**REPORT**

Thesis: Analysis of weather, location and holidays affecting the reception of Uber drivers in New York City.

**Descriptive analysis**

Description: The dataset was collected from 01/01/2015 to 30/06/2015, contains more than 10 million observations of 4 variables aggregated by hourly and by county, then combined with generated weather data out 29,101 observations on 13 variables.

● pickup\_dt: Observation period (quantitative)

● borough: The borough of NYC. (qualitative) – String

● pickups : Number of uber rides. (quantitative) – Integer

● spd: Wind speed in miles per hour. (quantitative) - Float

● vsb: Visibility in Miles to the nearest tenth. (quantitative) - Float

● temp: temperature in Fahrenheit. (quantitative) - Float

● dewp: Dew point in degrees Fahrenheit (quantitative) - Float

● slp: Sea level pressure. (quantitative) - Float

● pcp01: Liquid precipitate for 1 hour. (quantitative) - Float

● pcp06: Liquid precipitation for 6 hours. (quantitative) - Float

● pcp24: liquid precipitate 24 hours. (quantitative) - Float

● sd: Thickness of snow in inches. (quantitative) - Float

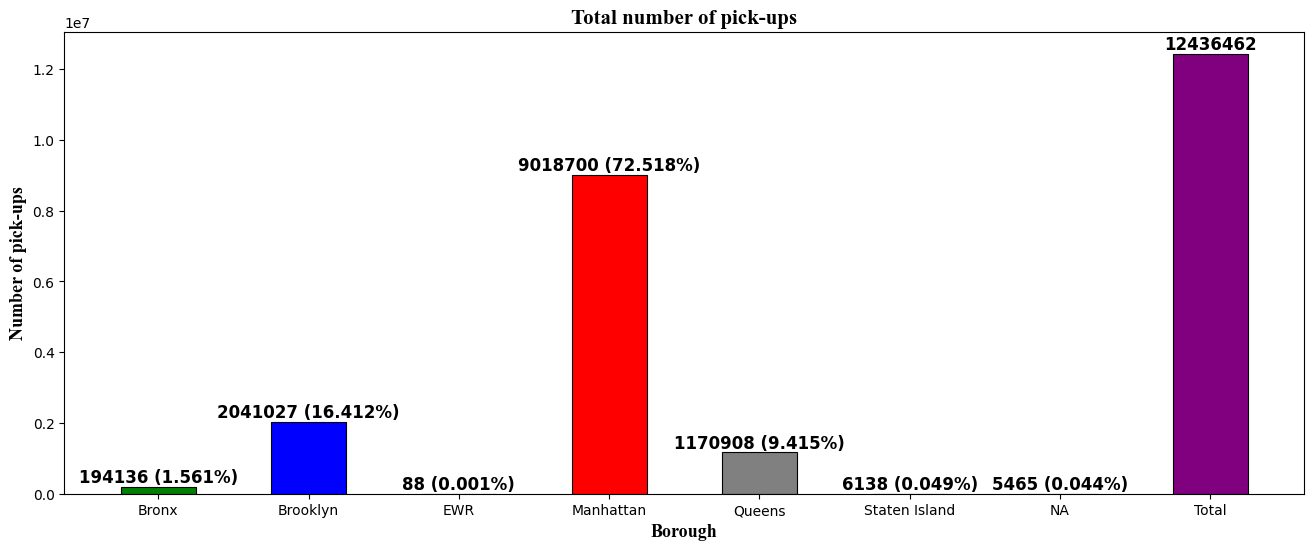
● hday: Is it a holiday (Y) or not (N) (qualitative) – String

Location (New York City), including counties:

