

Data Exploration Through Tableau

ISHA SINGH
ID: 20510570

01/31/2022

—
CS 360: DATA VISUALIZATION
—

ALARK JOSHI

LINK TO THE DATA SET

[https://www.kaggle.co
m/camnugent/californi
a-housing-prices](https://www.kaggle.com/camnugent/california-housing-prices)



INFORMATION ABOUT THE DATA SET

INFORMATION ABOUT THE DATA:

The California housing data set contains data from the 1990 U.S Census.

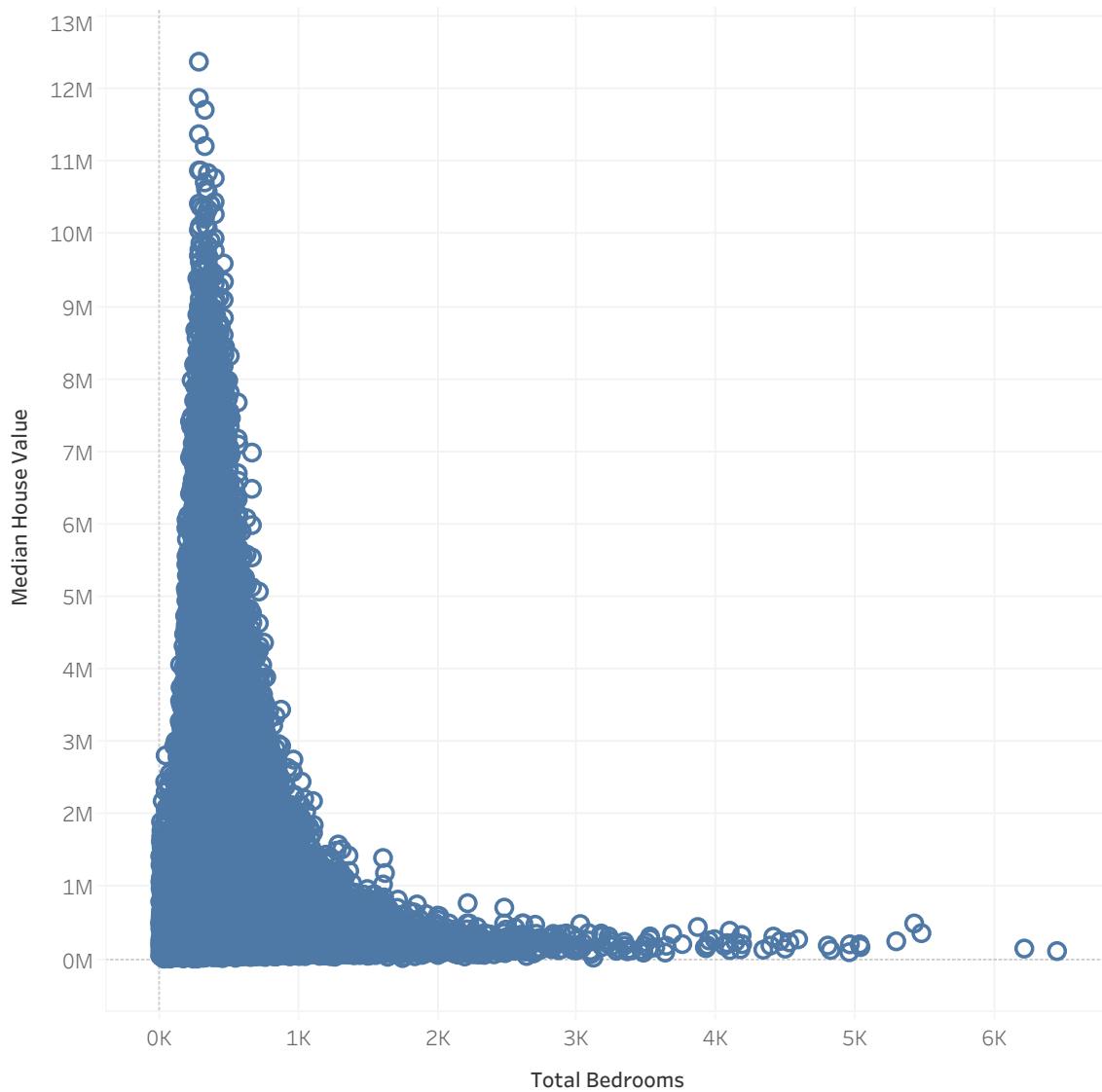
DETAILS ABOUT COLUMNS:

