Result Key

0: Less than an hour.

1: 1-2 hours

2: 2-3 hours

3: 3-5 hours

4: 5-24 hours (by the end of the day)

5: 24-168 hours (by the end of the week)

6: 168+ hours (more than a week)

The results from my program were not super accurate, but did seem to adhere to the trend represented. (Most complaints were handled in 5-24 hours.)

What I used:

Complaint Type, Descriptor, Community Board, Borough, Open Data Channel Type, Month.

I pulled the created the month column by pulling from the ‘created date’. I didn’t keep the created date because it was too specific to be helpful. Other columns were also overly specific (serial number, address name, latitude, longitude, etc.). And some had no variety at all (like the facility type, which was pretty much all ‘precinct’. The complaint types and descriptions seemed to have the most predictive power. Which makes sense, different issues require different solutions. I also dropped all of the rows that still had empty entries.