* Bronx
* EWR
* Brooklyn
* Manhattan
* NA
* Queens Staten
* Island

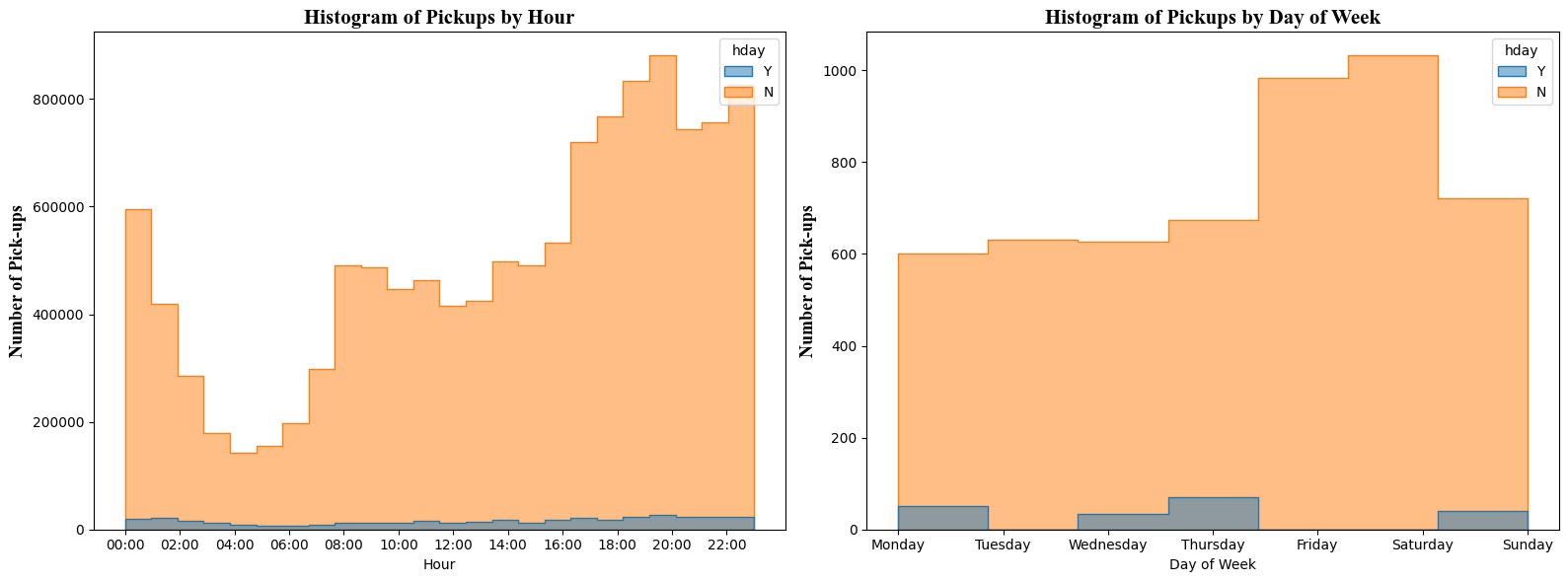
**Analysis of EDA:**

1. Total number of bus trips of each district

****

* We can see that the number of Ubers pickup services in Manhattan is very large and accounts for more than 72% of the total number of rides of the districts in New York City.
* From that, it is observed: The number of passengers, Uber pick-up service in Manhattan has a great influence on the total number of pick-up trips on New York City.

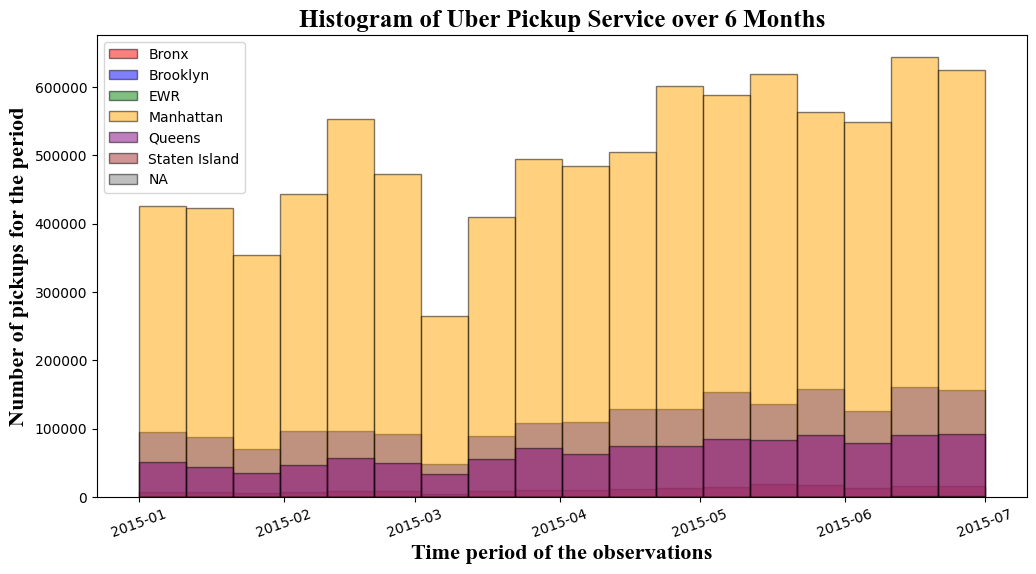
1. Number of bus trips by hour and between holidays and weekdays:



In general, the number of trips on weekends or holidays is much higher than on weekdays when it reaches 800,000 trips in 5 months in all districts. Customers usually pick up the car mainly in the evening around 20h and 22h. On weekdays, although the bus trip is not much, there is no difference in the time that customers book the car, all time frames are full of customers.

