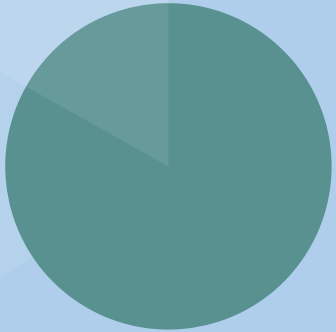
Predicting Divvy Stocking Requirements

Daniel Geihsler



The Problem

- Bikes not available at preferred starting points
- No available docks at destinations

Properly stocked stations are essential to customer retention

The Data

Trip Information

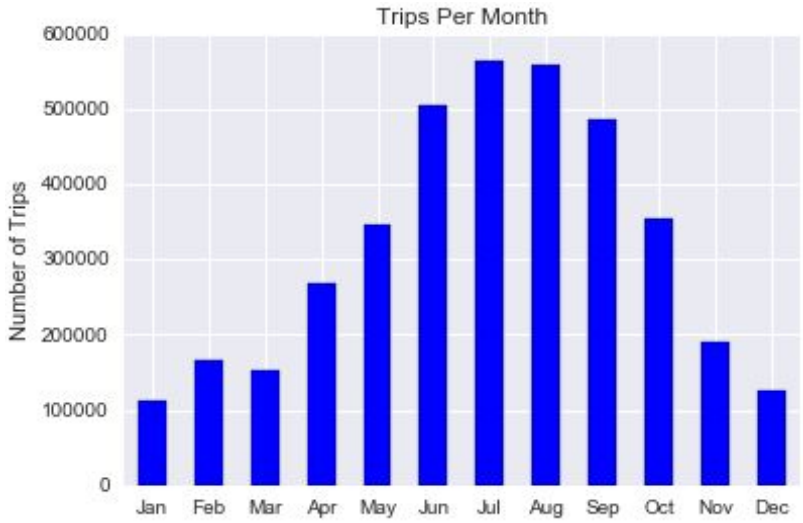
- All trips taken in 2017
- Unique Trip ID
- Start time and station, end time and station, type of user

Station Information

- Station data as of Q4 2017
- Station Location
- Station Capacity

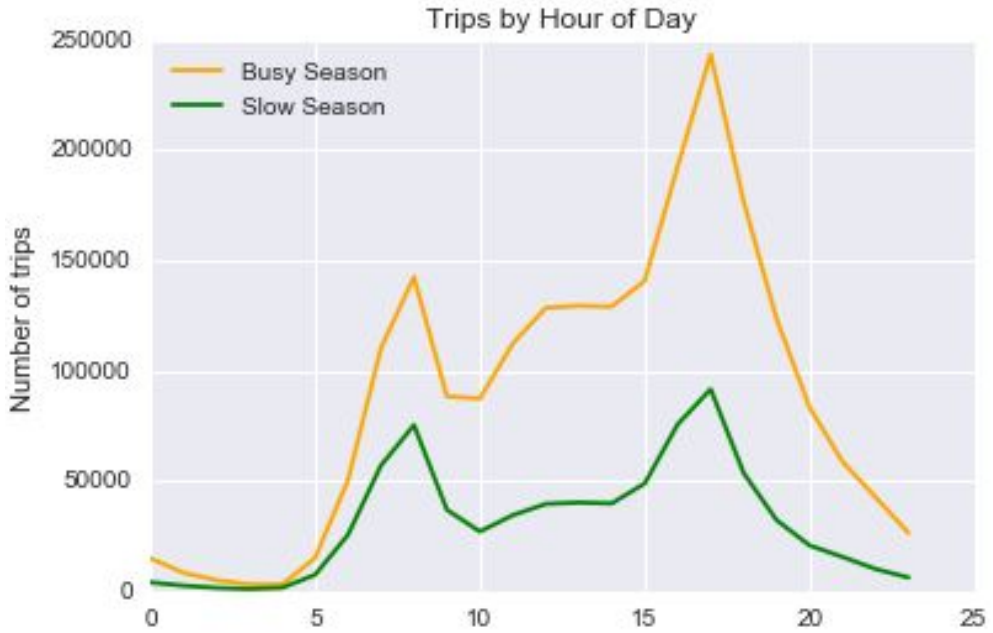
Rider Usage

- Busy Season:
  - June - September
- Slow Season:
  - January - March & November - December



