Analysis of

NYC Yellow Taxi data

Frank Wang

4/24/2016

NYC Yellow Taxi Data

This dataset includes trip records from all trips completed in yellow taxis in NYC in 2014.
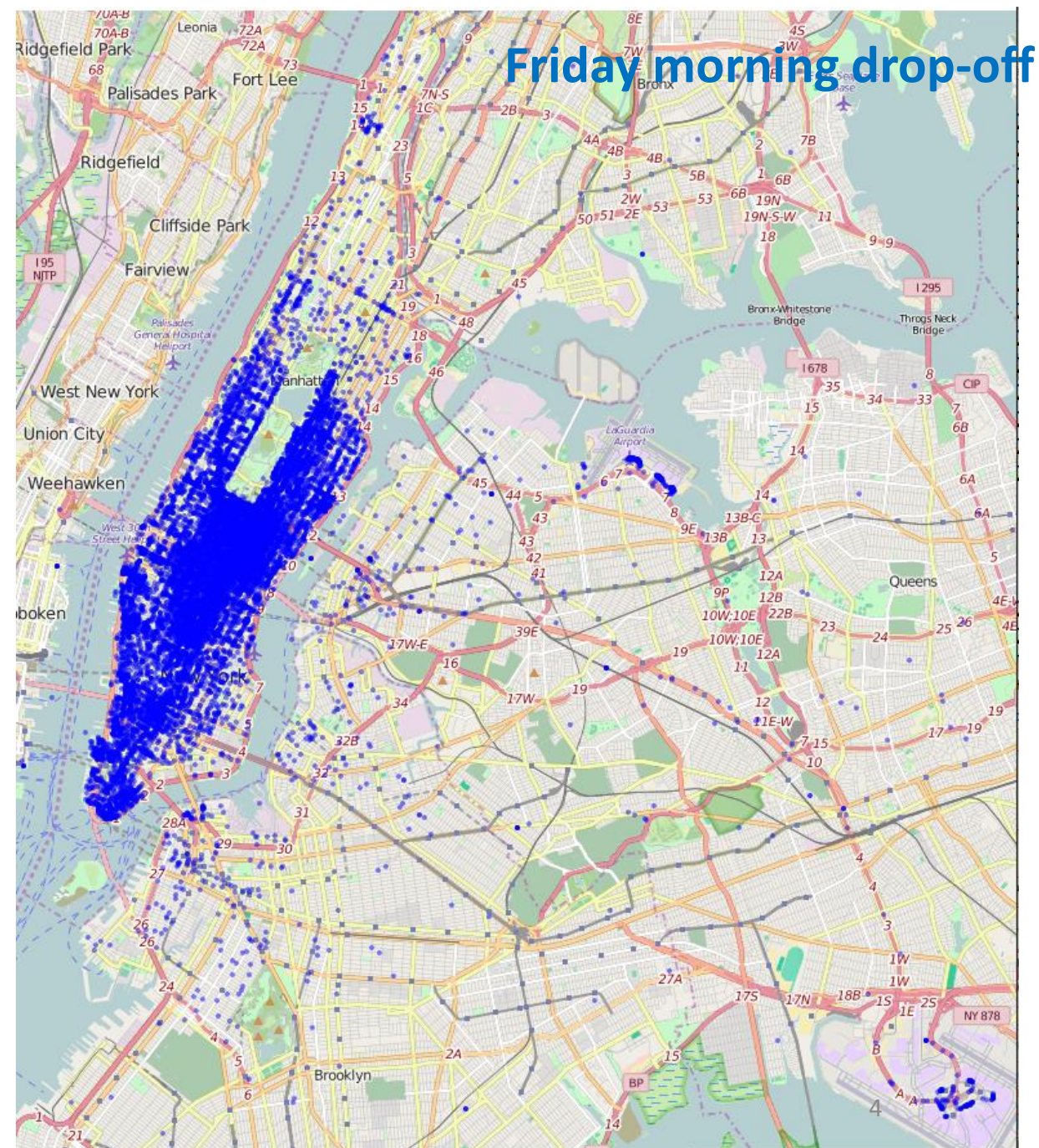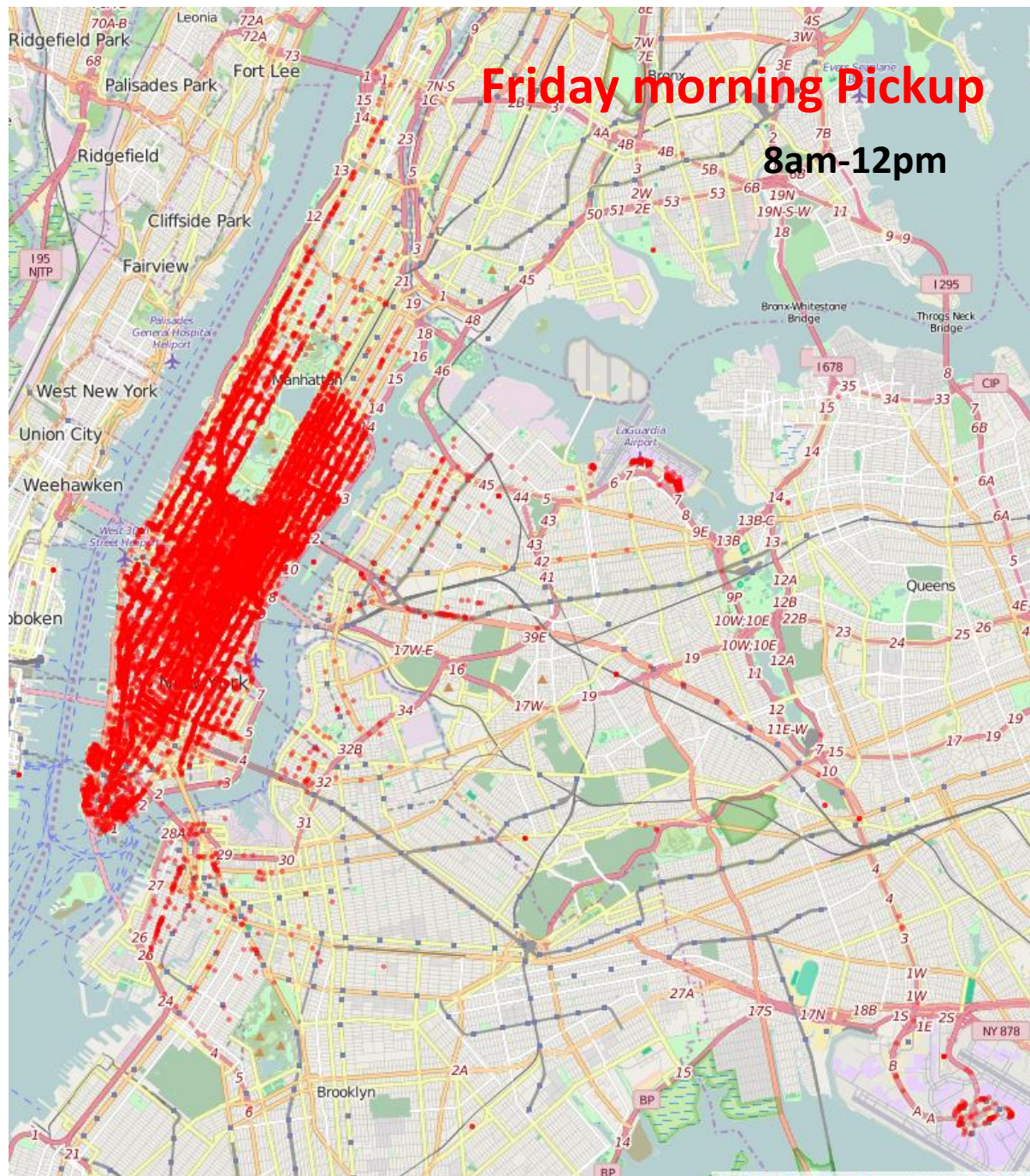
Records include fields capturing:
- ❏ pick-up and drop-off **dates/time**s
- ❏ pick-up and drop-off **locations**
- ❏ trip distances
- ❏ itemized **fares**
- ❏ rate types
- ❏ payment types
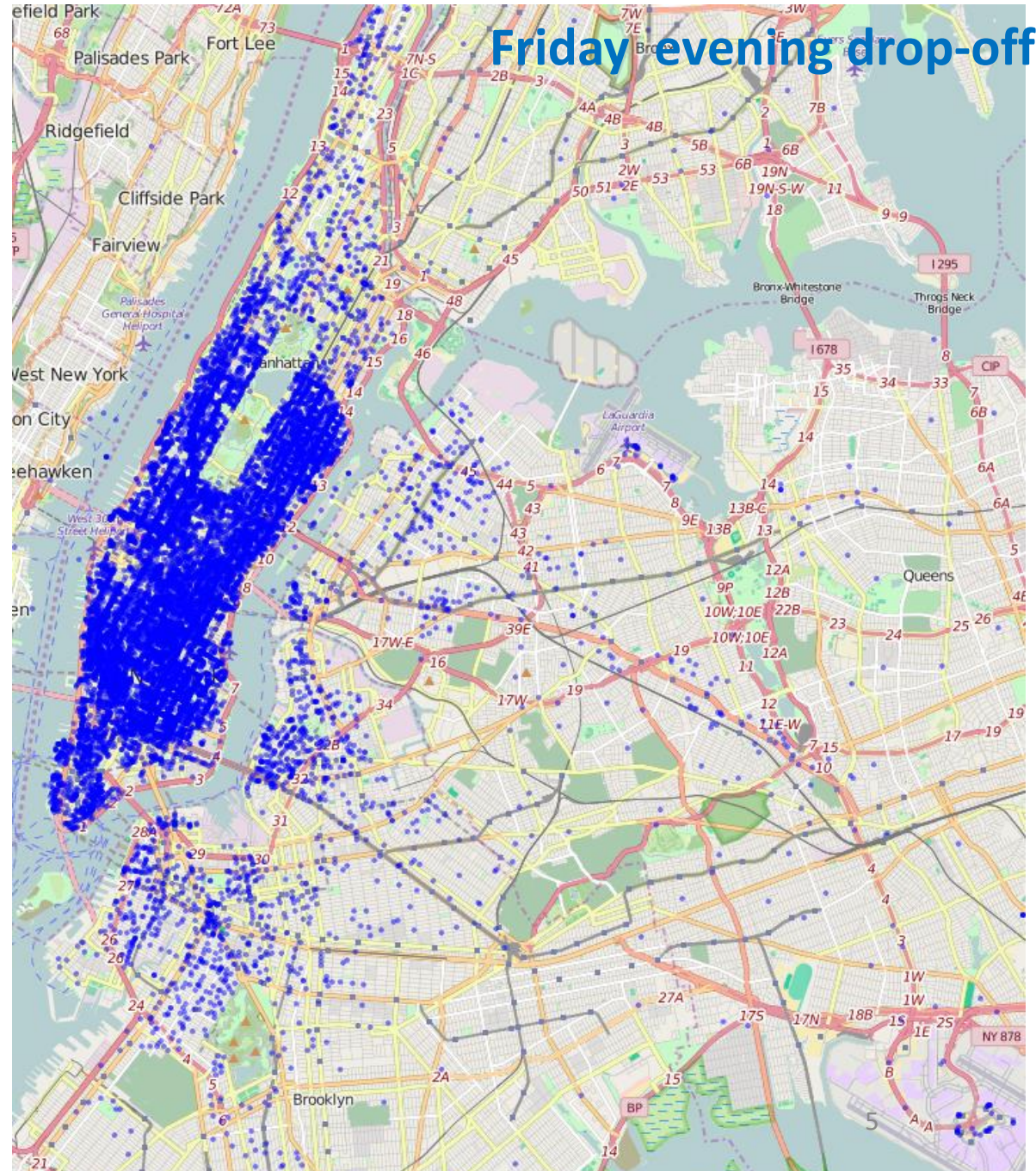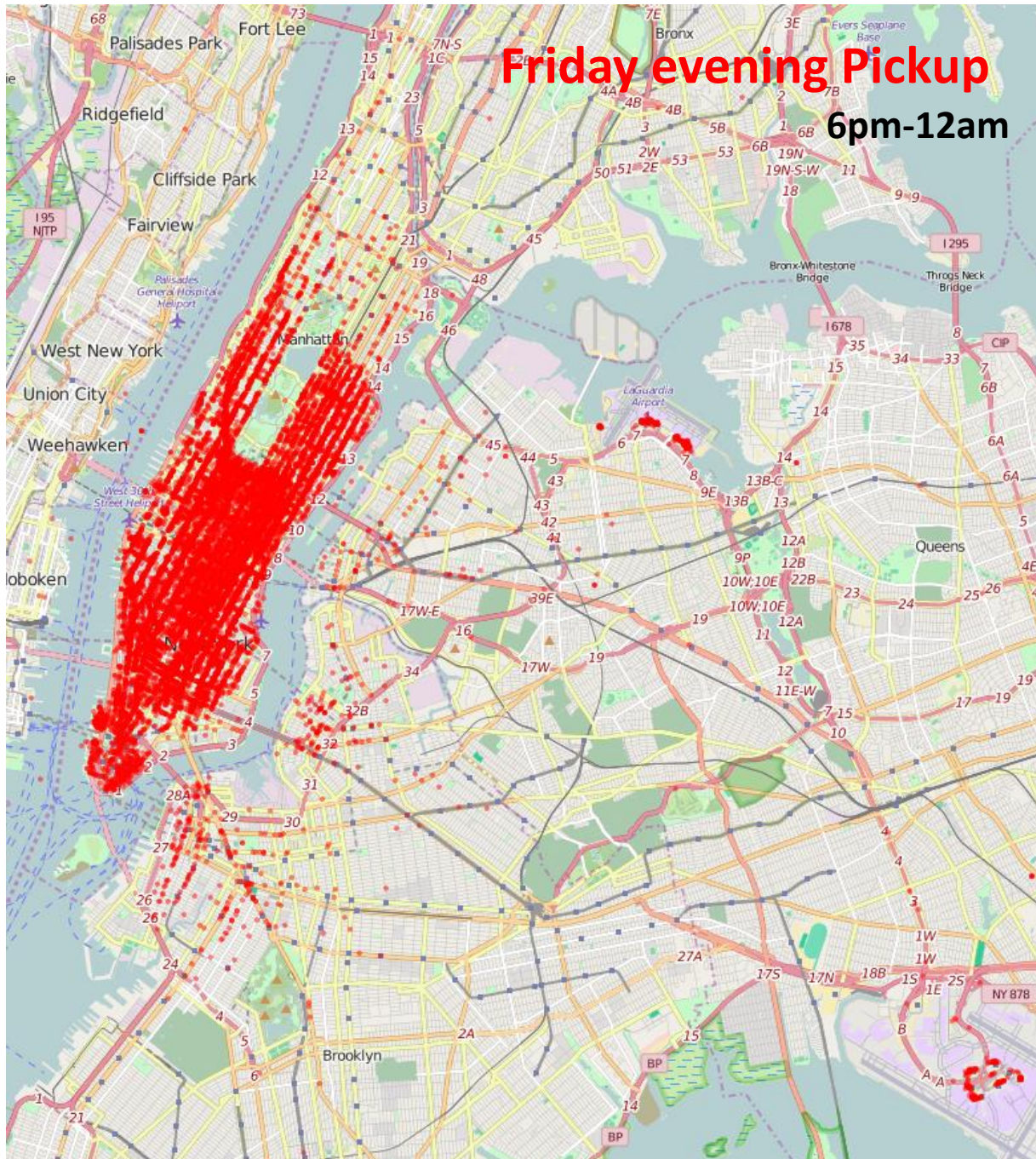- ❏ **passenger** counts

