Bikeshare Network Expansion and transportation in Pittsburgh

By: Justin Do, Yuru Zhang, Xinkai Li, Haoji Wang

Table of Contents

Background

- **POGO Bike Share**
  - Launched in 2015, providing eco-friendly transportation for locals and visitors.
- **Network Expansion**
  - Continuous expansion with more stations and bikes added each year.
- **Key Government Policies**
  - 2014: Announced the first protected bike lane in downtown, with 5 miles of lanes planned by 2016.
- **Bike(+) Plan**
  - 2020: Introduced a plan to expand bike infrastructure from 70 miles to 220 miles over the next decade.
- **Policy Goal**
  - Enhance cycling safety, convenience, and promote sustainable transportation.
  - This version is more streamlined for easy presentation.

Objective

- Understand/Analyze Trends in Ridership Using Available Data
- Analyze the Effect of Bike Infrastructure Changes on Bus Ridership
- Evaluate the Impact of New Bike Lanes on Ridership
- Study the Effect of Bike Share Network Expansion on Transportation Choices

POGO Data Cleaning

- **Download the past 12 months of POGO trip data and station locations from February to October 2023, along with on-road bicycle pavement markings.**
- **Convert and import Excel files into CSV format for Jupyter Notebook / R Studio.**
- **Reformat the April 2023 data to match the format of other trip data.**
- **Combine all 12 months of trip data into a single data frame for easier analysis.**

Data sets

- Healthy rider
  - Stations
  - Trip data
- POGOH
  - Stations
  - Trip data
- Monthly Bus riders
  - Average daily riders
- Monthly Bus stoppage
  - Average boarding on and off per stop
- Transit Stop Usage
- Bike pavement markings
  - Additional bike lanes added per year

Hypothesis

1. Is there a difference in overall quantity/duration/boarding of ridership of POGOH/Healthy rider and Bus riders in and outside of Oakland

a. HO_a: there is a difference in average duration for POGOH outside and inside Oakland.

b. H1_a: there is not a difference in average duration for POGOH outside and inside Oakland.

   i. HO_a1: There is a difference in average duration for POGOH outside and inside Oakland based on season.

   ii. HO_a2: There is not a difference in average duration for POGOH outside and inside Oakland based on season.

c. HO_b: there is a difference in average duration for Healthy Rider outside and inside Oakland.

d. H1_b: there is not a difference in average duration for Healthy Rider outside and inside Oakland.

   i. HO_b1: There is a difference in average duration for Healthy Rider outside and inside Oakland based on season.

   ii. H1_b2: There is not a difference in average duration for Healthy Rider outside and inside Oakland based on season.

e. HO_c: There is a difference in average daily riders for any given bus route for any given day outside and inside Oakland.

f. H1_c: There is not a difference in average daily riders for any given bus route for any given day outside and inside Oakland.

   i. HO_c1: There is a difference in average daily riders for any given bus route for any given day outside and inside Oakland based on season.

   ii. H1_c2: There is not a difference in average daily riders for any given bus route for any given day outside and inside Oakland based on season.

g. HO_d: There is a difference in average # of riders boarding off/on for any given bus stop for any given day outside and inside Oakland

h. H1_d: There is not a difference in average # of riders boarding off/on for any given bus stop for any given day outside and inside Oakland

   i. HO_d1: There is a difference in average # of riders boarding on/off for any given bus route for any given stop outside and inside Oakland based on season.

   ii. H1_d2There is not a difference in average # of riders boarding on/off for any given bus route for any given stop outside and inside Oakland based on season.

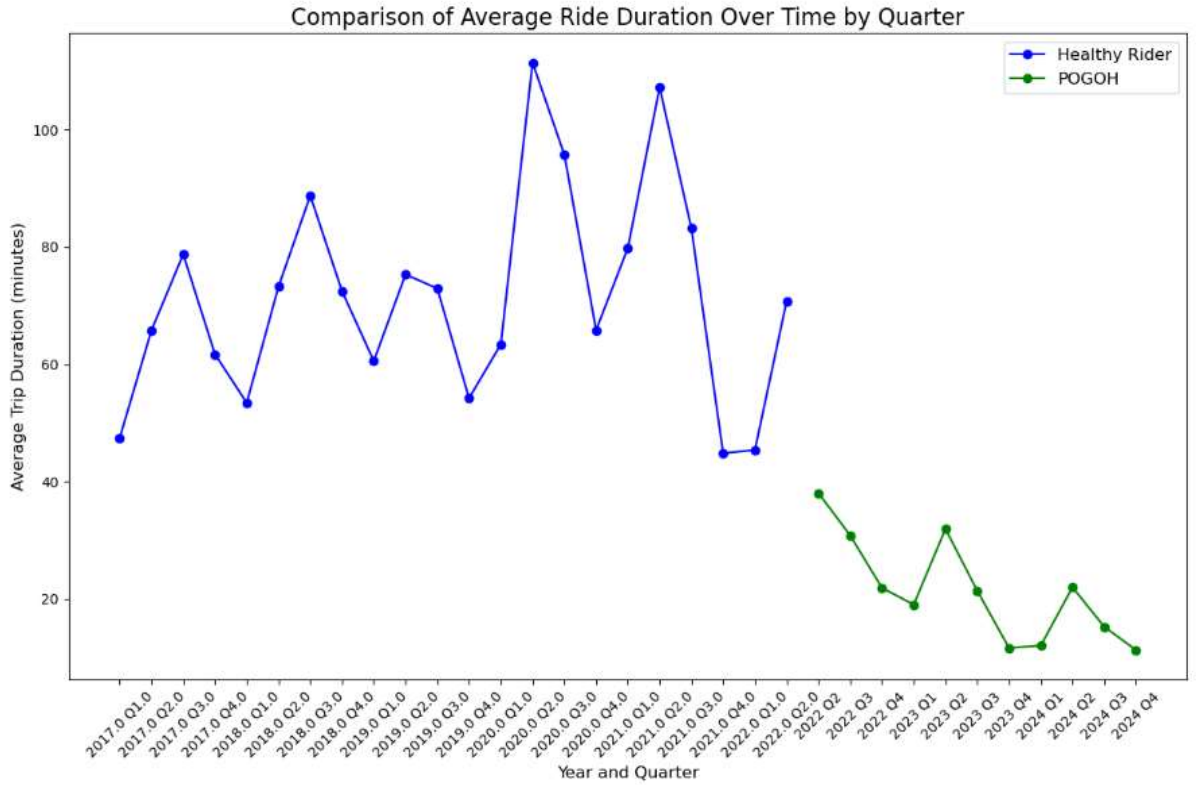
Hypothesis part 2

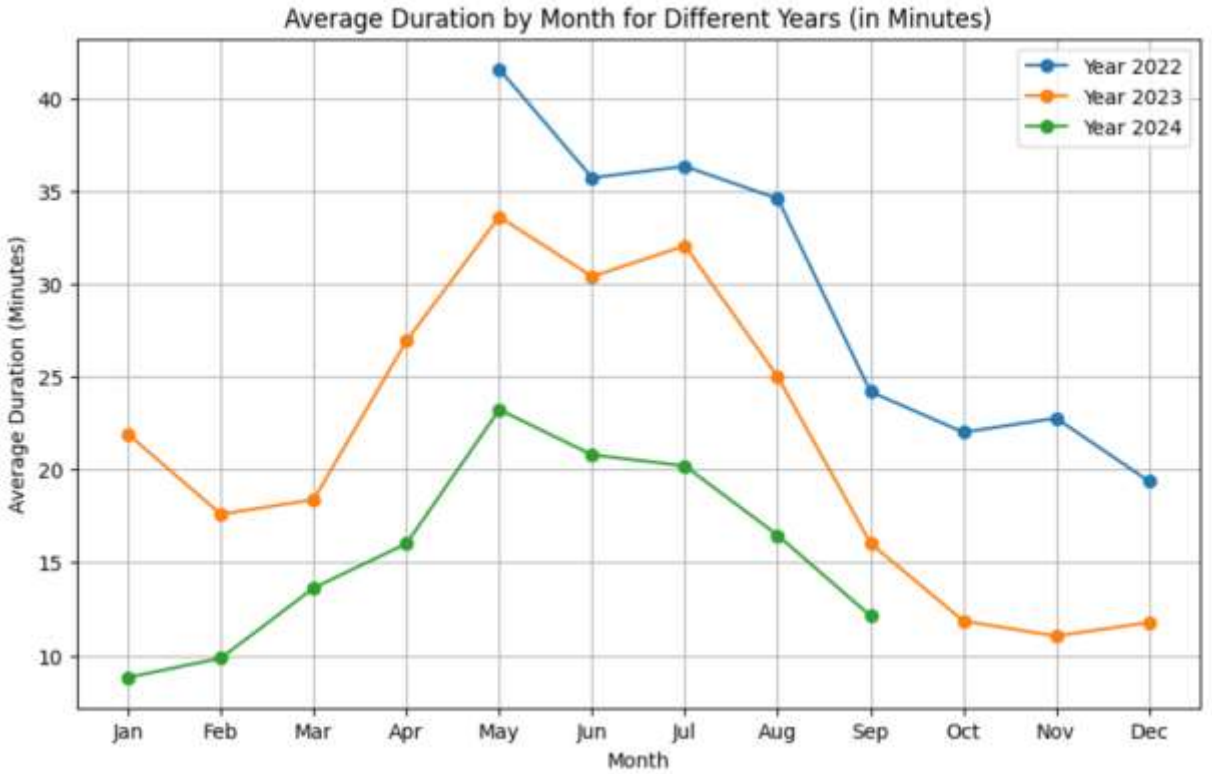
2. Introduction of new bike lanes & stations to bus ridership
    a. HO_e: The POGOH bike infrastructure does affect bus ridership in Pittsburgh
    b. H1_:e The POGOH bike infrastructure leads to a reduction in bus ridership in Pittsburgh
3. Introduction of new bike lanes near existing stations increases POGOH ridership
    a. HO_f: There is no significant effect in ridership before and after the addition of bike lanes
    b. H1_f: There is a significant effect in ridership with the addition of new bike lanes near existing stations
4. Introduction of new bike stations increases POGOH ridership at nearby stations
    a. HO_g: New bike stations does not have a significant effect on POGOH ridership at nearby stations
    b. H1_g: New bike stations does have a significant effect on POGOH ridership at nearby stations

POGOH & Healthy rider

	year	Total Stations	Stations Added
0	2015	50	50.0
1	2016	50	0.0
2	2017	52	2.0
3	2018	105	53.0
4	2019	114	9.0
5	2020	100	-14.0
6	2021	105	5.0
0	2022	37	37.0
1	2023	60	0.0

```
the average duration of bike ride from 2017-2024 Q1 is 37.52031769867208 minutes
the average duration of bike ride from 2017-2024 Q2 is 62.93017190583292 minutes
the average duration of bike ride from 2017-2024 Q3 is 47.42110760971055 minutes
the average duration of bike ride from 2017-2024 Q4 is 27.349986205268173 minutes
```