- **LONGITUDE**- A measure of how far west a house is; a more negative value is farther west
- **LATITUDE** - A measure of how far north a house is; a higher value is farther north
- **HOUSING MEDIAN AGE** - Median age of a house within a block; a lower number is a newer building
- **TOTAL ROOMS** - Total number of rooms within a block
- **TOTAL BEDROOMS**- Total number of bedrooms within a block
- **POPULATION** - Total number of people residing within a block
- **HOUSEHOLDS** - The total number of households, a group of people residing within a home unit, for a block
- **MEDIAN INCOME** - Median income for households within a block of houses (measured in tens of thousands of US Dollars)
- **MEDIAN HOUSE VALUE** - Median house value for households within a block (measured in US Dollars)

Reference: <https://developers.google.com/machine-learning/crash-course/california-housing-data-description>

VISUALIZATION #1

Total Bedrooms vs Median House Value



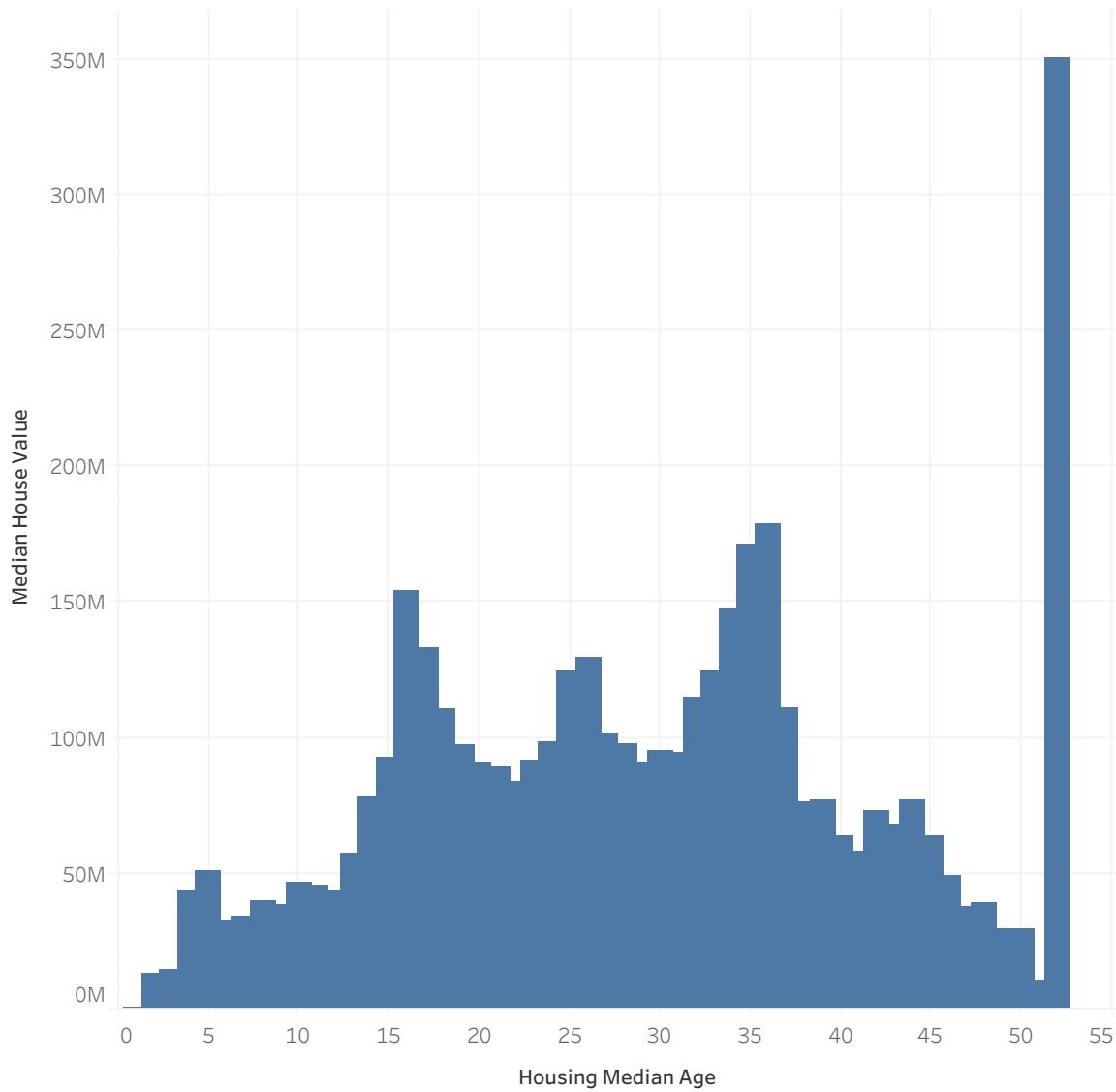
Total Bedrooms vs. Median House Value.

