<div align="center">

Data Visualization Final Project Proposal
Katon Minhas

</div>

I.   Background

As a California resident, I've been following the housing crisis in the state with a very close eye for many years now. While much of the US faces housing shortages, California is among the worst offenders. California currently holds the second fewest housing units per resident in the country, with an estimated 4 million new units required to return to a healthy housing market. This drastic difference between supply and demand has led to what researchers have described as three separate housing crises in the state. These are:

a) Severely increasing homelessness – rising rents caused by housing shortages push high-income people into units usually reserved for middle income, middle-income into those typically for low income, and low-income people onto the streets. California has the highest per capita homelessness rate in the country, and more than double the number of homeless residents as the next highest state.

b) Low-income poverty levels – Low-income residents devote more than half of their paychecks to rent, forcing people with previously decent-paying jobs to effectively live below the poverty line.

c) Decrease in new homeownership – Effecting middle and even upper-class residents, the median mortgage cost to buy a home in California is now three times the median income in the state.

II.   Project Description

For my project, I plan to produce visualizations pertaining to each of these three aspects of California's housing crisis. I will use Tableau (or Power BI) to create separate dashboards for each, with a variety of visualization types including maps and more traditional plots. The extent of the project depends on the data collection phase.

Some good datasets that have already been collected are pertaining to homelessness counts and shelter data. Datasets at https://www.hudexchange.info/resource/6802/2022-ahar-part-1-pit-estimates-of-homelessness-in-the-us/ provide a look at this information that I can join by year to housing construction data to create a look at correlation between the two factors.

The major data science portion of the project will involve projections into the future. By joining housing inventory data with homelessness data and housing cost data, I hope to use regression techniques to predict best- and worst-case scenarios of homeless populations and home prices going into the future based on specific policy measures. One idea is to have a slider where users can change the number of houses built per year parameter and see the ensuing effect on the predicted homeless population or increase/decrease in housing prices

over time should those units be built. In this way, my dashboard will become more interactive and advanced.

The resulting dashboards will be primarily published on social media. I hope to link the dashboards and post screenshots on my public Twitter account to generate interaction. There is a large housing advocacy presence on Twitter.

III.   Project Timeline

April 17   Revised Proposal + Continue Dashboarding
April 24   Continue Dashboarding + Begin final paper
May 1   Final Dashboarding + Continue final paper
May 8   Final Paper complete
May-June Post results publicly