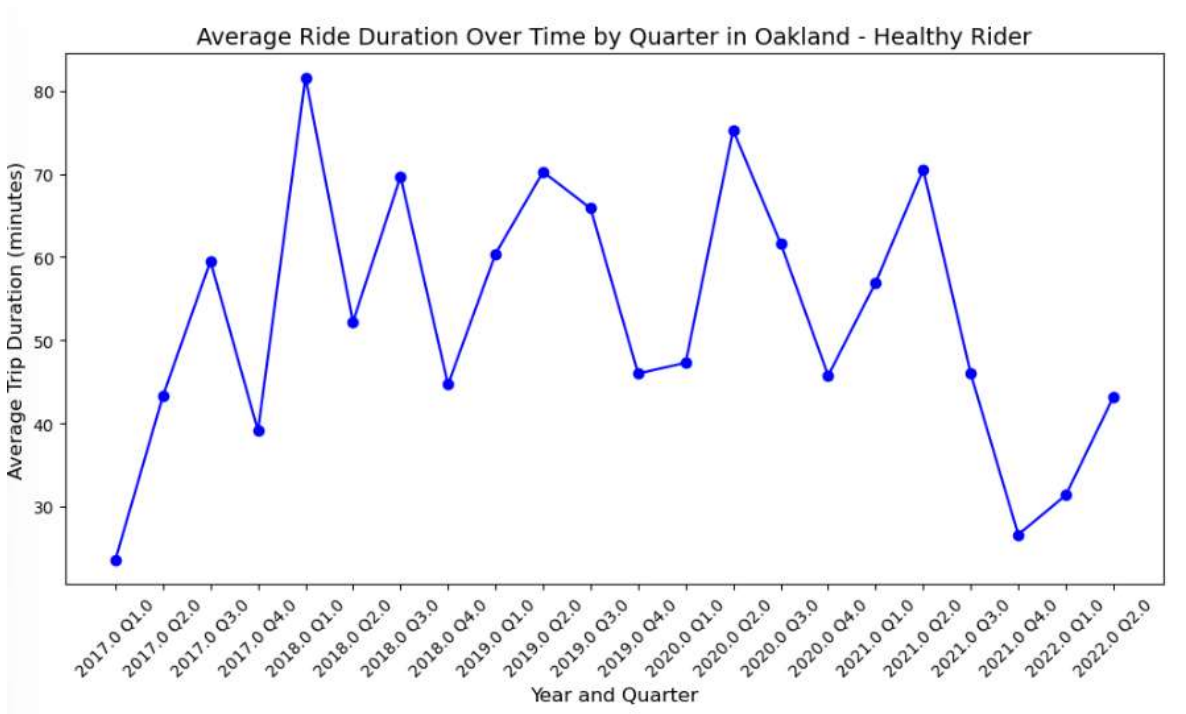
Healthy Rider

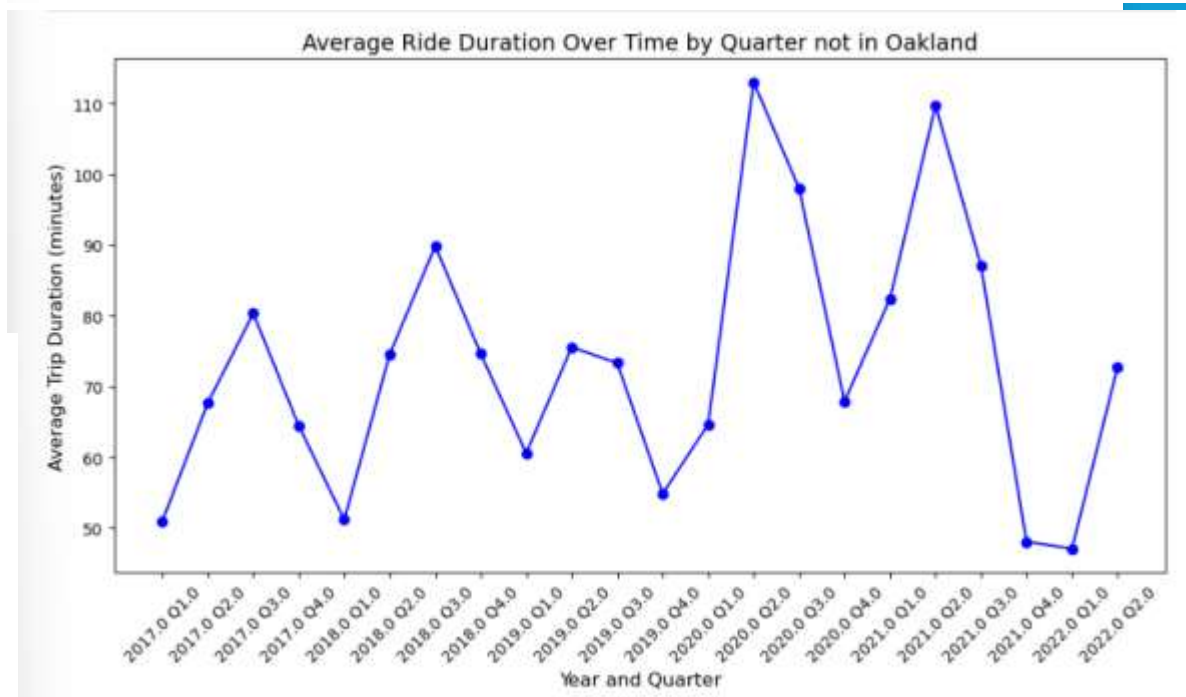
```
The average duration of Q1 Healthy riders within oakland is 47.00077550474662 minutes
The average duration of Q2 Healthy riders within oakland is 61.390866290018835 minutes
The average duration of Q3 Healthy riders within oakland is 58.044006951469406 minutes
The average duration of Q4 Healthy riders within oakland is 36.3611922005571 minutes
```

```
The average duration of Q1 Healthy riders not within oakland is 63.14483590165863 minutes
The average duration of Q2 Healthy riders not within oakland is 89.45684759033313 minutes
The average duration of Q3 Healthy riders not within oakland is 85.79569953960295 minutes
The average duration of Q4 Healthy riders not within oakland is 59.369845310485815 minutes
```

```
#S Milvale Ave / Center Ave
#O'Hara st / University Pl
#Robinson St / Allequippa St
#Oakland Ave / Fifth Ave
#Forbes Ave / Coltart Ave
# Coltart Ave / Zulema St
#atwood St/ Bates St
#Boulevard of Allies / Parkview Ave
#Forbes ave / S Bouquet St
#Forbes ave / Schenley Drive
#Schenley Drive Extension / Schenley drive
#Fossie way / Forbes Ave
#Center Ave / N Dithridge St
#N Neville St / Ellsworth Ave
#Technology Drive
```

	year	Total Stations	Stations Added
0	2015	50	50.0
1	2016	50	0.0
2	2017	52	2.0
3	2018	105	53.0
4	2019	114	9.0
5	2020	100	-14.0
6	2021	105	5.0



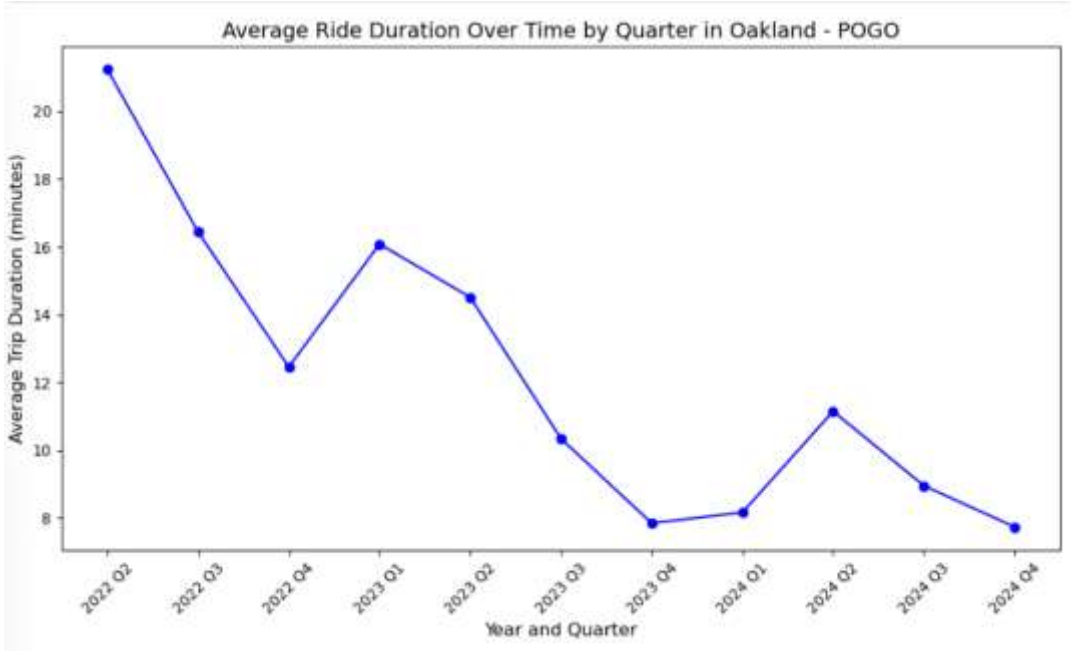


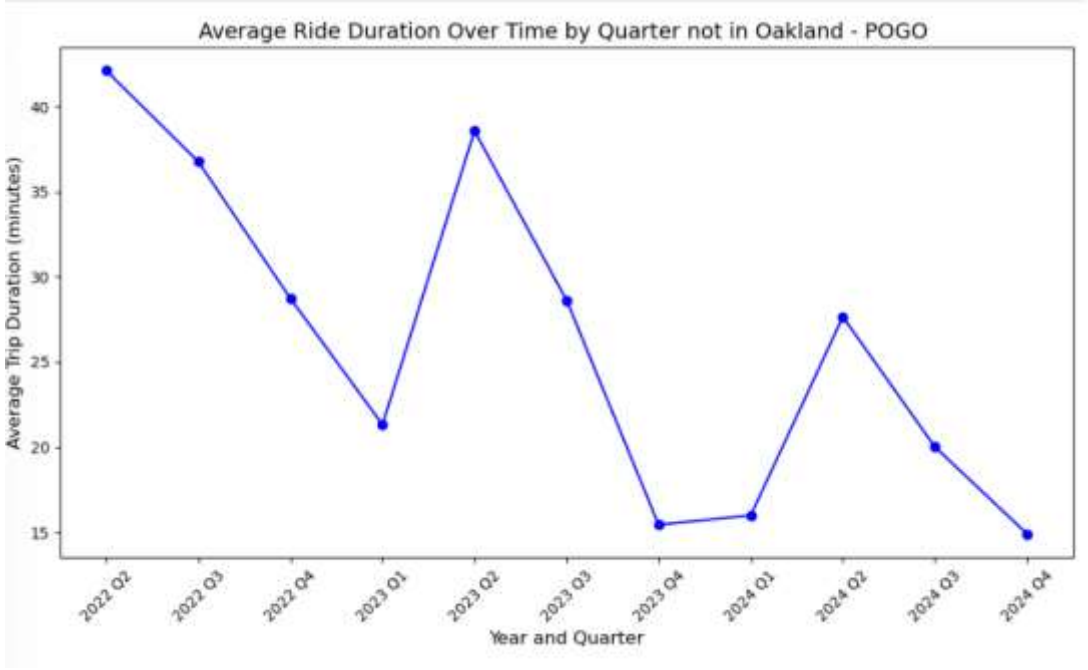
POGOH

```
The average duration of Q1 Pogo riders within oakland is 10.080798340561058 minutes
The average duration of Q2 Pogo riders within oakland is 11.75690890790136 minutes
The average duration of Q3 Pogo riders within oakland is 9.45819135052428 minutes
The average duration of Q4 Pogo riders within oakland is 8.27060722113833 minutes


The average duration of Q1 Pogo riders not within oakland is 17.558208860393925 minutes
The average duration of Q2 Pogo riders not within oakland is 29.438102652883217 minutes
The average duration of Q3 Pogo riders not within oakland is 22.01725436587108 minutes
The average duration of Q4 Pogo riders not within oakland is 17.073157267030524 minutes
```

```
#S Milvale Ave / Center Ave
#O'Hara st / University Pl
#Robinson St / Allequippa St
#Oakland Ave / Fifth Ave
#Forbes Ave / Coltart Ave
# Coltart Ave / Zulema St
#atwood St/ Bates St
#Boulevard of Allies / Parkview Ave
#Forbes ave / S Bouquet St
#Forbes ave / Schenley Drive
#Schenley Drive Extension / Schenley drive
#Fossie way / Forbes Ave
#Center Ave / N Dithridge St
#N Neville St / Ellsworth Ave
#Technology Drive
```

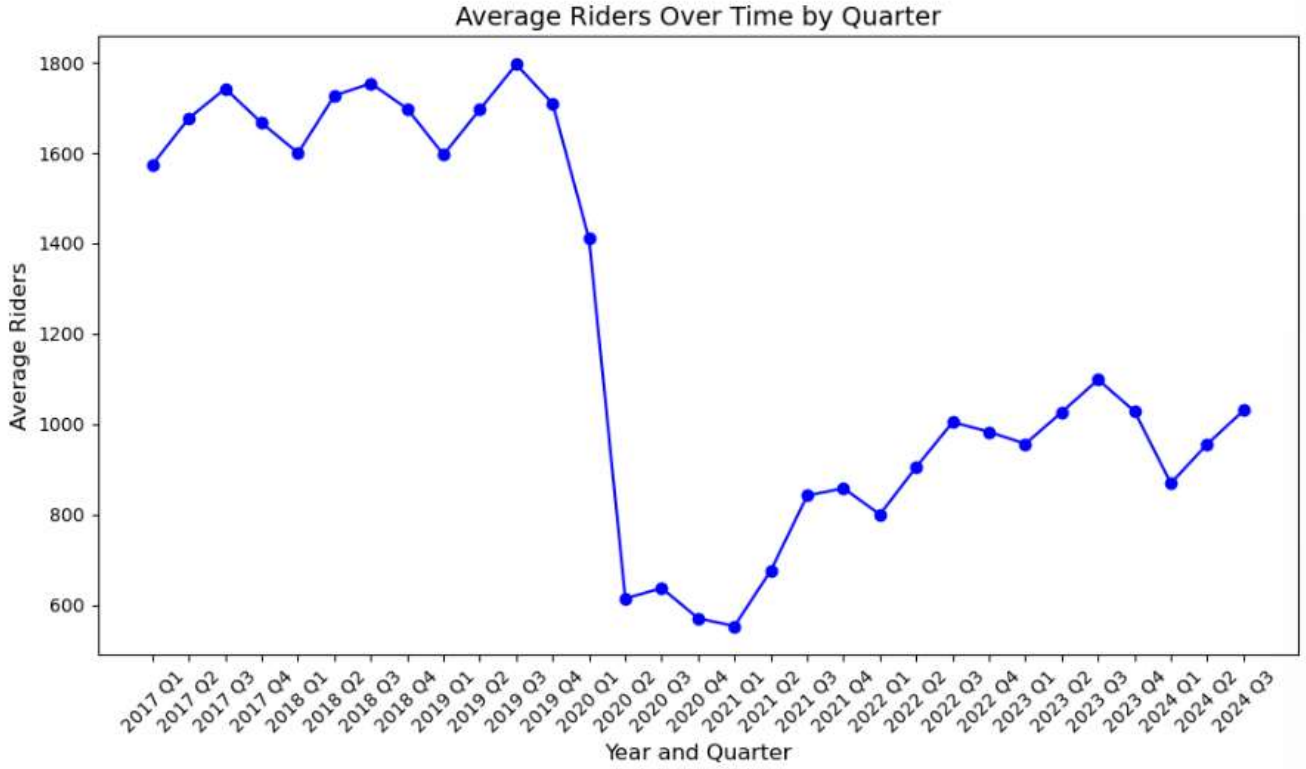
	Year	Total Stations	Stations Added
0	2022	37	37.0
1	2023	60	0.0





Overall Bus ridership from 2017-2024

```
the average amount of daily riders for Q1 from 2017 to 2024 is 1163.2434447729995 Riders
the average amount of daily riders for Q2 from 2017 to 2024 is 1155.8342379150213 Riders
the average amount of daily riders for Q3 from 2017 to 2024 is 1233.6305719253987 Riders
the average amount of daily riders for Q4 from 2017 to 2024 is 1209.4740036366554 Riders
```