#The trip data was not created by the TLC, and TLC makes no representations as to the accuracy of these data.
Only 3 days dataset (1M) was used in this study. It is a challenge to handle such huge dataset in a PC
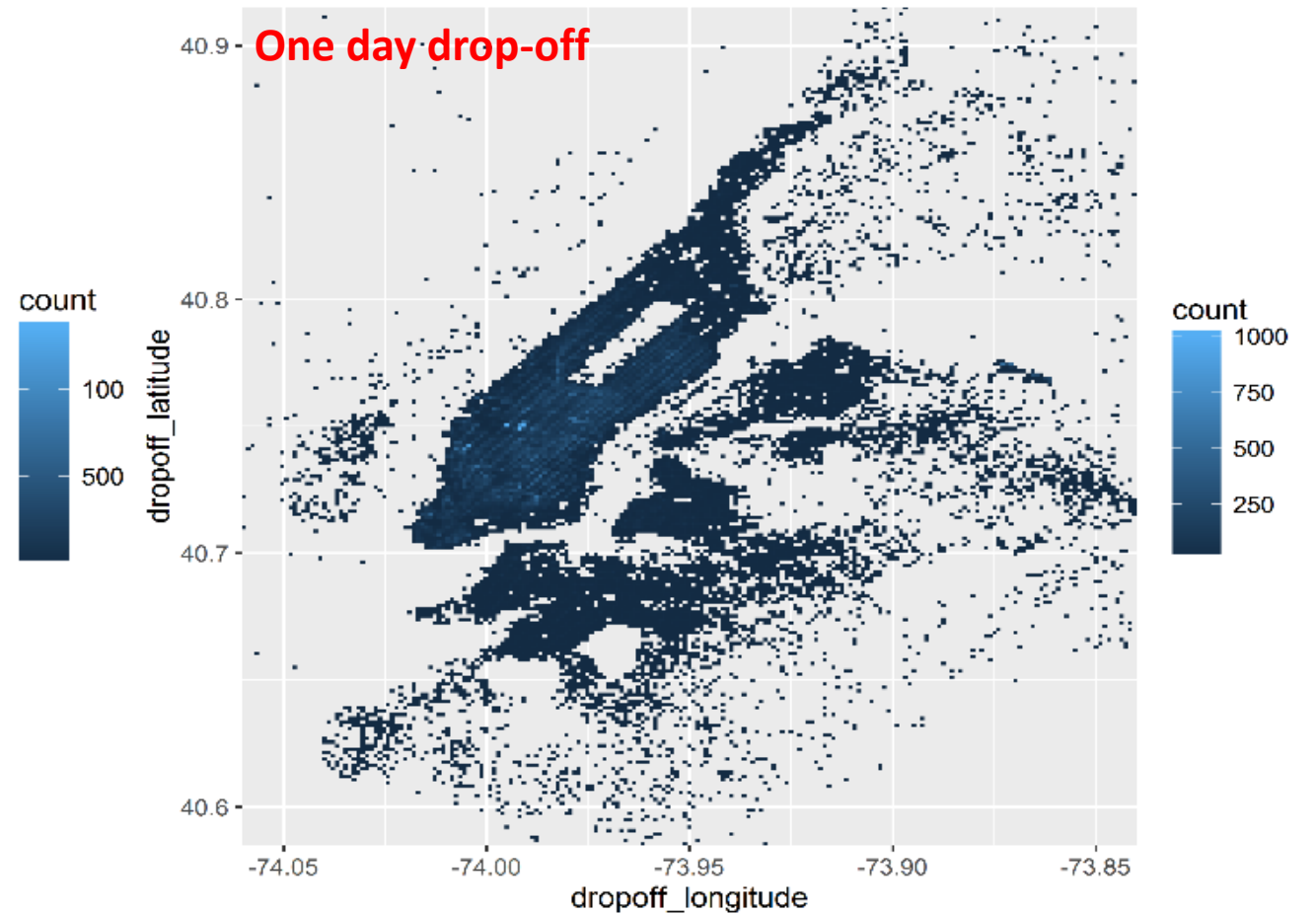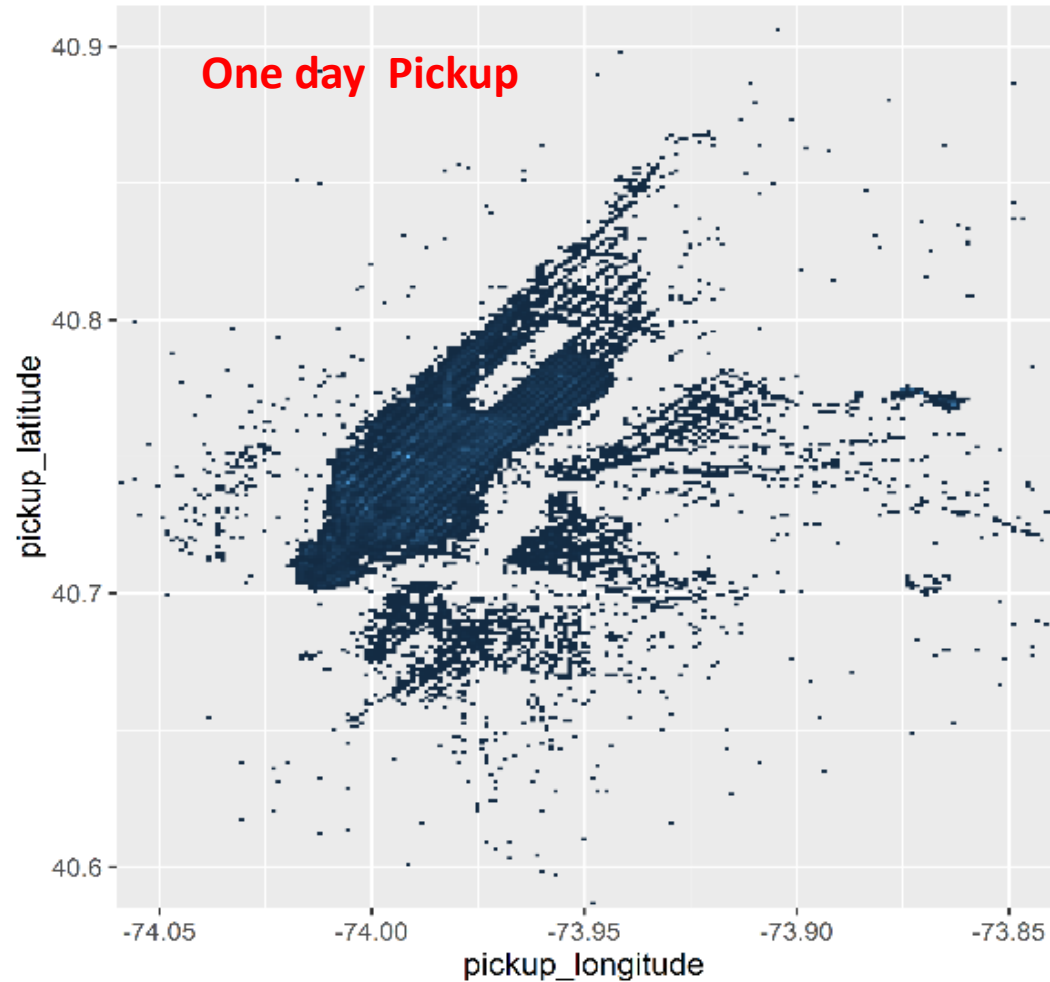
Explore the pattern in location and time
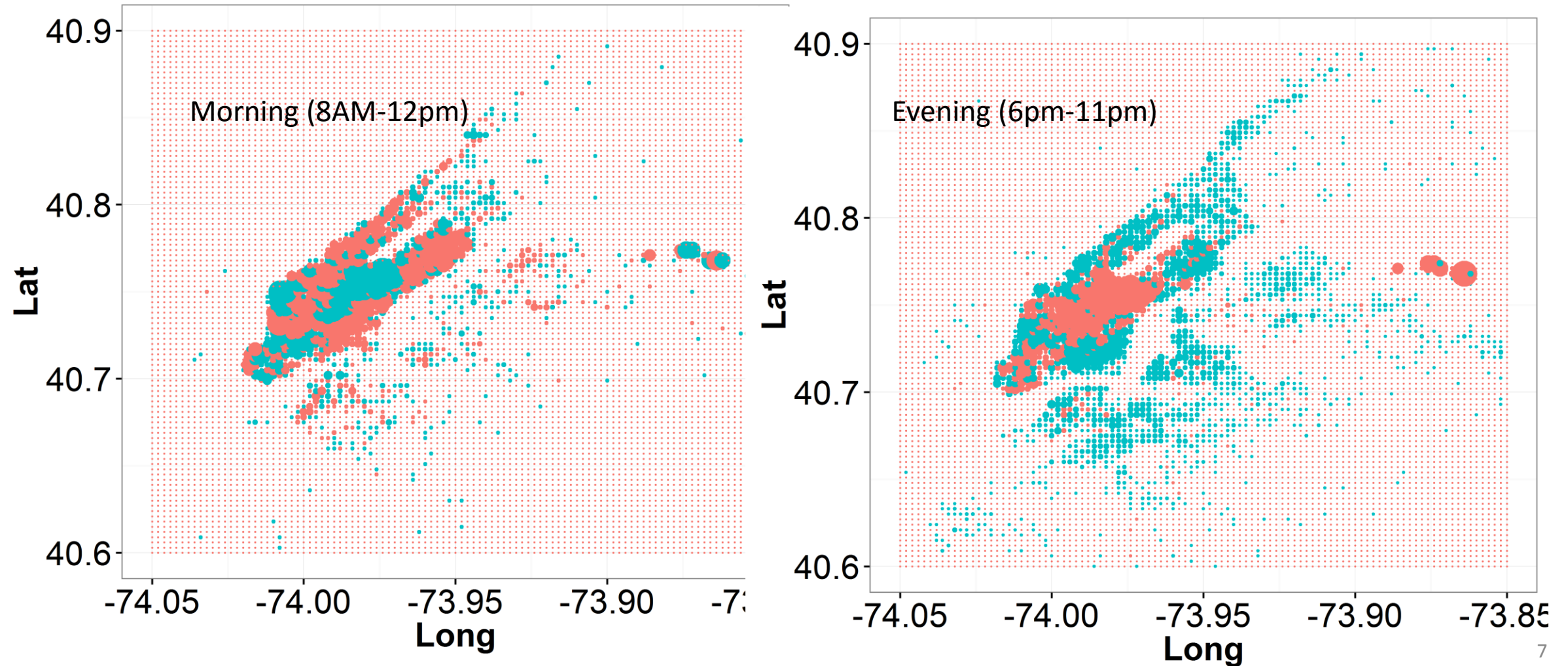
❑Where people go

