

---

---

# Citibike NYC Data Analysis

— R Project - Jackie Pham —

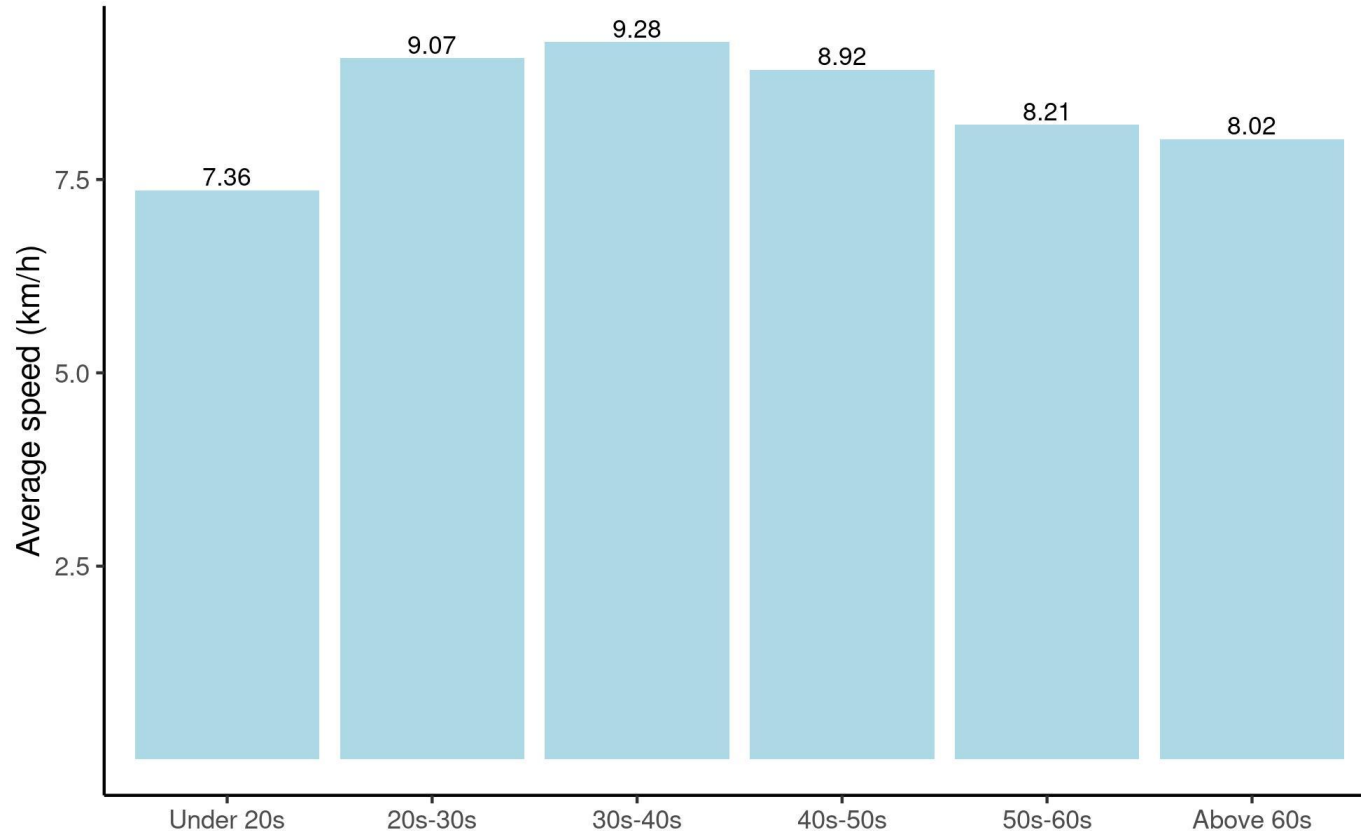
---

---

# Background

- This project is to explore the data associated with the New York City bike share program, Citi Bike.
- There are over 850 Citi Bike stations in New York City; users check a bike out from a starting station and then dock that bike at a different station when they reach their destination.
- Hereby dataset contains information about individual trips for October 2020.
- Number of total trips: 2,248,869

# Average speed of different age groups



The 20-30 yo group has the highest average speed (9.28 km/h)

It's understandable that group aged under 20s is at the bottom, even lower than above 60s group because the dataset only includes from 16 yo + users, while above 60s group includes people up to 75 yo.

