NYC Car Accidents

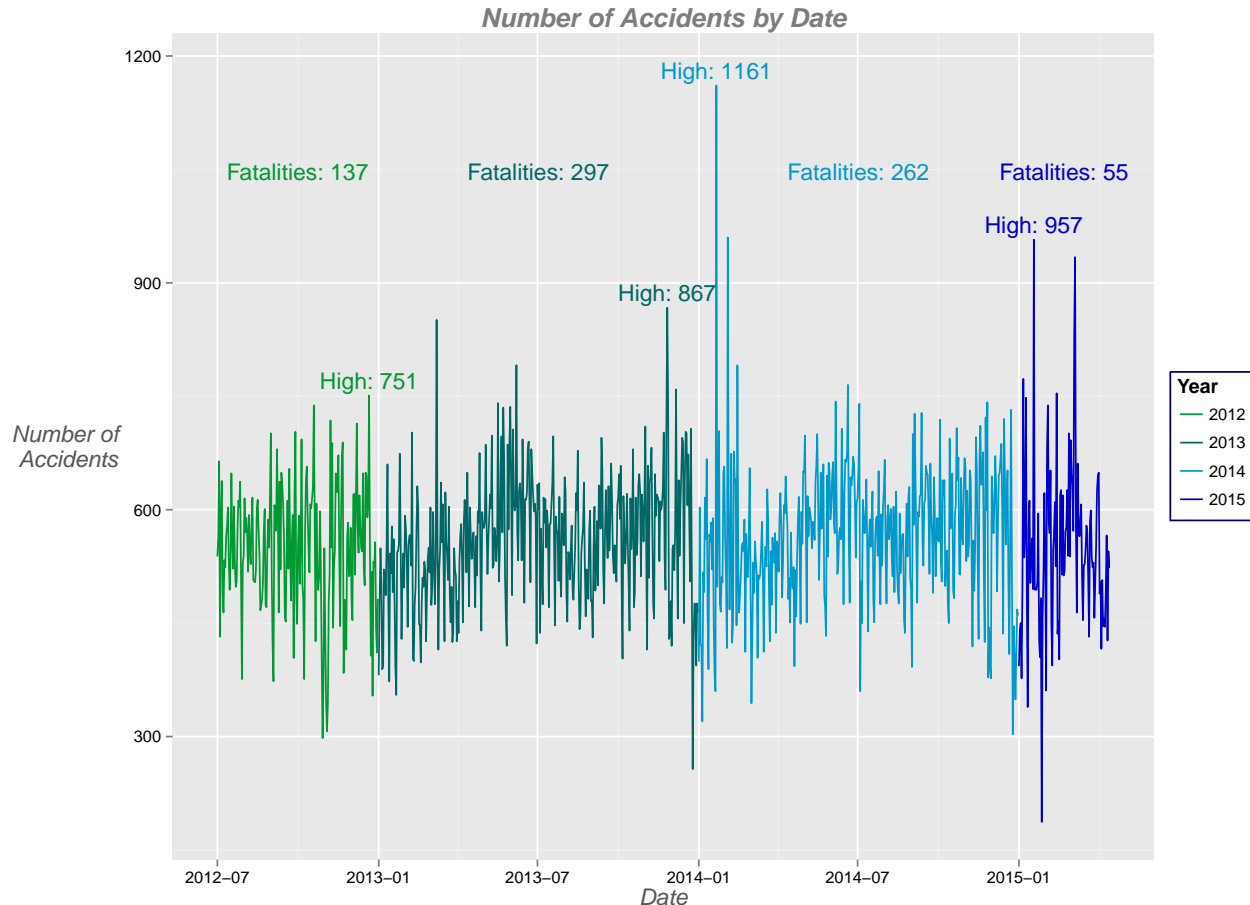*Paul Jimenez*

Intro

I decided to explore the New York City Motor Vehicle Collisions data set using the r packages `dplyr`, `ggplot2`, and `ggmap`. Below is my report with the discoveries I made. **IMPORTANT NOTE:** There were 89,878 missing latitude longitude pairs at the time I analyzed this data set. My analysis with `ggmap` therefore only uses what was availabe to me at the time.
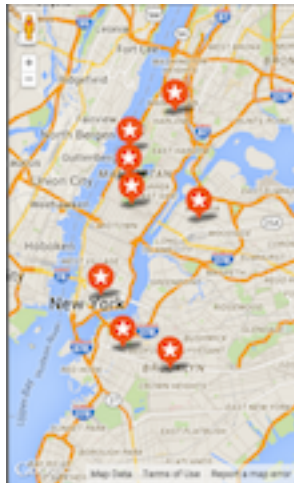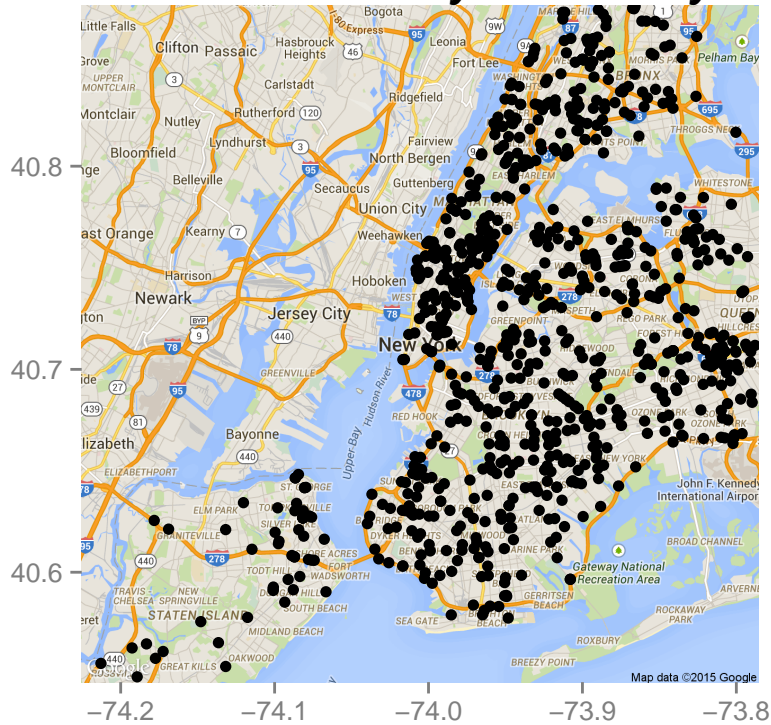
Examining accidents by date

I aggregated all accidents that occurred by date from July 1, 2012 - April 14, 2015. For each unique year, I also included the number of fatalities that occurred and the highest amount of accidents that occurred on one day. Below is the graph of my results

After I saw this graph, I noticed that the two highest peaks that occur for 2014 and 2015 happen around the same date. In fact, the peak for 2014 occurs on **January 21st** and the peak