❑Time dependence

❑Taxi at Airport

❑Tip

Friday morning Pickup

8am-12pm

Friday morning drop-off

Friday evening Pickup
6pm-12am

Friday evening drop-off

More people take taxi when they go home

One day Pickup

One day drop-off

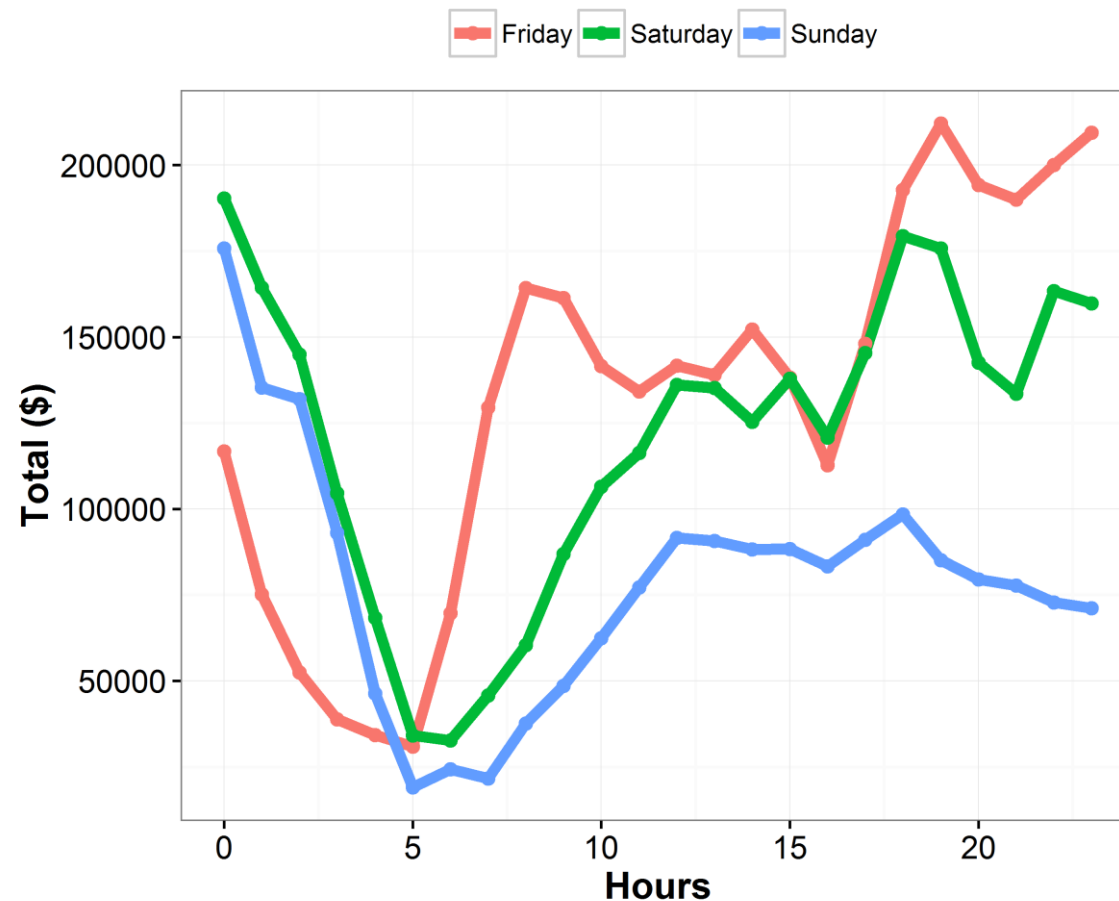# Net flow

At each location, calculate the difference between pickup and drop-off
red more pickup, blue more drop off