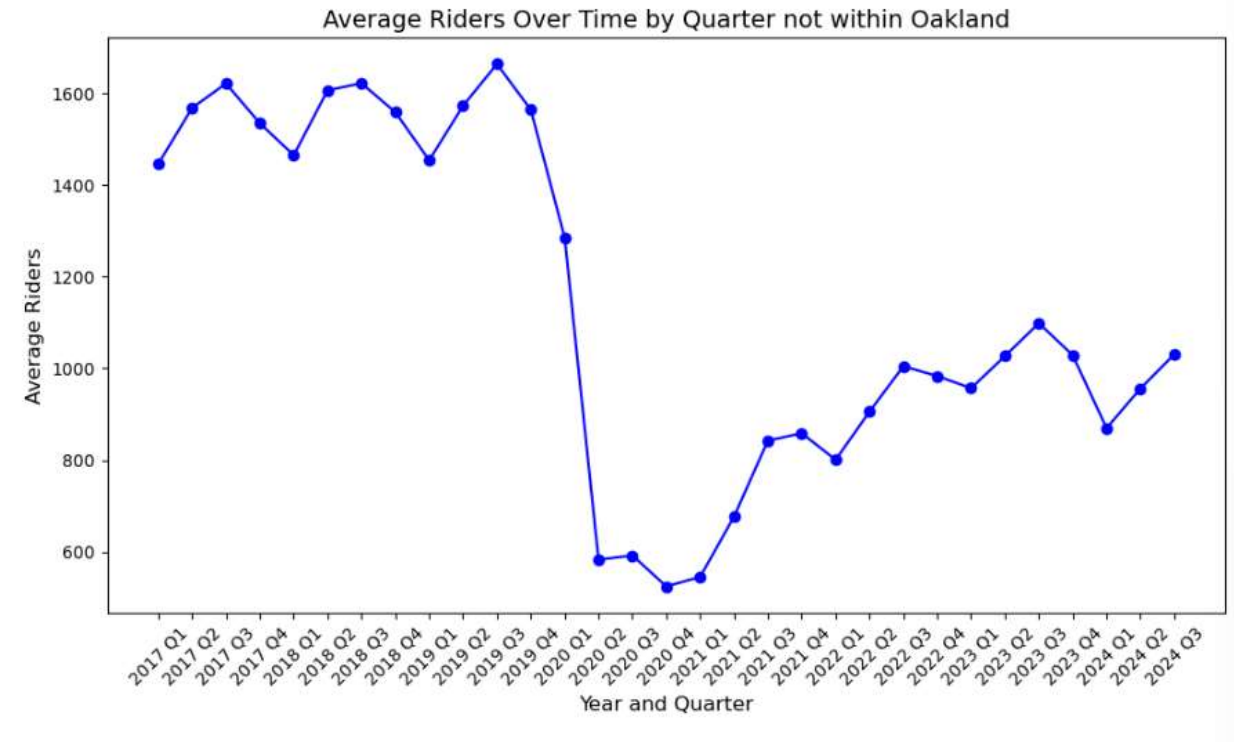


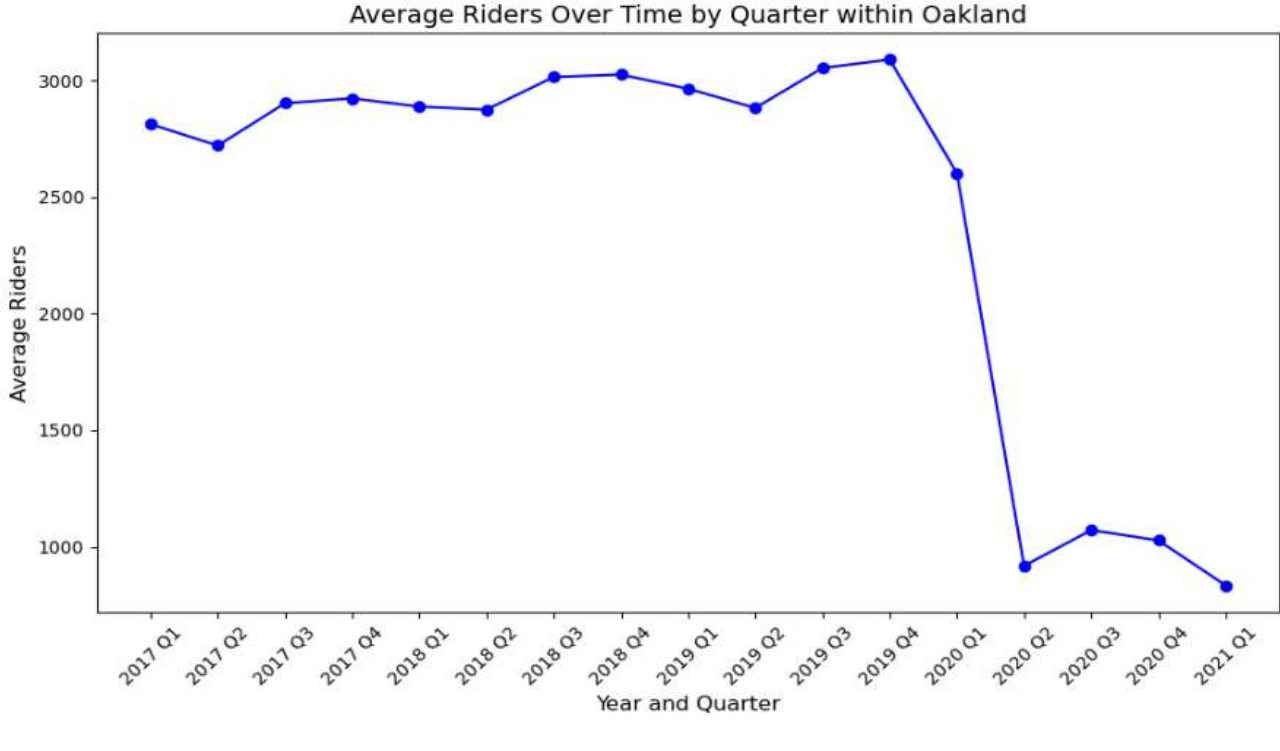

```
#54 - Northside-Oakland-South side
#61A - North Braddock (passes through Oakland)
#61B - Braddock-Swissvale (passes through Oakland)
#61C - McKeesport-Homestead (passes through Oakland)
#61D - Murray (passes through Oakland)
#75 - Ellsworth (passes through Oakland and Shady side)
#81 - Oak Hill (local oakland area)
#P3 - East Busway-Oakland (mentions Oakland within the name)
#cleaning data set and remove all NAN values from any of the rows.
```

Daily bus riders on any given bus route on any given day

```
The average riders for Q1 for any given bus route on any given day within the oakland area is 2664.205126000482 riders
The average riders for Q2 for any given bus route on any given day within the oakland area is 2348.825956808343 riders
The average riders for Q3 for any given bus route on any given day within the oakland area is 2510.3721986406154 riders
The average riders for Q4 for any given bus route on any given day within the oakland area is 2516.1002617935424 riders
```

```
The average riders for Q1 for any given bus route on any given day not within the oakland area is 1083.953633019941 riders
The average riders for Q2 for any given bus route on any given day not within the oakland area is 1097.821746887606 riders
The average riders for Q3 for any given bus route on any given day not within the oakland area is 1171.6710518053956 riders
The average riders for Q4 for any given bus route on any given day not within the oakland area is 1136.3140180261744 riders
```





Average on and off boarding

```
#print(Monthly_Updating_Bus_Stop_Usage['route_name'].unique())
#looking at all of the route names and seperating two data sets for inside and outside oakland
#54 - Northside-Oakland-South side
#61A - North Braddock (passes through Oakland)
#61B - Braddock-Swissvale (passes through Oakland)
#61C - McKeesport-Homestead (passes through Oakland)
#61D - Murray (passes through Oakland)
#75 - Ellsworth (passes through Oakland and Shady side)
#81 - Oak Hill (Local oakland area)
#P3 - East Busway-Oakland (mentions Oakland within the name)

#print(Monthly_Updating_Bus_Stop_Usage['datekey'].unique())
#there is only data of
#2019 - 09 - september Q3
# 2020 - 01 - January Q1
# 2020 - 09 - Septemver Q3
# 2021 - 04 - April Q2
#see the average boarding on/off for the entire data set
print('the average number of riders boarding per stop is',Monthly_Updating_Bus_Stop_Usage['avg_ons'].mean())
print('the average number of riders boarding per stop is',Monthly_Updating_Bus_Stop_Usage['avg_offs'].mean())
```

```
the average number of riders boarding per stop is 6.1394704920988445
the average number of riders boarding per stop is 6.117055338053056
```

```
Monthly_boarding_oakland_ids = ['054', '061A', '061B', '061C', '061D', '075', '081', 'P3']
Monthly_Boarding_oakland = Monthly_Updating_Bus_Stop_Usage[Monthly_Updating_Bus_Stop_Usage['route_name'].isin(Monthly_boarding_oakland_ids)]
Monthly_Boarding_nonoakland = Monthly_Updating_Bus_Stop_Usage[~Monthly_Updating_Bus_Stop_Usage['route_name'].isin(Monthly_boarding_oakland_ids)]

#Looking at the average
print('the average number of riders boarding on within oakland is',Monthly_Boarding_oakland['avg_ons'].mean())
print('the average number of riders boarding off within oakland is',Monthly_Boarding_oakland['avg_offs'].mean())

print('the average number of riders boarding on not within oakland is', Monthly_Boarding_nonoakland['avg_ons'].mean())
print('the average number of riders boarding off not within oakland is', Monthly_Boarding_nonoakland['avg_offs'].mean())
```

```
the average number of riders boarding on within oakland is 44.487042119916694
the average number of riders boarding off within oakland is 44.328374583575965
the average number of riders boarding on not within oakland is 6.077961606587996
the average number of riders boarding off not within oakland is 6.055764427595076
```

Hypothesis 1

1. Is there a difference in overall quantity/duration/boarding of ridership of POGOH/Healthy rider and Bus riders in and outside of Oakland
    a. HO_a: there is a difference in average duration for Healthy Rider outside and inside Oakland.
    b. H1_a: there is not a difference in average duration for Healthy Rider outside and inside Oakland.
        i. HO_a1: There is a difference in average duration for Healthy Rider outside and inside Oakland based on season.
        ii. H1_a2: There is not a difference in average duration for Healthy Rider outside and inside Oakland based on season
    c. HO_b: there is a difference in average duration for POGOH outside and inside Oakland.
    d. H1_b: there is not a difference in average duration for POGOH outside and inside Oakland.
        i. HO_b1: There is a difference in average duration for POGOH outside and inside Oakland based on season.
        ii. HO_b2: There is not a difference in average duration for POGOH outside and inside Oakland based on season.
    e. HO_c: There is a difference in average daily riders for any given bus route for any given day outside and inside Oakland.
    f. H1_c: There is not a difference in average daily riders for any given bus route for any given day outside and inside Oakland.
        i. HO_c1: There is a difference in average daily riders for any given bus route for any given day outside and inside Oakland based on season.
        ii. H1_c2: There is not a difference in average daily riders for any given bus route for any given day outside and inside Oakland based on season.
    g. HO_d: There is a difference in average on/off boarding for any given bus stop for any given day outside and inside Oakland
    h. H1_d: There is not a difference in average on/off boarding for any given bus stop for any given day outside and inside Oakland
        i. HO_d1: There is a difference in average # of riders boarding on/off for any given bus route for any given stop outside and inside Oakland based on season.
        ii. H1_d2There is not a difference in average # of riders boarding on/off for any given bus route for any given stop outside and inside Oakland based on season.

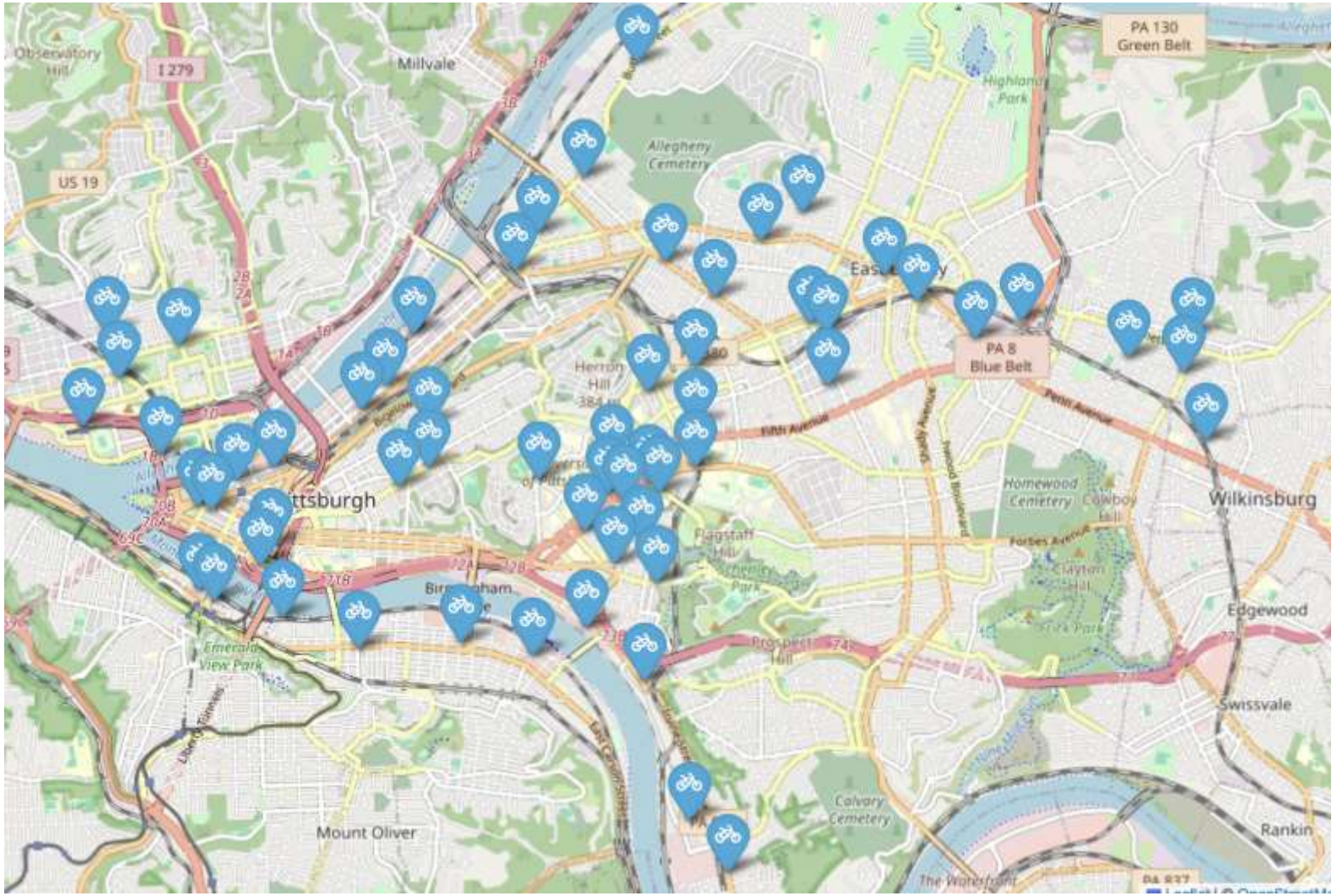
Hypothesis 2

- Introduction of new bike lanes & stations to bus ridership
  a. HO_e: The POGOH bike infrastructure does not affect bus ridership in Pittsburgh
  b. H1_e: The POGOH bike infrastructure leads to a reduction in bus ridership in Pittsburgh

POGOH Bike Station Effect

The influence of new bicycle stations on the ridership of nearby buses was studied.