for 2015 occurs on **January 18th**. I looked into this and discovered that *Dr. Martin Luther King Jr. Day* was on **January 20th** for 2014 and **January 19th** for 2015. This may imply that there is an influx of people coming into New York City for the weekend or more people are driving around to go events. In any case, the data from the past two years indicate that this weekend in particular is a time to be very carefull driving.

The image below shows where the car accidents occurred on January 21st 2014. The points are very much evenly distributed across the city aside from some clusters of points located in the Downtown Manhattan area. Below that is an image of events planned for MLK day that I found from a website. The events were from 2013 but I am assuming they would be in a simiar location for 2014. **Note:** 202 of the 1161 accidents are not on the map due to not having coordinates to plot.
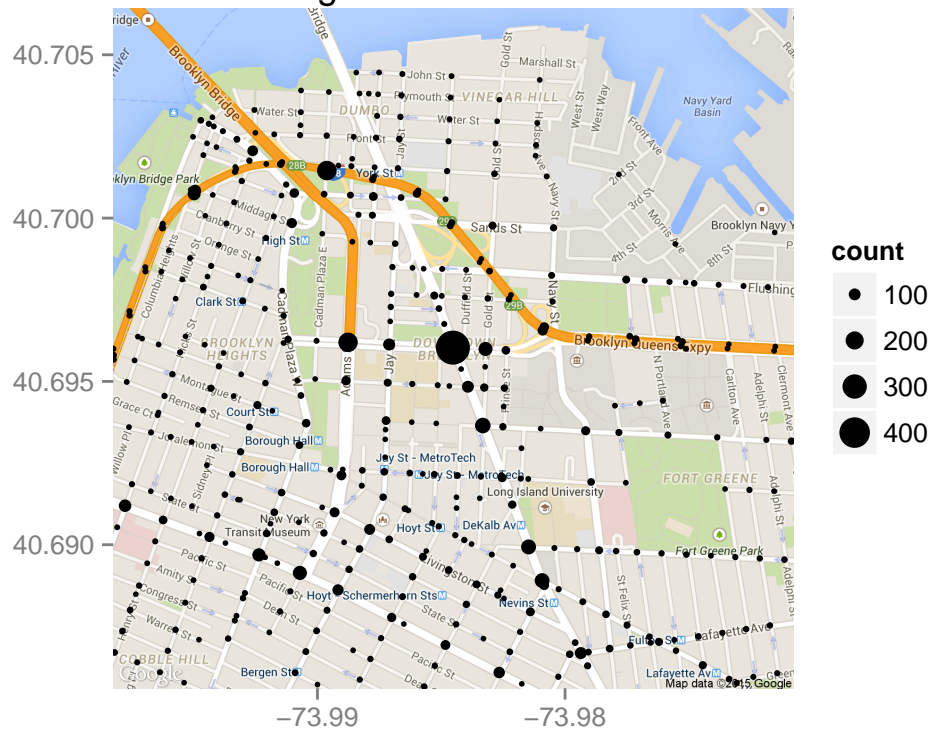
Accidents in New York City on January 21st 2014