# Average speed by age and gender (1)

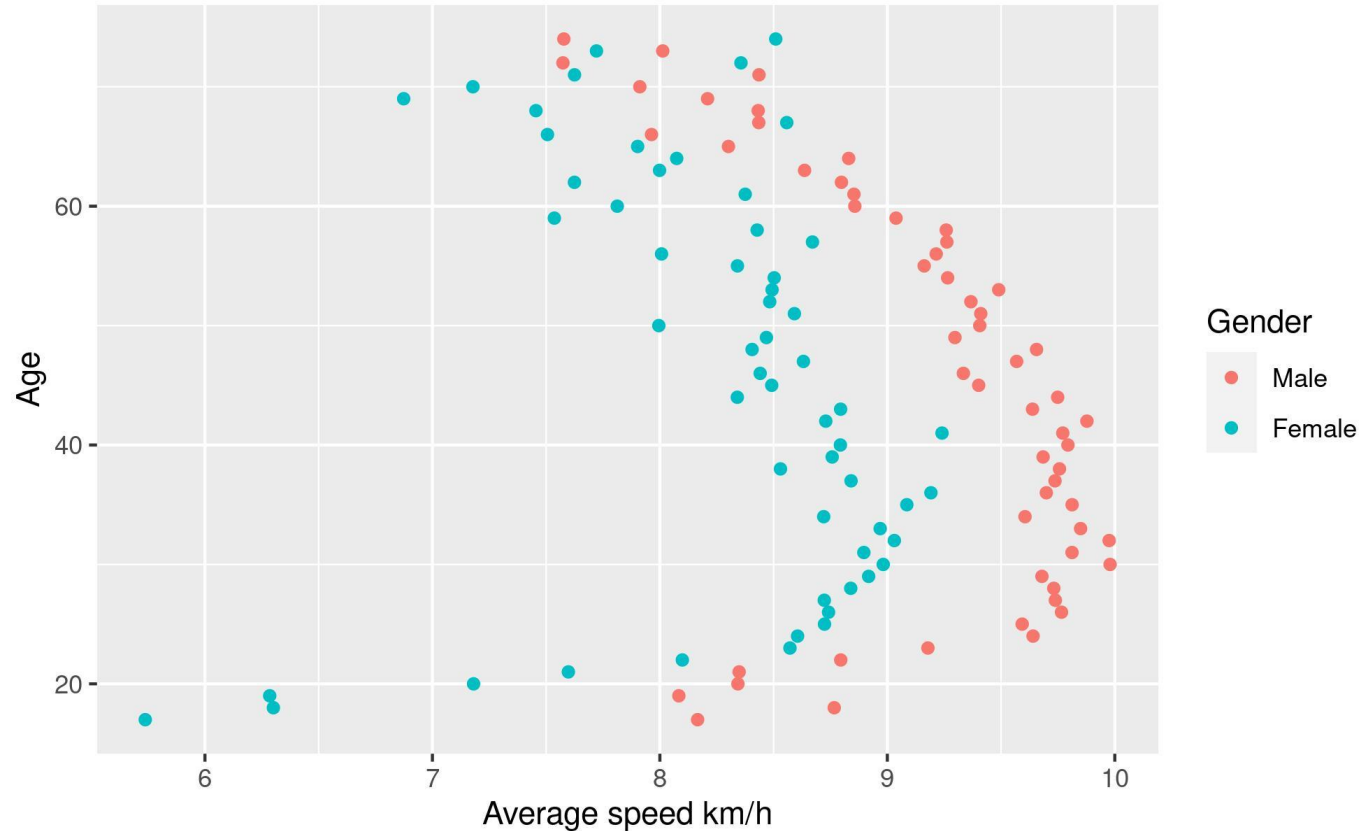


Male users, on average, have higher average speed than female users in all groups.

