Visualizing Uber Ride Data

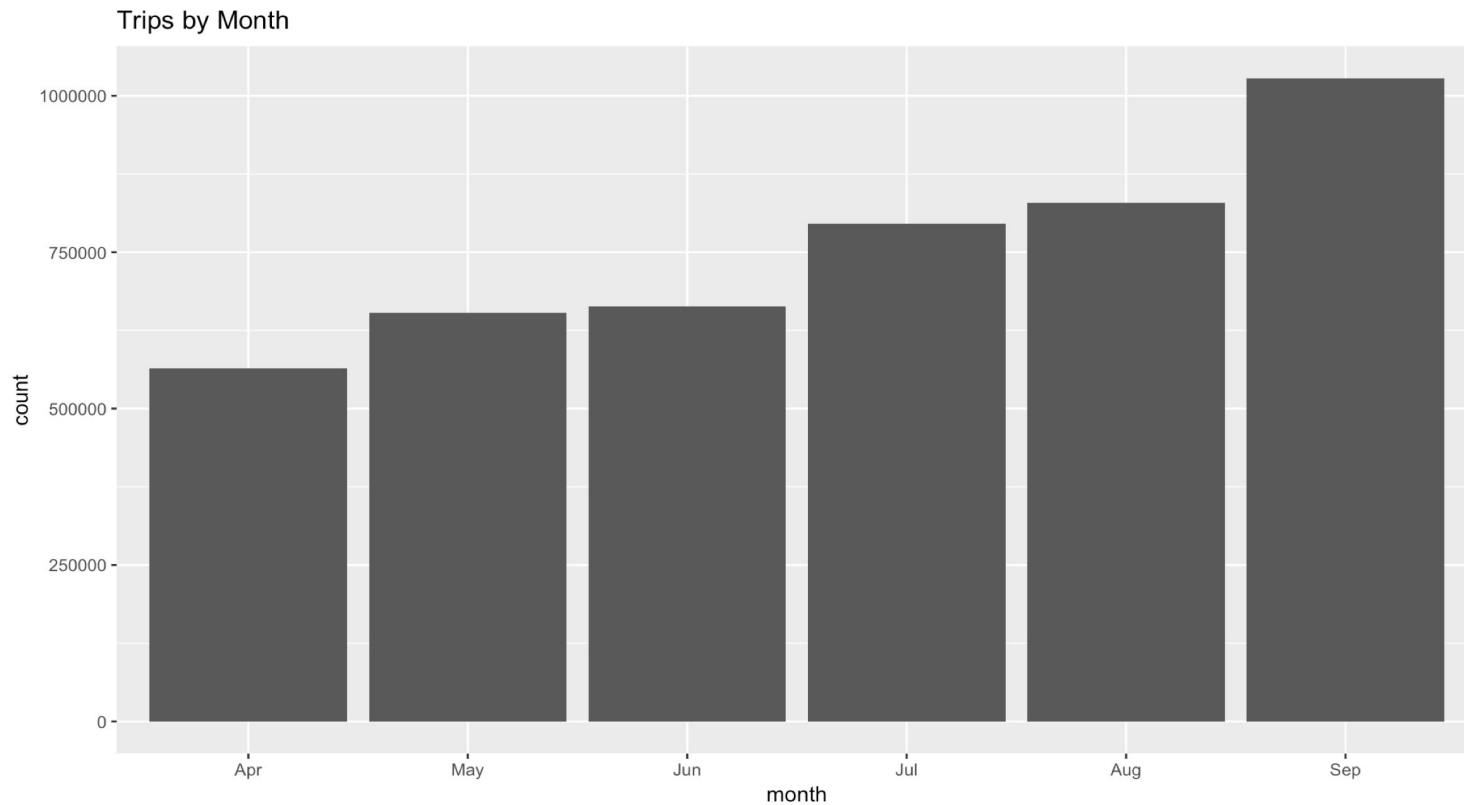BST 270 Final Presentation
10-17-17
Chris Hilger

