Analysis of Citi Bike Data

Objective:-

- The objective of this Project is to aggregate the data found in the Citi Bike Trip History Logs to build a data dashboard, story, or report. To find the answers related to rider, riding, bike, and Destination etc.

Problem Statement:-

Since 2013, the Citi Bike Program has implemented a robust infrastructure for collecting data on the program's utilization. Through the team's efforts, each month bike data is collected, organized, and made public on the Citi Bike Data webpage.

However, while the data has been regularly updated, the team has yet to implement a dashboard or sophisticated reporting process. City officials have a number of questions on the program.

Data Description:-

(This data file contains only September month data.)

- Trip Duration (seconds)
- Start Time
- Stop Time
- Start Station ID
- Start Station Name
- End Station ID
- End Station Name
- End Station Latitude
- End Station Longitude
- Bike ID
- User Type (Customer = 24-hour pass or 3-day pass user; Subscriber = Annual Member)
- Birth Year

- Gender (F = Female, M = Male)
- Age(Year)
- Year(September 2019 & 2020)

Data Insertion, Preprocessing, clustering and modeling

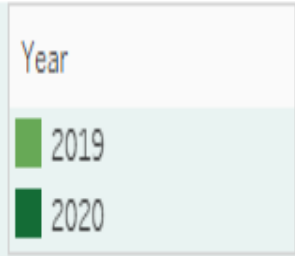
- The csv file of the data has been inserted in Tableau.
- Created the Bottom N and Top N parameter.
- Created Bin for the Age.
- Clustered the some columns and created set in tableau.

ANALYSING THE REPORTS & DASHBOARDS

Tableau Public - City_Bike_Analytics

Pivot Table

Year	Count of Tripduration	Avg. Tripduration	Tripduration	Avg. Age
2019	49,216	811	39,926,046	37
2020	53,833	1,576	84,829,288	38

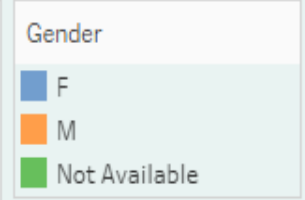


Key Insight

- In 2019 Count of Trip duration, average Trip duration and total Trip duration is less compared to 2020.
- Total Trip duration is 11090.57 Hours in 2019 and 23563.69 Hours in 2020.

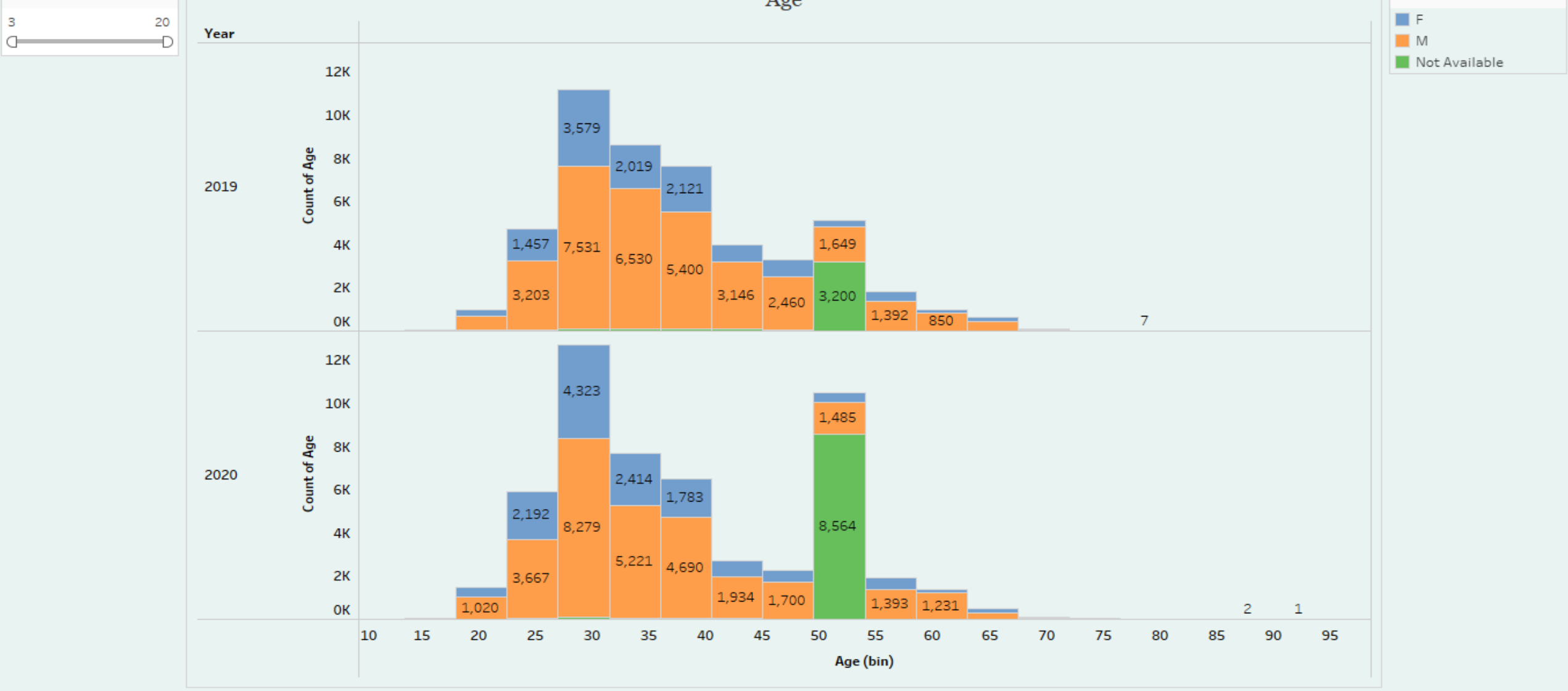
Categories

Year	Usertype	Gender	Count of Tripduration	Avg. Tripduration	Tripduration	Avg. Age
2019	Customer	F	1,611	2,788	4,491,271	30
		M	3,004	1,515	4,549,781	31
		Not Available	2,773	2,639	7,316,923	50
	Subscriber	F	10,438	550	5,739,154	36
		M	30,408	572	17,396,753	38
		Not Available	982	440	432,164	43
2020	Customer	F	4,031	4,127	16,636,771	30
		M	6,185	1,858	11,489,458	30
		Not Available	8,333	2,970	24,745,235	51
	Subscriber	F	9,842	858	8,441,323	36
		M	24,882	925	23,006,456	38
		Not Available	560	911	510,045	43

