Final Project Report

Mike Pendleton

West Texas A@M University

The data sources I have leveraged for the project are <https://data.lacity.org/Public-Safety/LAPD-Calls-for-Service-2019/r4ka-x5je> and <https://data.lacity.org/Public-Safety/LAPD-Calls-for-Service-2020/84iq-i2r6>. This way we can look to see if there is any variance from one year to the other.

What locations are the calls coming from in LA county? What exactly is call type text, and how is it used? Are dispatchers putting the code in manually as the data types are all text fields, including the date fields?

The first data that’s easy to visualize is that there are several discrepancies when loading the dataframes. For instance in the area\_occ there is Van Nuys, 77th street for expected values; however, there are also references to outside and central. The next example is in the call type text, there are references to man, woman, other. These should be corrected for the City of LA as this is difficult to know what these mean, but I will divulge further on this in the Graduate Enhancement Essay.

Call type text field is what describes the reason for the dispatch call. Code 6 is police dispatched though this does not tell us why. Other descriptions are clear like traffic stop and dom viol. One of the more curious thoughts I have on the data set is why there does not seem to be a relational database design.

With all of the fields being text fields, then it would be difficult as a dispatcher to select data. There are several call text codes, and a more feasible method would be to have a drop down that would then populate both the code and call type fields.

I have concluded though that this is more due to the json data availability and not the database that LA County is using. When I attempted to change the data types in the json converter in mysql workbench for the sql import, zero columns of data was imported. This then makes more sense as the json data for API, is more to have a way for others to review the data outside of the organization.

I also created a time series according to locations expecting to see more of a curve as more calls are with code 6 than others. Out of 1000 records in 2020 there were around 480 calls for code 6 and in 2019 there were around 370. The bar graphs reflect this as this looks like there is more of a correlation to crime then population as lately there have been more people moving out of California. Though there is a high probability that more people moved in to LA from other areas such as the San Bernardino area.

The scatter plot shows times and cities symmetrically so with all data in the data frame there is not much to determine from either year, outside of the fact that the majority of calls for both years was code 6. The scatter plot does visualize an outlier in regards to the location. Outside is consistent with not much white space. No other location has that consistency, as all other locations have plenty of white space not matching times, as the outside location does.

The tools utilized are matplotlib, pandas, python. These were required in order to run the api’s to collect and input the data to dataframes. As discussed earlier, with the discrepancies I think that the answers to this would be more internal and will provide my insights on the enhancement essay.

Roadmap:

| Project Release | N | N + 1 | N + 2 | N + 3 |
| --- | --- | --- | --- | --- |
| Timeframe |  |  |  |  |
| Key Deliverables | Jupyter Notebook file  report  hostinfo with all info  Database |  |  |  |
| Business Sponsor(s) | West texas A@M |  |  |  |
| Business Constituency | Dr Babb |  |  |  |
| Business Processes | Python |  |  |  |
| Business Applications | Pandas, api |  |  |  |
| Data Subjects |  |  |  |  |
| Source Systems | Visual studio |  |  |  |
| Service Level Agreements (SLAs) |  |  |  |  |
| Applications Retired |  |  |  |  |
| Technology Required | PC, software for programming |  |  |  |
| Budget Allocation |  |  |  |  |
| Project Team (IT & Business) |  |  |  |  |
| Participating Resources  (outside of direct project team) |  |  |  |  |