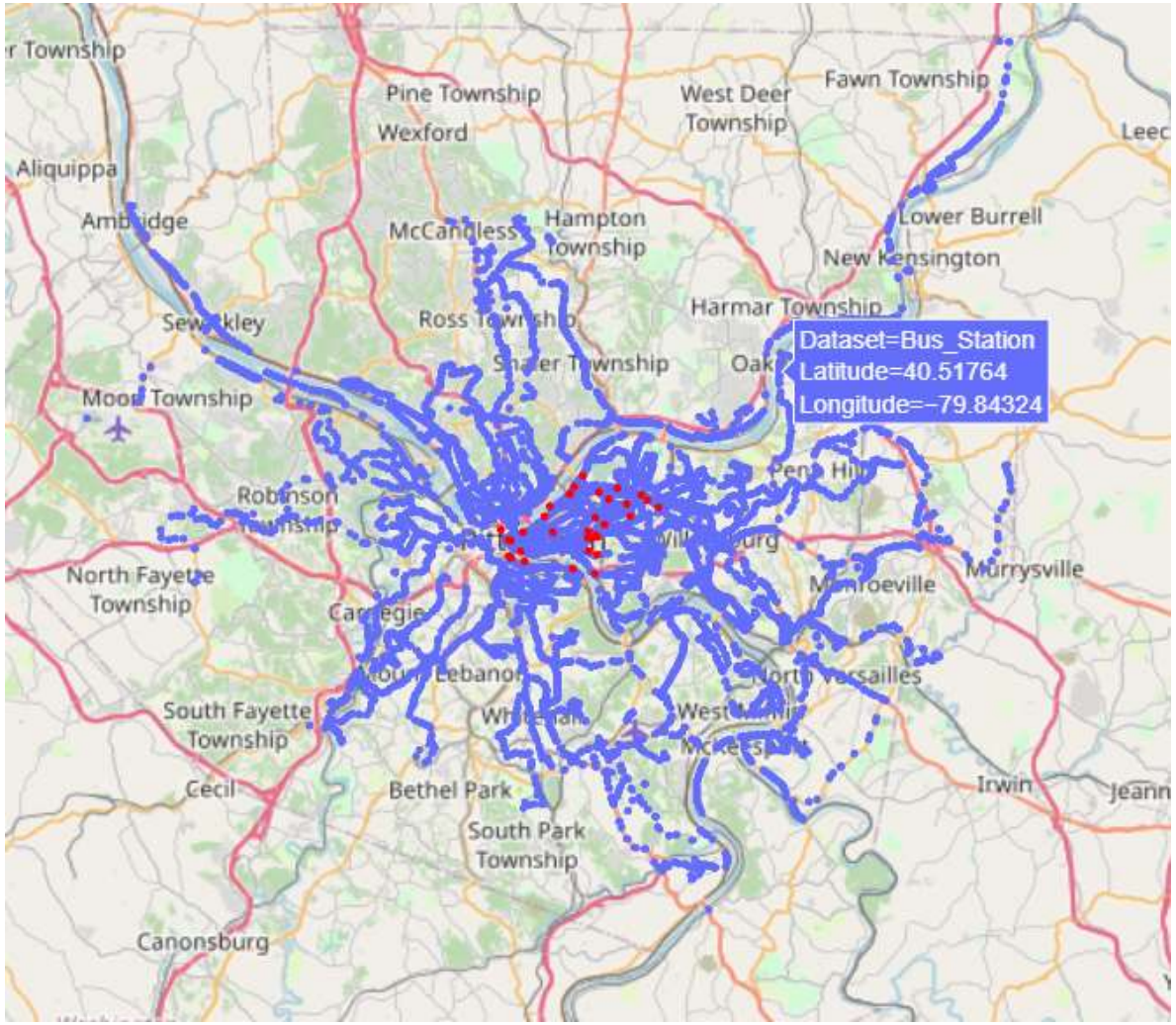
```
import pandas as pd
import folium

# Load the bike station and bus stop data
bike_station_data = pd.read_csv('pogoh-station-locations-october-2024.csv')
bus_stop_data = pd.read_csv('Monthly_Updating_Bus_Stop_Usage.csv')

# Selecting relevant columns for bike stations and bus stops
bike_station_locations = bike_station_data[['Name', 'Latitude', 'Longitude']]
bus_stop_locations = bus_stop_data[['stop_name', 'latitude', 'longitude']].copy()

# Converting any potential data type issues for latitude and longitude columns
bus_stop_locations.loc[:, 'latitude'] = pd.to_numeric(bus_stop_locations['latitude'], errors='coerce')
bus_stop_locations.loc[:, 'longitude'] = pd.to_numeric(bus_stop_locations['longitude'], errors='coerce')

# Drop rows with invalid coordinates
bus_stop_locations = bus_stop_locations.dropna(subset=['latitude', 'longitude'])

# Create a base map centered around Pittsburgh
pittsburgh_map = folium.Map(location=[40.4406, -79.9959], zoom_start=12, tiles='OpenStreetMap')

# Limit markers to avoid performance issues
MAX_MARKERS = 100

# Adding a sample of bike stations to the map
for _, row in bike_station_locations.head(MAX_MARKERS).iterrows():
    folium.Marker(
        location=[row['Latitude'], row['Longitude']],
        popup=f"Bike Station: {row['Name']}",
        icon=folium.Icon(color='blue', icon='bicycle', prefix='fa')
    ).add_to(pittsburgh_map)

# Adding a sample of bus stops to the map
for _, row in bus_stop_locations.head(MAX_MARKERS).iterrows():
    folium.Marker(
        location=[row['latitude'], row['longitude']],
        popup=f"Bus Stop: {row['stop_name']}",
        icon=folium.Icon(color='red', icon='bus', prefix='fa')
    ).add_to(pittsburgh_map)

# Save the map to an HTML file
pittsburgh_map.save('pittsburgh_bike_bus_city_map.html')
```



The current location of the bicycle station in Pittsburgh

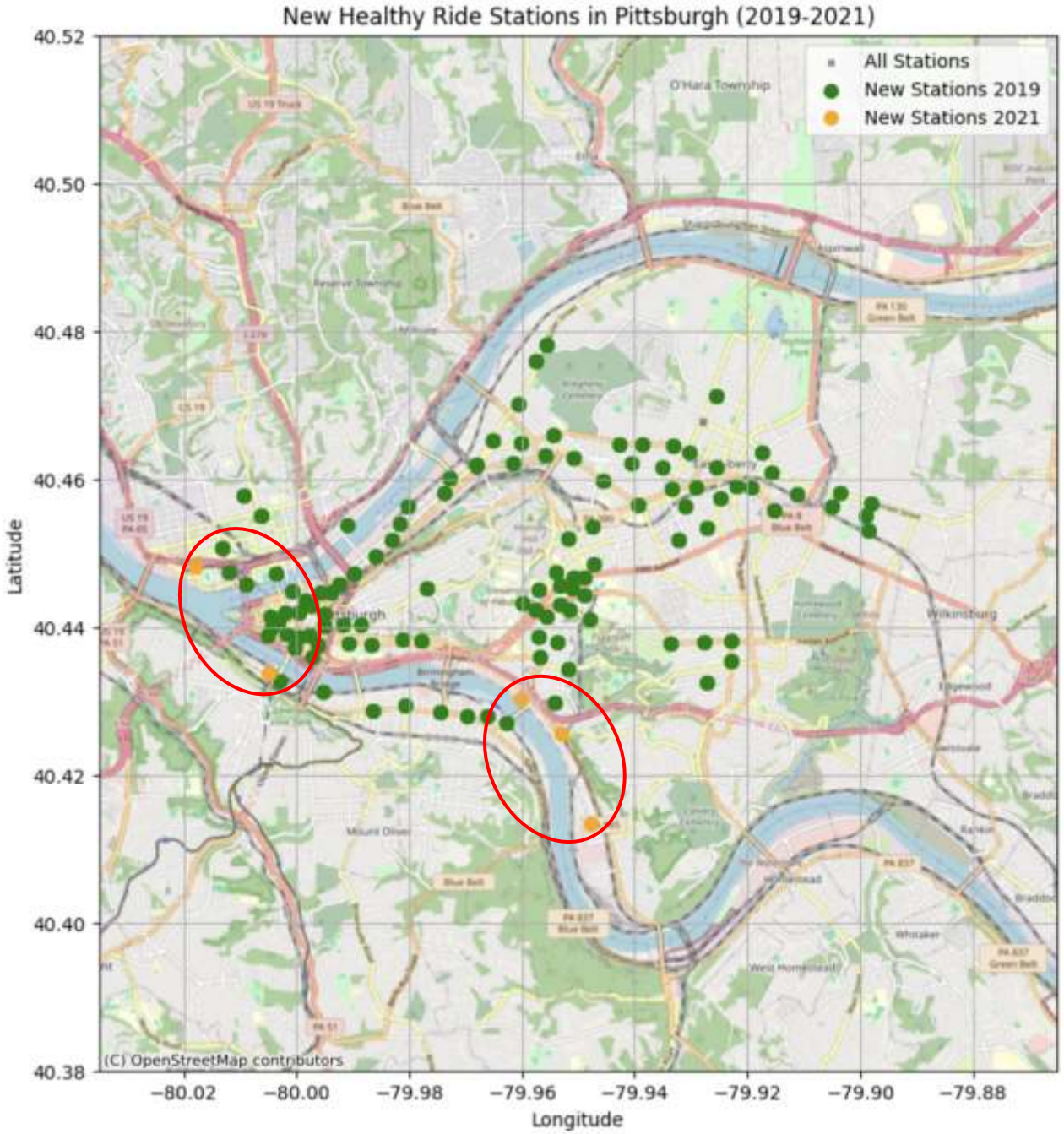
Bus and POGOH Stations

The bus stations and bicycle stations in Pittsburgh are shown on the map on the right. It can be found that the bicycle stations are mainly concentrated in the center of Pittsburgh, and are more obvious in the urban area and around the school.



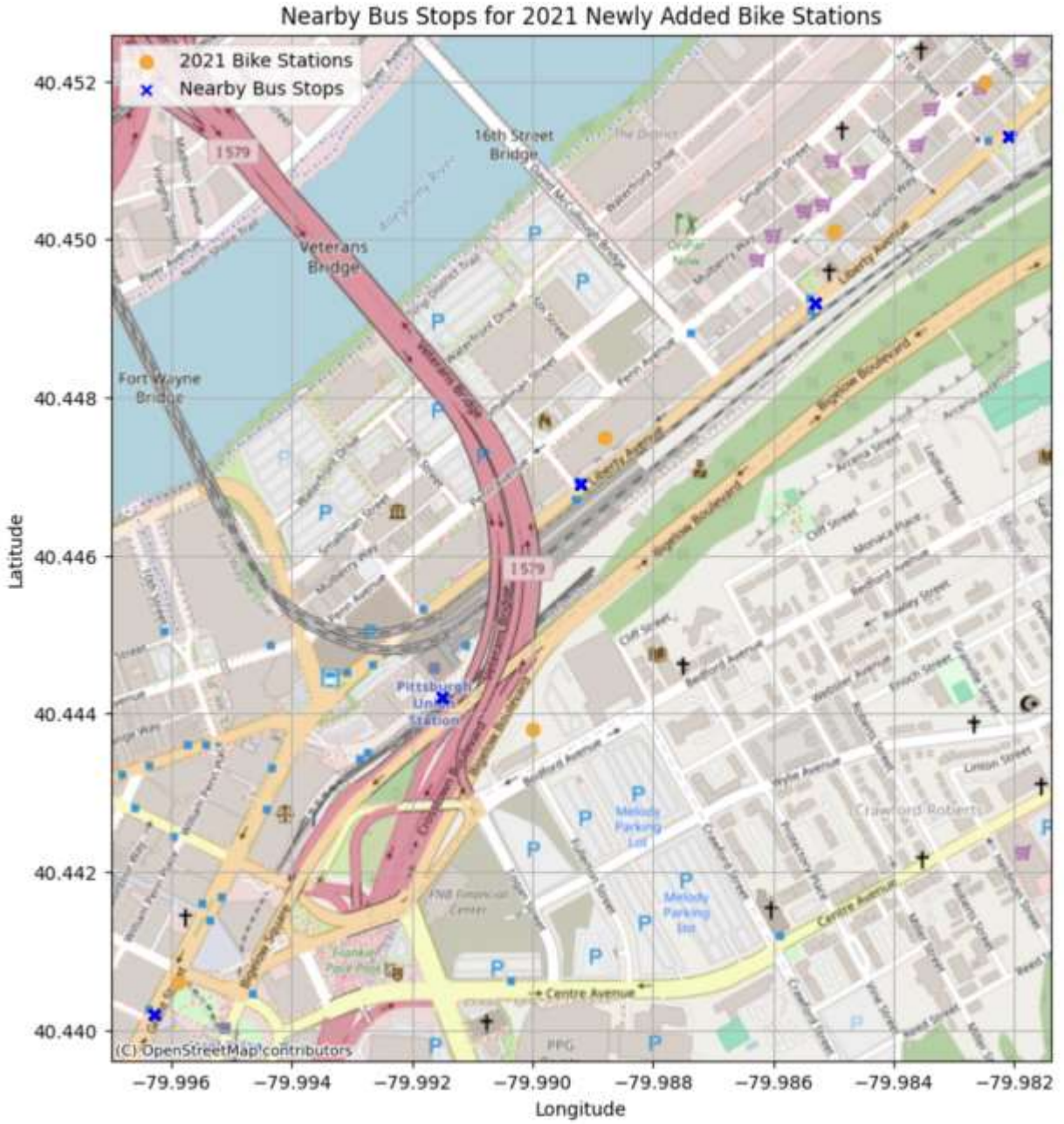
New Bike Station 2019-2021

Choosing to plot the growth of bicycle stations in the city of Pittsburgh over the three years from 2019 to 2021, the image shows in 2021 the addition of several new stations along the river and downtown in addition to the original number in 2019.