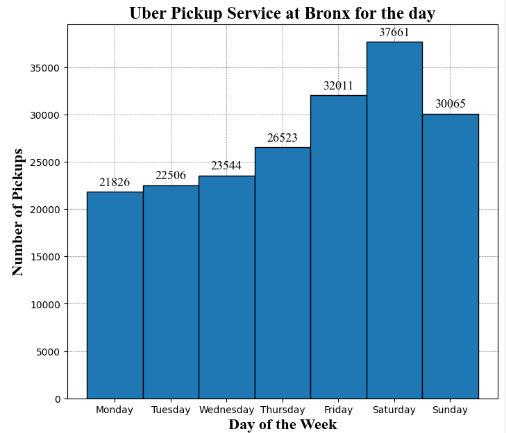
On public holidays/weekends, Friday and Saturday have more customers using the service than other things with over 1000 trips. However, for weekdays, Friday and Saturday are very few customers, mainly people use the uber service on Monday, Thursday and Sunday.

1. Number of trips within 6 months:

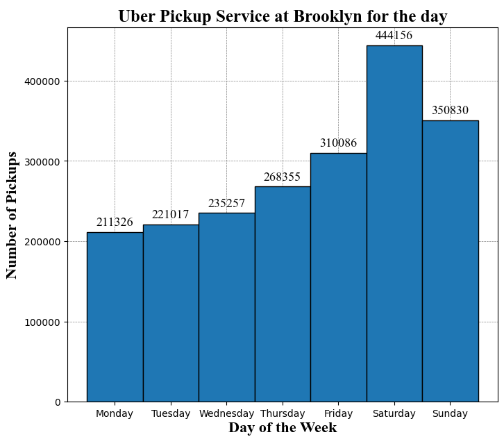


Within 6 months, Manhattan district leads in the number of bus trips with more than 600,000 trips (June-2015). Next is NA district with the peak time of nearly 200,000 trips. In third place is Queens, with nearly 100,000 trips. Next is Staten Island with about 30,000 trips. The remaining counties are much less than those mentioned above.

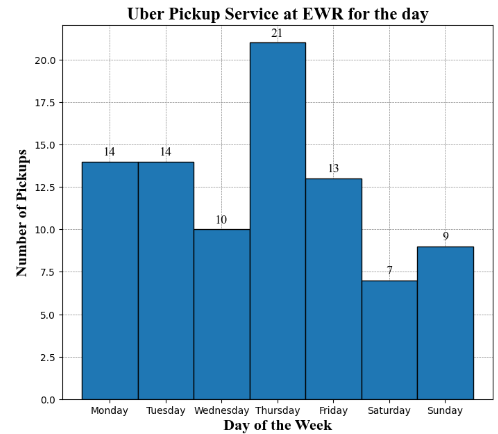
1. Number of trips in a week of each county in 6 months:



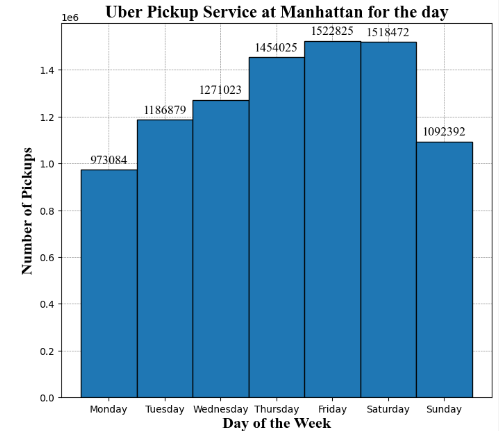
* BronX District, has the highest number of trips on Saturday with 37,661 trips.



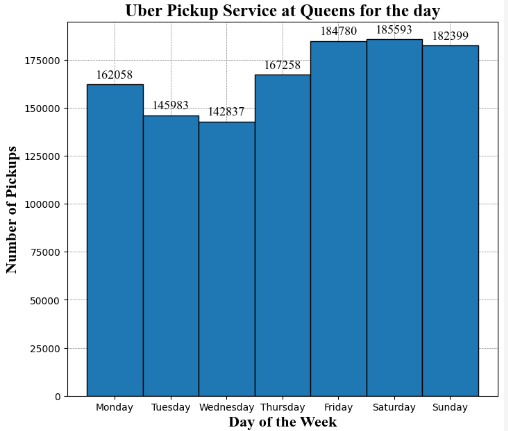
* Brooklyn District, has the highest number of trips on Saturday with 444,156 trips.



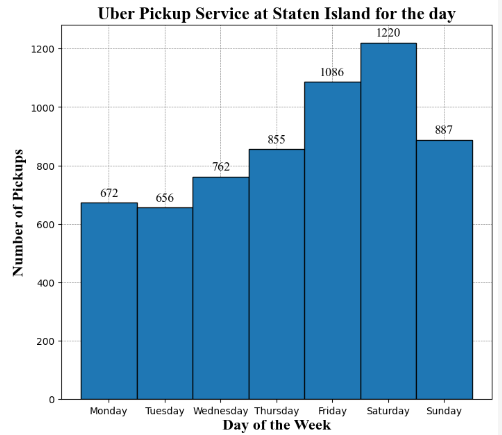
* EWR District, has the highest number of bus trips on Thursday with 21 trips.



