Traffic Relations

Nicholas Sugarman, David Nyberg, Eric Ha

Project Description

We intend to investigate traffic data to find if there is a correlation between seasons and the traffic on different types of roads.

For example we may find that rural roads increase in traffic over winter months.

Prior Work

- **American Transportation Research Institute (ATRI)**: transportation research
  - http://atri-online.org/
- **Google Maps**: rerouting method to avoid traffic, location services pull in data
- **ArcMap:** allow users to build maps, can add traffic control extension, they use historical traffic data which transfers to analysis of predictive traffic

Our Dataset

- Dataset: https://www.kaggle.com/jboysen/us-traffic-2015/feed
  - compiled by US Department of Transportation (over 7 million data points of daily traffic data including locations, road types, time of day, traffic volumes, etc.)

- Dataset has been downloaded to every member's machines

Proposed Work

- **Data Cleaning:** clean data to get rid of any Null values or incorrect values in the dataset.
- Reduce our dataset to only interesting variables.
  - Clean out outliers (for events that cause high traffic)
- **Data Reduction:**
  - Attribute Subset Selection: remove irrelevant attributes: traffic flow direction, station_id
- **Data Transformation:**
  - Normalization of traffic volume for areas so they are on even scale
- **Data Visualization:**
  - Display a multi-histogram to show different seasons on a select road and their traffic volumes and compare

Tool list

- Python as our primary programming language to perform data mining
  - Includes libraries: pandas, matplot, numpy, etc
- Github for project tracking
- Visualization tool (Python, excel, or another external source)

Evaluation

Based on our results we can determine if there is a correlation between traffic on different types of roads based on the month.

Using this result we could potentially plan construction so that it occurs during low stress periods.