More people go home by taxi

Time dependence
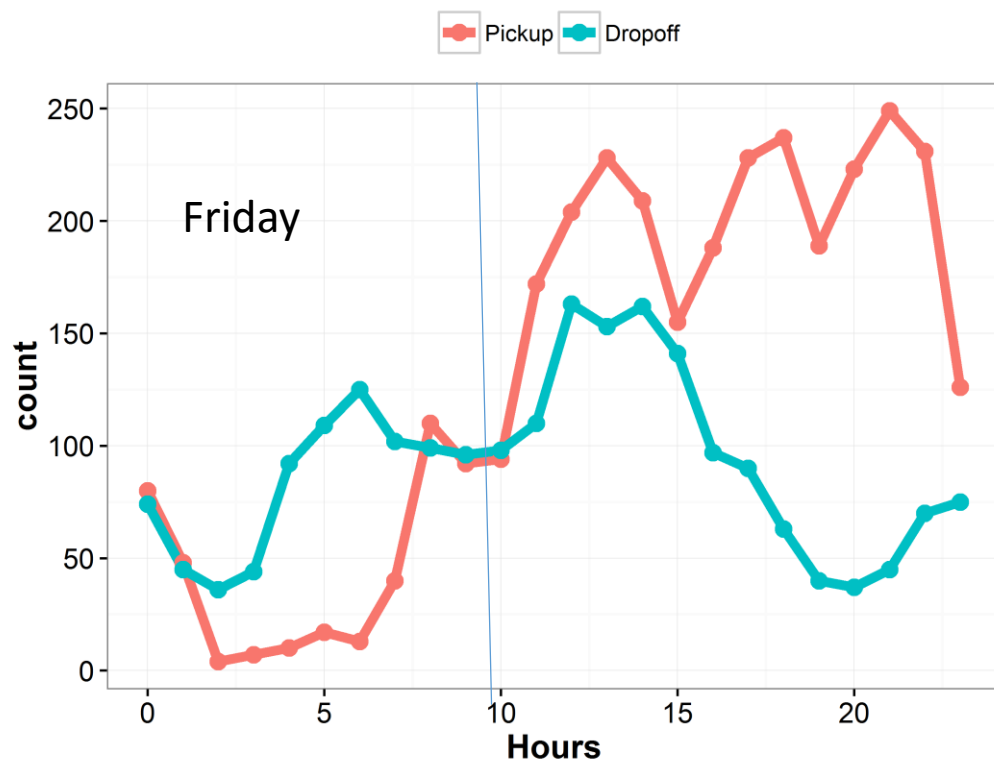
❑ There is minimum activity at 5AM for both workday and weekend
❑ There is peak right at 8AM for work day, but no such rush hour on weekend
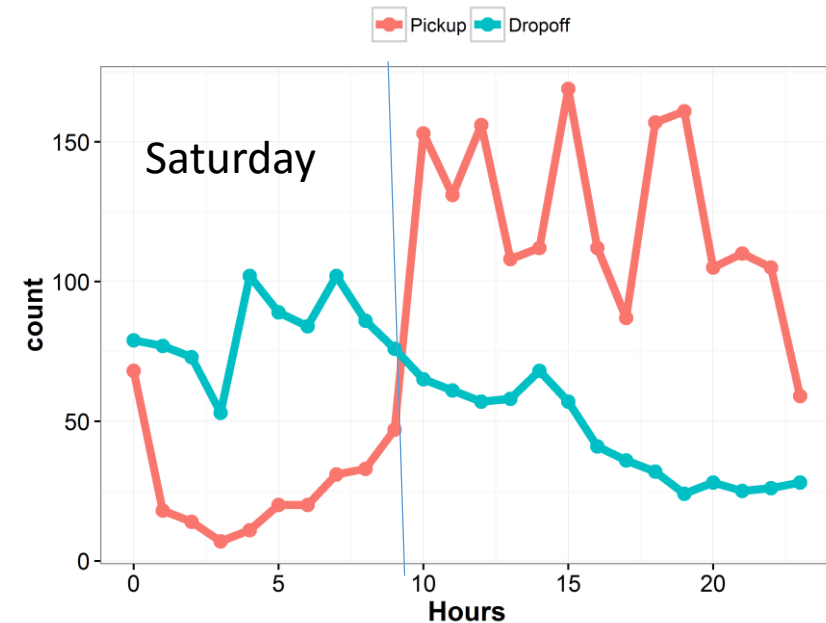❑ There is weak activity (local minimum) at 4pm (difficult time to find a taxi)