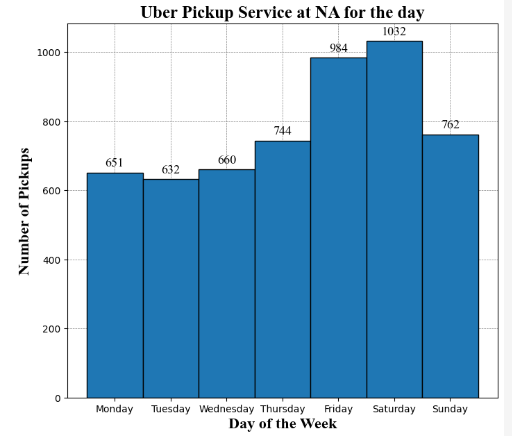
* Manhattan District, has the highest number of trips on Friday with 1,522,825 trips.



* Queens District, has the highest number of trips on Saturday with 185,593 trips..

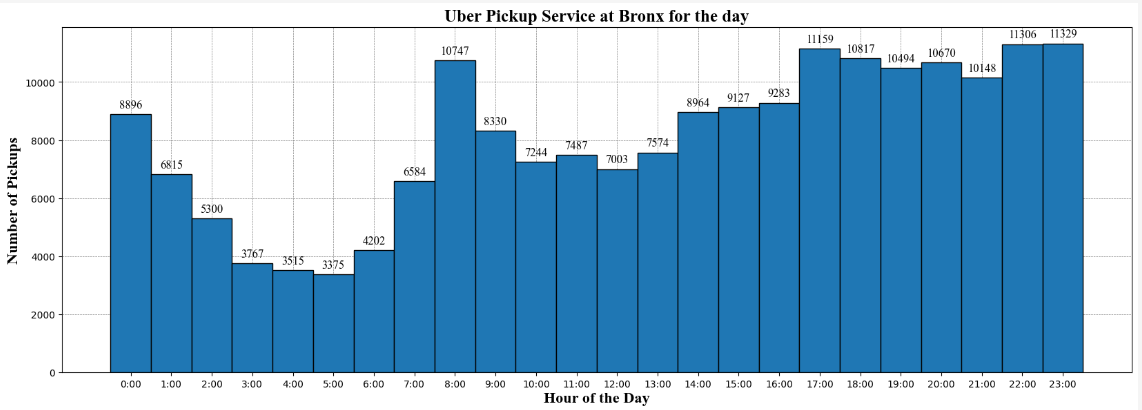


- Staten Island District, has the highest number of trips on Saturdays with 1220 trips.

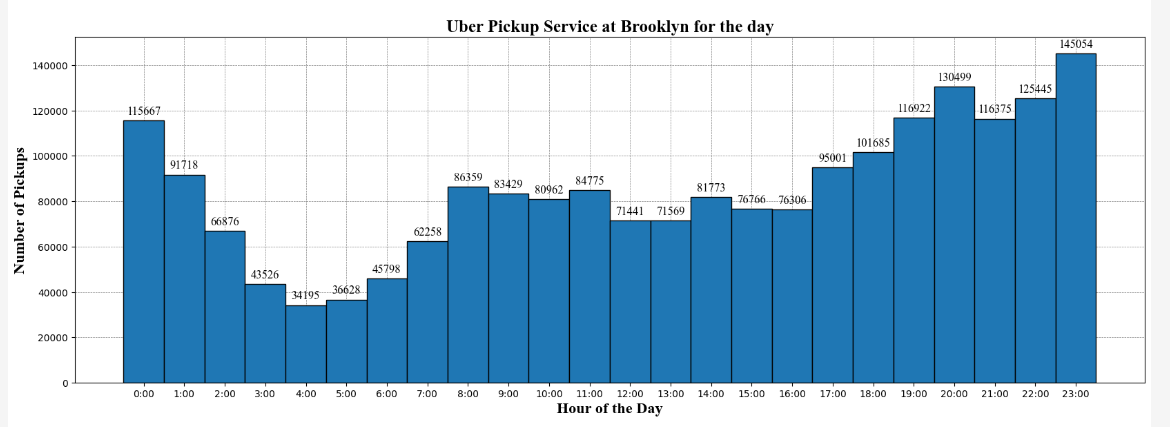


* NA District, has the highest number of bus trips on Saturday with 1032 trips.