Relevant Bus Station

Look for a nearby bus stop based on the new bike stop found. Hypothesis 2 is analyzed based on the passenger flow changes of bus stations in the region shown in the statistical figure.

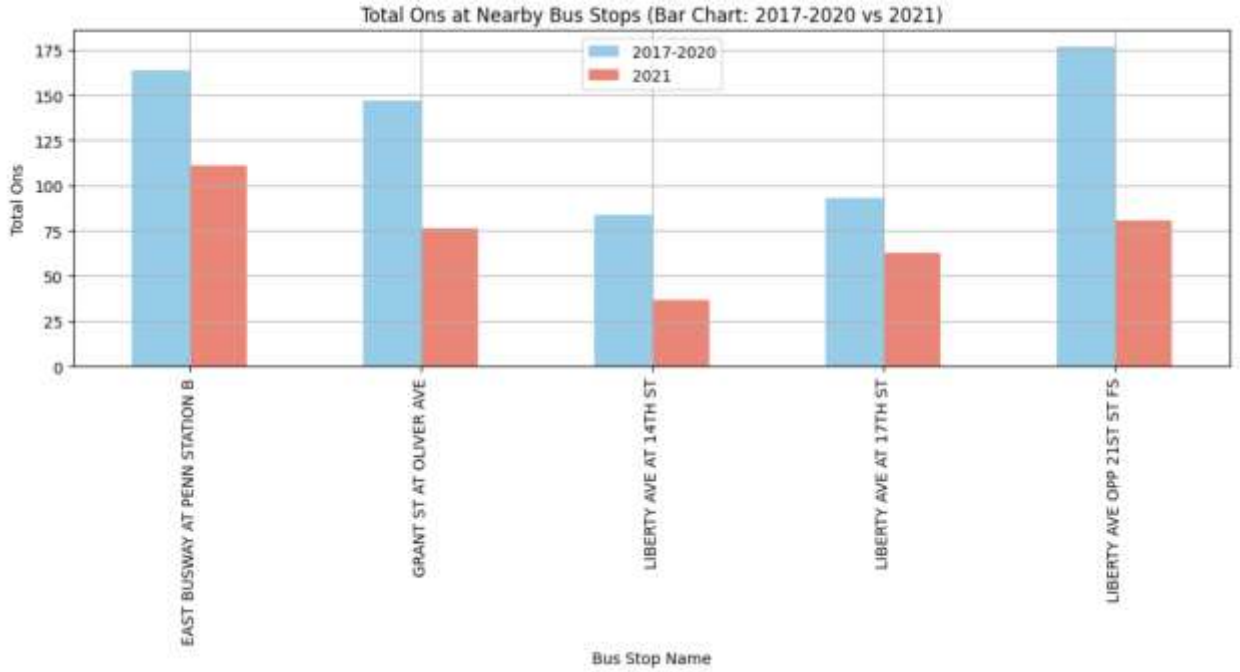


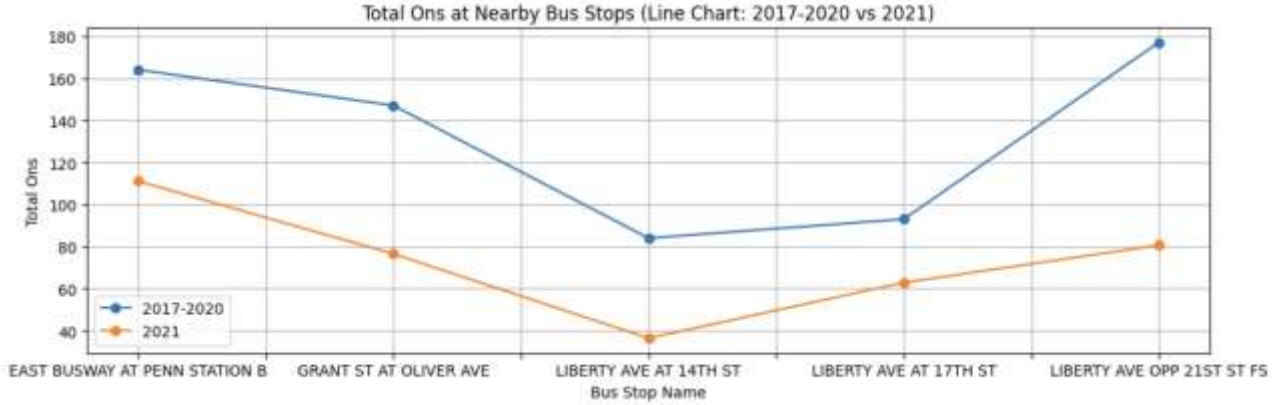
New Bike Station 2022-2024

Plot the increase of bicycle stations in the city of Pittsburgh in the nearly three years from 2023 to 2024 based on the number of stations in 2022, especially in the period of 2023, the number of increases is relatively large.

Relevant Bus Station Ridership

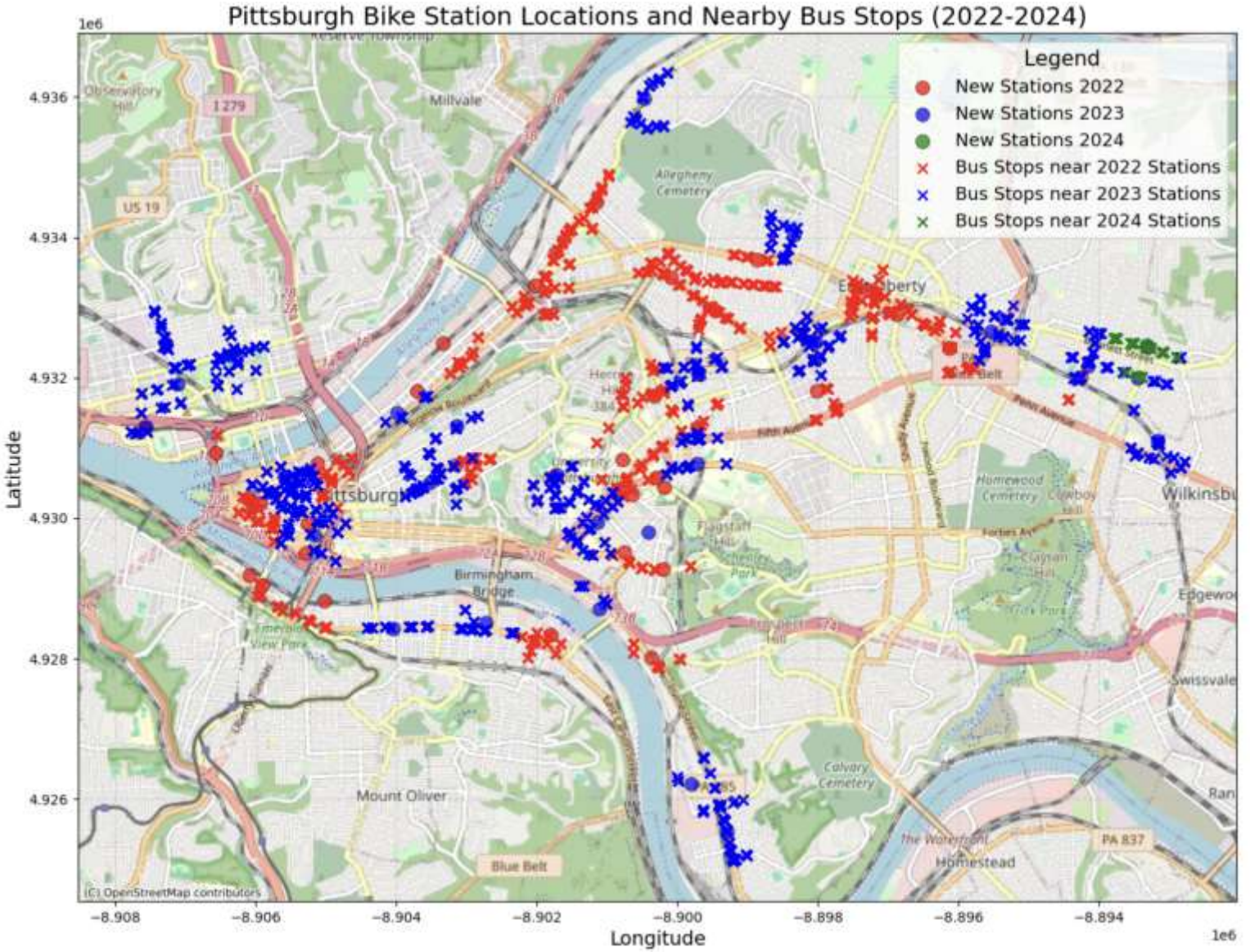
It can be seen that the passenger flow changes of buses near the new bike station in the past three years, especially in 2021. There is a downward trend, which indicates to some extent that bicycle stations affect bus ridership.





Relevant Bus Station

Due to the lack of specific bus station ridership in 2022-2024, it is impossible to accurately see the impact of bicycle stations on it. But based on the new bike stop found, look for a nearby bus stop. It can be seen that the new bicycle stations in the past three years have a great relationship with the bus stations in the urban area, the University of Pittsburgh and the surrounding densely populated areas (near large supermarkets, etc.).

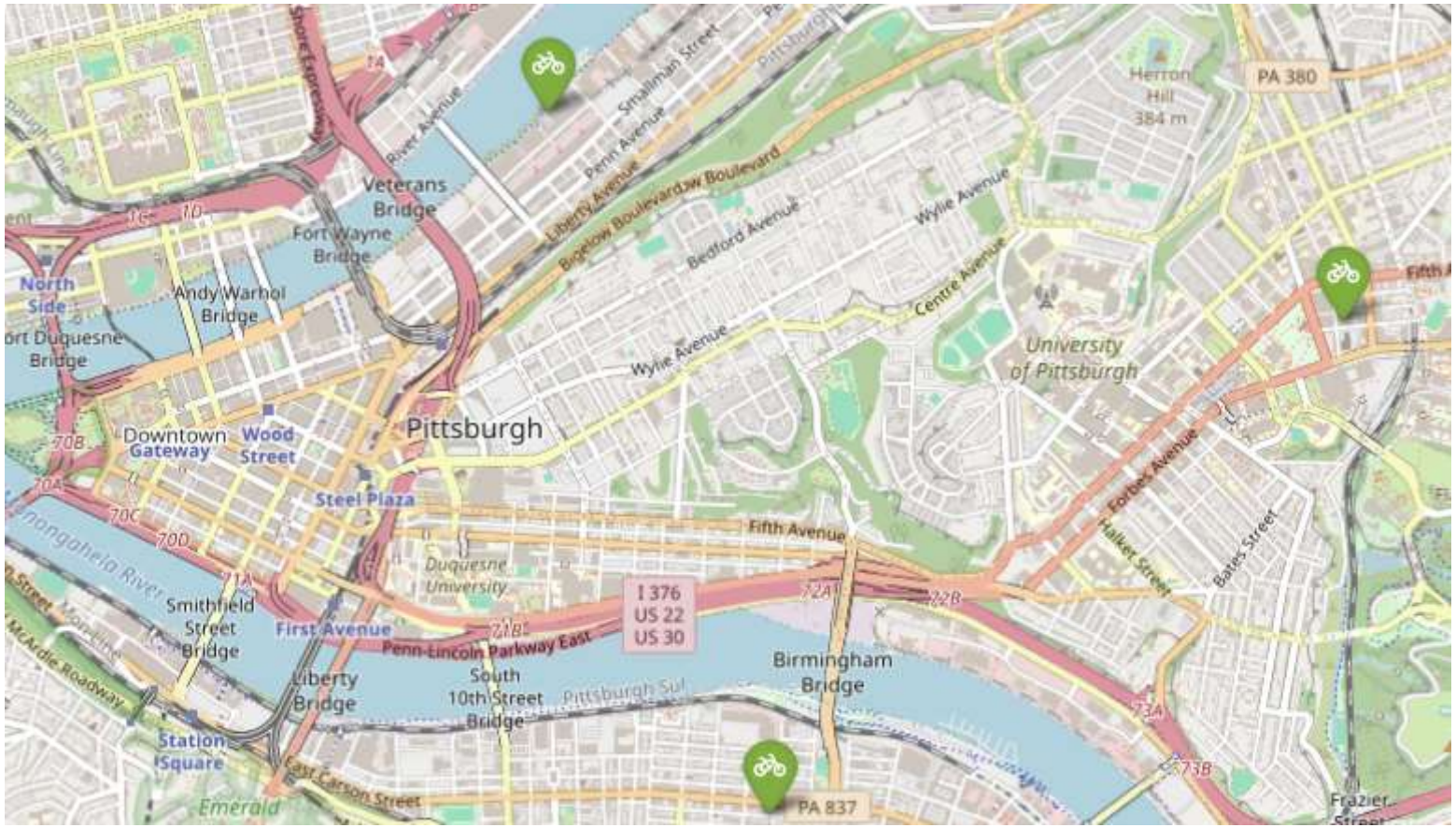


New POGOH Lane Effect

Compared with bicycle stations, the distribution of bicycle lanes is relatively scattered and discontinuous. Several bicycle lanes are selected to study their influence on the surrounding bus ridership.