Taxi at LaGuardia Airport

- More people leave at morning and arrive at evening
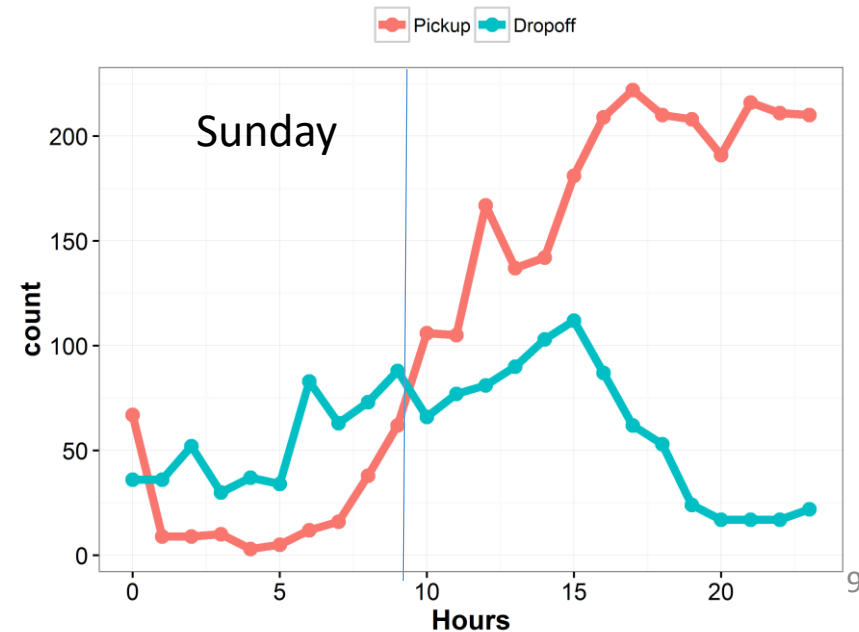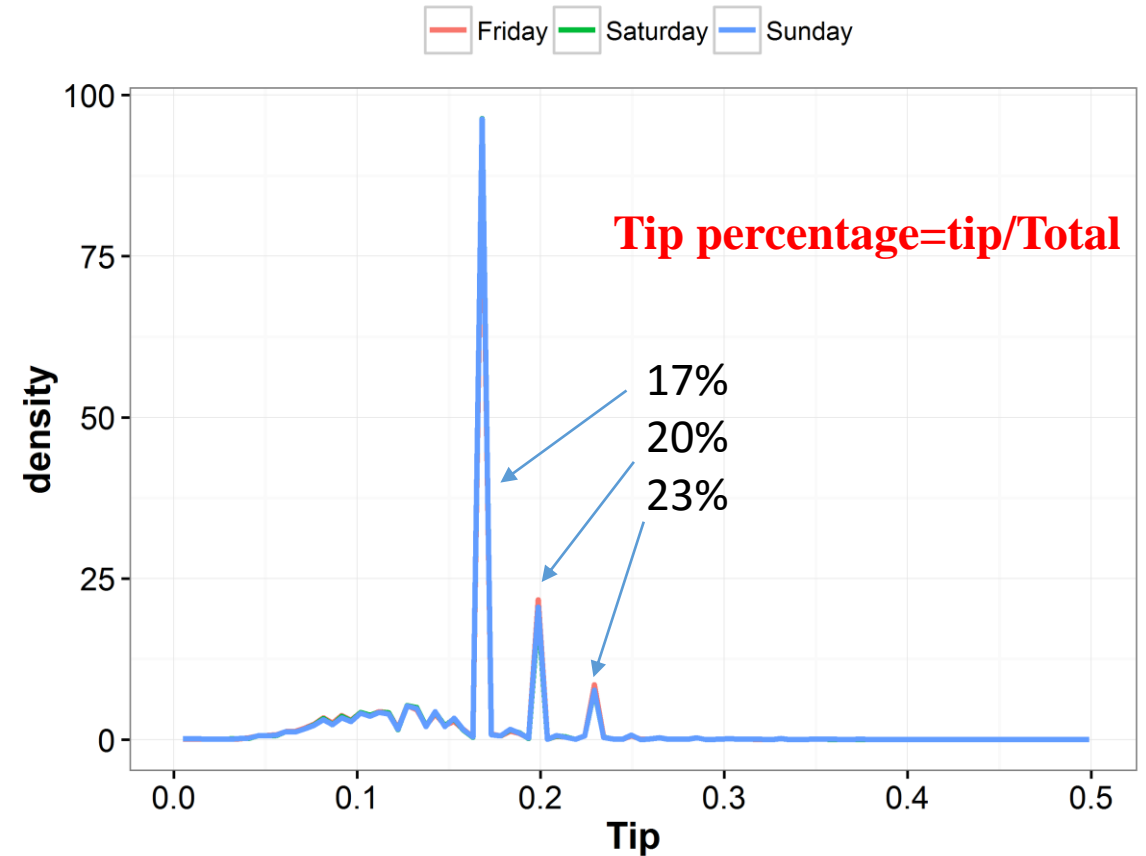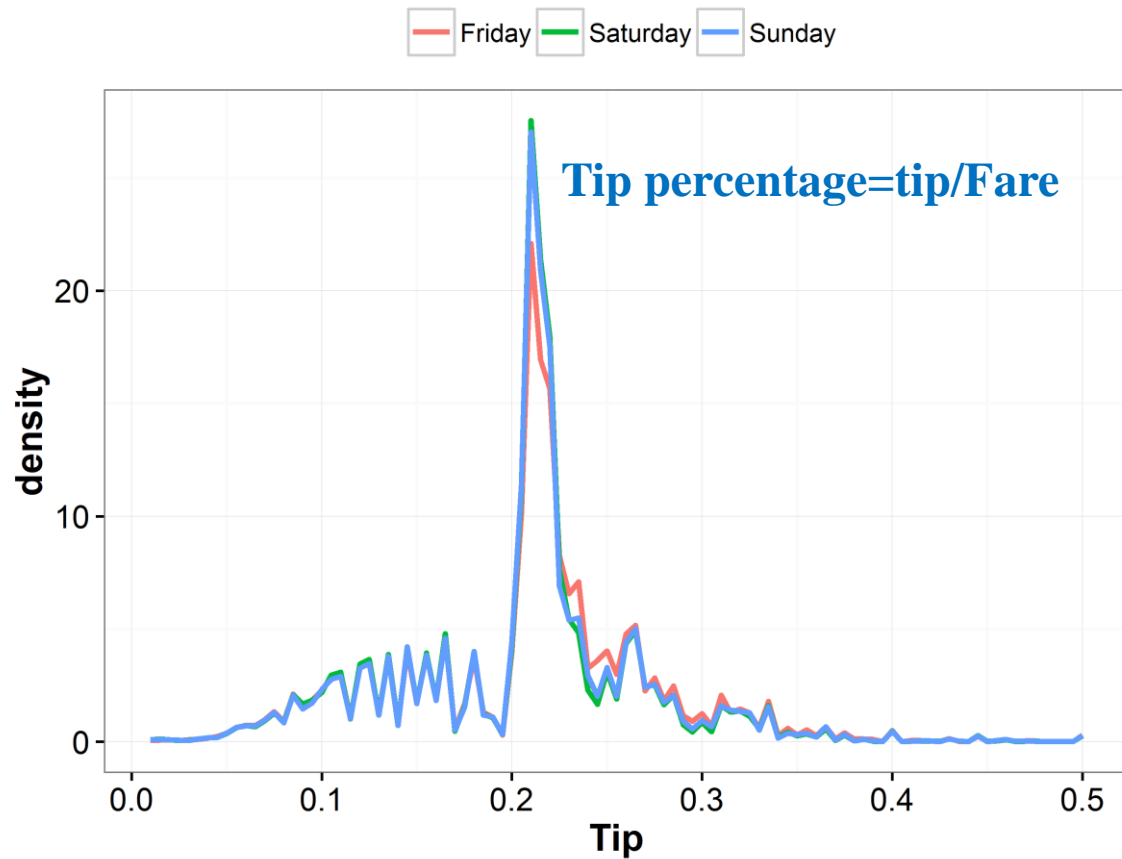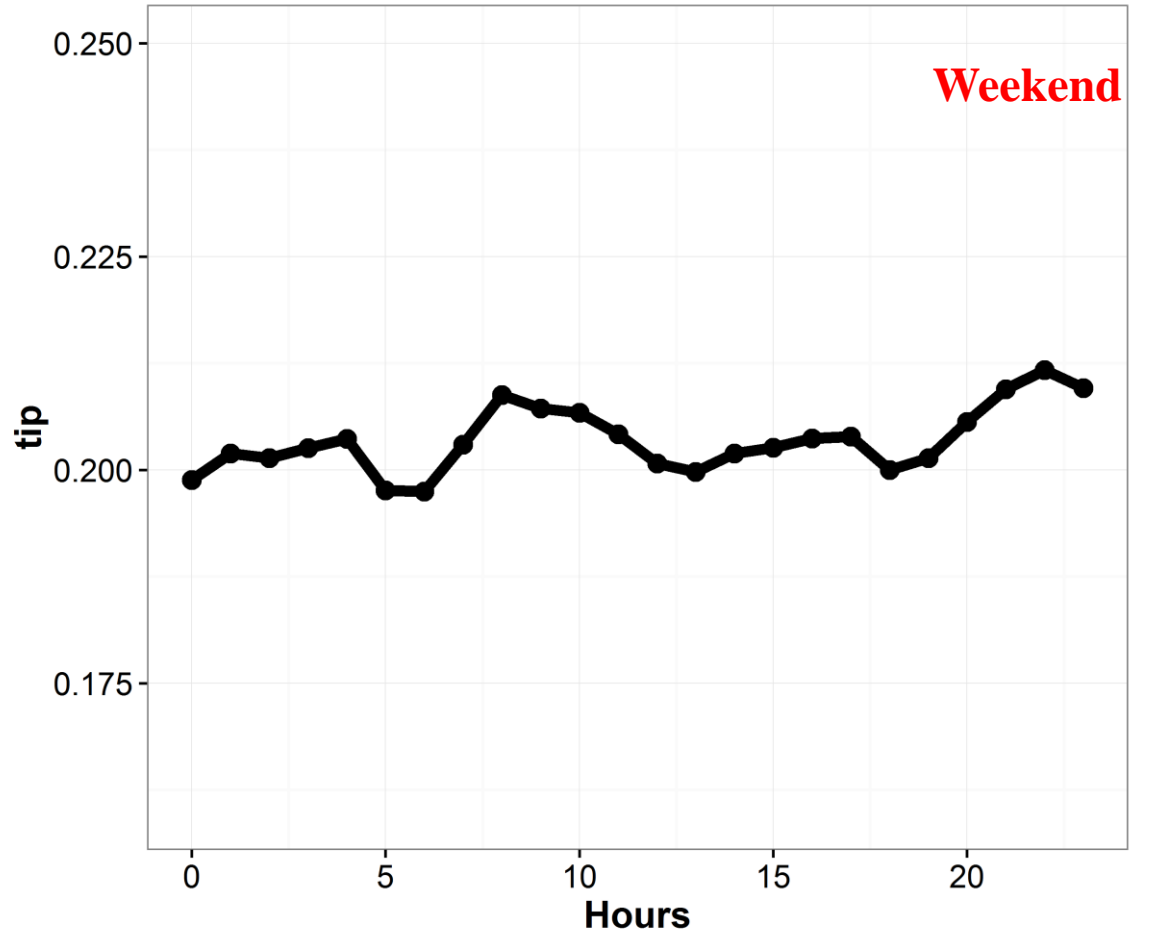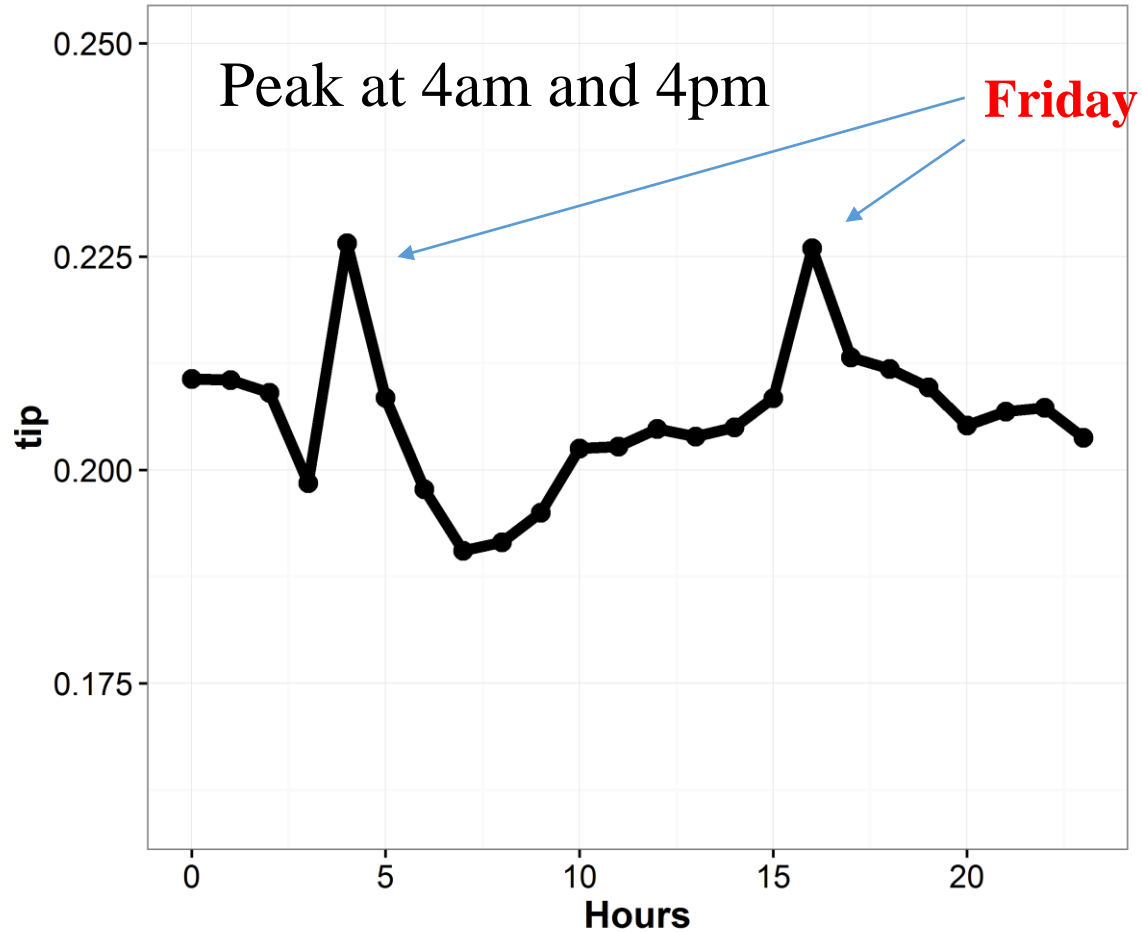- More people take taxi from airport