My hypothesis was to figure out if there is positive correlation between total number of bedrooms within a block and median house value within a block. To find a solution to the hypothesis, I created the scatterplot above which represents total bedrooms vs medium house value. Unfortunately, my hypothesis was not particularly

correct. The reason for that was because there is a negative correlation between total bedrooms and median house value.

VISUALIZATION #2

Housing Median Age vs Median House Value

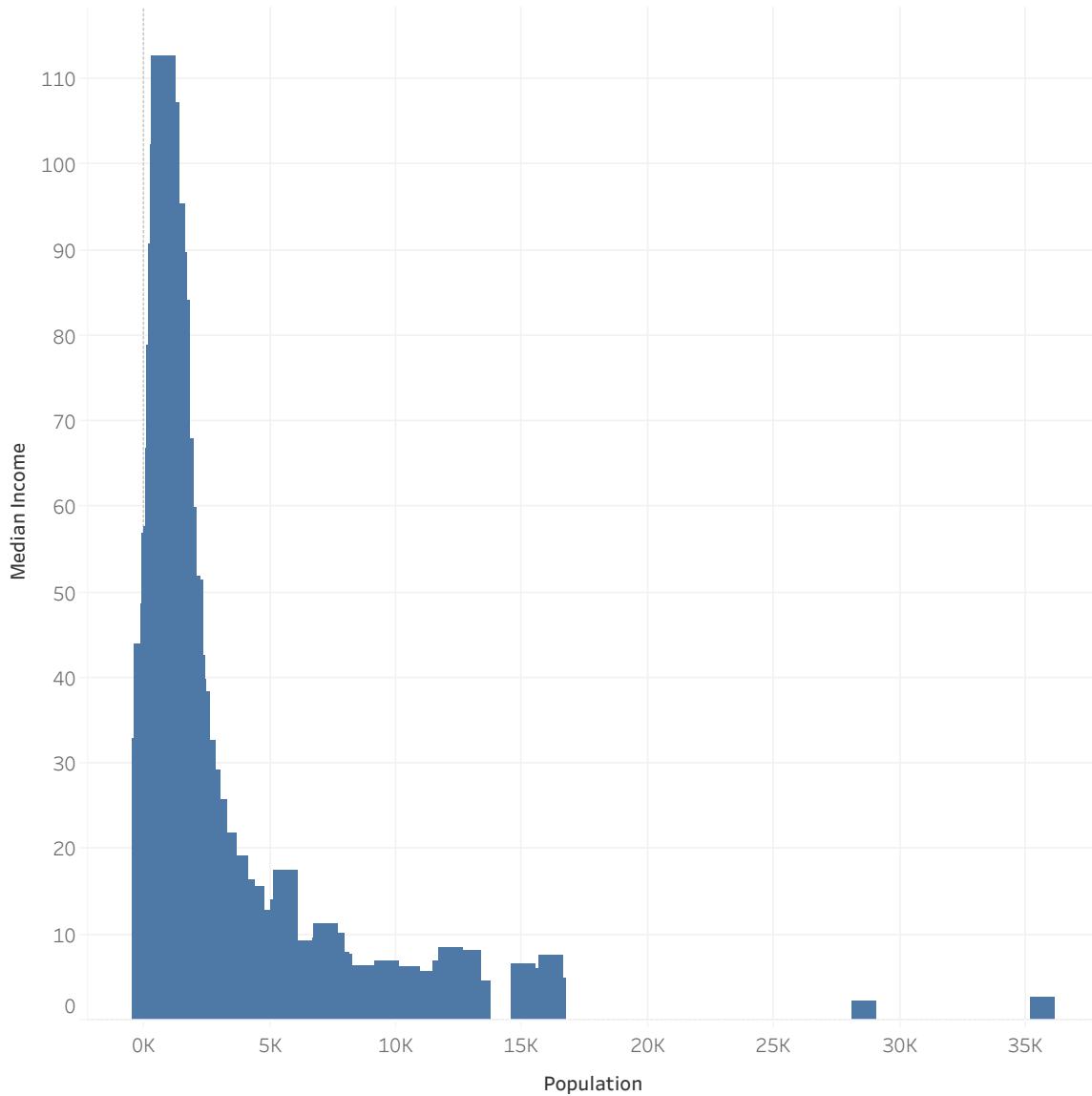


Housing Median Age vs. Median House Value. Details are shown for Median House Value.

My hypothesis was to figure out if there is some sort of relationship amongst the housing median age and median house value and seemed like that it is somewhat symmetric but the older the house is the more expensive it becomes. In that case, the statement is true as a lot of houses have a better pricing value when years go by.