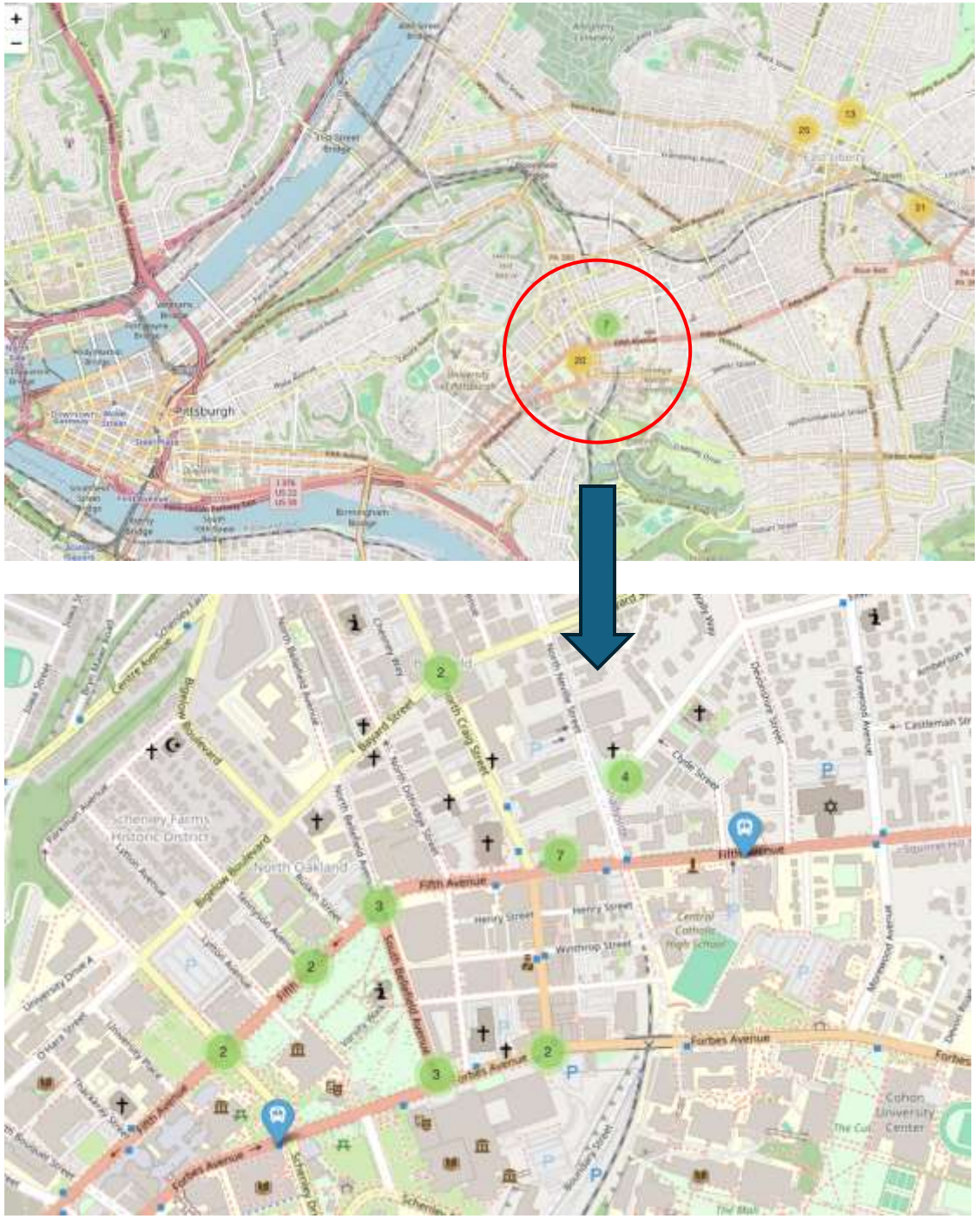
New POGOH Lane

Based on the relevant web sites, we selected several new bike lanes from 2019 to 2021, which are mainly in the area around the Pittsburgh metropolitan area.



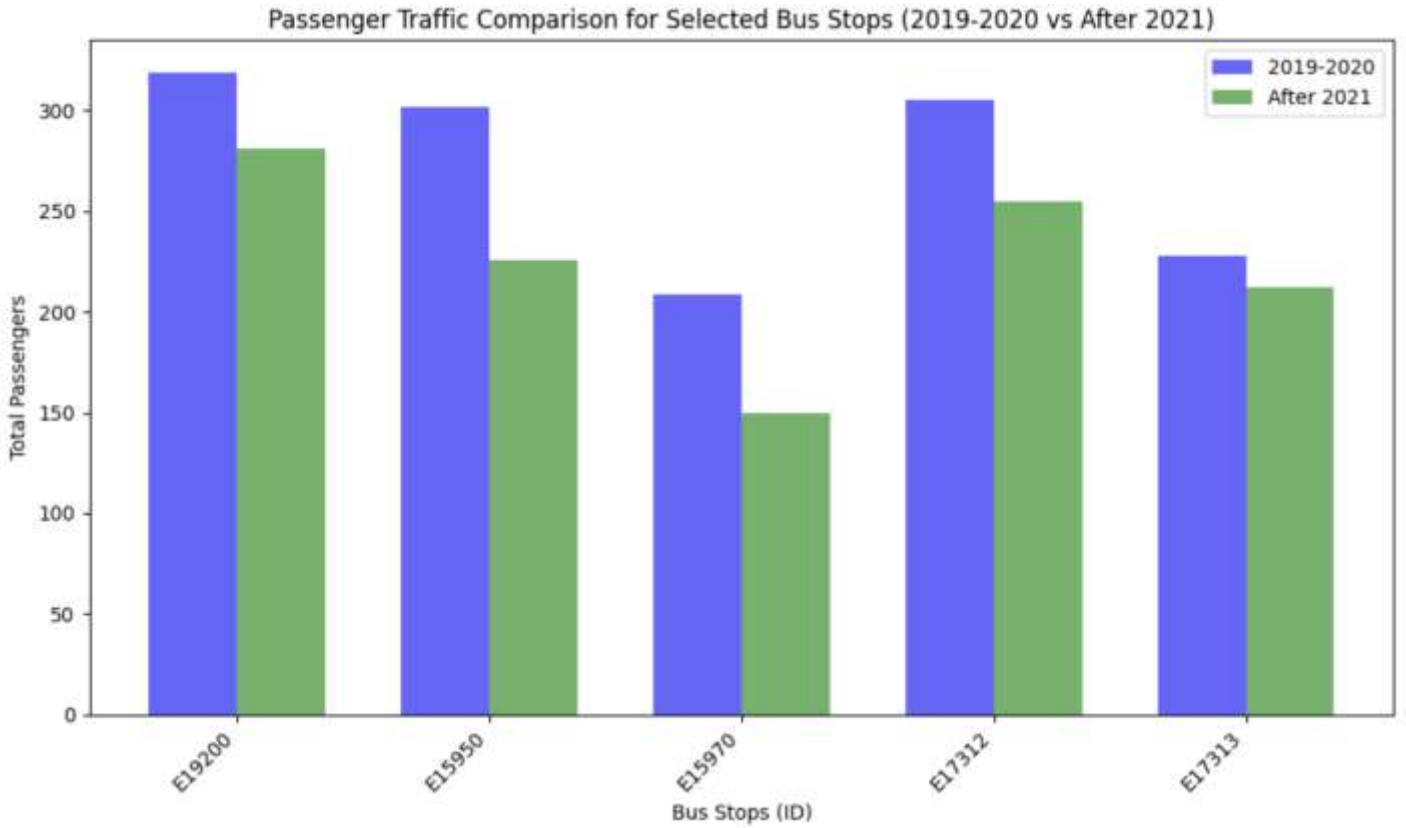
Relevant Bus Station

Taking the new bike lane near the University of Pittsburgh as an example, the bus stop within a radius of 500 meters is found, so as to study the impact of the new lane on passenger flow.



Relevant Bus Station Ridership

It can be seen that similar to the effect of the new bicycle station, the bus ridership near the new bicycle lane changes. There is a downward trend, especially with a new bike lane near the university, some students may choose to ride bikes instead of buses. This shows that to some extent, bike lanes affect bus ridership.



Regression Result

Current function value: 4.838832
Iterations 4

Poisson Regression Results

Dep. Variable:	Ridership	No. Observations:	5
Model:	Poisson	Df Residuals:	1
Method:	MLE	Df Model:	3
Date:	Thu, 05 Dec 2024	Pseudo R-squ.:	0.1279
Time:	13:21:20	Log-Likelihood:	-24.194
converged:	True	LL-Null:	-27.743
Covariance Type:	nonrobust	LLR p-value:	0.06883

	coef	std err	z	P>|z|	[0.025	0.975]
const	5.4293	0.066	81.981	0.000	5.300	5.559
Post	0.2019	0.089	2.261	0.024	0.027	0.377
Treatment	0.1892	0.079	2.403	0.016	0.035	0.344
Interaction	-0.2675	0.117	-2.289	0.022	-0.497	-0.038

- The post-intervention effect was significant (Post coefficient 0.2019 and significant), meaning that the addition of bike stops had a significant effect on ridership at certain bus stops.
- The Treatment effect was significant (the treatment coefficient was 0.1892 and significant), indicating that the change of some stations had a positive effect on the overall passenger flow.
- The interaction effect is significant and negative, indicating that the combination of different factors has opposite effects on passenger flow, and it may be necessary to further investigate which combination of factors leads to this negative effect.

Hypothesis 2

- Introduction of new bike lanes & stations to bus ridership
  a. HO_e: The POGOH bike infrastructure does not affect bus ridership in Pittsburgh
  b. H1_e: The POGOH bike infrastructure leads to a reduction in bus ridership in Pittsburgh

In the current data set on bicycles and buses, it can be found that the increase of bicycle stations and lanes can reduce the passenger flow of nearby bus stations to a certain extent and relieve certain pressure, especially in the peak period of bus passenger flow. And to make more accurate predictions, you need more data.

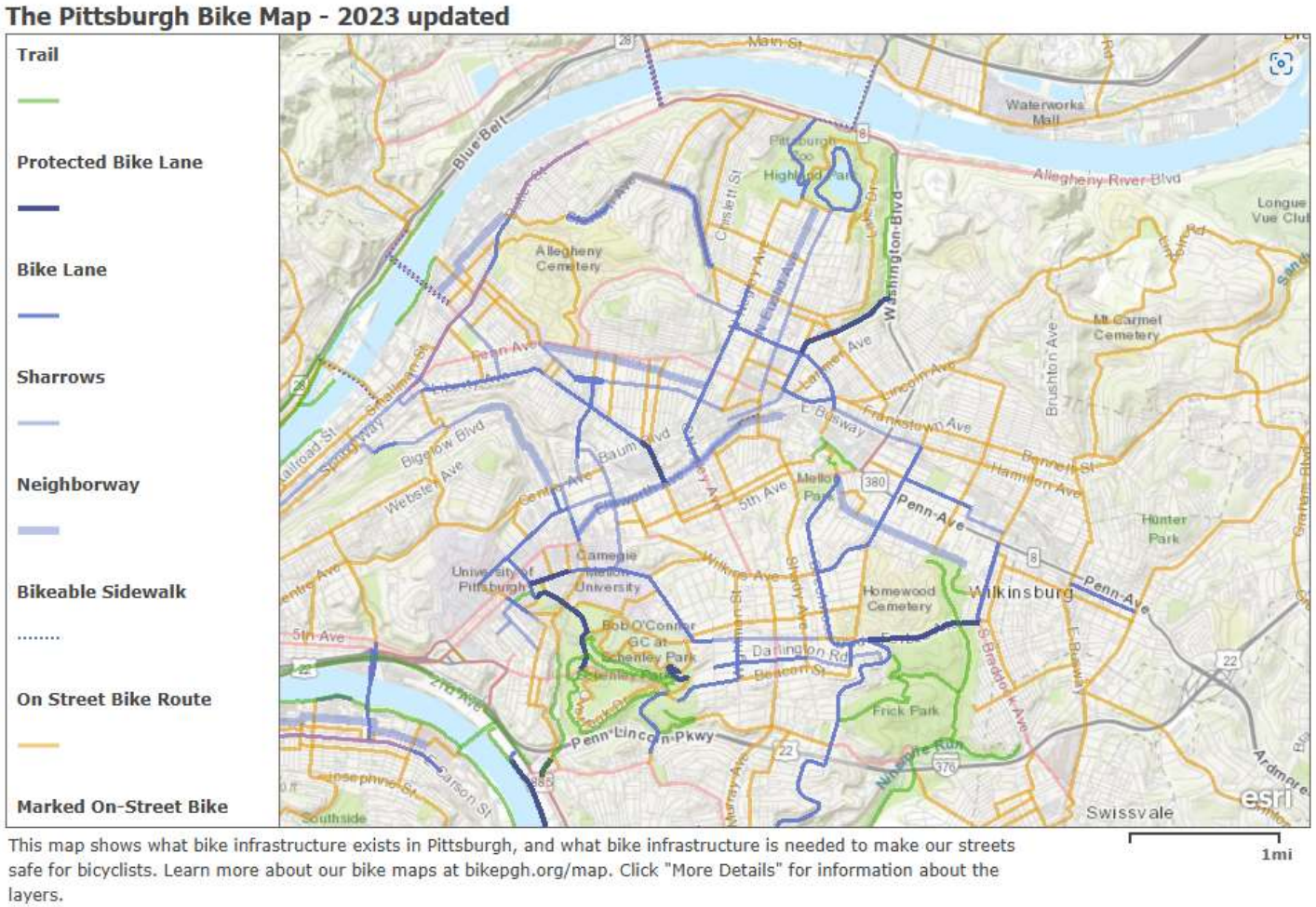
Hypothesis 3

- Introduction of new bike lanes near existing stations increases POGOH ridership

a. HO_f: There is a significant effect in ridership with the addition of new bike lanes near existing stations
b. H1_f: There is no significant effect in ridership before and after the addition of bike lanes