Dataset Overview

- 4MM + Uber rides in NYC during 2014
- Date, Time, Lat/Long of pickup location
- Stored in multiple csv files
- Obtained via FOIL
- Accessed via Kaggle
  https://www.kaggle.com/fivethirtyeight/uber-pickups-in-new-york-city?sortBy=null&group=datasets
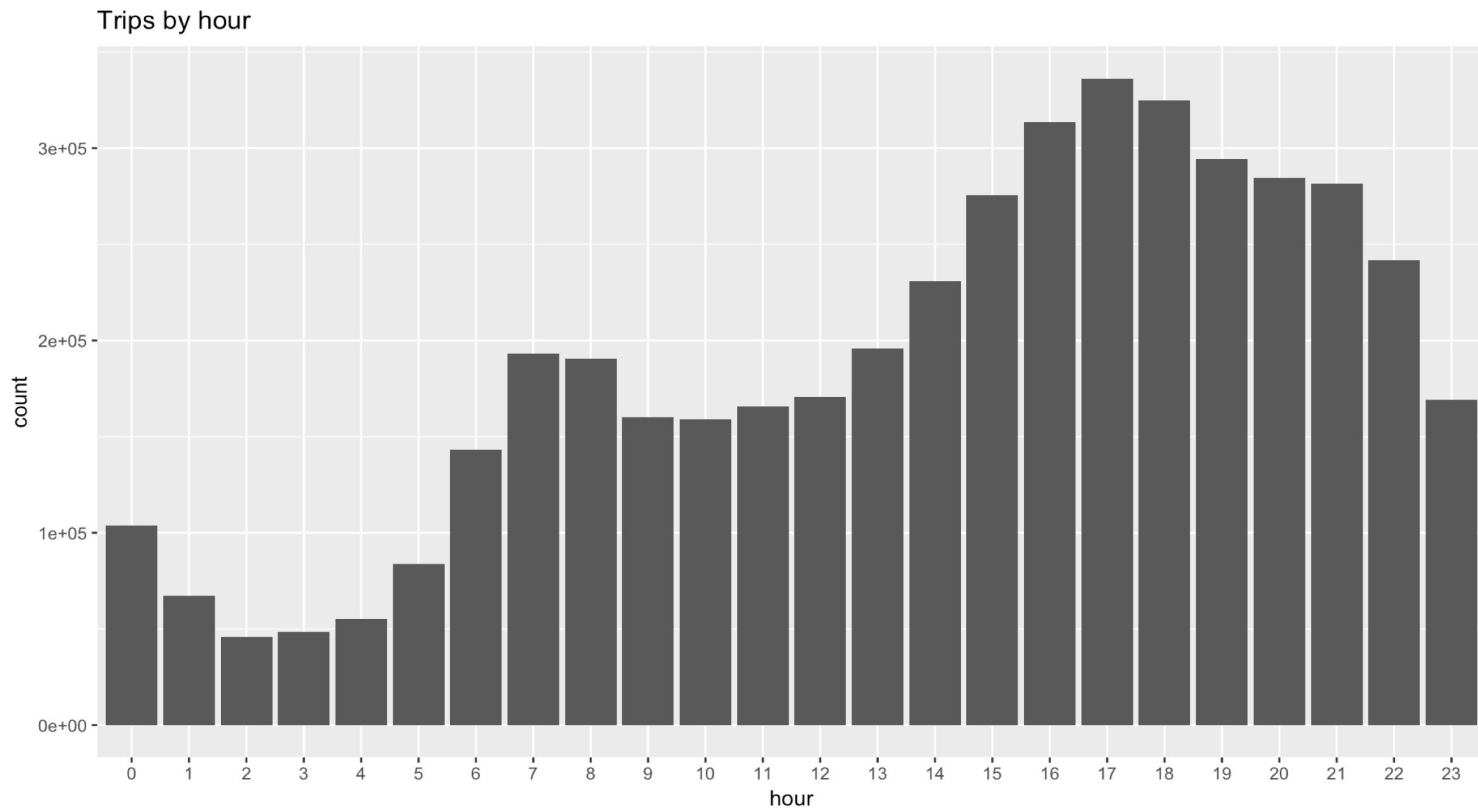
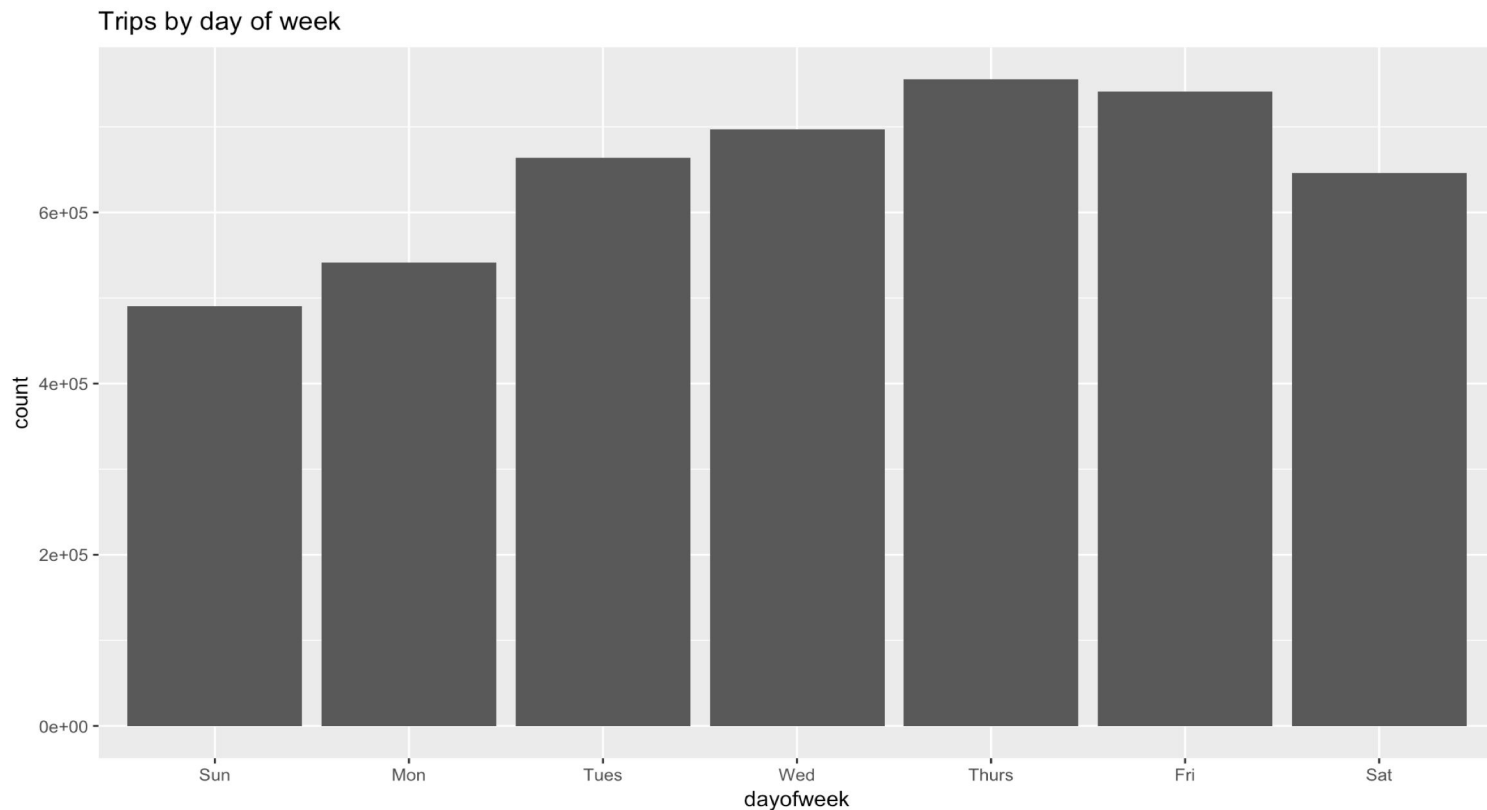# Uber growing rapidly in 2014

