COMS E6998 Digitally Mediated Storytelling Reflection Document

Jayson Ng and Abdullah Al-Syed

Our process for working together was very simple. We would work remotely and have an in-person meeting once a week to go over the tasks we need to do accomplish by our next meeting. We used github to collaborate our coding efforts and addressed any issues we had as a team. From our experience, gathering data and using tabula to convert PDF tables into a readable format was greatly appreciated. However, what we would have done differently in this respect was to start data collection earlier so that we can obtain more data points to further improve our system. Another area that we found worked well was the data set that we used. Initially we tried using a much more general data set provided by New York City that include a huge array of aspects that made it somewhat difficult to pinpoint specific areas in which would want to use when making our report. Through NYP Comp-Stat we can easily develop a story since it contains all data pertaining to the week as well as historical data.

It would also have been better to use a relational database system to extract and store the data, in which SQL queries could have been used to quickly obtain the necessary information rather than relying on very specific regular expressions that had to be tailored to the specific format of the data.