Data Science For Good

Leverage Your Skills to Help Your City!

Source: https://www.thestar.com/news/gta/2017/08/19/arbitration-gives-toronto-firefighters-a-new-collective-agreement.html

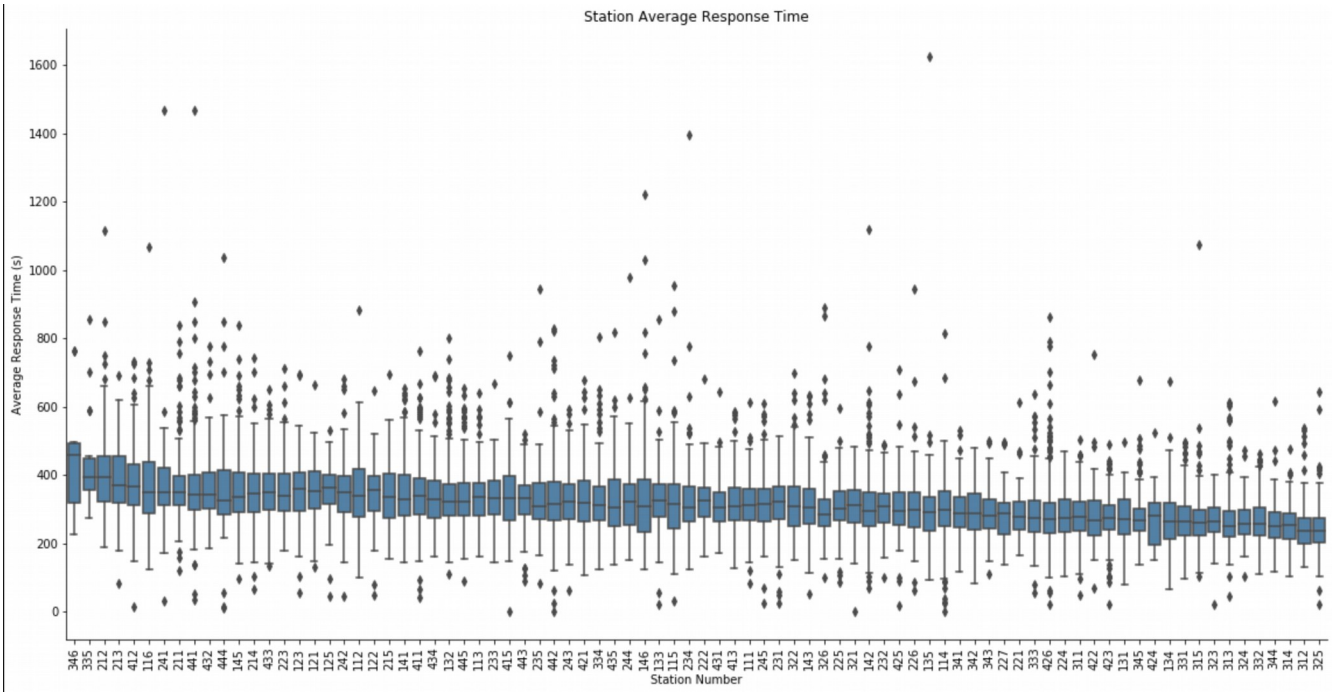
More and more major cities through-out North America are posting open data sets with a wide variety of information. As data scientists, we're always on the look out for ways to learn something new or practice skills – why not 'practice' and also add value your community? Who doesn't want their city to be more safe and efficient?

As part of a Udacity course, I was tasked to do analysis on a dataset to answer some questions and write a blog post. I always struggle with creativity in coming up with ideas to tackle on open ended projects, but I truly wanted to do something I could tie to my home city – Toronto, Ontario. I opted to analyze our city's fire service data for efficiencies and opportunities.