Keys to Reproducibility

- Accessing correct files
- Sampling data
- GGPlot installation
- Proper date & time parsing/manipulation
- Commenting

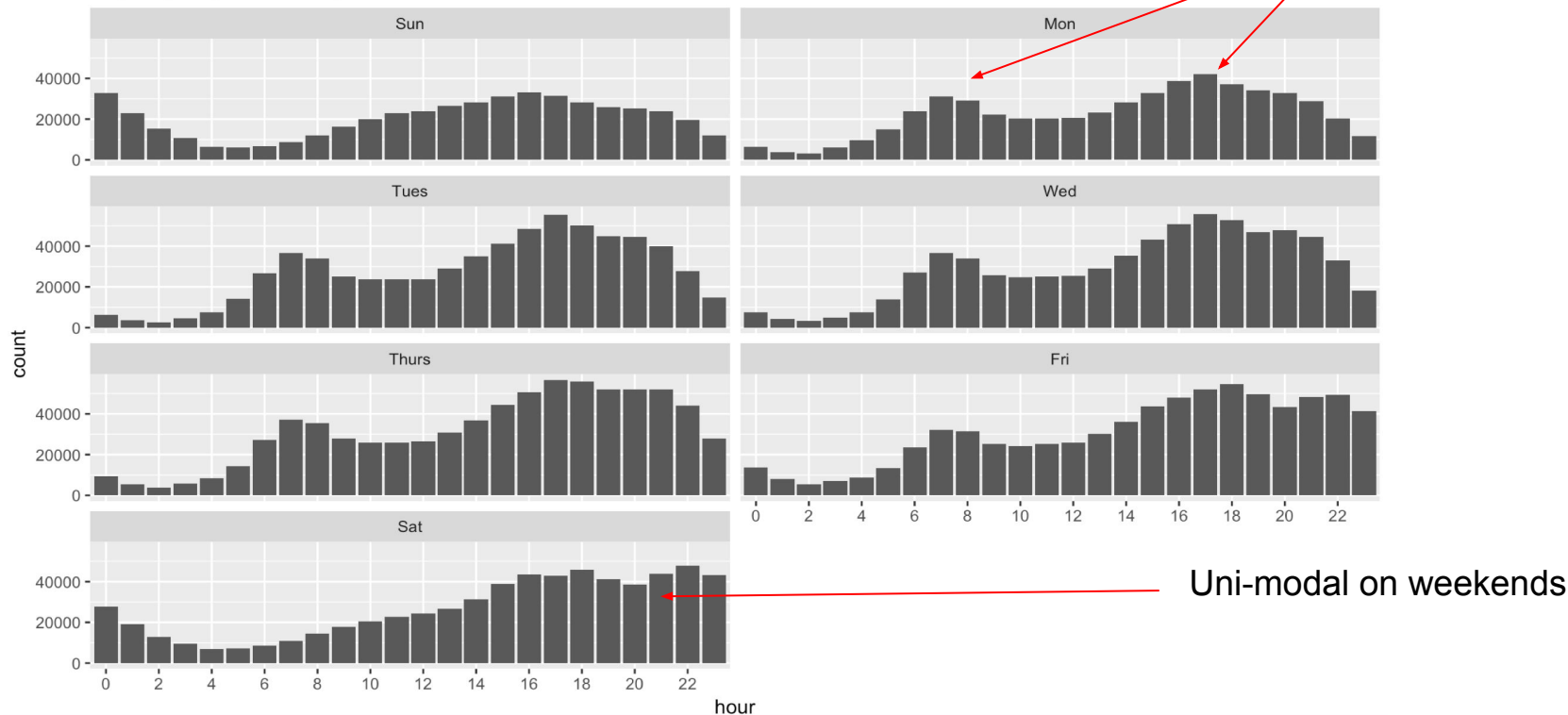# Significant variability in demand over time of day