VISUALIZATION #3

Population vs Median Income

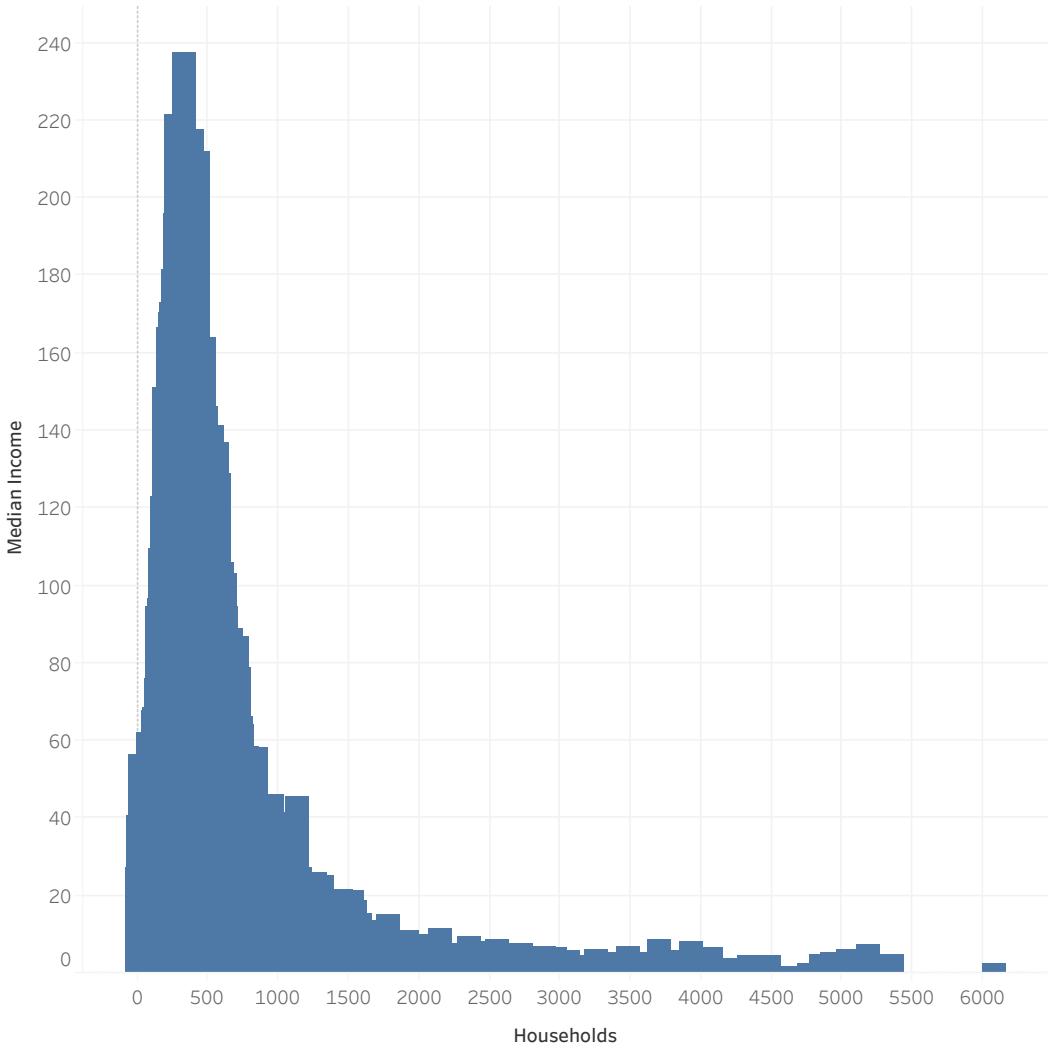


Population vs. Median Income.

My third hypothesis was to figure out on if there is some sort of relationship amongst the population with the median income and it seemed that the graph was nearly skewing to the right so possibly that can mean that the mean is more than the mode. There are as well couple outliers, which lie around 28K and 26K.

