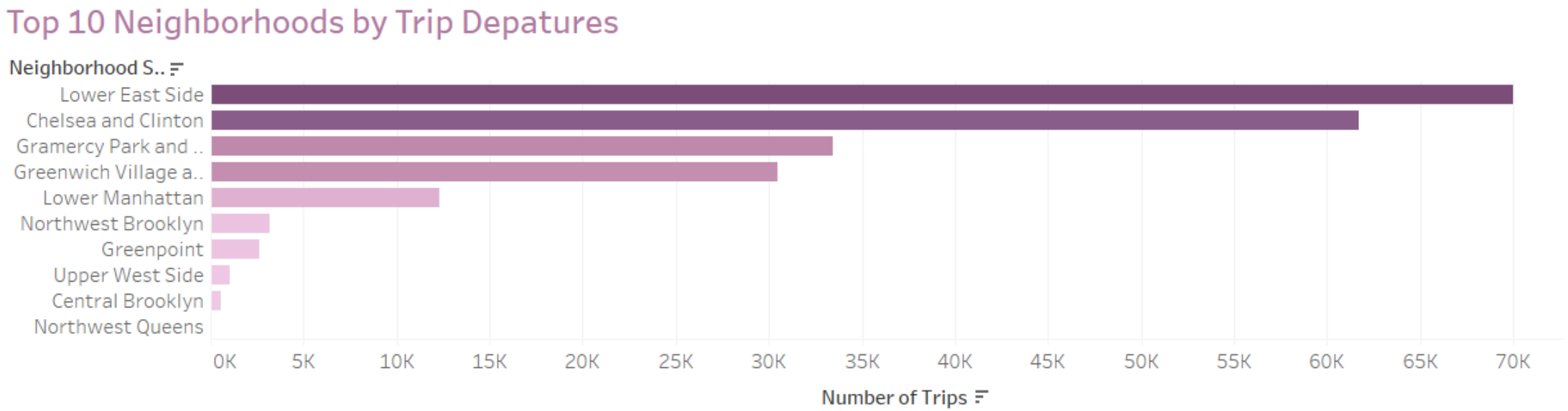
# Where Do Bike Trips Begin? – By Borough (Manhattan Leads the Way)




| Borough Start | Usertype | |
|---|---|---|
| Brooklyn | Customer | 1480 |
| | Subscriber | 5029 |
| Manhattan | Customer | 16802 |
| | Subscriber | 194450 |
| Queens | Customer | 1 |
| | Subscriber | 43 |

•Manhattan clearly stands out as the most active borough for bike trips, far surpassing Brooklyn and Queens. This confirms a strong concentration of cycling activity in Manhattan, likely due to its population density, tourism, and cycling infrastructure.

The bar chart is based on a 5-year period to highlight borough-level trends, but Queens does not appear visually due to its extremely low trip count. The adjacent table shows 2-year data, which makes it possible to display Queens despite its limited activity. This design choice ensures both readability and completeness




- This chart shows the ten neighborhoods with the highest number of bike trip departures over the year. Lower East Side and Chelsea and Clinton rank at the top, with a significantly higher volume of trips compared to other neighborhoods. This indicates strong cycling activity in these areas, likely due to their urban density, infrastructure, or overall attractiveness.

# Key Findings

## User Behavior

- Most riders are **subscribers**.
- Their usage **increases significantly from May to September**, peaking in **July**, then **drops sharply in winter** (November to February).
- **Non-subscribers follow the same seasonal trend**, but with fewer trips.

## Weather Impact

- **Higher temperatures** lead to more bike usage.
- **Rainy days** result in a significant drop in trip volume.

