

Project Problem and Hypothesis

- What's the project about? What problem are you solving?
 - Overall negative effects of decreasing affordable housing on a community.
 - There is a lack of will to search for creative ideas for incentivizing increasing the supply of affordable housing and keep the wealth generated by land rents within the community instead of losing it to non-local speculators.
- Where does this seem to reside as a machine learning problem? Are you predicting some continuous number, or predicting a binary value?
 - Predict the near term absurd outcomes for those who earn below the income median.
 - Possibly the displacement effects, distance from employment, commute time, percentage of income towards rent/mortgage.
 - These all have continuous values so I expect to use linear regressions
- What kind of impact do you think it could have?
 - Inform policy decisions on alternative ways to create revenue for either basic income or residential housing vouchers.
- What do you think will have the most impact in predicting

the value you are interested in solving for?

- The granularity of the data available will affect the confidence and p-value greatly. If i'm only able to obtain summary data then the number of observations will be too low to generate a model which has a significant predictive capability.

Datasets

- Description of data set available, at the field level.
 - <http://www.zillow.com/research/data/> - Zillow data
 - Relevant data on price per square foot, density per zip code, increase year over year
 - <http://sfrb.org/annual-eviction-reports> - SF Rental Board statistics
 - Shows numbers and types of evictions per month at the city level
 - <http://www.antievictionmappingproject.net/ellis.html>
 - Ellis act evictions, by month and location since 1998

Domain knowledge

- What experience do you already have around this area?
 - I'm actively engaged in civic hack-a-thons and basic income projects. My father obtained his masters in

economics and became a professor of social work. I've always had an interest in the policies used to shape community outcomes.

- Does it relate or help inform the project in any way?
 - I am driven to point out alternative policy proposals that have not received the attention they deserve in my opinion. I'm dedicated to proving the positive outcomes of economies which focus on the well-being of the entire populace.
- What other research efforts exist?
 - Many research projects have focused on specific negative outcomes of an economy which only works for the wealthiest. Like the Berkeley Urban Displacement Project (<http://www.urbandisplacement.org/>) which shows the effects of displacement on the poorest, but not how that displacement also affects the community by driving up poverty, exhaustion from commute times, displacement from work, etc. We need to show that when everyone is taken care of the myriad of negative outcomes avoided has a network effect. This network effect has the possibility to be sustainable and will pay forward as a positive feedback into our bay area economy.

Project Concerns

- What questions do you have about your project? What are you not sure you quite yet understand?
 - How to obtain the key data from two projects related to displacement (Urban Displacement and Anti-Eviction mapping). Both websites do not have open links to download the raw data which drives their visualizations. I do not understand how many people impacted we may be missing and how large of a factor that is.
- What are the assumptions and caveats to the problem?
 - What data do you not have access to but wish you had?
 - I wish I had access to more granular information on evictions and people who wish they could find housing nearer to employment but can't find affordable housing (opportunity exclusion effects)
 - What is already implied about the observations in your data set? For example, if your primary data set is twitter data, it may not be representative of the whole sample.
 - That we have all the data on how people are affected as a metric, but the problems of unaffordable housing situations go beyond

measurements found in eviction statistics

- What are the risks to the project?
 - What's the cost of your model being wrong? (What's the benefit of your model being right?)
 - The cost of being wrong is implementing policy which will have a negative outcome in terms of housing. Perhaps it may drive further inequality and exclusion (for example, properties used only for AirBnB rentals may proliferate under a land value tax system)
 - Is any of the data incorrect? Could it be incorrect?
 - It is incorrect in capturing the extent of displacement as many affected may have been stealth tenants (i.e. sublets, couch surfers or similar)

Outcomes

- What do you expect the output to look like?
 - Heat maps over time showing displacement effects overlaid with indicators of real estate/rental increases and projections into the near future based on current trends.
- What does your target audience expect the output to look like?

- Maps of the Bay area showing shifting populations based on economic level.
- What gain do you expect from your most important feature on its own?
 - Showing that decreasing the percentage of income paid to rent will decrease displacement.
- How complicated does your model have to be?
 - Complicated in that I have to find similar geographic summaries of the data between housing, rental, evictions and displacement
- How successful does your project have to be in order to be considered a "success"?
 - To be considered a success, it should convince the average audience member there is a correlation in speculative housing bubble property prices/rents and displacement.
- What will you do if the project is a bust (this happens! but it shouldn't here)?
 - Have a few drinks at the pub.