

FINAL PROJECT PROPOSAL

ELMADDIN KARIMOV



AUSTIN, TX

For the datasets, I will take csv files from several sources. From the Austin Data page at data.austintexas.gov, I file for parks, restaurant inspection scores, crime incidents, etc.

The datasets include zip codes, so that I could aggregate all my indicators of quality of life and group them together by zip code. Basically, I will try to

measure quality of life through numbers and metrics for Austin Zip Codes and compare them to each other. Also, I will download Austin Crime data set from Kaggle.

Data Sources:

1. <https://data.austintexas.gov/browse?category=Health+and+Community+Services>
2. https://www.kaggle.com/jboysen/austin-crime#austin_crime.csv

Proposed Transformation:

- Download datasets, clean them of any irrelevant columns and drop duplicates, forming data frames.
- After exporting these data frames into SQL workbench tables, then develop one final aggregate table to show all quality of life metrics.

Potential analysis on the datasets:

- What zip codes have the cleanest restaurants on average?
- What zip codes see more crime vs. others?
- What zip codes have the most parks?