Capstone Project 1 Proposal

Bay Area Housing

The objective of this project is to research a trend in the San Francisco Bay Area housing market using the Zillow API. Trends to research could be how are prices affected by region, zip code, and/or city over time or explore how different types of housing (single-family, condominiums, etc.) prices have grown over time. Are there any areas that tend to be more expensive due to local landmarks/businesses?

This problem is especially relevant to home-seekers looking for affordable housing in the Bay Area as I've personally had 4 different families move in next door only to move out within a year. This project could also be useful for politicians and tech companies as the housing market is deeply impacted by both being a rapidly growing tech hub.

After briefly looking over the Zillow API, it seems to be a reliable source for obtain all the data I need. If that fails, or I run into problems, Quandl.com also has an API for housing data which I could possibly fall back on.

My proposal for this project is to explore how housing has grown, and continues to grow, by zip code for the San Francisco Bay Area. I'll be checking if there are any statistical price differences among those neighborhoods for my data analysis phase. It's quite possible that houses are more expensive because their localized to big businesses, public transportation, or urban areas. When it comes to modeling the project, I believe I'll be using an order of regression or a clustered model.

When I'm finished with this project, I'll be delivering my code, slide-deck, and possibly even a heatmap in the form of geographical maps or a quick clip/GIF. I'm hoping that this project can help visualize patterns that drive housing prices up and possibly come up with

methods to alleviate rising costs. One drastic decision that could be made is to centralize businesses to a specific area, making housing prices correlated by how close they are to them. Although impractical, it's an example of what decision-makers could theorize after looking at this project.