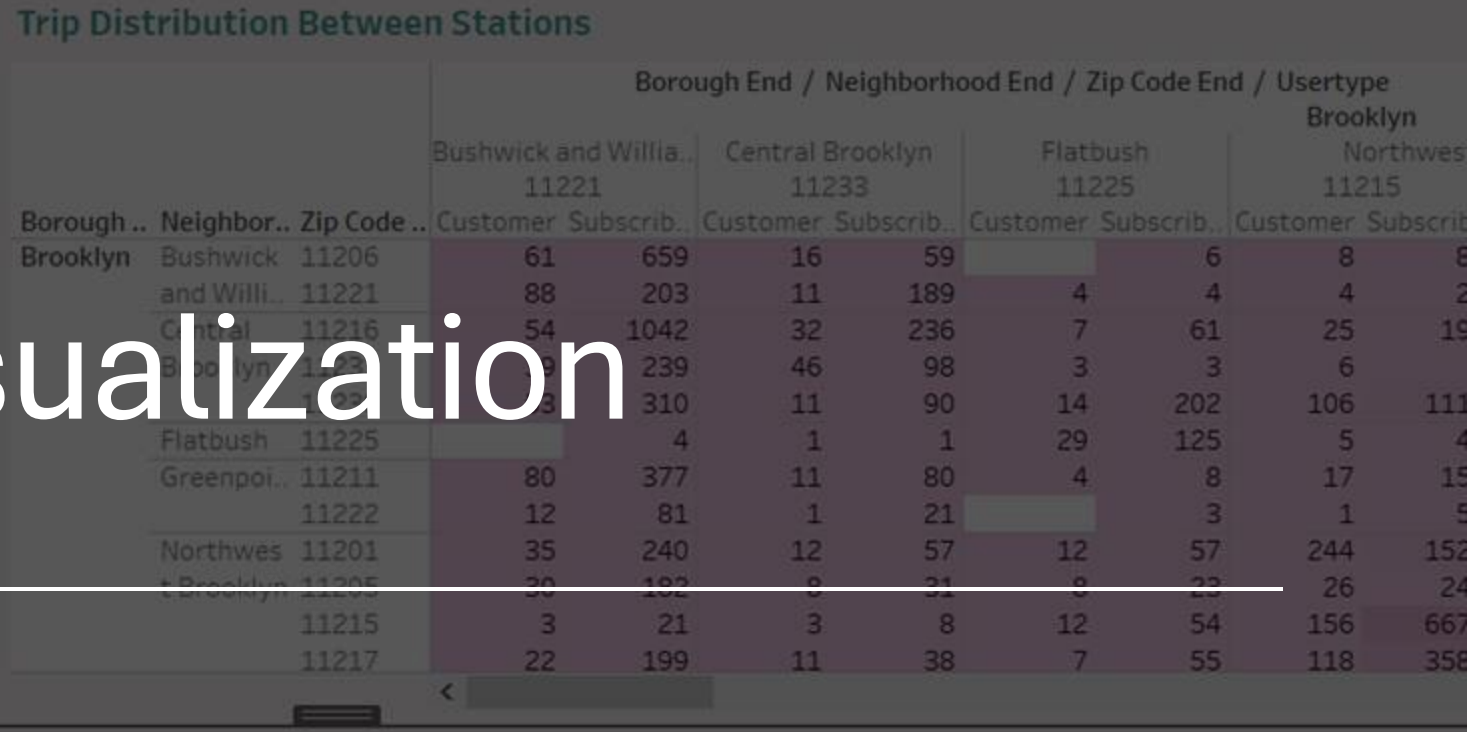
## Geographic Insights

- **Certain zones**, especially *Lower East Side* and *Chelsea & Clinton*, show the **highest cycling activity**.
- These are key areas for strategic infrastructure investment.

# Strategic Recommendations

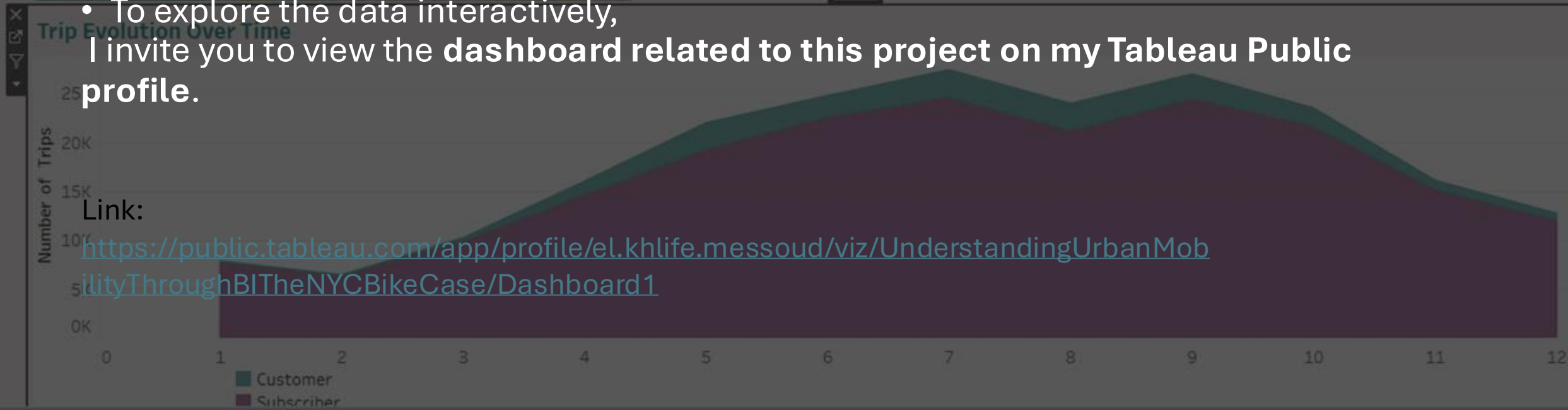
- Launch promotions in April and maintain them throughout the summer season (May to September).
- Offer discounts or rain-friendly gear on rainy days to encourage continued usage.
- Continuously reinforce stations in high-demand areas with more bikes during summer.
- Maintain targeted offers to convert casual riders into subscribers throughout the high season (May to September).

# Interactive Visualization


- To explore the data interactively,

I invite you to view the **dashboard related to this project on my Tableau Public profile**.




Link:
https://public.tableau.com/app/profile/el.khlife.messoud/viz/UnderstandingUrbanMobilityThroughBITheNYCBikeCase/Dashboard1