VISUALIZATION #4

Households vs Median Income

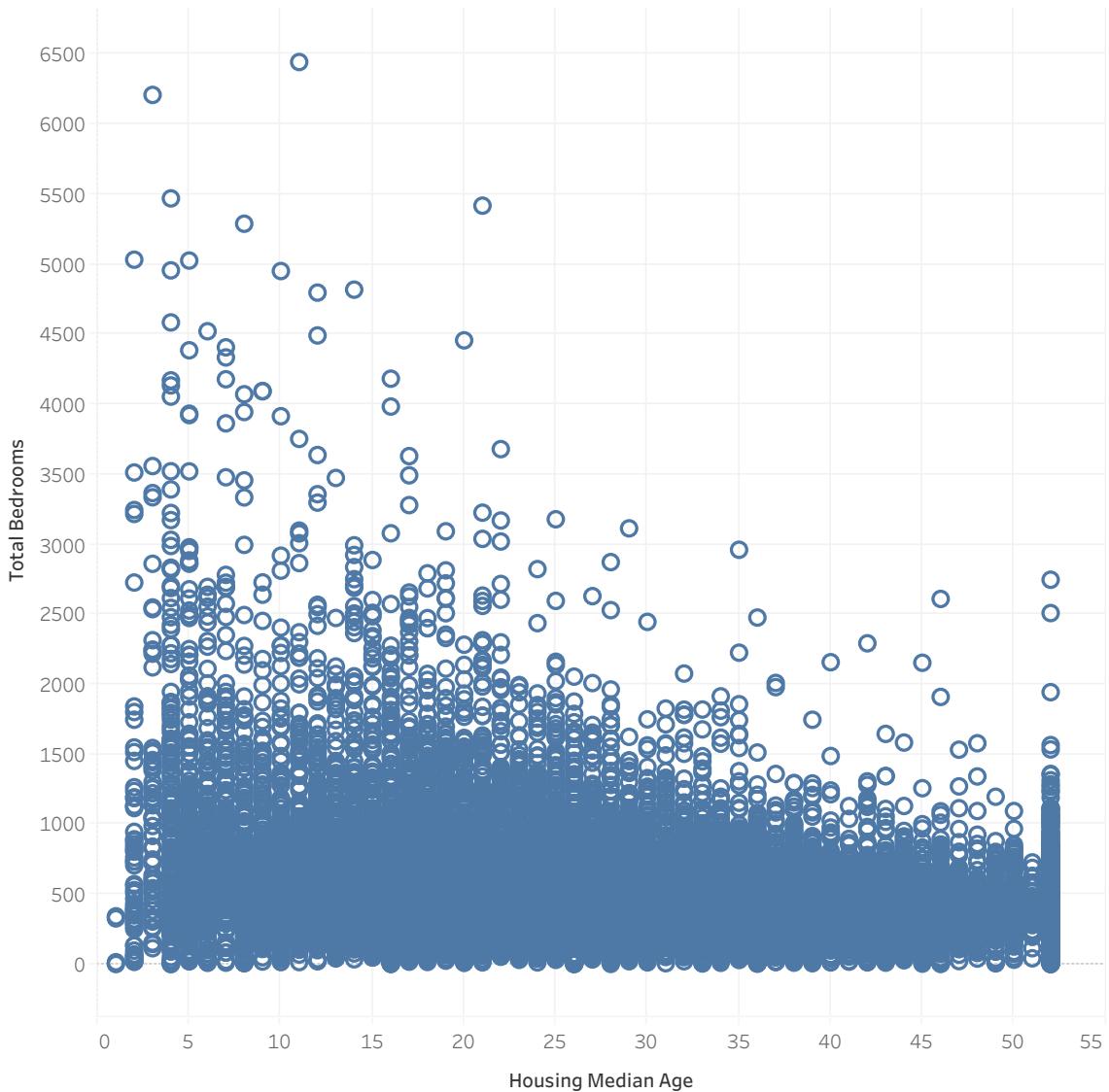


Households vs. Median Income.

The following graph above shows how the less the number of households, the more the median income there is. It seems to be weird as I believe that a greater number of households should have a high increase in median income as well. Alas, my hypothesis was not knowingly correct, but does possibility can be he more the people in the family, the more the expenditure and therefore less income.

VISUALIZATION #5

Housing Median Age vs Total Bedrooms



Housing Median Age vs. Total Bedrooms.

Sometimes with the age a lot of people tend to make more rooms in their houses. That is why I wanted to see if how many bedrooms are there when a specific block's home gets older. In that case, we can see that my hypothesis was wrong as there is not a set factor in a decision where people tend to have more rooms when home gets older. Therefore, my hypothesis is wrong and there is no association.