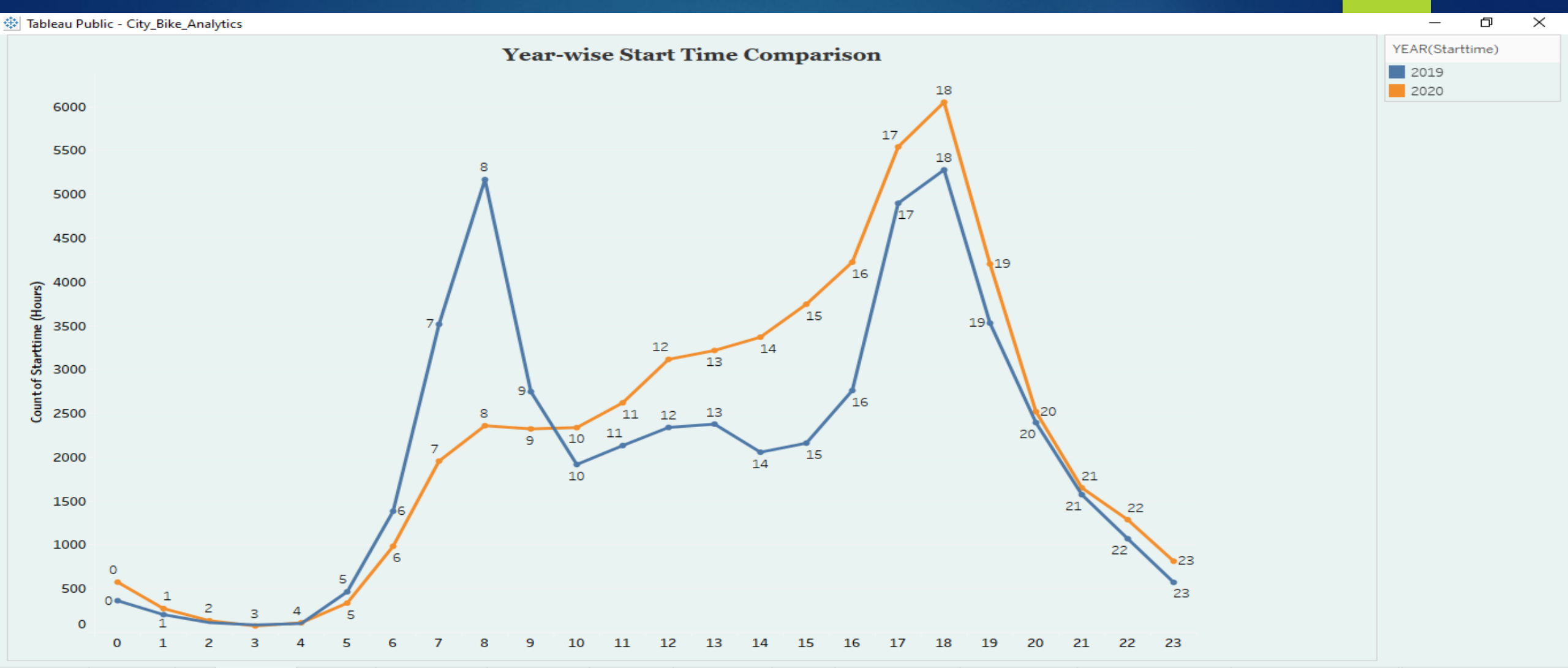


From the table-

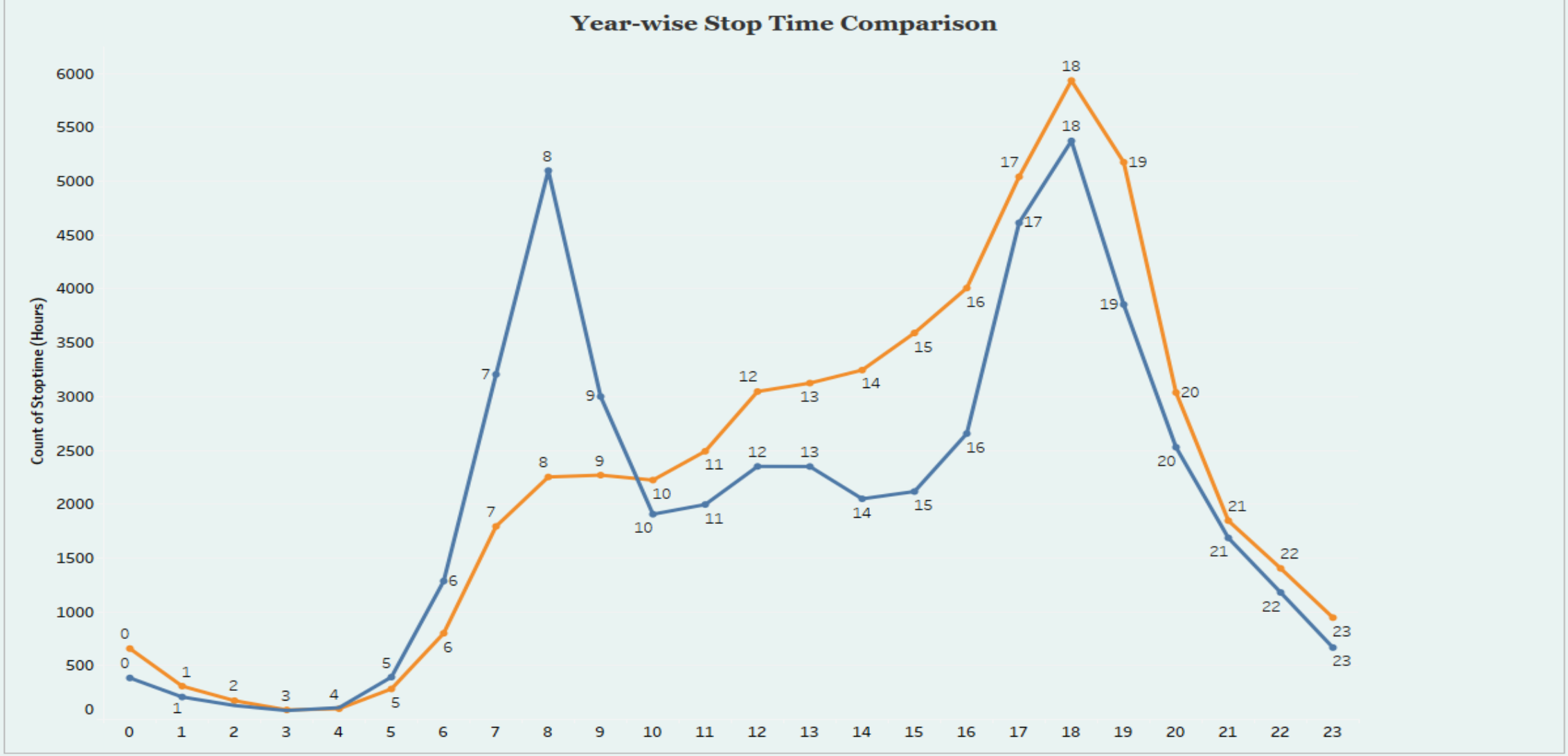
- Customer whose gender is unknow their total Trip duration and average age is higher than the male and female.
- Male and Female Subscriber have taken maximum number of rides than the customer in both years.
- Total Trip duration of Male and Female subscribers is more than the Customer in 2019 but in 2020 female’s customers total trip duration id higher than the subscriber.
- Average age of the unknown gender is higher than the male and female customers and subscrIbers.



