# Predicting Traffic Accident Likelihoods in Seattle

Ming Huang

# Why?

- Approximately 10,000 collision related traffic accidents a year according to the 2013 Traffic Report
- Being able to estimate traffic accident likelihoods will allow for better resource allocation of 911 response teams
- Allows for better awareness of risky situations

# Goals of this Project

- Use publically available and relevant data
- Identify predictive features and models
- Create an automatable data pipeline
- Build a live app reporting likelihoods

# Overall Project Summary

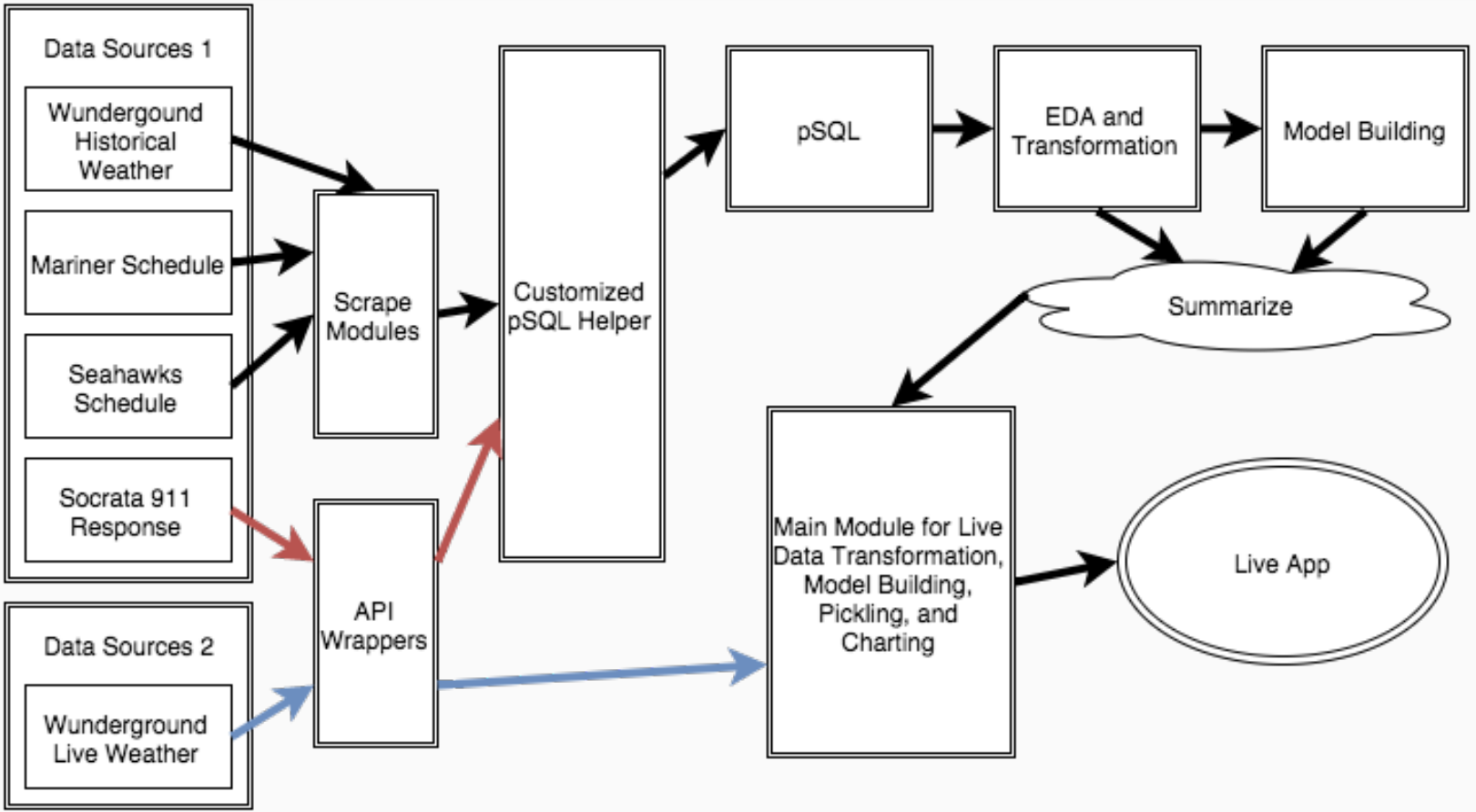
**Workflow:**

- Build scrapers, API wrappers, and data processing modules to collect data.
- Feature engineer through the use of python and sql.
- Evaluate multiple models select the most suitable.
- Build data pipeline models capable of automation and flexible model selection.
- Build a live app to report current conditions using the above model.

**Data Sources:**

- Socrata 911 Incident Response records
- Wunderground historical weather data
- Wunderground real time weather API
- Mariners game schedule
- Seahawks game schedule