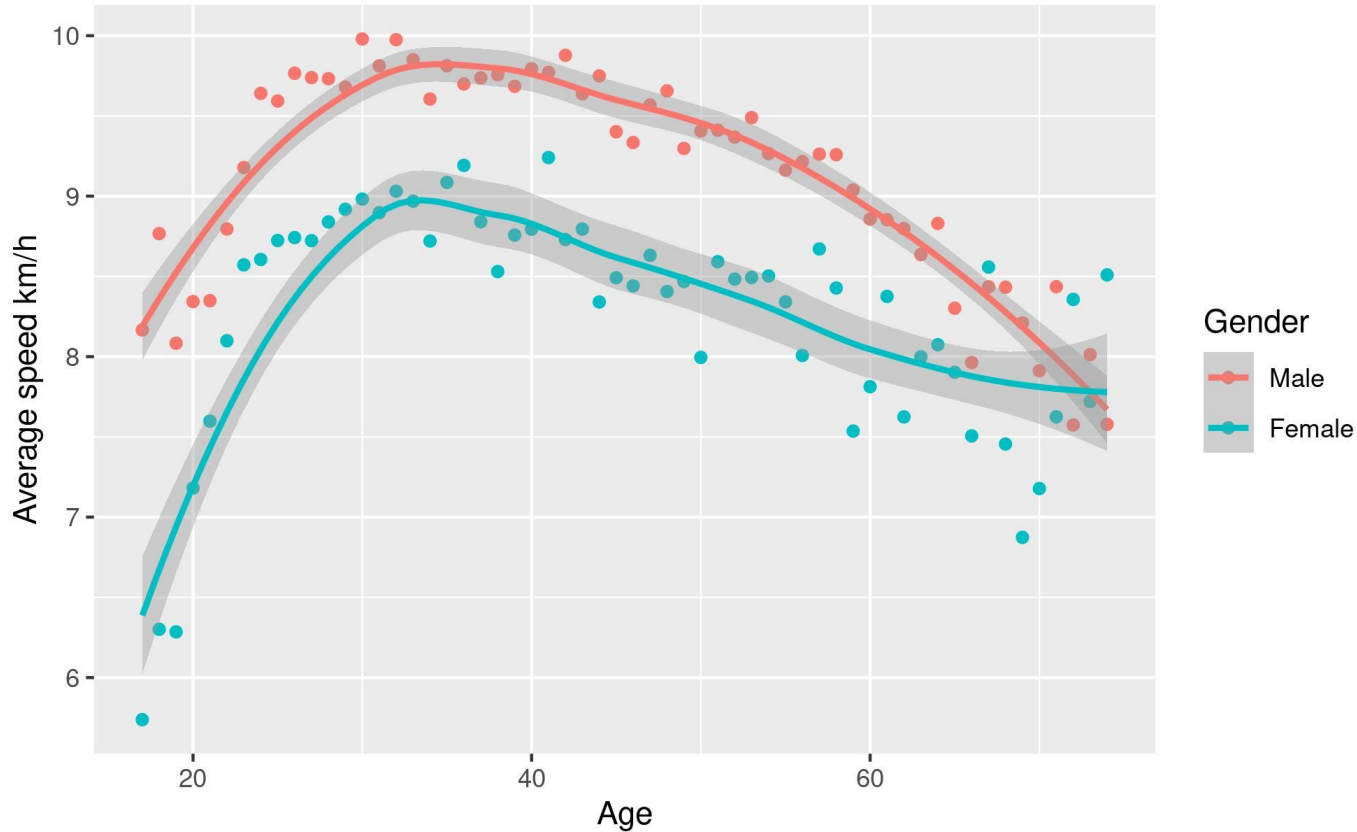
When doing more calculation we find that average speed of male group is approximately 8.846 km/h while female group's is 7.941km/h (1.114 times higher)

Users with an unknown gender do not follow any specific pattern. It is likely that there isn't enough data to properly visualise those users.