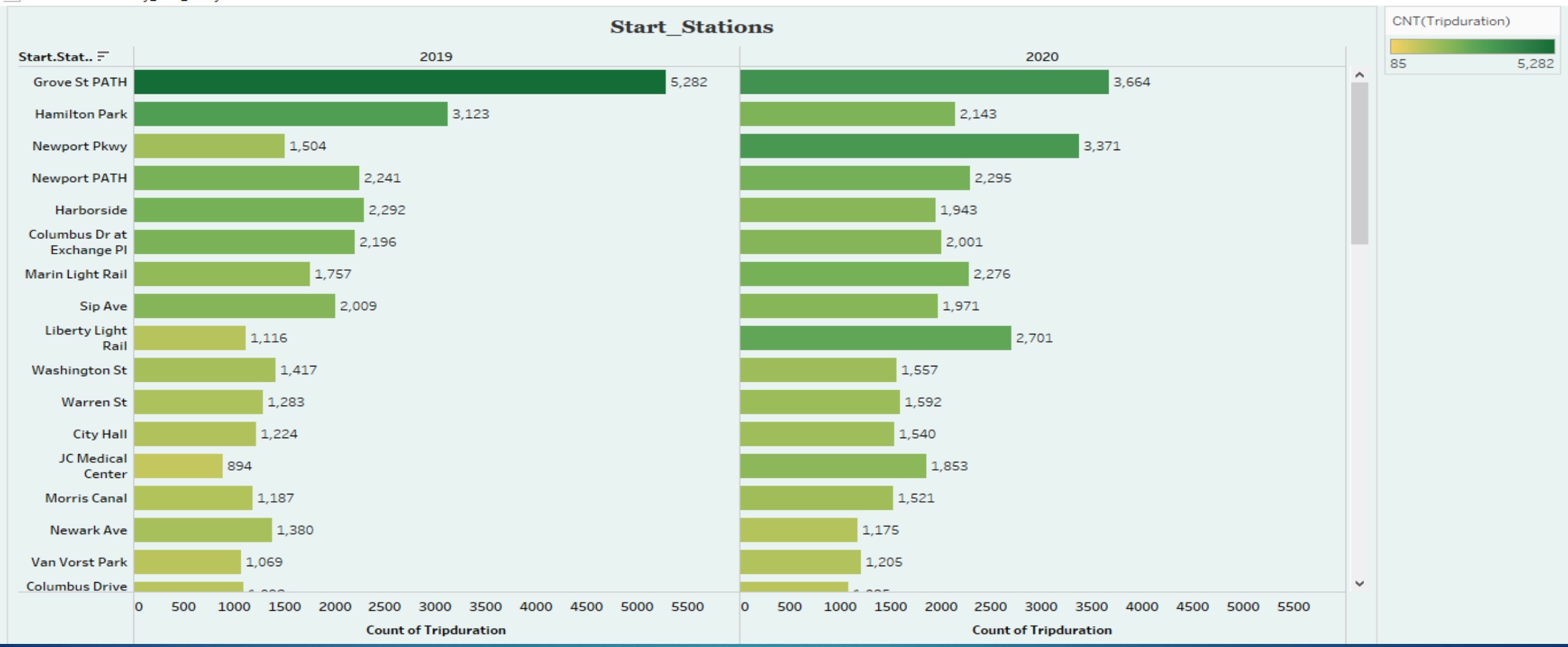
Nearly 27 and 5o old years riders have taken maximum number of rides.



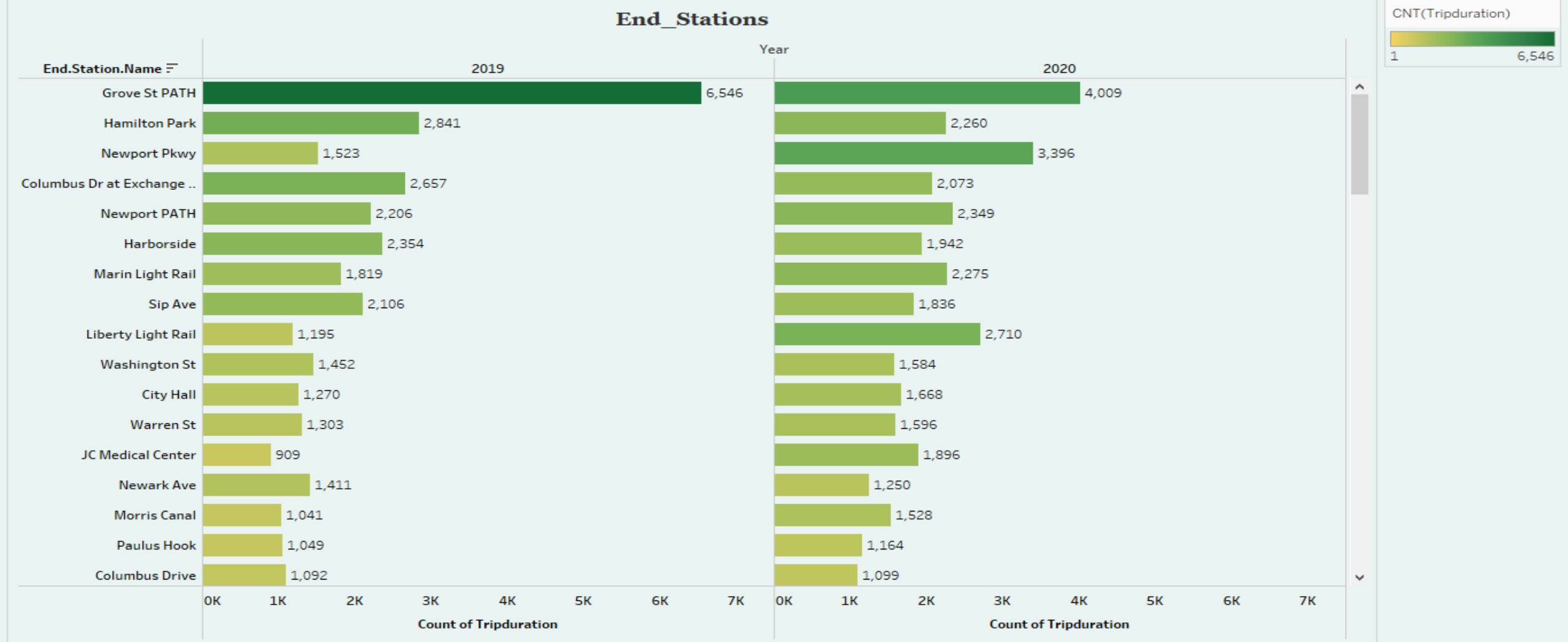
- At 8 am and 6 pm is the peak hour when the maximum number of people started their ride.
- At 8 am in 2019 people started their ride much more than the 2020



- At 8 am and 6 pm is the peak hour when the maximum number of people ended their ride.
- At 8 am in 2019 people ended their ride much more than the 2020.