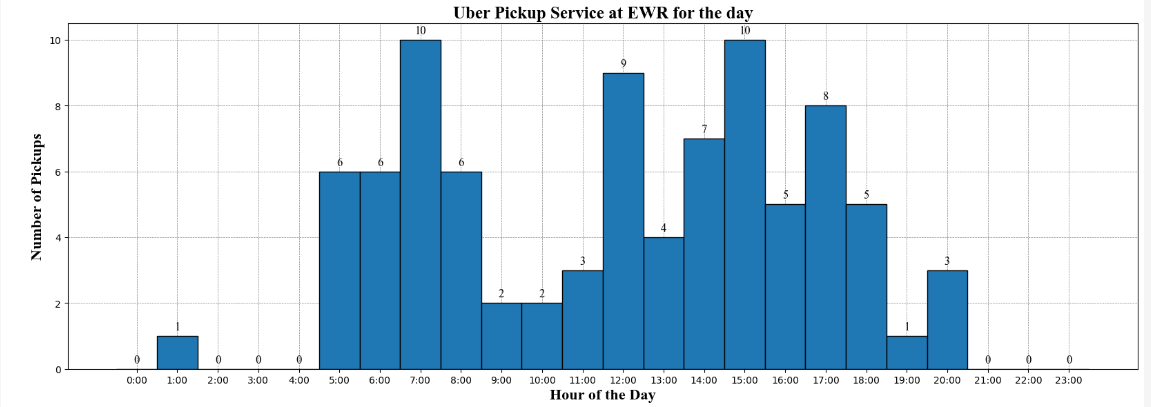
1. Number of bus trips by hour of each district:



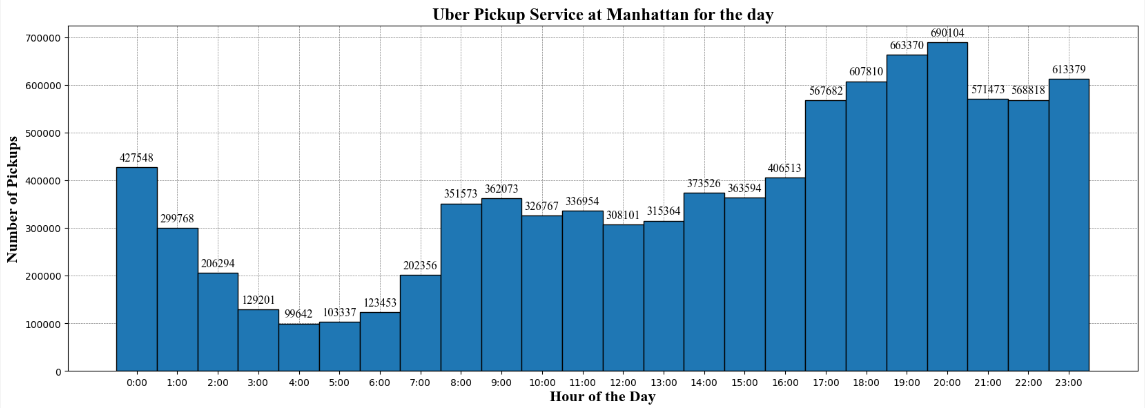
In the Bronx, customers often catch the bus to go to work at 8 am and at 5 pm at the end of the shift with the number of trips of 10,747 and 11,159 respectively. After 17:00, the number of car rides also increased steadily.



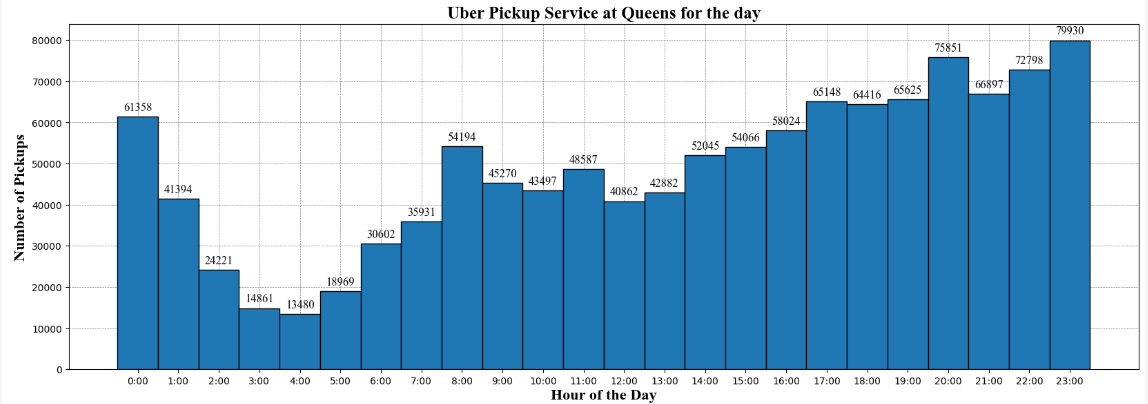
Brooklyn District, customers also catch the bus at 2 peak hours at 8 am and 5 pm. At 23:00, the demand for the service is much higher.