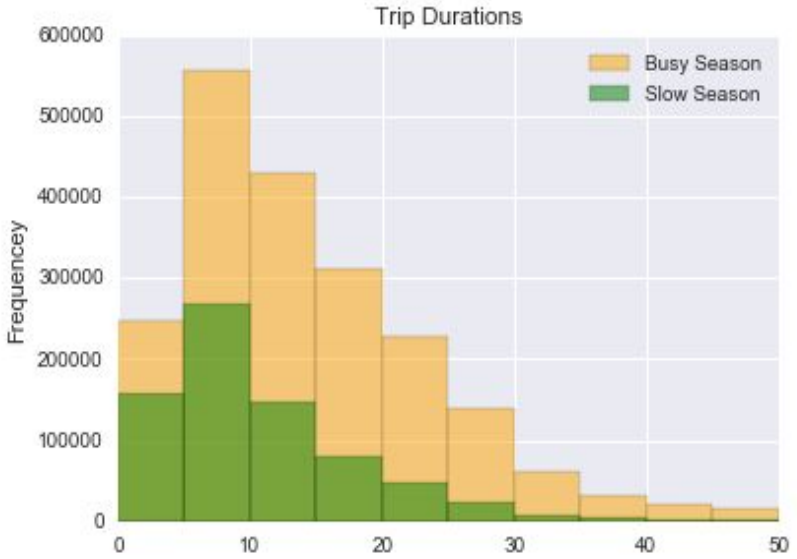
Rider Usage - Number of Trips

- More trips on weekdays in slow season
- More trips on weekdays in busy season
- Peaks in number of trips come during commuting hours (8am and 5pm)



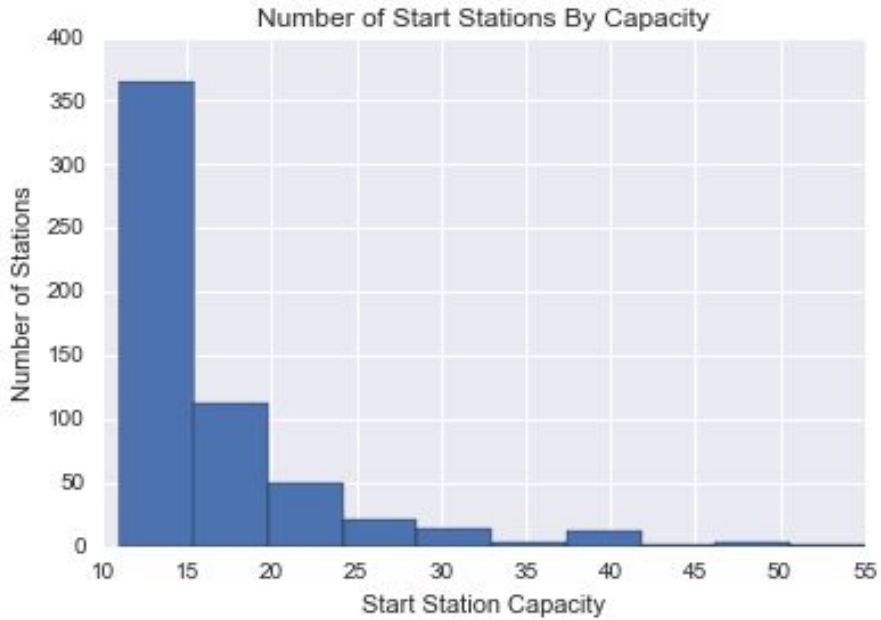
Rider Usage - Average Trip Duration

- Average trip is 15.92 minutes
- Average trips are longer in the busy season vs slow season
- Distribution of trip durations follow similar shape in busy and slow seasons
- Average trips are longer on weekend days throughout the year