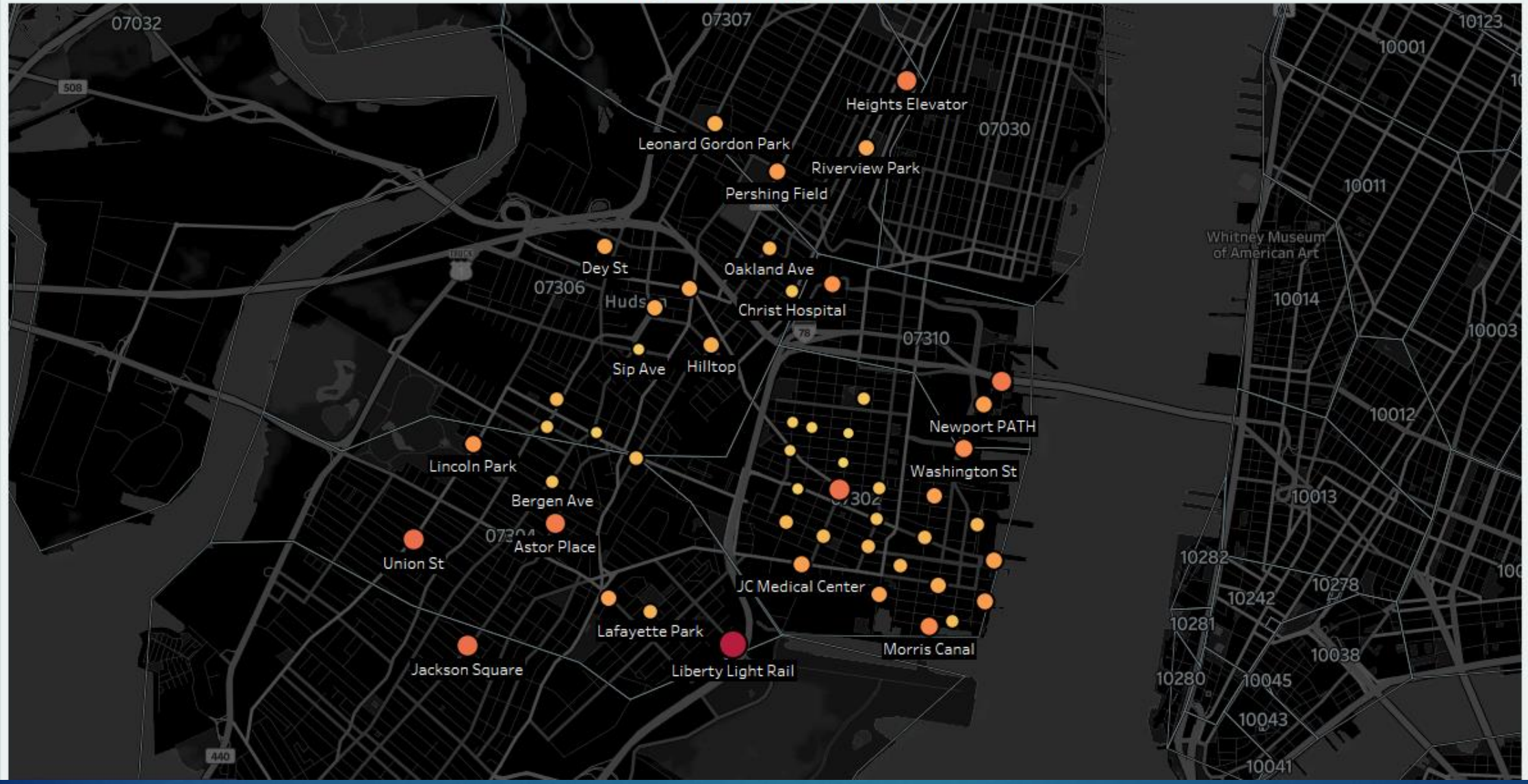


Grove St PATH is the top Start Station from where people started their ride.



Grove St PATH is the top End Station from where people Ended their ride.