Passager pickup and dropoff at LaGuardia Airport

- Many people (36%) don't give tip or missing data there?

  Marjory people used this formula: Tip percentage =Tip/Total

Tip percentage=tip/Fare

Tip percentage=tip/Total

17%
20%
23%

Min. 1st Qu. Median Mean 3rd Qu. Max. 0.000125 0.133300 0.166700 0.157700 0.166700 0.961000
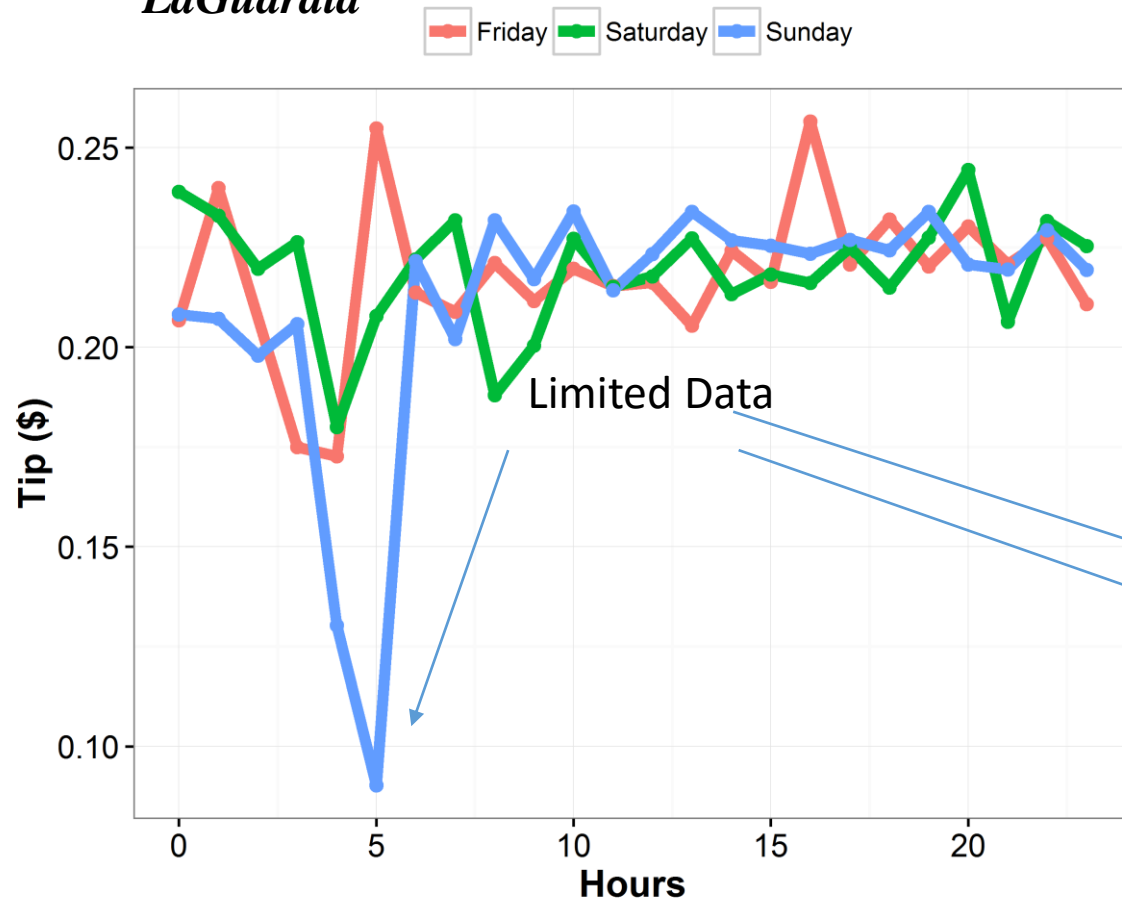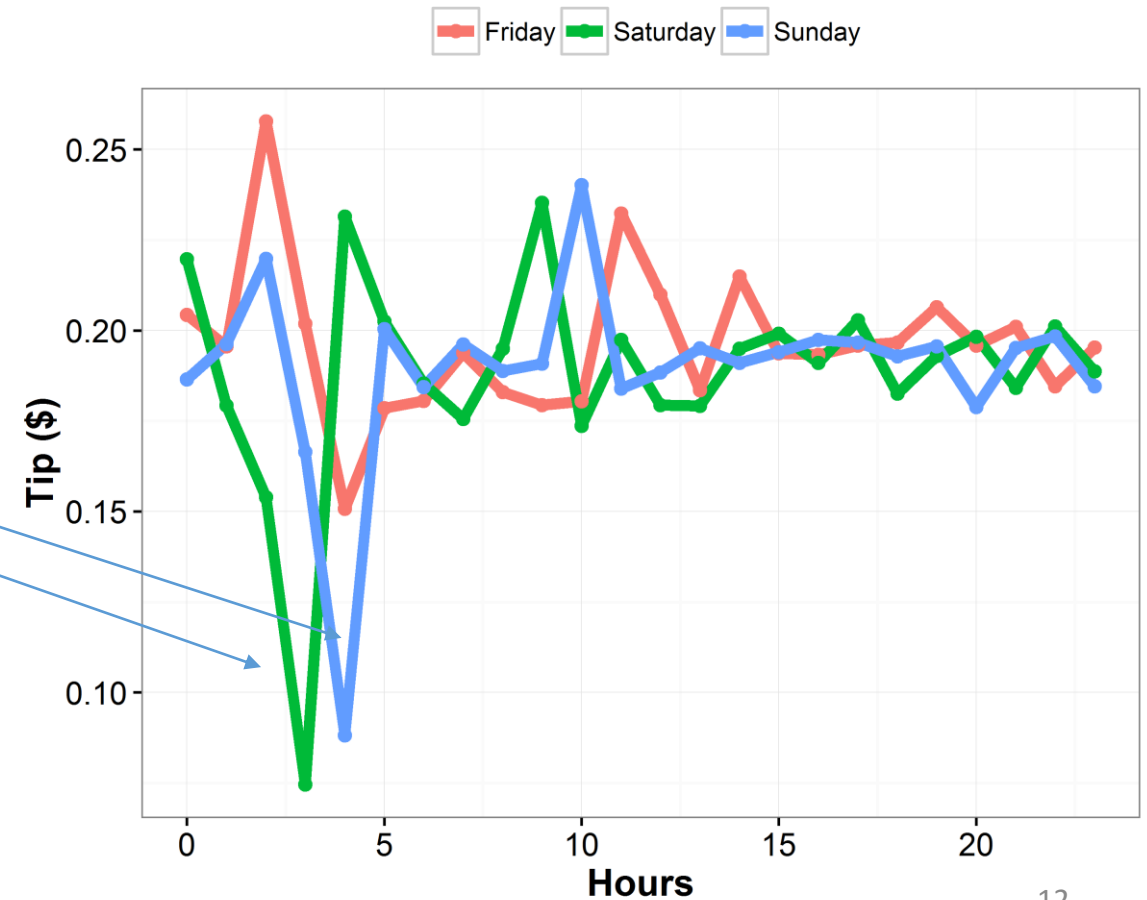
Tips variation with hours