Trips by hour

Some variability in demand over day of week

Trips by day of week

Digging deeper...
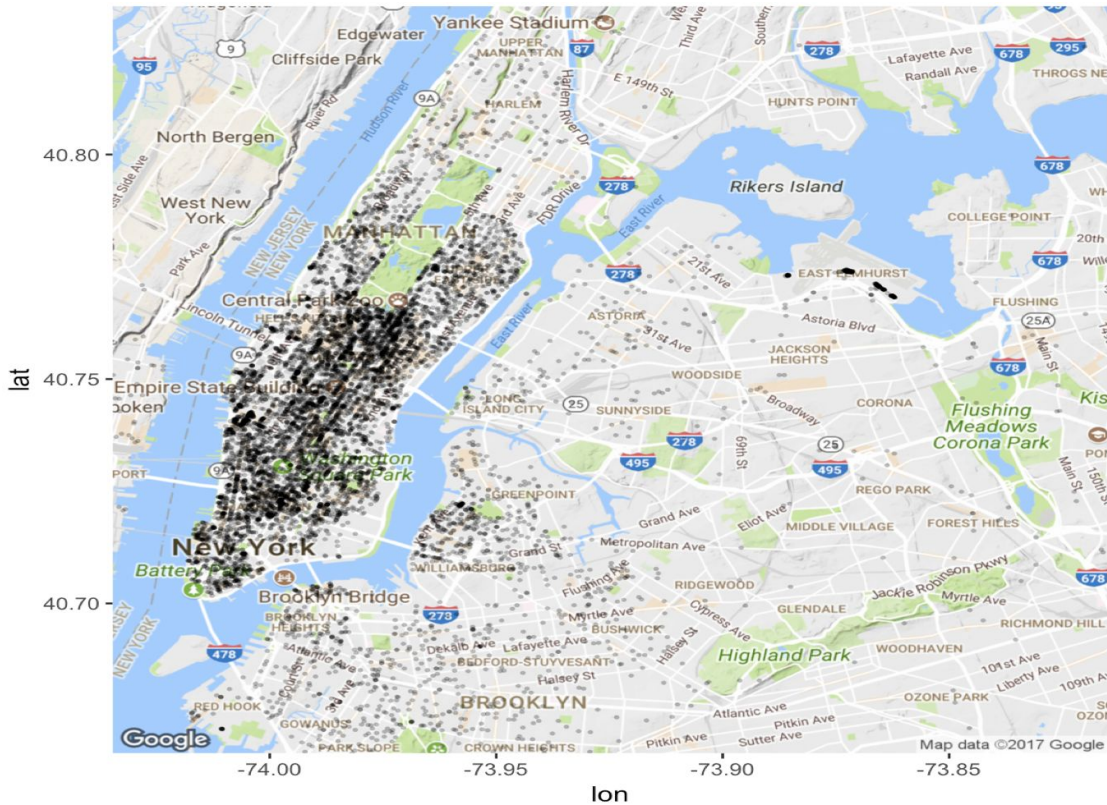

Trips by day of week by hour

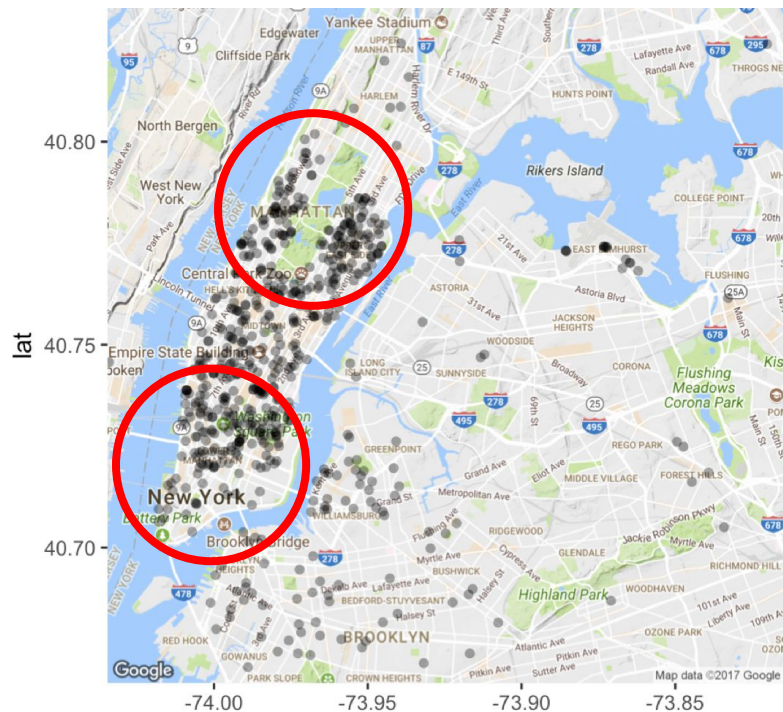# Do pickup locations vary by time/day of week?

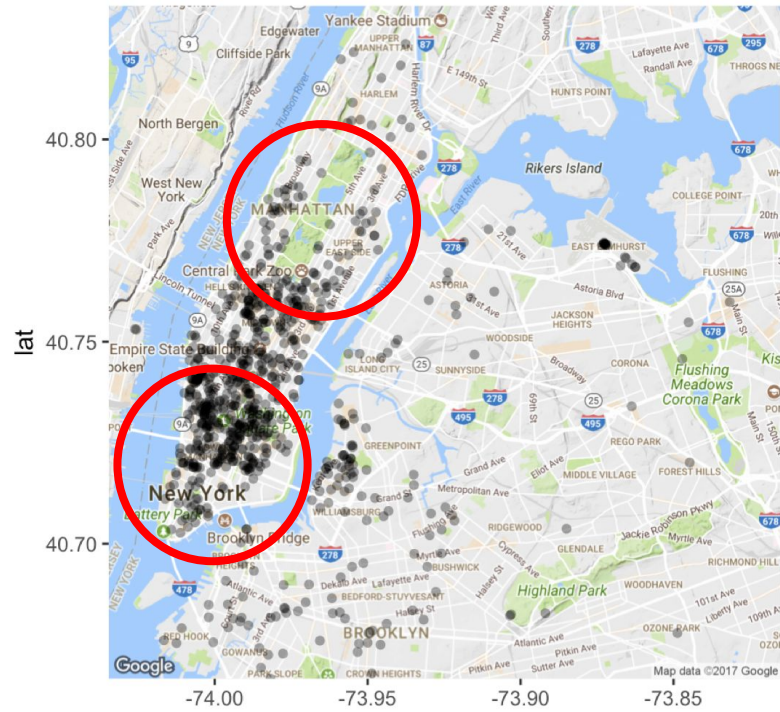# Visualizing Pickup Locations

Random sample of 10K
points from 4MM

Uber Pickups Shift from Upper to Lower Manhattan

Monday - Wednesday Morning
6:00 - 8:00 AM

Thursday - Saturday Evening
9:00 - 11:00 PM

Questions???

Special Thanks:

Kaggle
FiveThirtyEight
Abhijith Asok
Heather Mattie