Map Of Starting Journey



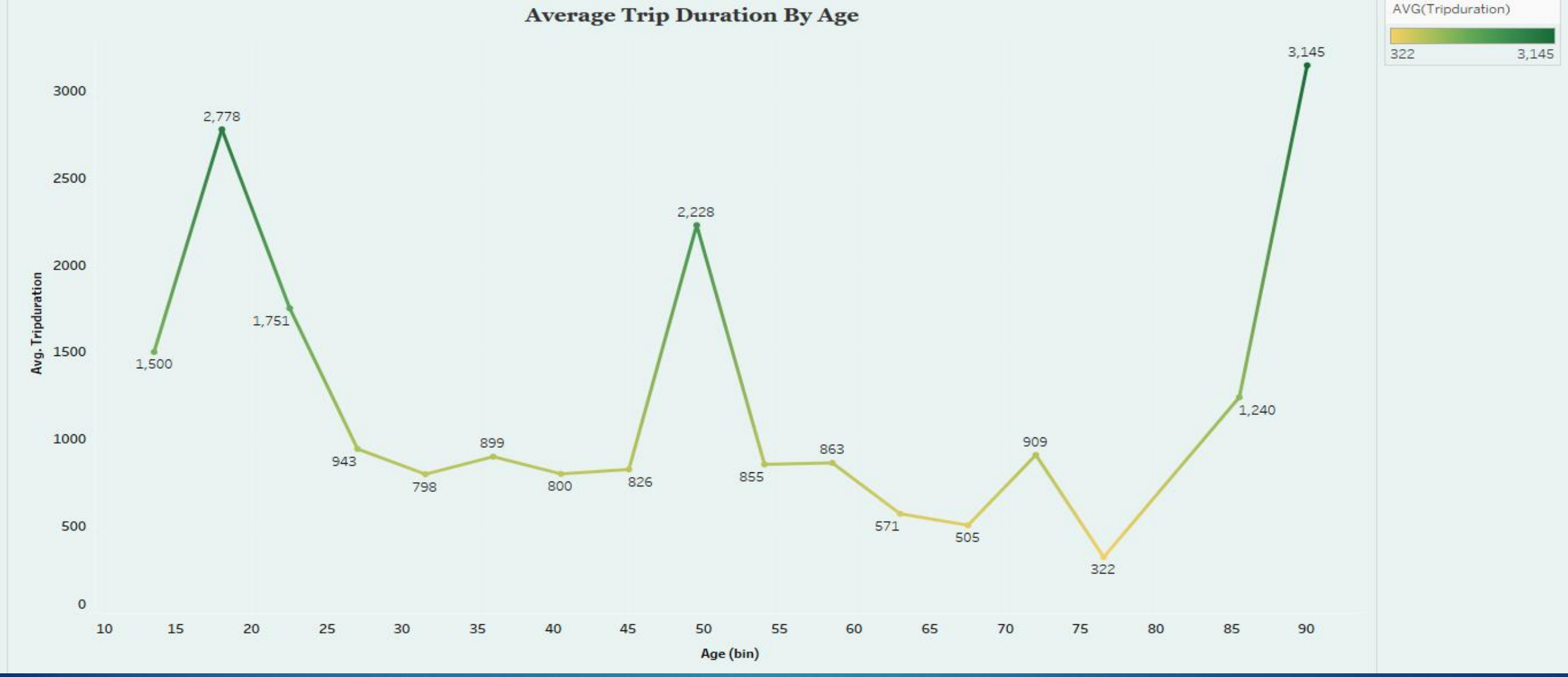
Gender

- (All)
- F
- M
- Not Available

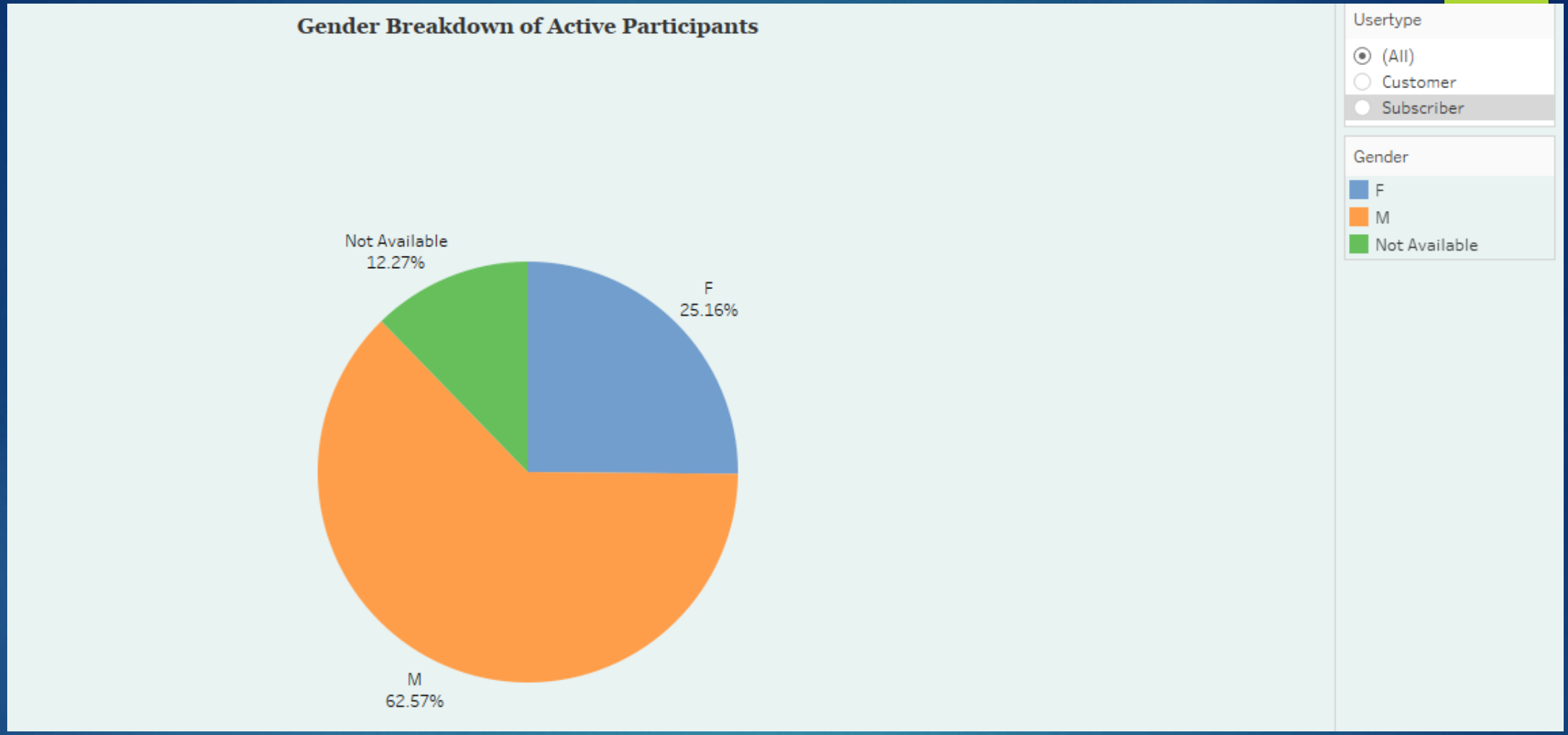
AVG(Tripduration)

541 3,568

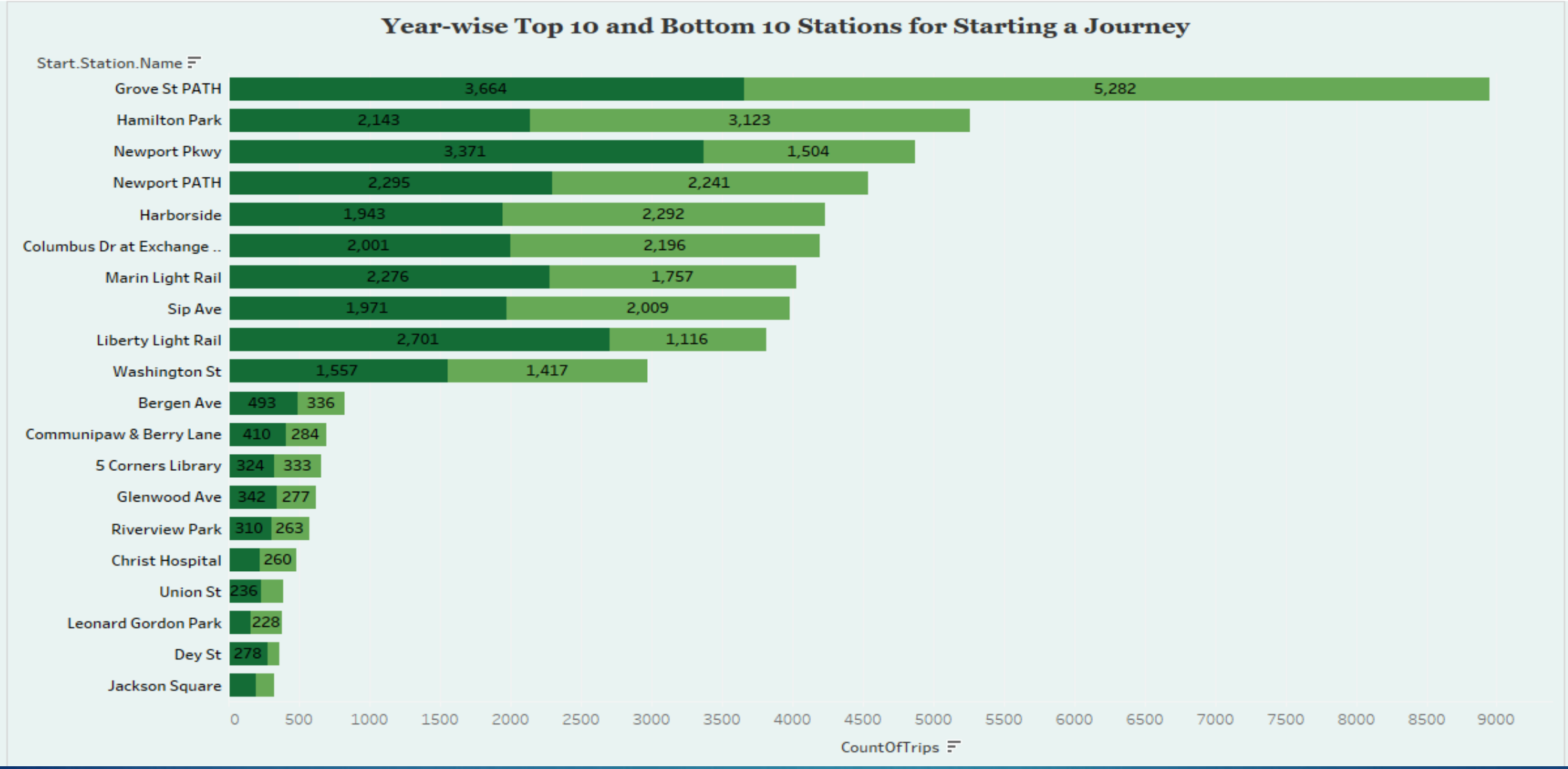
Liberty light Rail followed by Union St and Jackson Square are the starting stations from where average trip duration is maximum.