# Average speed by age and gender (2)



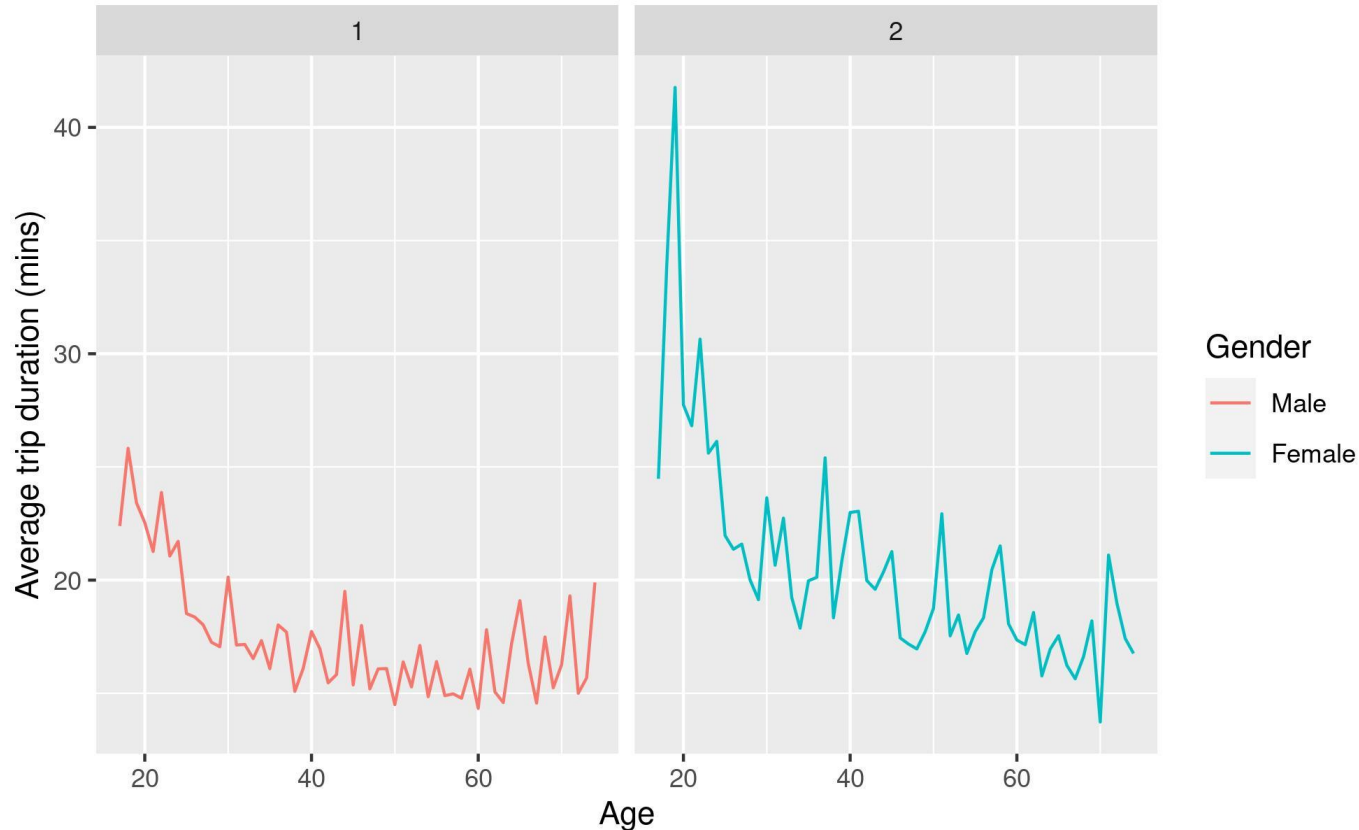
# Average speed by age and gender (3)



Is the average speed affected by age and gender? YES!

This plot also shows that both groups reach the peak of their speed around the age 35-40 yo.