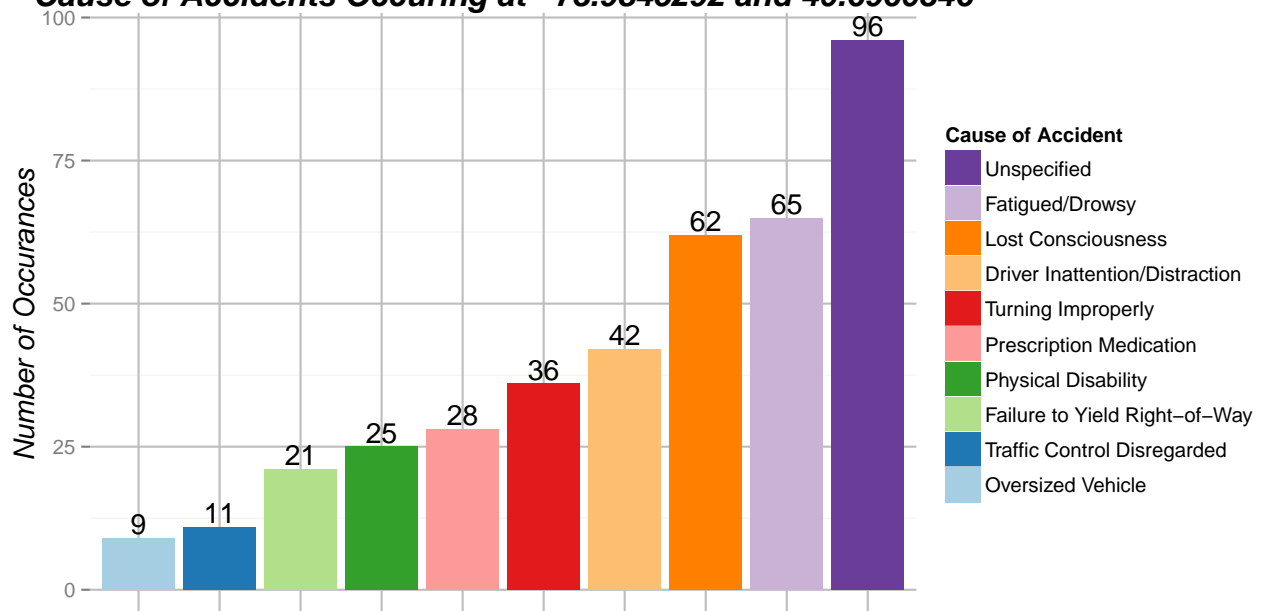### Exploring Coordinate Pairs with Highest Amounts of Accidents

I decided to get a closer look at where the highest amount of accidents were occurring.

Accidents in New York at longitude −73.9845292 and latitude 40.6960346

The highest amount of accidents occurs on an intersection after coming off a highway. Notice also that many accidents occurred right when the highway empties out onto a street. I was then interested in the main causes of the accidents that occurred at that intersection.

Cause of Accidents Occuring at −73.9845292 and 40.6960346
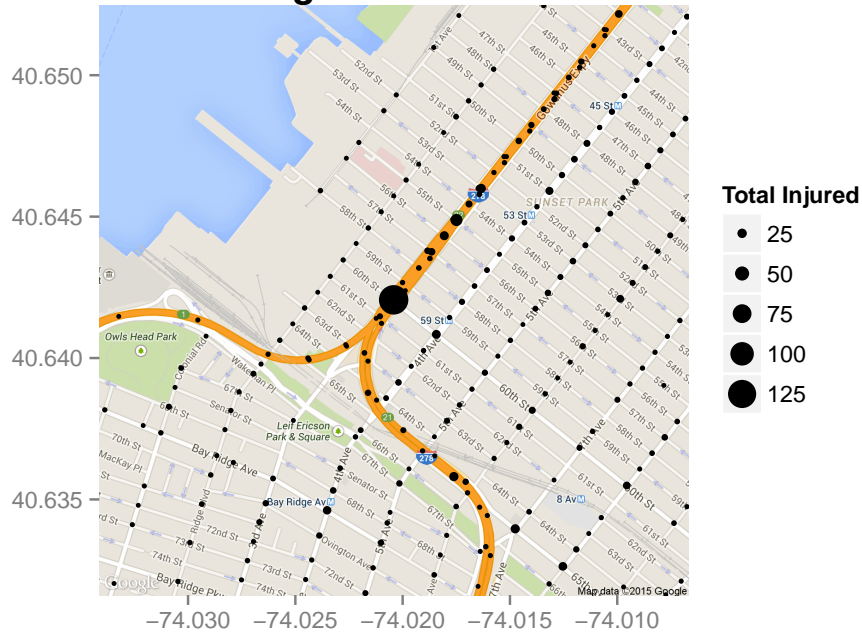

Nothing too revealing for the causes. Unspecified could be further explored to see if the actual reasons for these accidents happening are related to the structure of the intersection i.e. not clear traffic signs, confusing layout, etc. However, the high amount of accidents associated with Fatigued/Drowsy and Lost Consciousness **is concerning**.

Exploring Areas of High Injury

This was perhaps the most interesting observation I discovered. Here is the plot of the area where the second highest amount of injuries occurred in NYC

Injuries from Accidents at longitude –74.02042 and latitude 40.6420578

As you can see, these accidents occur right when two separate freeways merge. The high amount of injuries is most likely due to the speed at which these cars are going once they begin to merge.