VISUALIZATION #6

Total Bedrooms VS Median House Value

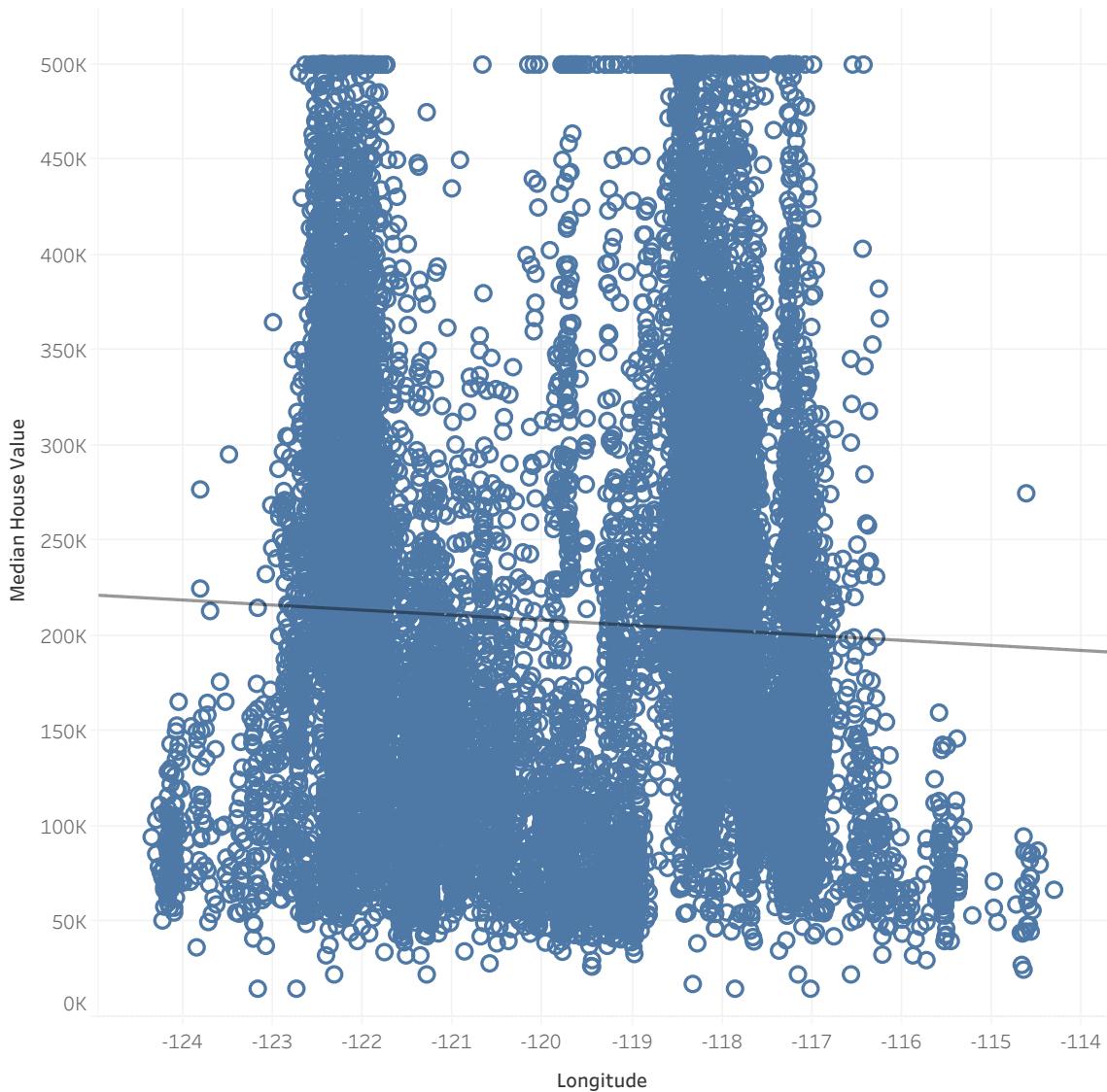


Total Bedrooms vs. Median House Value.

I wanted to see if the total bedrooms and median house value have some relationship. As you can see from the visualization above, there is no correlation between them, so which means total bedrooms has no “association” with median house value. The correlation coefficient value was 0.002 percent. That was very shocking as I thought that more the bedrooms, the more the square feet would be needed which means the more the value of the house.