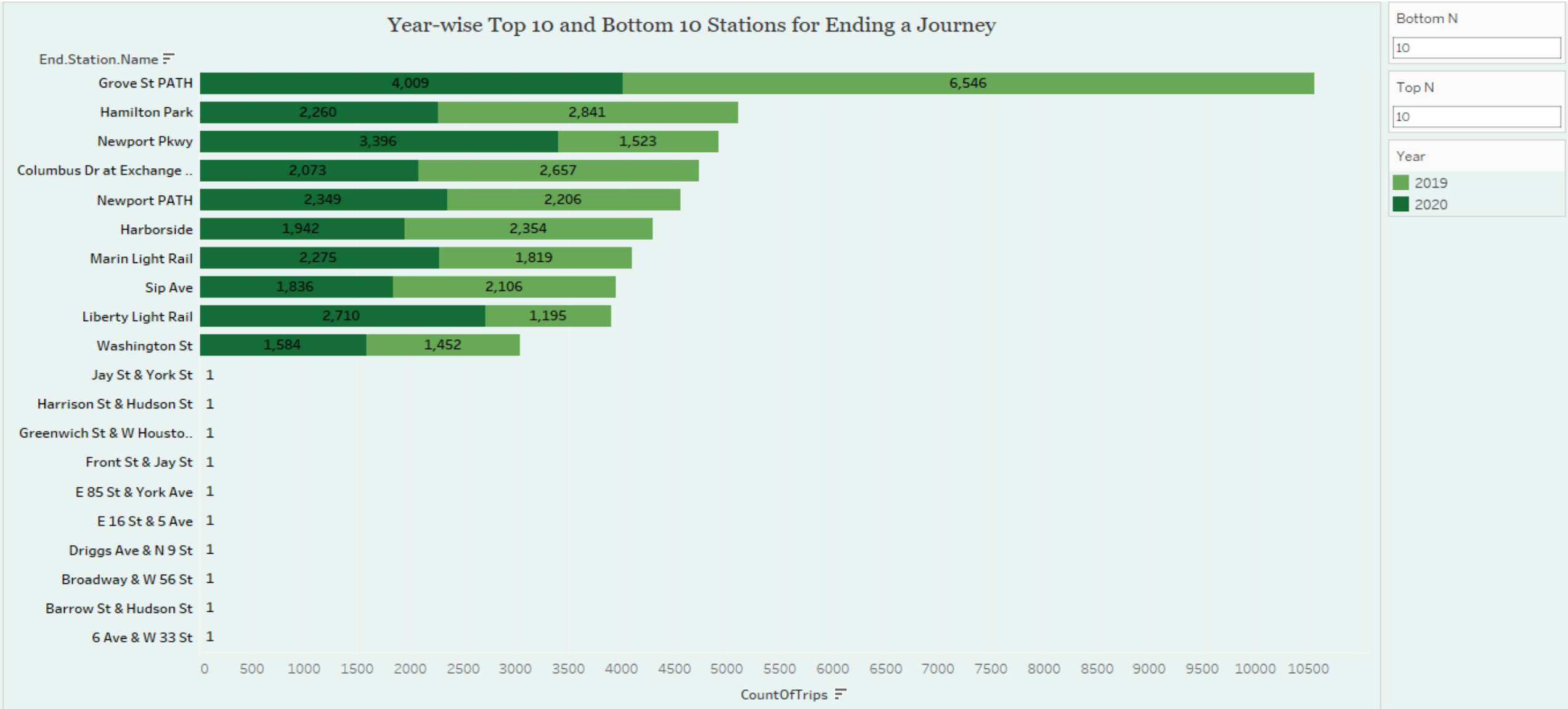
90 followed by 18 and 50 years old riders average trip duration is higher compared to other ages.



- Male with 62.57% have taken maximum number of rides.
- Female have taken only 25.16% and 12.27 % of rides . 12.27 % can be male, female or transgender.



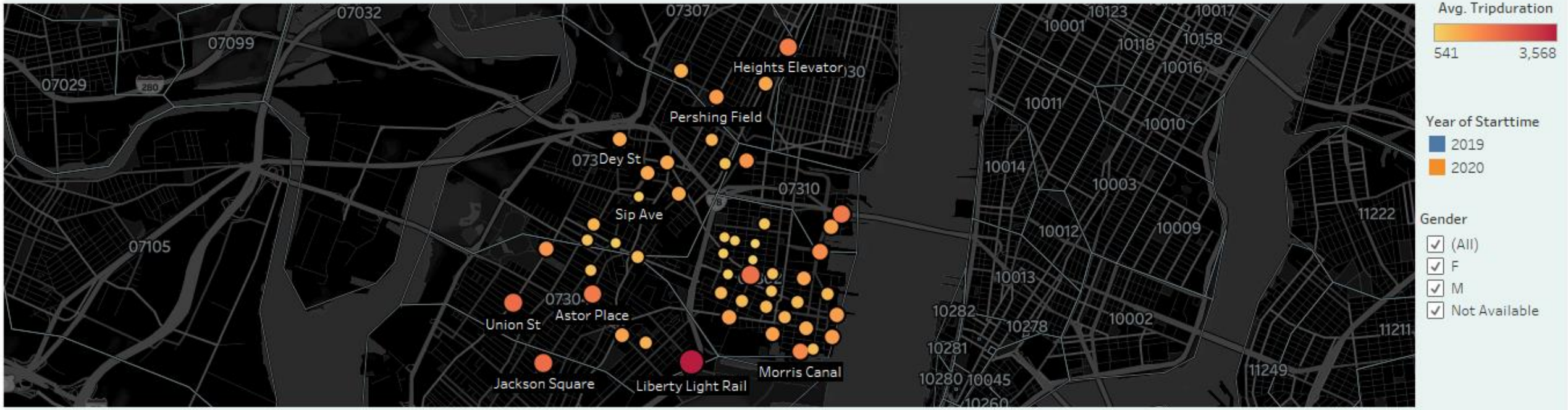
Grove St Path followed by Hamilton Park Newport Pkwy are the top Starting stations from where the maximum numbers of ride have been taken and Jackson Square is the station from where minimum numbers of ride have been taken.