Pitts Bike Lanes

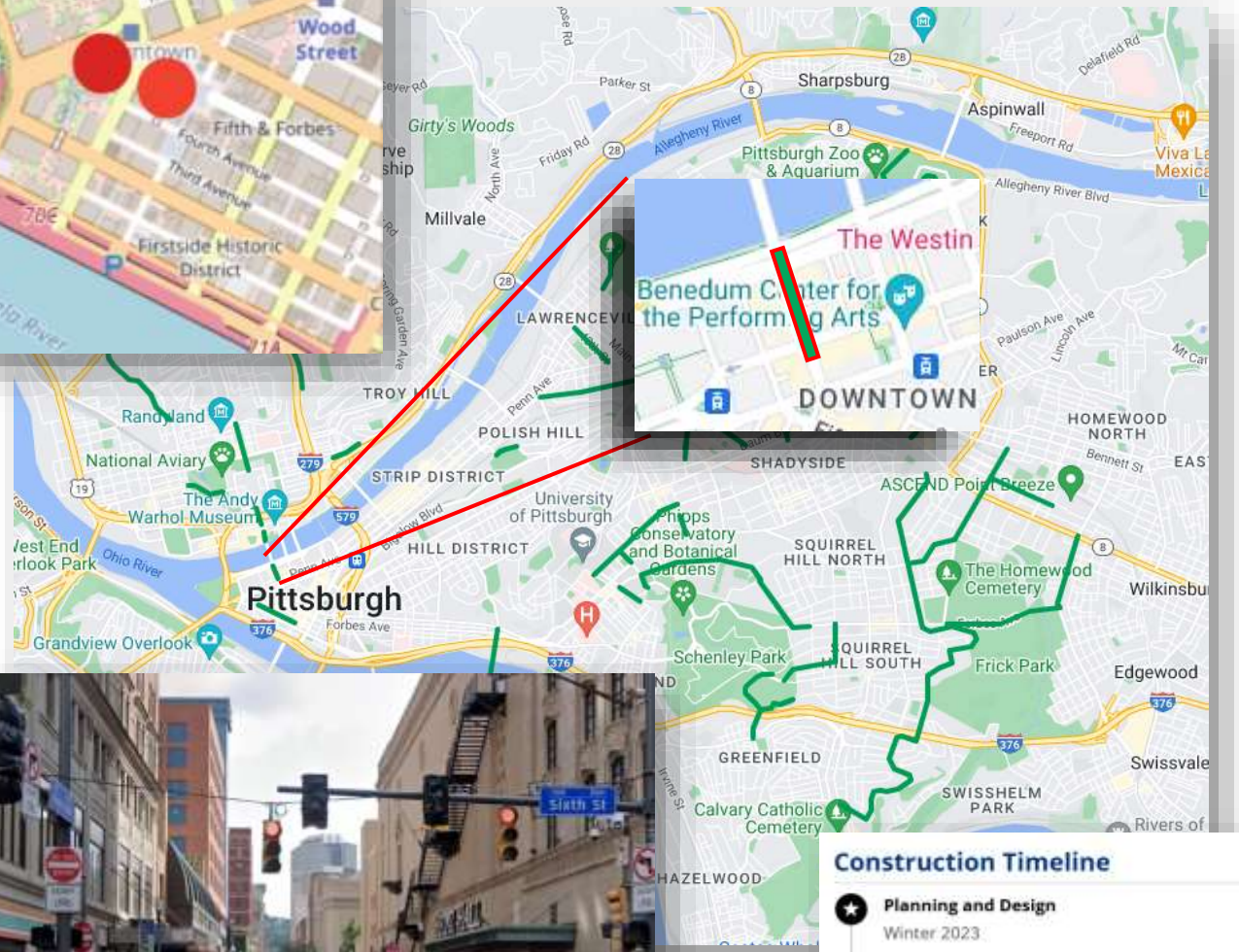
Pittsburgh redesigned Allegheny Circle, turning a one-way ring road into a two-way street with bike lanes, addressing past urban renewal issues.



New Bike Lanes

A new bike lane was constructed on 6th Ave by summer 2023. We plan to analyze ridership behavior near this bike lane to test our hypothesis.

In the end, we identified two bike stations adjacent to this bike lane, as shown on the right.



Control Subjects

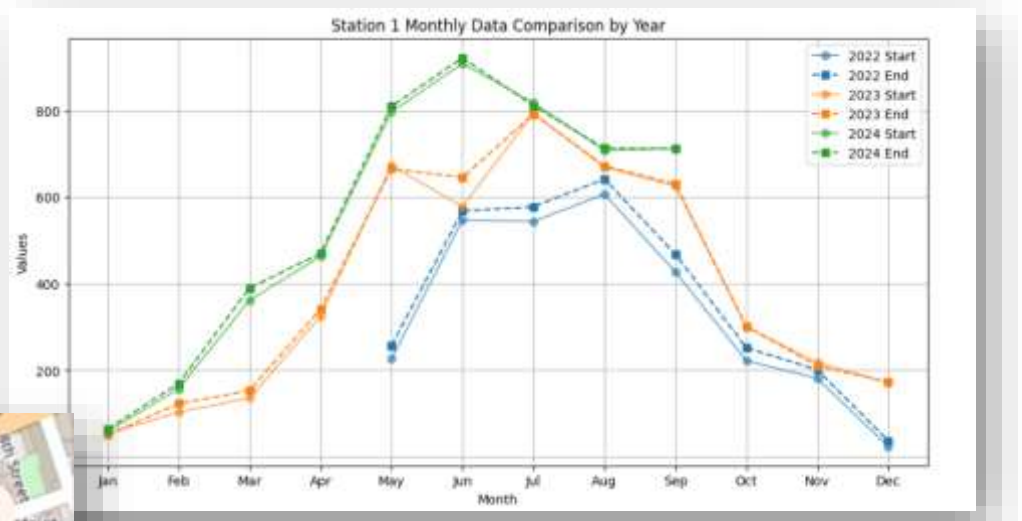
Federal ave is an old bike lane which is near to 6th new bike lane. Federal ave is chose as control subjects A.

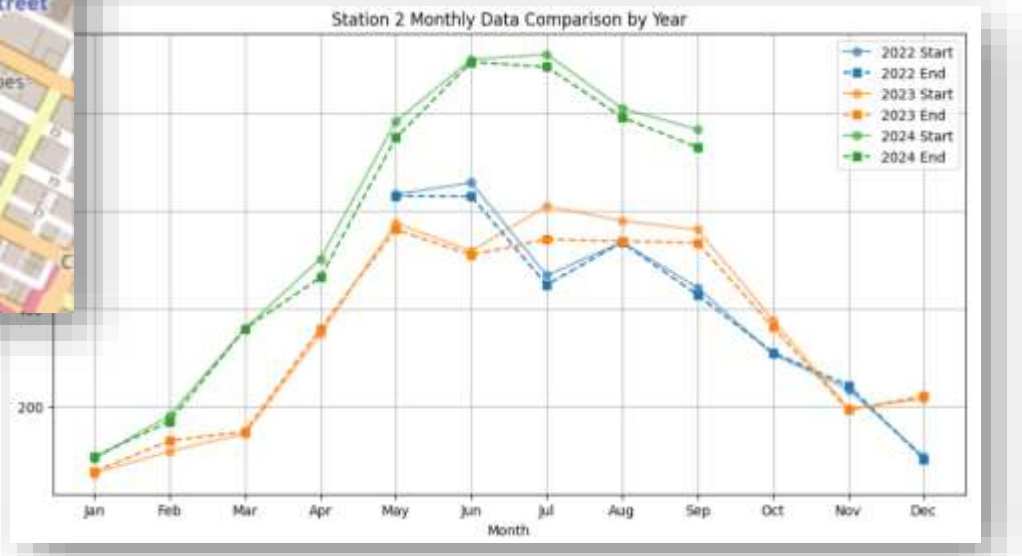
10th is a street without bike lane. 10th is chosen as control subjects B.

6th Ridership

In this research, the number of starts and ends at each station is used as the number of passengers at that station.

From the monthly ridership graph, it is clear that POGOH riders in two stations increase with the year.





POGOH

DID

- The Difference-in-Differences (DiD) method was used to assess the impact of the intervention (e.g., installation of new bike lanes) on the number of cyclists. By comparing changes in ridership before and after the intervention, between the intervention and control groups.

$$y_{it} = \beta_0 + \sum_m \beta_m \cdot \text{Month}_m + \beta_1 \text{Treatment}_i + \beta_2 \text{Month}_m \times \text{Treatment}_i + \epsilon_{it}$$

DID

- Treatment is used to mark the intervention and control groups.
- Post is used to mark the months before and after the intervention. Interaction is the
- Interaction term between Treatment and Post and is used to measure the effect of the intervention.

Station	Month	Ridership	Treatment	Post	Interaction
N	May-22	437.5	1	0	0
N	Jun-22	601.25	1	0	0
N	Jul-22	510.25	1	0	0
N	Aug-22	580.75	1	0	0
N	Sep-22	442	1	0	0
N	Oct-22	273.25	1	0	0
N	Nov-22	215.5	1	0	0
N	Dec-22	63	1	0	0
C_A	May-22	633.5	0	0	0
C_A	Jun-22	645	0	0	0
C_A	Jul-22	459.5	0	0	0
C_A	Aug-22	537	0	0	0
C_A	Sep-22	436.5	0	0	0
C_A	Oct-22	309.5	0	0	0
C_A	Nov-22	240	0	0	0
C_A	Dec-22	95.5	0	0	0
C_A	Jan-23	65.5	0	0	0
C_A	Feb-23	120	0	0	0
C_A	Mar-23	147	0	0	0
C_B	May-22	241.5	0	0	0
C_B	Jun-22	557.5	0	0	0
C_B	Jul-22	561	0	0	0
C_B	Aug-22	624.5	0	0	0
C_B	Sep-22	447.5	0	0	0
C_B	Oct-22	237	0	0	0
C_B	Nov-22	191	0	0	0
C_B	Dec-22	30.5	0	0	0

Regression Analysis Summary

SUMMARY OUTPUT

regression statistics	
Multiple R	0.166693503
R Square	0.027786724
Adjusted R Square	-0.007353515
standard error	256.2944059
observed value	87

variance analysis

	df	SS	MS	F	Significance F
regression analysis	3	155823.2101	51941.07002	0.790738051	0.502456222
residual	83	5452006.268	65686.82251		
total	86	5607829.478			

	Coefficients	standard error	t Stat	P-value	Lower 95%	Upper 95%	Lower limit 95%	Upper limit 95%
Intercept	403.5294118	43.954128	9.180694286	2.83994E-14	316.106419	490.9524045	316.106419	490.9524045
X Variable 1	5.31422E-14	76.13078289	6.98038E-16	1	-151.4210652	151.4210652	-151.4210652	151.4210652
X Variable 2	85.92892157	68.3294688	1.257567534	0.212075045	-49.97564367	221.8334868	-49.97564367	221.8334868
X Variable 3	-1.04995E-13	118.3501116	-8.87158E-16	1	-235.393612	235.393612	-235.393612	235.393612

POGOH

DID Model Overview

•**Multiple R = 0.167**, indicating a weak linear relationship between dependent and independent variables.
•**R Square = 0.0278**, explaining less than 3% of the variance in the dependent variable.
•**Adjusted R Square = -0.0074**, suggesting the model has little to no predictive capability.
•**Standard error = 256.29**, indicating large deviations between predicted and actual values.
•**Non-significant variables**: All independent variables have **p-values > 0.05**, with **X Variable 1** and **X Variable 3** having **p = 1**, showing no impact on the dependent variable.

POGOH

GLM

- This research choose to use **Generalized Linear Models (GLM)** to analyze the ridership data next:
- The data is count-based (ridership), making Poisson regression a suitable choice for modeling discrete data.
- GLM allows us to account for both the intervention effect and group differences, providing clear insights into whether the intervention impacted ridership.

$$\log(\mu_i) = \beta_0 + \beta_1 \cdot \text{Post}_i + \beta_2 \cdot \text{Treatment}_i + \beta_3 \cdot (\text{Post}_i \times \text{Treatment}_i)$$

GLM

- **Pseudo R-squared**: 0.02366 (indicates a small model fit)
- **Log-Likelihood**: -7251.9
- **Model Significance (LLR p-value)**: 7.114e-76
- **Post Intervention Effect**: Significant (coef = 0.1930)
- **Treatment Effect**: No significant impact (coef ≈ 0)
- **Interaction Effect**: No significant impact (coef ≈ 0)

```
Optimization terminated successfully.
         Current function value: 83.354899
         Iterations 4
                          Poisson Regression Results
==============================================================================
Dep. Variable:              Ridership   No. Observations:                   87
Model:                        Poisson   Df Residuals:                       83
Method:                           MLE   Df Model:                           3
Date:                Thu, 05 Dec 2024   Pseudo R-squ.:                0.02366
Time:                        12:03:24   Log-Likelihood:                -7251.9
converged:                       True   LL-Null:                       -7427.6
Covariance Type:            nonrobust   LLR p-value:                 7.114e-76
==============================================================================
                 coef    std err          z      P>|z|      [0.025      0.975]
------------------------------------------------------------------------------
Intercept      6.0002      0.009    702.824      0.000       5.984       6.017
Post           0.1930      0.013     15.358      0.000       0.168       0.218
Treatment   2.504e-16      0.015   1.69e-14      1.000      -0.029       0.029
Interaction 8.936e-17      0.022    4.1e-15      1.000      -0.043       0.043
==============================================================================
```

GLM
Model Overview

- **Intervention effect**: There was a significant increase in the number of cyclists after the intervention, suggesting that the intervention had a positive impact on the number of cyclists.
- **Experimental group effect**: There was no significant difference between the experimental group and the control group, suggesting that the changes in the experimental group itself were not sufficient to independently explain the increase in cycling.
- **Model performance**: The Poisson regression model showed a significant effect of the intervention, but the model fit was relatively low, suggesting that further improvements to the model or consideration of other factors may be required.