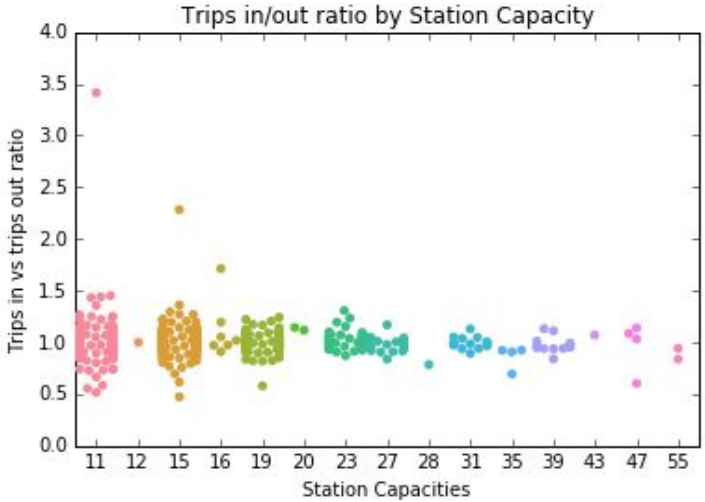
Stations

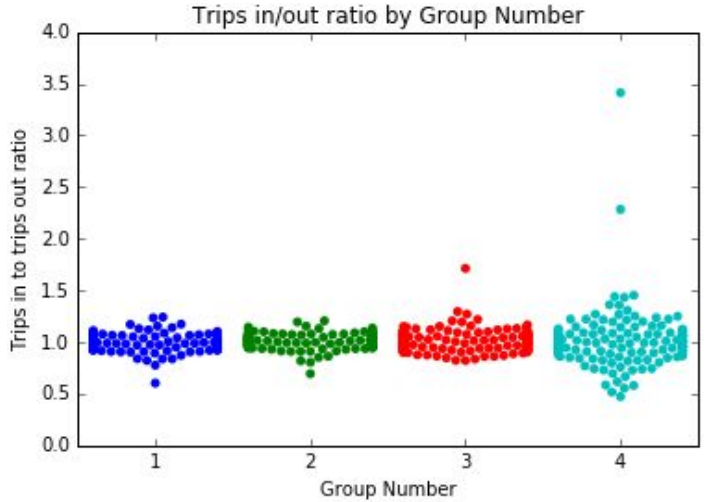
- 585 Stations
- 62% of stations have capacities between 10 and 15
- Average station capacity is larger for higher trafficked stations
- Average Station Capacities:
  - Group 1: 27.58
  - Group2: 17.69
  - Group 3: 15.28
  - Group 4: 13.64



Stations - Trips in / Trips out

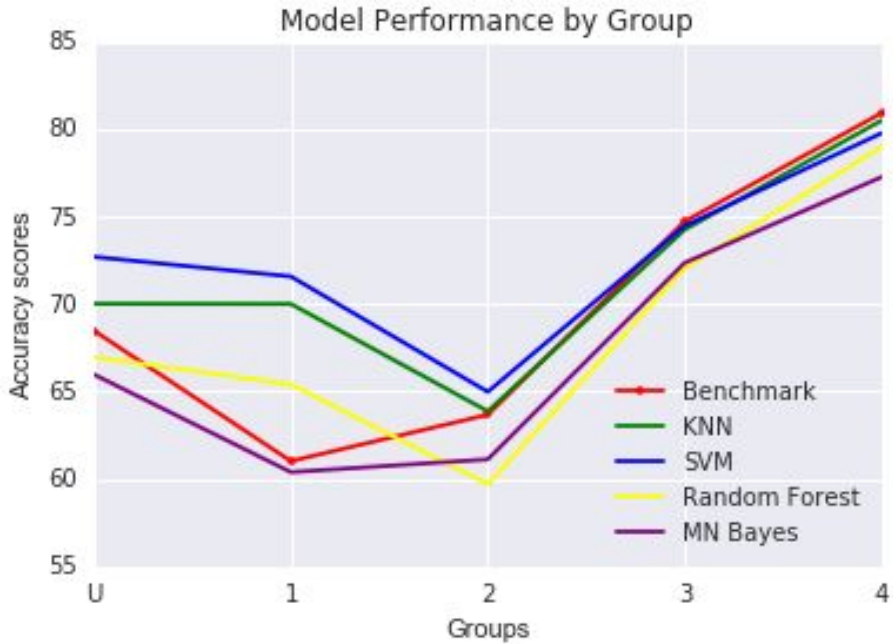
- Similar average ratio between groups
- Lowest trafficked group has highest variance
- Lowest capacity stations have highest variance





Modeling

- Average ratio for each station for 2017 was used as the benchmark
- Models trained to predict stocking class
  - Range of ratios with similar stocking suggestions
- SVM and KNN models outperform benchmarks in groups 1 and 2, and align with benchmarks in groups 3 and 4
  - Groups based on station capacity (1 being the lowest)



Recommendations

- Stocking Classes and Recommendations:

Ratio's and Recommendations:	
0.0 – 0.49	30% or more bikes docked than open slots
0.5 – 0.79	15% more bikes docked than open slots
0.8 – 1.19	Equal number of bikes docked and open slots
1.2 – 2.49	15% more open slots than bikes docked
2.5 +	30% or more open slots than bikes docked

- SVM model should be used to predict stocking class for all station groups
- Improvements:
  - Weather Data
  - Multiple Daily Stocking Times