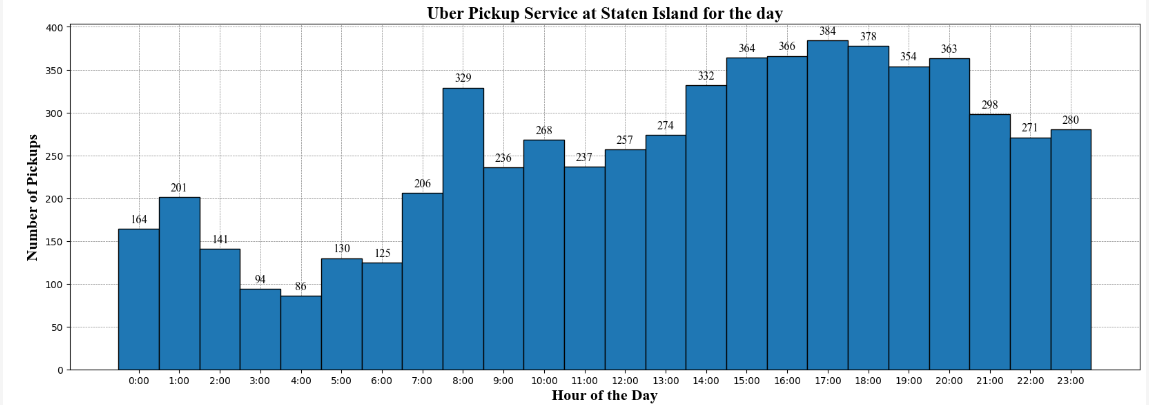
EWR district, customers do not use uber much, only about 10 people catch the bus at 7:00 am and 15:00 pm. Customers in EWR district have no need to use Uber when the number of rides after 9pm is 0 rides.



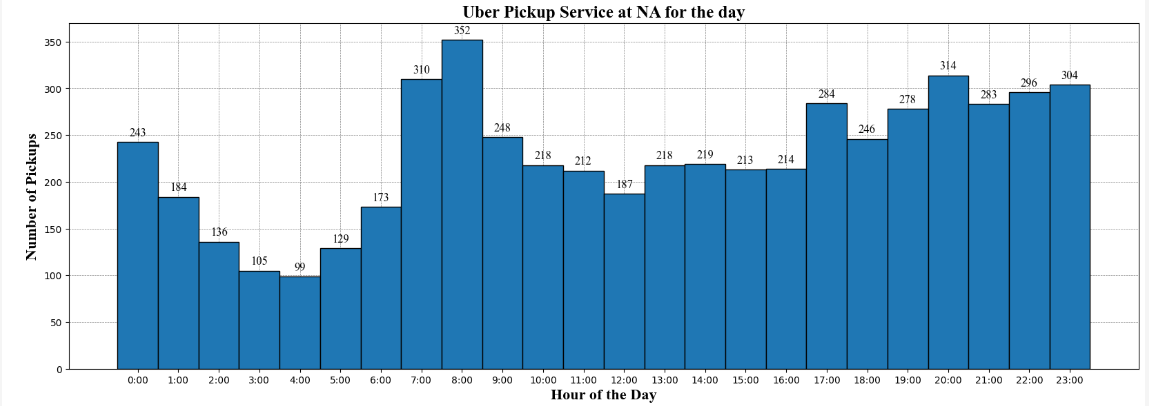
In the densely populated Manhattan district, we can see that the number of rides is significantly higher than that of other districts in each time frame, the highest is still at 8 pm with 690,014 trips.



In Queens District, the number of rides increases by the hour after 18:00.

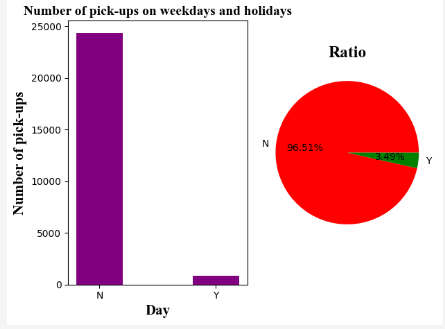


* On Staten Island, customers seem to be less likely to use Uber's car service late at night after 9pm. Customers often book cars from 15:00 to 20:00 with nearly 400 trips.