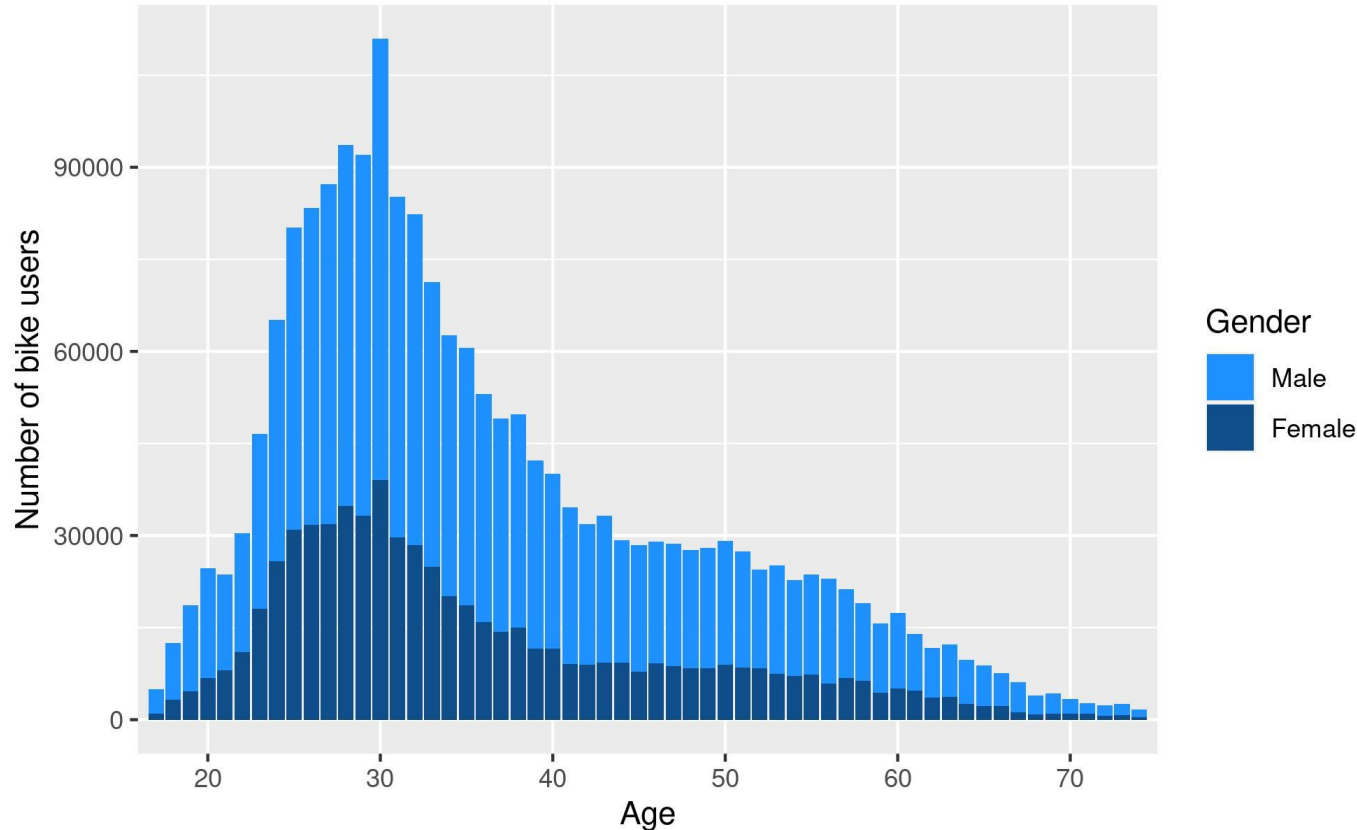
# Average trip duration by age and gender



Female riders on average spend more time on their trips compared to their male counterpart. This can be explained by their speed which tends to be slower than male's.

The trip duration of female aged 20 yo especially high. This is quite interesting as when looking at the previous chart we find that the speed of this group is not that low! Could be because they love to cycle, hence, cycle longer distance?