Airport trip tip

❑ Weekday is better than weekend

LaGuardia

Friday Saturday Sunday
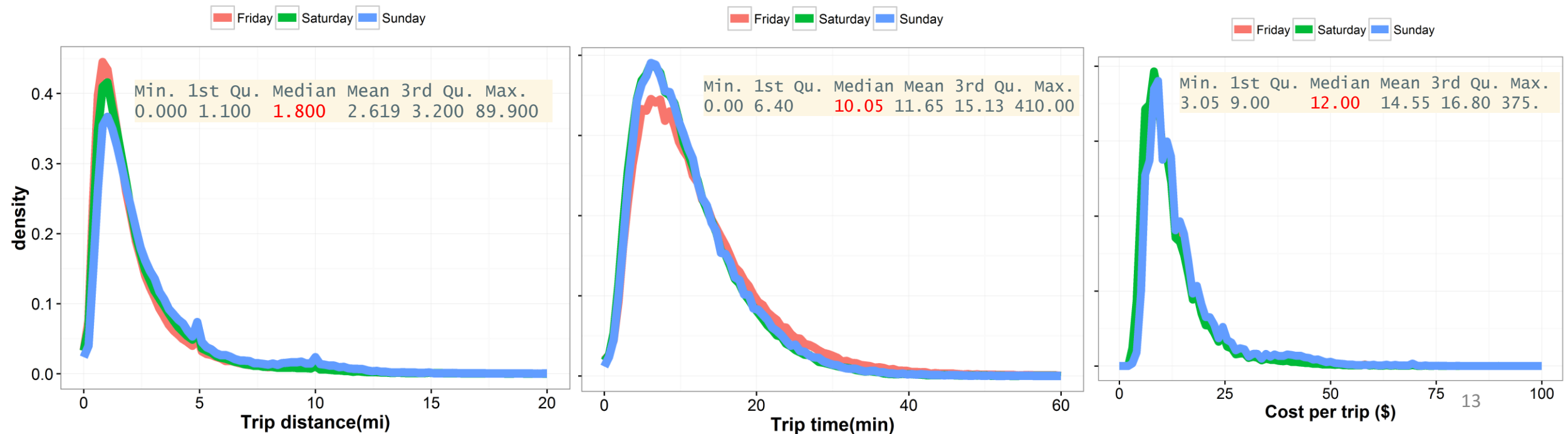
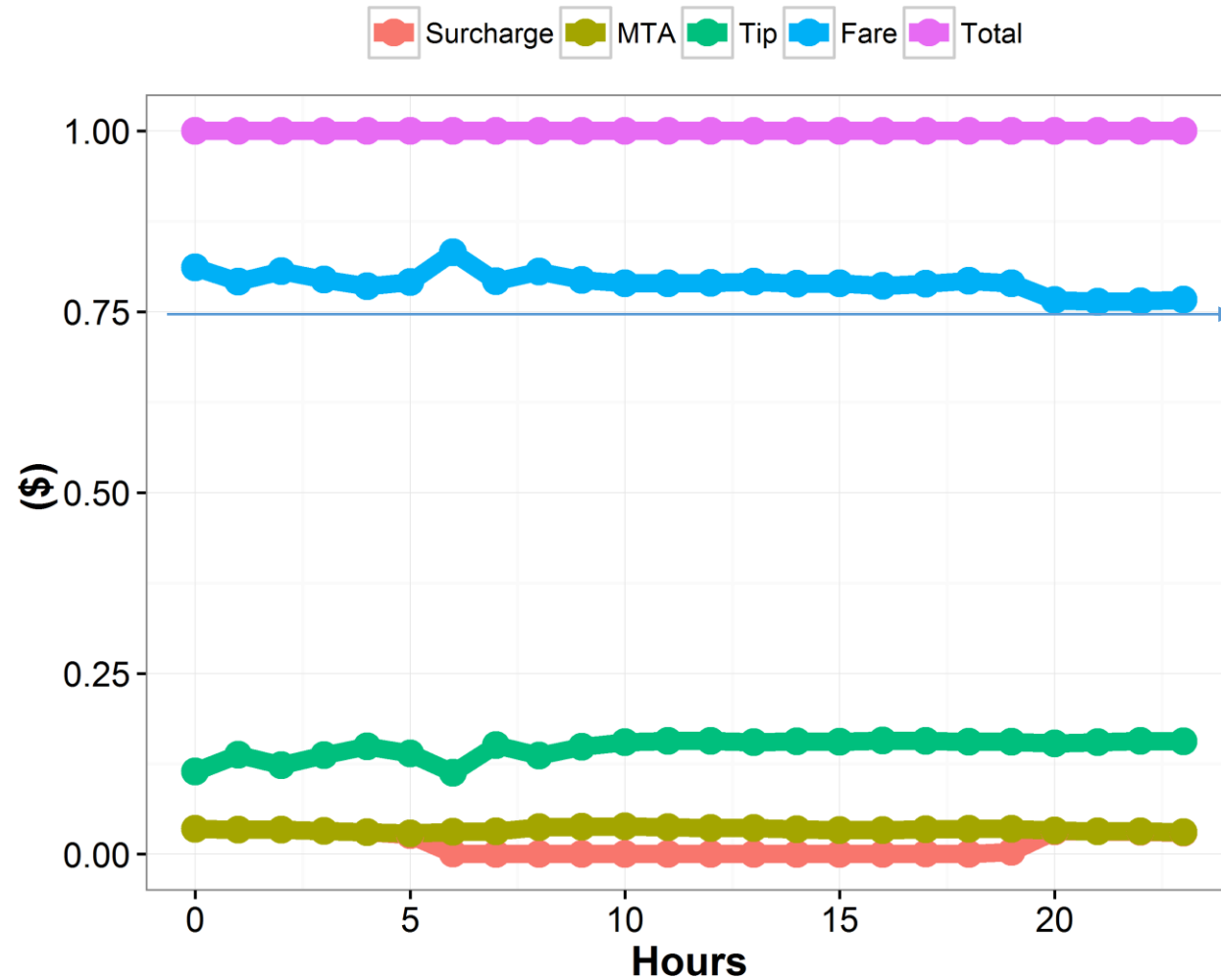Limited Data

JFK

Friday Saturday Sunday

Trips: time, distance and cost

❏ There median of trip distance, time and cost are: 1.8mile, 10 minutes and $12

Cost variation with hours

Passenger pay more
at evening time

Summary

- More people take taxi to go home at evening time.
- There is minimum activity at 5am
- It is difficult to find a taxi at 4pm, especially weekday
- People pay more tip at 4am and 4pm