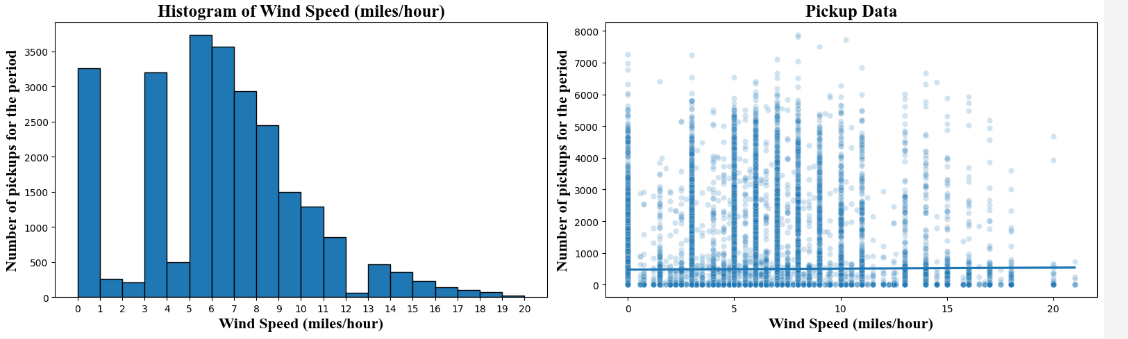
In Na district, customers often use Uber car service at 8:00 am with 352 trips and 5:00 pm with 284 trips. The period from 20:00 to 00h also has many customers using the service.

1. Compare bus trips between holidays/weekends and weekdays:



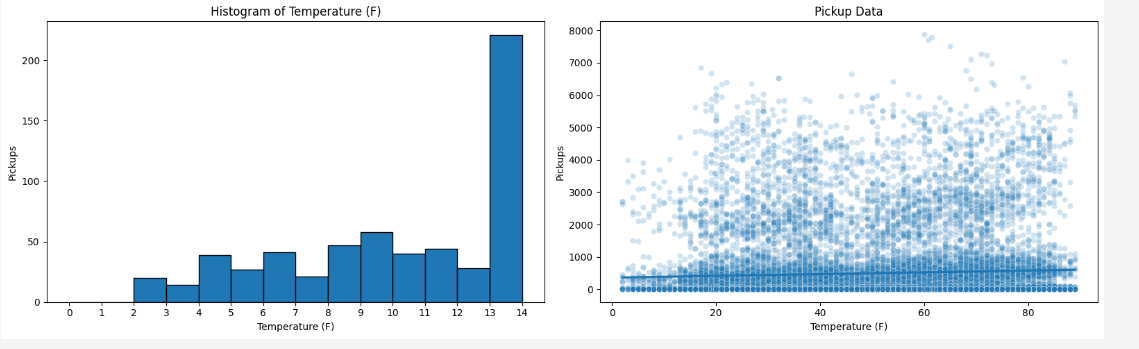
Looking at the chart, you can see a clear difference between weekdays and holidays/weekends, when there are nearly 25,000 trips on holidays, accounting for 96.51%, and less than 5000 trips on weekdays, accounting for 3.49%.

1. Number of bus trips of each district in windy weather:

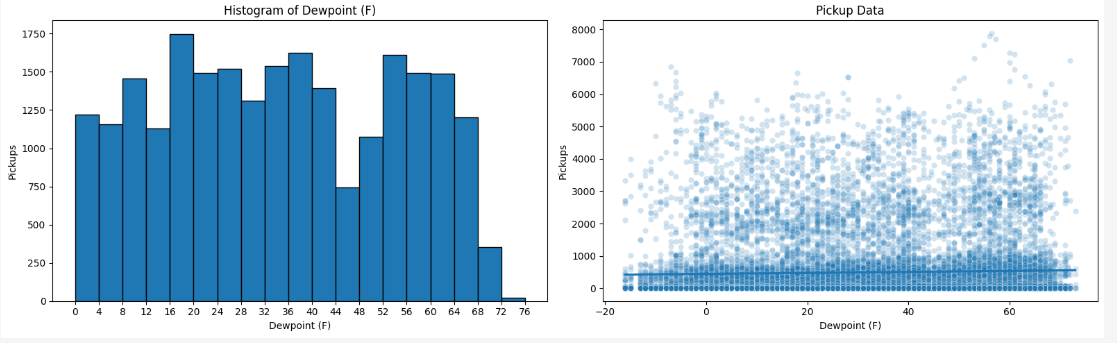


Wind speed level 0, level 3, level 6, level 7, level 8 occurs most often and it doesn't seem to affect the number of rides.