# Distribution of Citibike users' gender



At all age groups the number of male users are always higher than female users, especially at the age of around 30 yo.

Unsurprisingly the groups aged 25 yo - 35 yo show the highest number of users for both male and female.



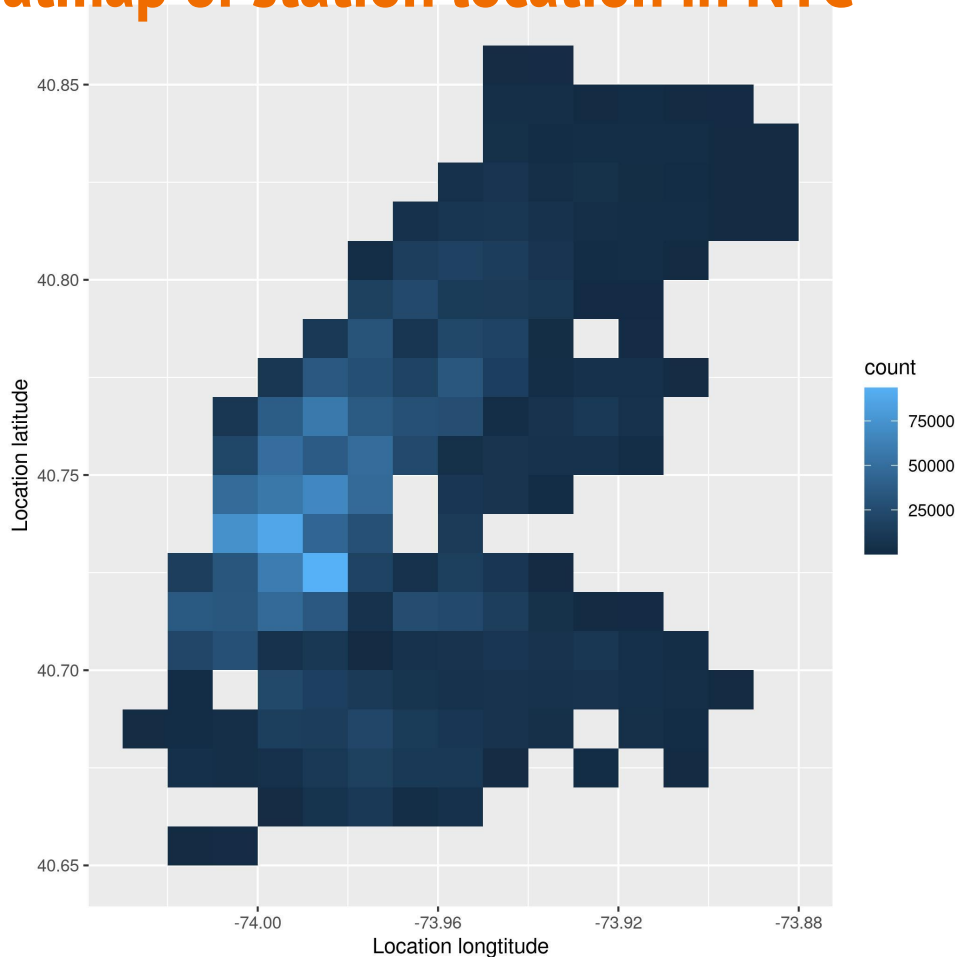
# Distribution of Citibike user types by age and gender



The numbers of subscriber of all ages are much higher than those of customers.

However at group aged 51 yo, the number of customer is exponentially higher than that of subscriber. It's likely that there is an error in data that needs to be re-checked, or perhaps there was an event causing the jump.

# Heatmap of station location in NYC



The map illustrates NYC which shapes Manhattan, Brooklyn, and Queens. The rectangle in Manhattan with no stations; that's Central Park.

Manhattan area observed high number of trips (more than 75000 trips recorded)

## Distance between 2 stations

Max distance between 2 stations is 21.21627km which links 14St & 7Ave St and W181 St & Riverside Dr

Min distance is 0.03677km links Pershing Square North and Pershing Square South