Grove St Path Ending Station records highest numbers of rides

Dashboard -1

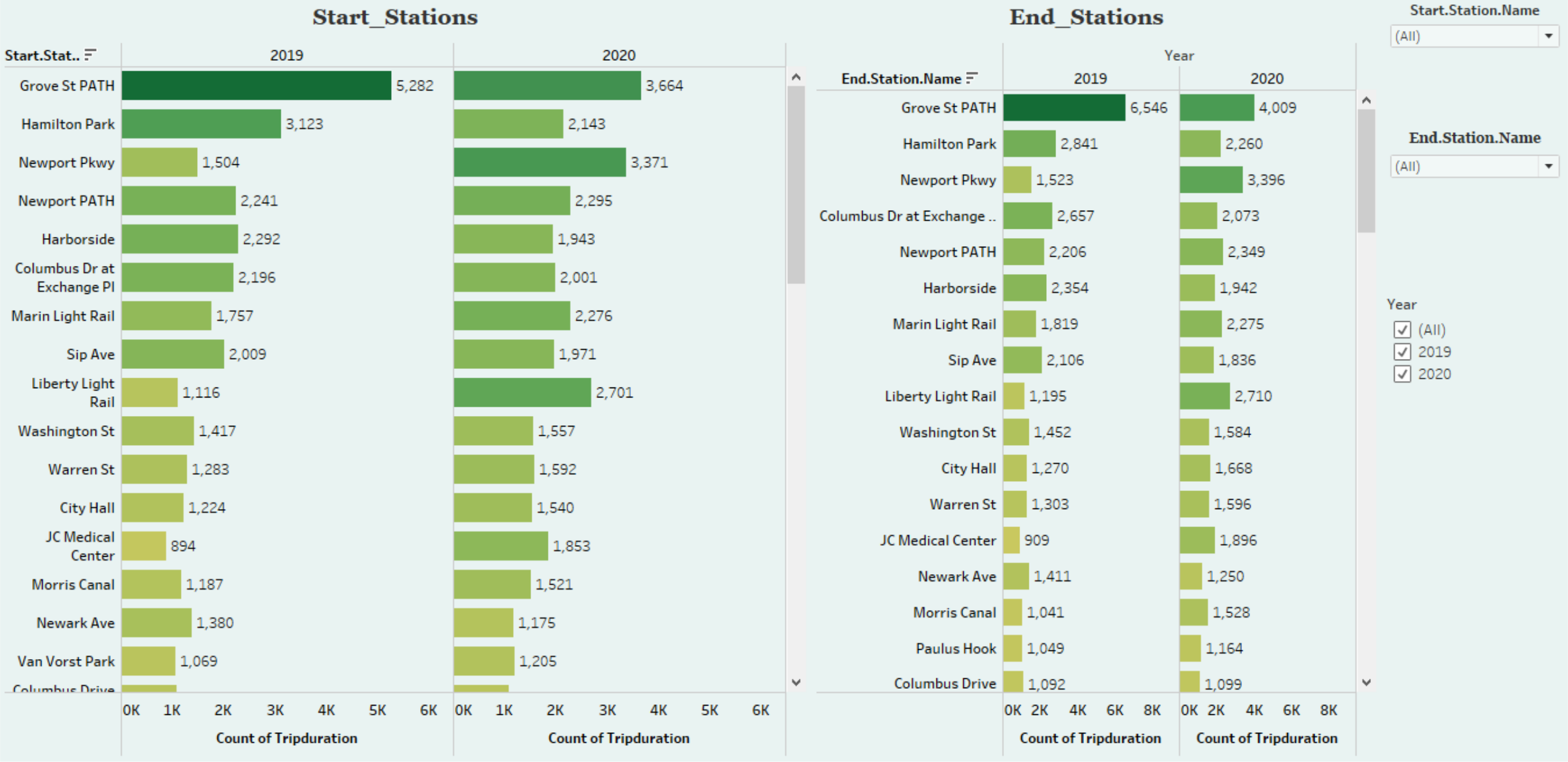
Map Of Starting Journey



Year-wise Stop Time Comparison

Year-wise Start Time Comparison

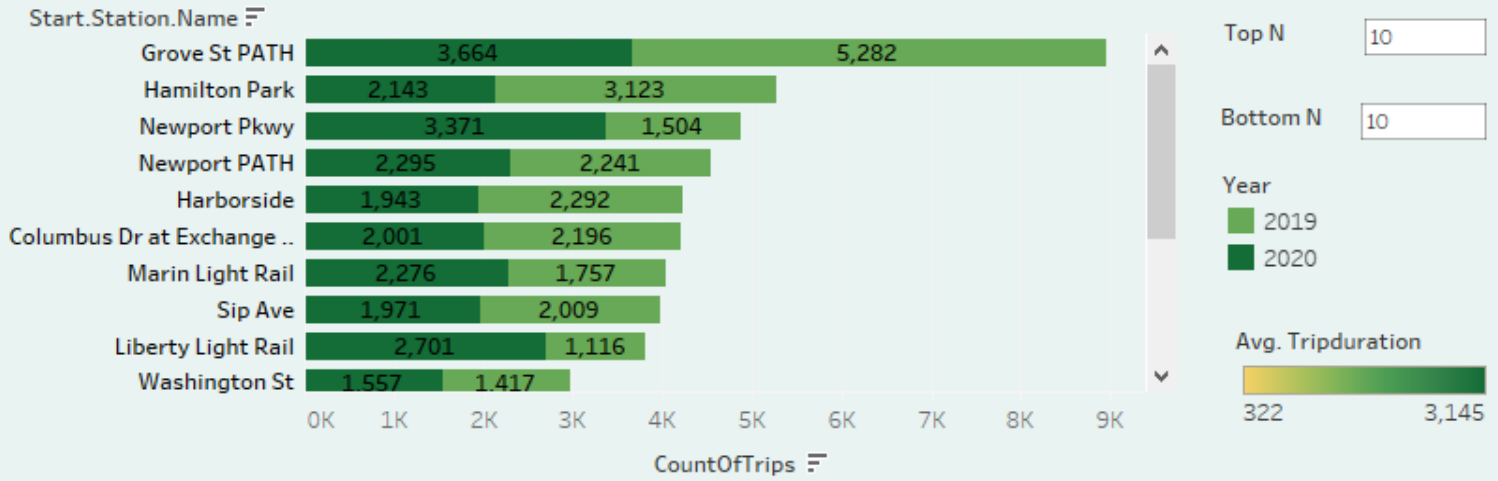
Dashboard-2


Start_Stations
Start.Stat..
2019
2020
Grove St PATH 5,282 3,664
Hamilton Park 3,123 2,143
Newport Pkwy 1,504 3,371
Newport PATH 2,241 2,295
Harborside 2,292 1,943
Columbus Dr at Exchange Pl 2,196 2,001
Marin Light Rail 1,757 2,276
Sip Ave 2,009 1,971
Liberty Light Rail 1,116 2,701
Washington St 1,417 1,557
Warren St 1,283 1,592
City Hall 1,224 1,540
JC Medical Center 894 1,853
Morris Canal 1,187 1,521
Newark Ave 1,380 1,175
Van Vorst Park 1,069 1,205
0K 1K 2K 3K 4K 5K 6K
Count of Tripduration
End_Stations
Year
End.Station.Name
2019
2020
Grove St PATH 6,546 4,009
Hamilton Park 2,841 2,260
Newport Pkwy 1,523 3,396
Columbus Dr at Exchange .. 2,657 2,073
Newport PATH 2,206 2,349
Harborside 2,354 1,942
Marin Light Rail 1,819 2,275
Sip Ave 2,106 1,836
Liberty Light Rail 1,195 2,710
Washington St 1,452 1,584
City Hall 1,270 1,668
Warren St 1,303 1,596
JC Medical Center 909 1,896
Newark Ave 1,411 1,250
Morris Canal 1,041 1,528
Paulus Hook 1,049 1,164
Columbus Drive 1,092 1,099
0K 2K 4K 6K 8K
Count of Tripduration
Start.Station.Name
(All)
End.Station.Name
(All)
Year
(All)
2019
2020