Hypothesis 3

- Introduction of new bike lanes & stations to bus ridership

a. HO_e: The POGOH bike infrastructure does not affect bus ridership in Pittsburgh

b. H1_e: The POGOH bike infrastructure leads to a reduction in bus ridership in Pittsburgh

Hypothesis 3 could not be rejected by now and new factors should be added in the model.

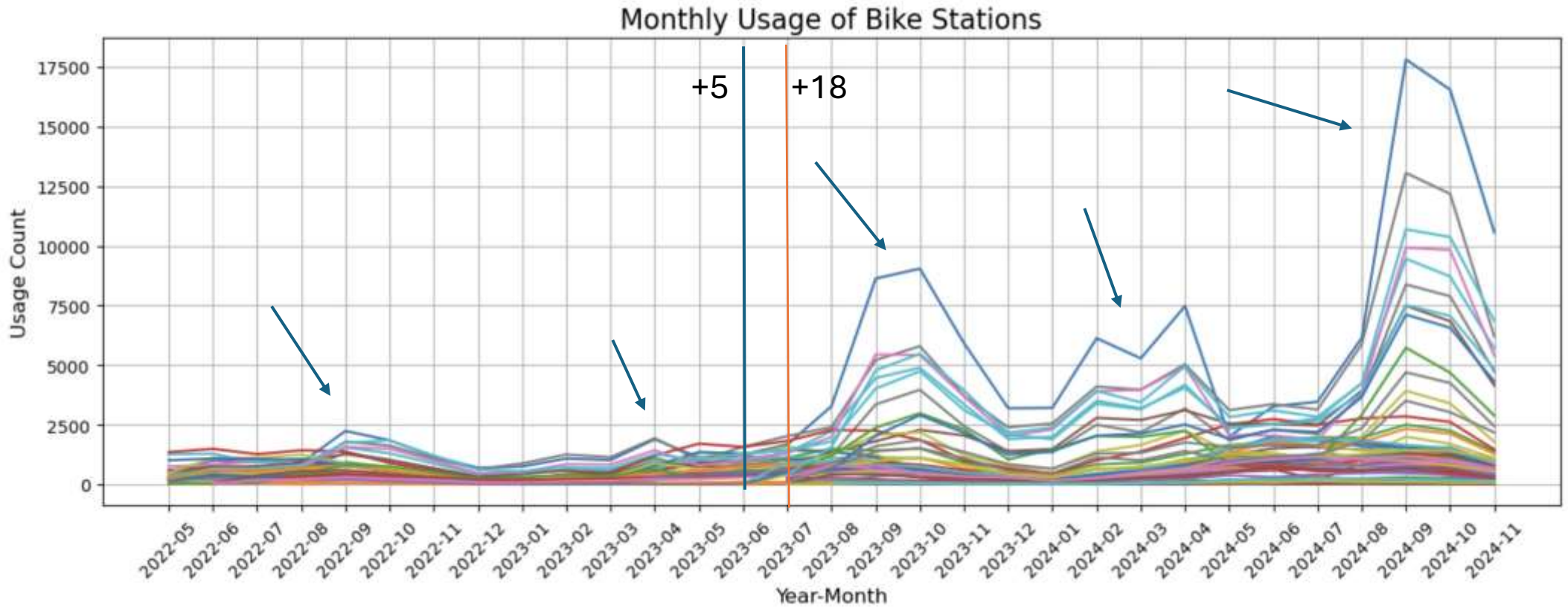
Hypothesis 4

- Introduction of new bike stations increases POGOH ridership at nearby stations

a. HO_h: New bike stations does not have a significant effect on POGOH ridership at nearby stations

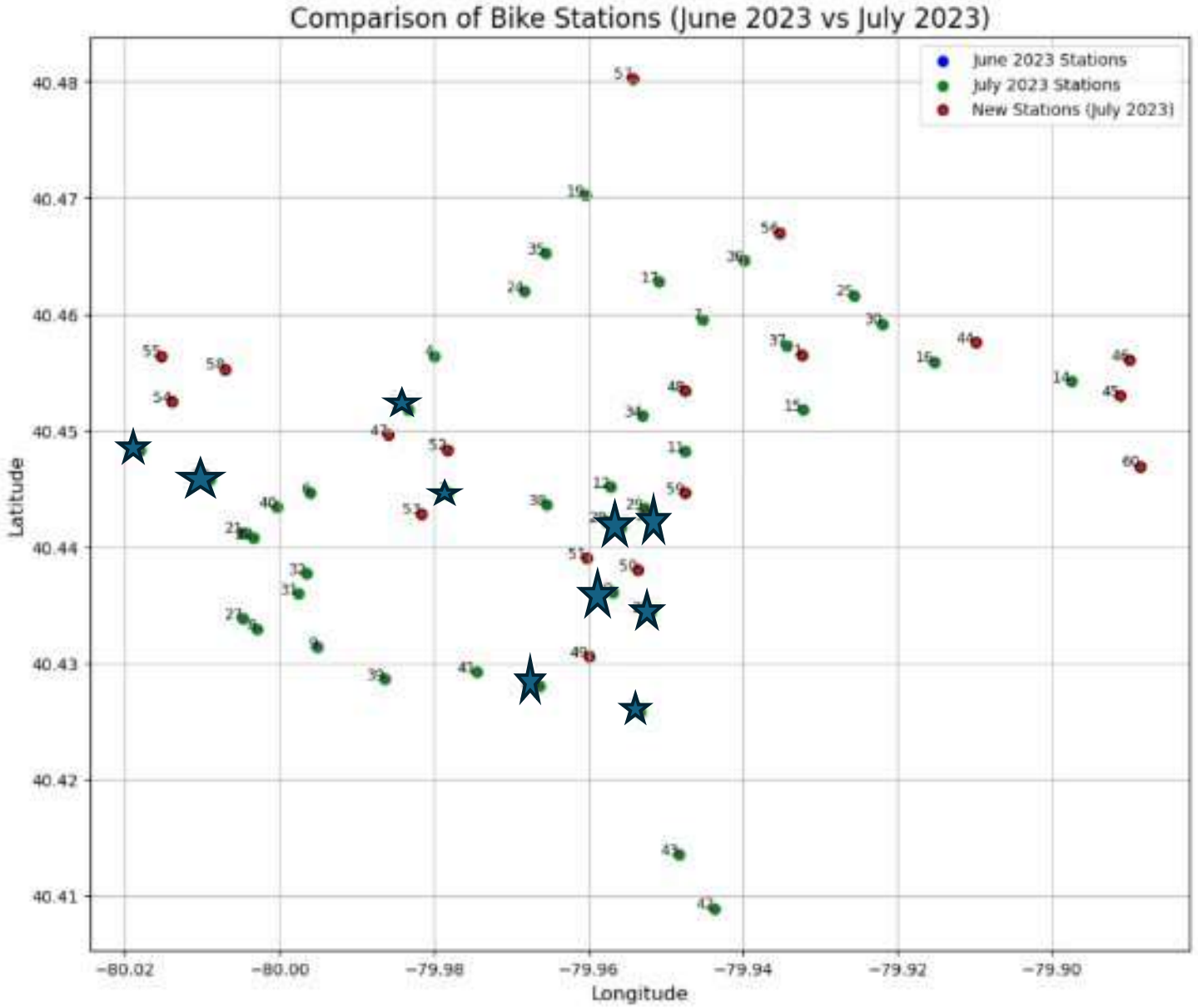
b. H1_h: New bike stations does have a significant effect on POGOH ridership at nearby stations

Insight of the total POGO usage



Nearby stations of new stations

Marked stations:
[5,23,
18,3,
29,13,33,10,20,
26,2,etc]



Also used a DiD test

- We compared the difference before and after there were new stations added nearby
- DiD test mark those stations as 1 and others as 0
- To see if these "1" stations change among all these stations

Result:

1.Treatment:post:441,
p value=0.23 > 0.05, no stong evidence

2. Treatment:-46:usage of the affected station is slightly lower than others

3.Post: 330,after adding new stations, usage of all stations has increased

OLS Regression Results

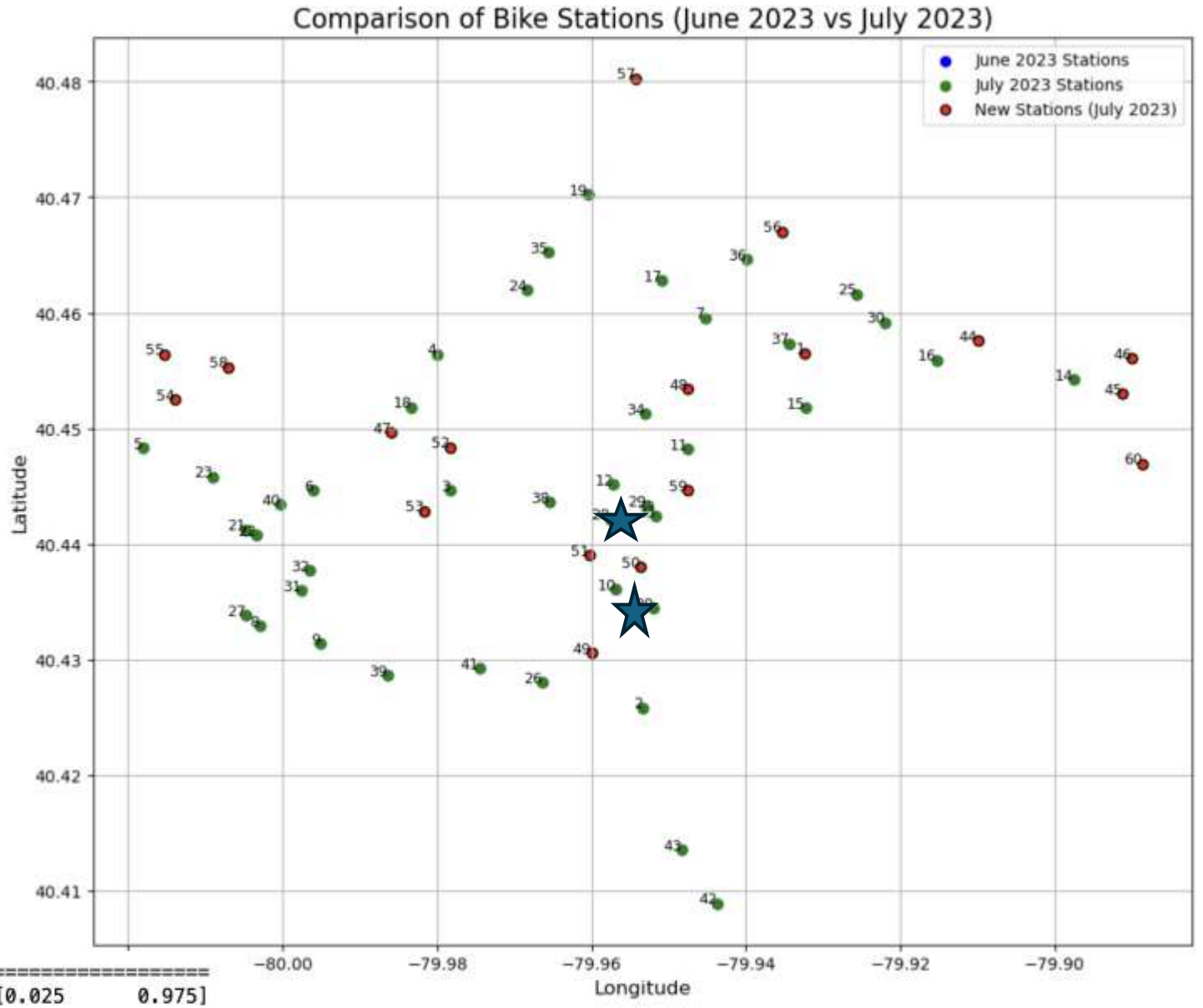
Dep. Variable:	Total_Count	R-squared:	0.022
Model:	OLS	Adj. R-squared:	0.020
Method:	Least Squares	F-statistic:	11.71
Date:	Thu, 05 Dec 2024	Prob (F-statistic):	1.39e-07
Time:	11:34:52	Log-Likelihood:	-13415.
No. Observations:	1541	AIC:	2.684e+04
Df Residuals:	1537	BIC:	2.686e+04
Df Model:	3		
Covariance Type:	nonrobust		

	coef	std err	t	P>|t|	[0.025	0.975]
Intercept	468.0000	281.497	1.663	0.097	-84.159	1020.159
Treatment	-46.3023	359.160	-0.129	0.897	-750.799	658.194
Post	330.0347	286.012	1.154	0.249	-230.980	891.050
Treatment:Post	441.4626	367.318	1.202	0.230	-279.035	1161.960

Omnibus:	1528.944	Durbin-Watson:	1.368
Prob(Omnibus):	0.000	Jarque-Bera (JB):	75768.704
Skew:	4.755	Prob(JB):	0.00
Kurtosis:	36.009	Cond. No.	25.5

If we only look into the effects on part of the stations:

Treatment_post = 1936, while P is nearly 0.



	coef	std err	t	P>|t|	[0.025	0.975]
Intercept	450.2000	177.251	2.540	0.011	102.520	797.880
Treatment	−74.5000	468.962	−0.159	0.874	−994.374	845.374
Post	337.7142	181.211	1.864	0.063	−17.733	693.161
Treatment:Post	1936.1033	484.281	3.998	0.000	986.183	2886.024

POGOH

Hypothesis 4

- Introduction of new bike stations increases POGOH ridership at nearby stations

a. HO_h: New bike stations does not have a significant effect on POGOH ridership at nearby stations

b. H1_h: New bike stations does have a significant effect on POGOH ridership at nearby stations

Generally speaking, new biking stations have some effect on nearby stations,

Some of the specific stations (oakland Pitt) are strongly affected – the usage increased significantly

Thank you
Any questions?