# Visualized Project Summary

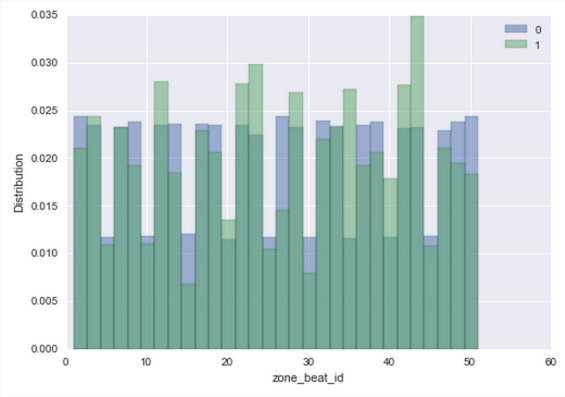


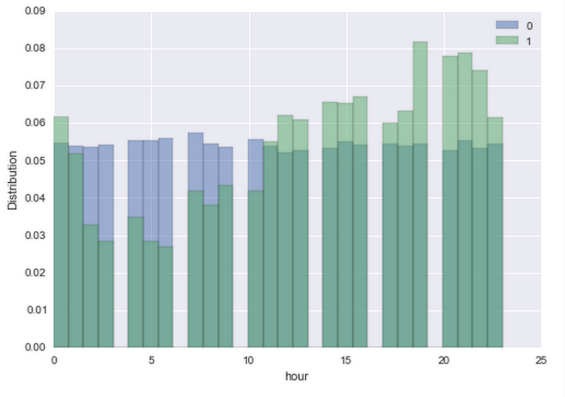
# Current Findings

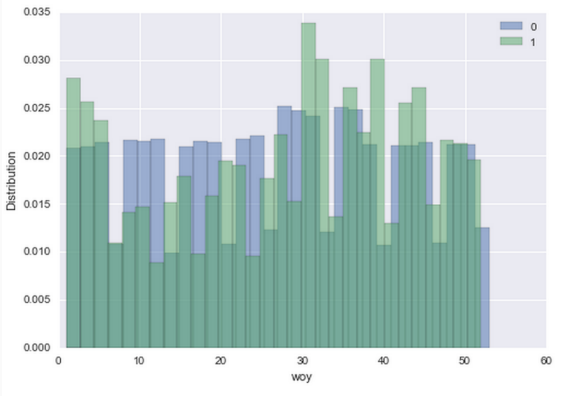
- Lots of unaccountable randomness
- Imbalanced classifiers
- Random Forest performed the best overall

# Current Findings continued

- There are noticeable patterns
- Still plagued with randomness

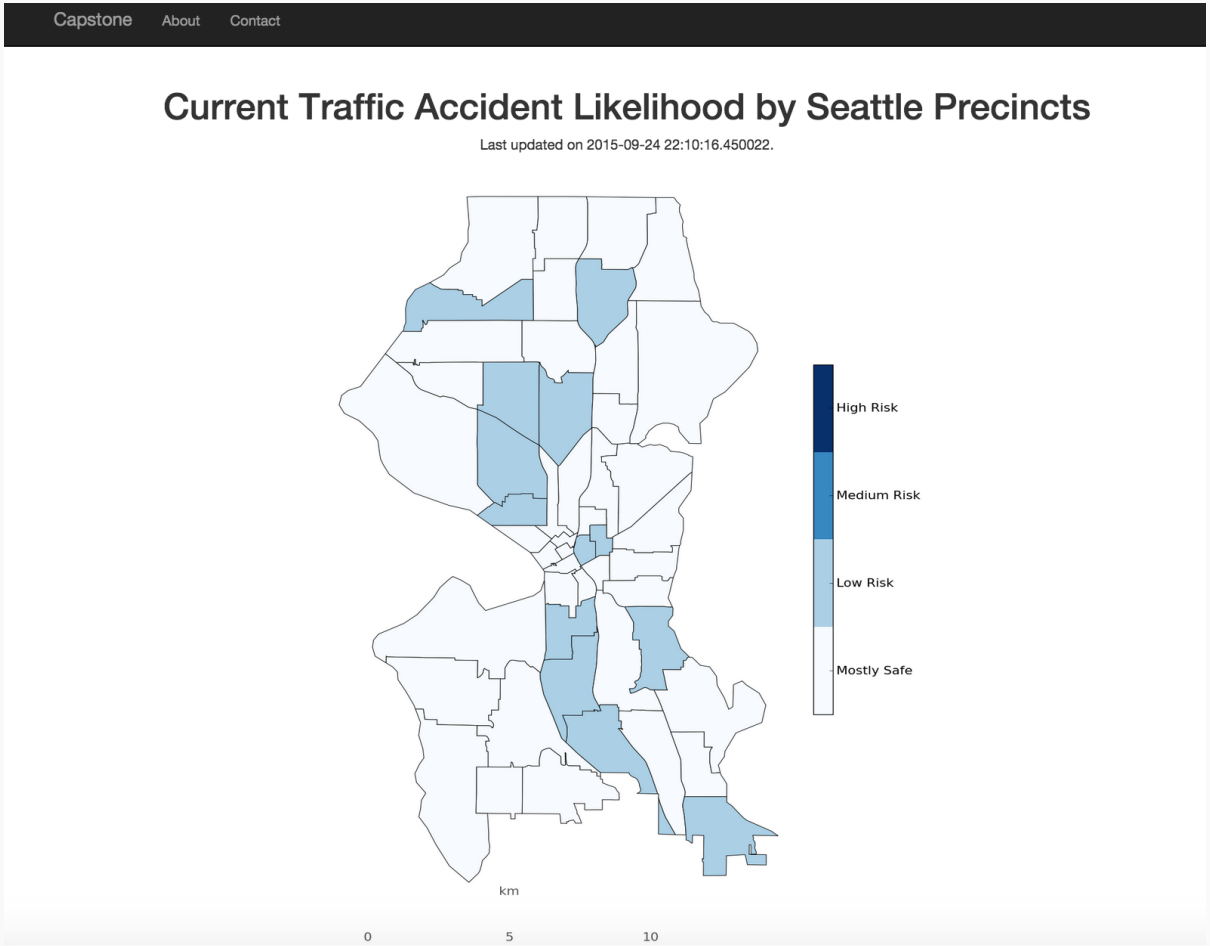


# My App

- Choropleth of Seattle precincts shaded by the rate of accident risks.
- Updates automatically every 60 minutes by using the most recent Wunderground weather report



# Next Steps

- Add and improve on existing features
- Find ways to account for randomness
- Implement non-conventional models
- Improve App functionality and design

# Thanks!

Ming Huang:

mnghuang@gmail.com
https://github.com/mnghuang