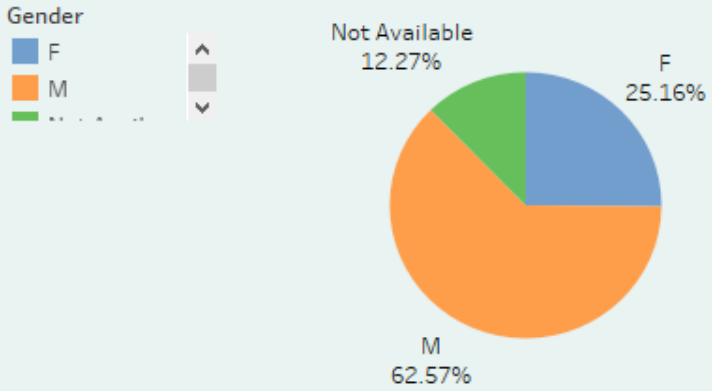
Dahboard-3

Average Trip Duration By Age

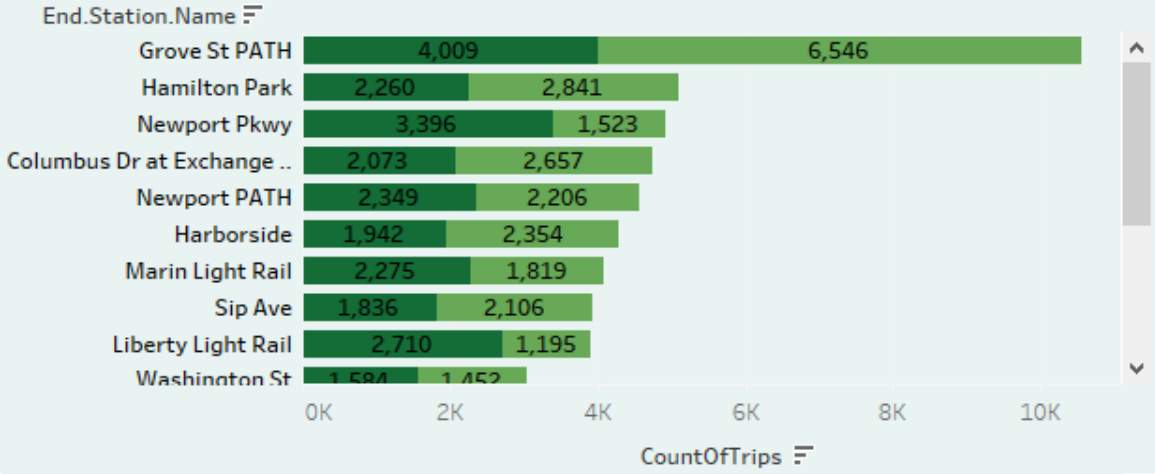
Year-wise Top 10 and Bottom 10 Stations for Starting a Journey



Gender Breakdown of Active Participants

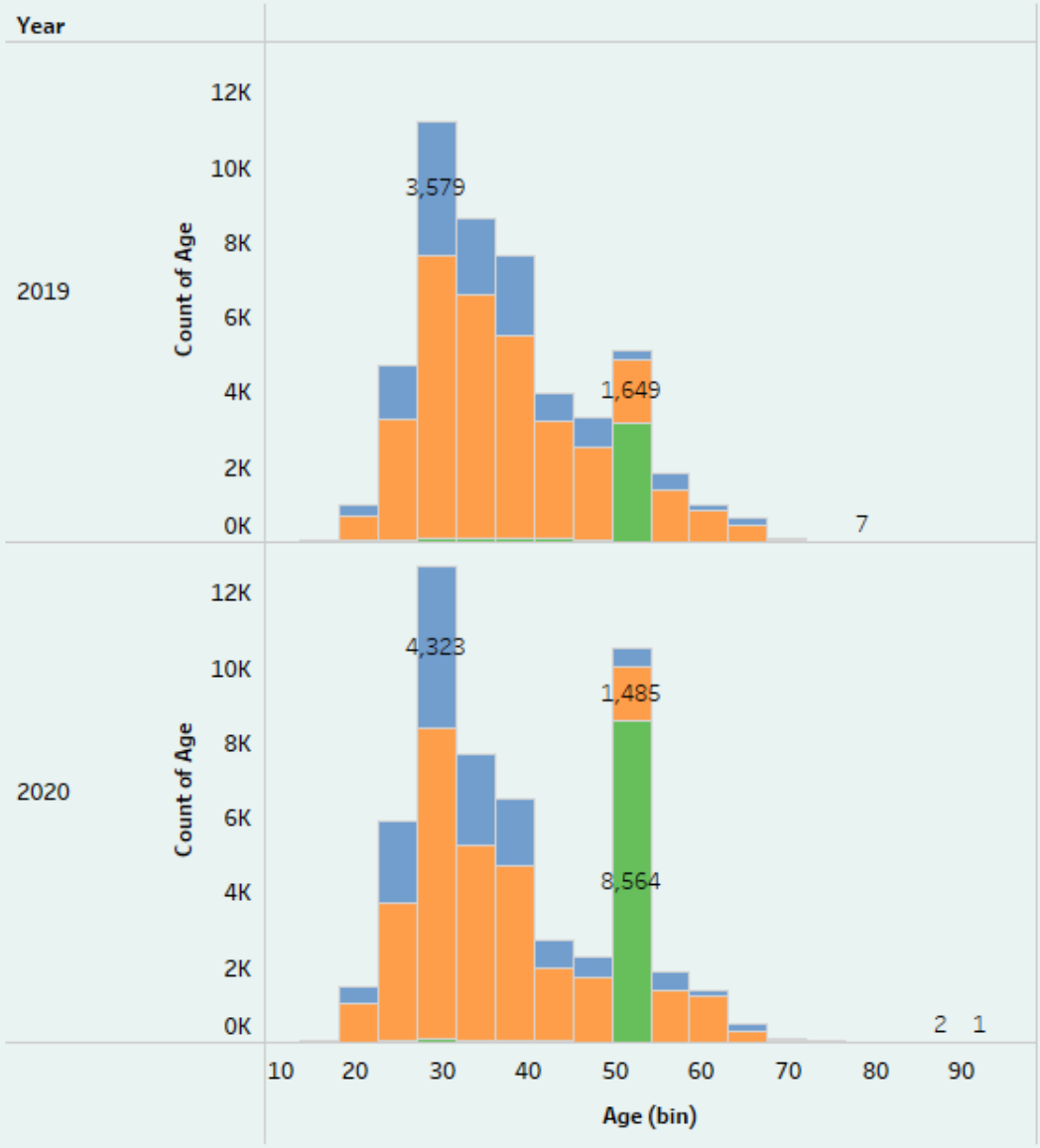


Year-wise Top 10 and Bottom 10 Stations for Ending a Journey



Dashboard-4

Age



Categories

Year	Usertype	Gender	Count of Tripdura..	Avg. Tripduration	Tripduration	Avg. Age
2019	Customer	F	1,611	2,788	4,491,271	30
		M	3,004	1,515	4,549,781	31
		Not Availab..	2,773	2,639	7,316,923	50
	Subscriber	F	10,438	550	5,739,154	36
		M	30,408	572	17,396,753	38
		Not Availab..	982	440	432,164	43
2020	Customer	F	4,031	4,127	16,636,771	30
		M	6,185	1,858	11,489,458	30
		Not Availab..	8,333	2,970	24,745,235	51
	Subscriber	F	9,842	858	8,441,323	36
		M	24,882	925	23,006,456	38
		Not Availab..	560	911	510,045	43

Pivot Table

Year	Count of Tripduration	Avg. Tripduration	Tripduration	Avg. Age
2019	49,216	811	39,926,046	37
2020	53,833	1,576	84,829,288	38

Gender
- F
- M
- Not Available

Year
- 2019
- 2020

Year
- (All)
- 2019
- 2020

Gender
- (All)
- F
- M
- Not Available

Usertype
- (All)
- Customer
- Subscriber

oTime | Start_Stations | End_Stations | Start_Map | TripDuration | Gender | Starting Journey | Ending Journey | Yearwise Analysis | Station Comparison | Top-N Metrics | **Distribution & Pivot Table**