The main questions I had were;

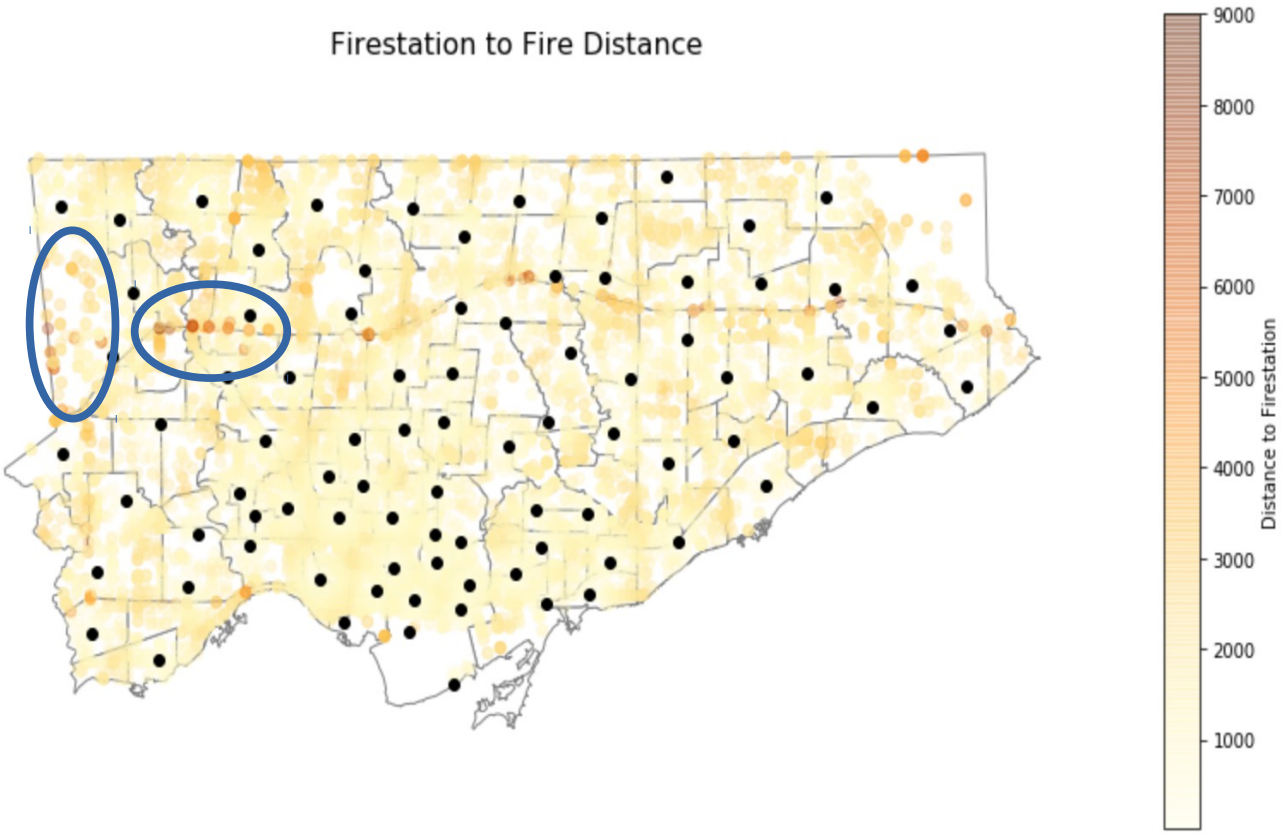
1. Is there a large variability in response time by the different stations?
2. Is the spatial distribution of fire stations appropriate?
3. What's the most important feature to predict response time to fires?

With respect to response time, all stations were generally consistent. There were some outliers that needed to be filtered out, but it's safe to say that across the city there are relatively consistent response time to fires (which is great news). The distributions for average response time for each station are shown below.



When it comes to the distribution of fire stations, visually it looks as though the density of fire stations aligns with the density of fires. The average euclidean distance to fire for all stations is less than 2km.

The above being said, there are some outliers when looking at all fires in the data set. These are likely instances of a different station other than the closest by proximity responding to the fire. Whats interesting is these instances seem to be clustered.



As can be seen from the above map, in the upper left area there are a number of fires that had notably long ‘response distances’ associated with them. Even more notable is that these fires seem to have fire stations in immediate proximity to them.

Further analysis from a model trained to predict response time given a number of other factors showed that Distance from Firestation has the highest impact on the response time of the fire. Given this (which may not be surprising), the area noted above should be focused on for fire response availability to ensure the closest station is always responding.

Now – what are you going to do to support your community?