1. Number of bus trips per county based on outdoor temperature:

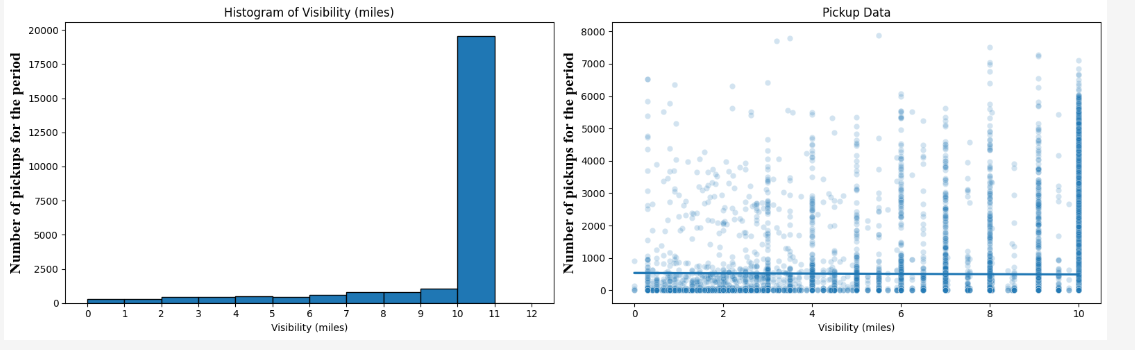


The common temperature in the districts is from 13 - 14 degrees F and the temperature is from 2 degrees F to 14 degrees F, customers still use Uber car service.

1. Số lượng chuyến xe của mỗi quận dựa vào sương mù: 

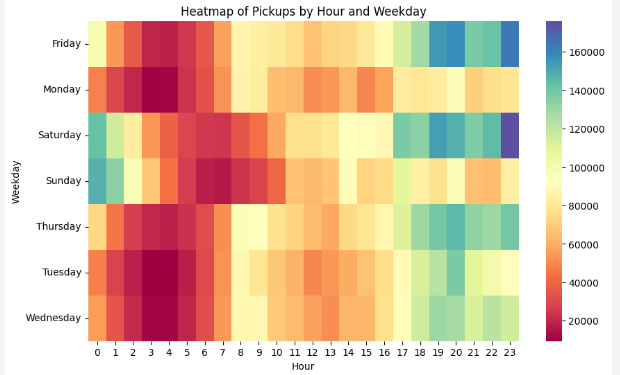
Fog density at YNC is frequent and has little impact on Uber service except when the fog is too dense above 72.

1. Number of bus trips per district based on visibility:



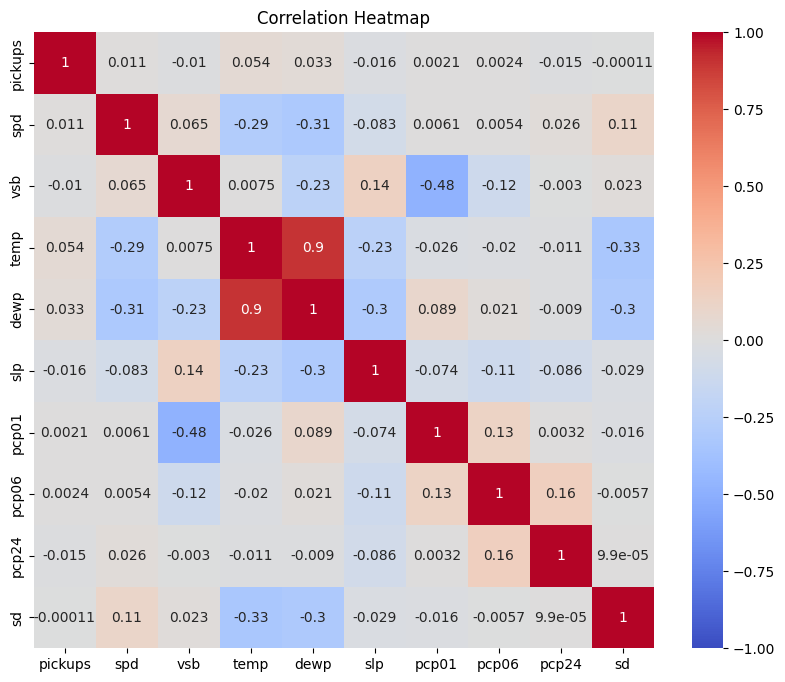
Visibility of 10-11 miles is ideal for safe travel.

1. Number of trips based on time and day of the week:



From the heatmap, it can be seen that Monday around 19:00 is the busiest time for Uber car service with over 100,000 trips. On Tuesday, at 20:00, the highest number of trips is about 140,000 trips. Wednesday is also from 19:00 onwards with a high number of bus trips. As for the two weekends, Saturday and Sunday, customers prefer Uber car service at midnight from 23:00 to 00:00 with over 140,000 trips in all districts.

12. Data Correlation



The correlation ranges from -1 to 1. Columns with correlations closer to 1 show that those columns have a positive impact on each other. On the contrary, the closer the correlation points are to -1, the more negatively those columns have a negative impact on each other.

Examples of columns that have a positive impact on each other are: dewp - temp