VISUALIZATION #7

Longitude VS Median House Value

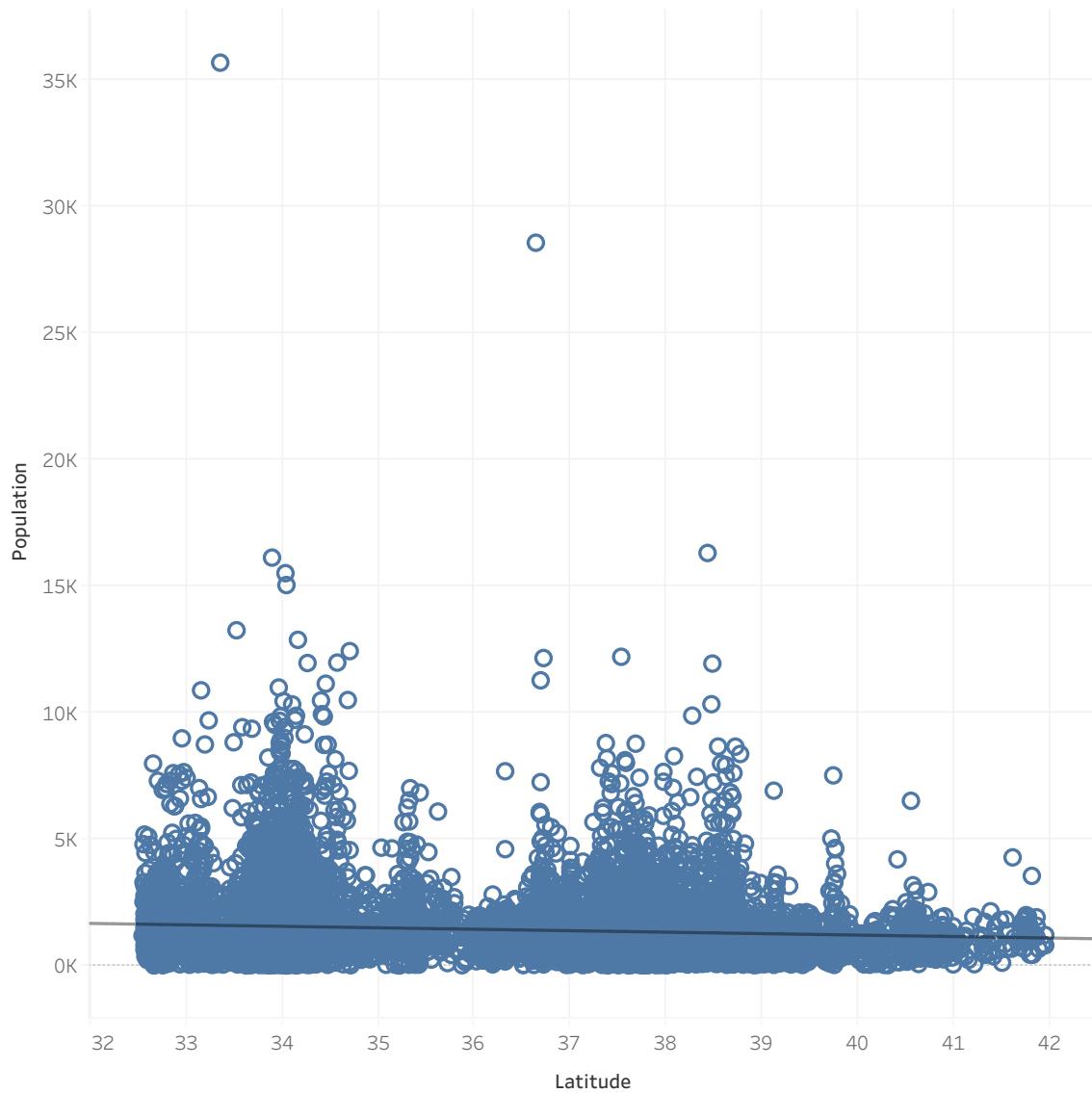


Longitude vs. Median House Value.

I was curious on understanding if there is any connection because longitude and median house value. For example, the closer proximity to the beach, the more expensive the house can be but sadly, it does not look like it. The correlation coefficient was very less: 0.002 percent, which means there was no association between those.

VISUALIZATION #8

Population Vs Latitude



Latitude vs. Population.

I wanted to see if there is some sort of association on if there is anywhere a more distinct population where more people prefer residing and it seemed to not be that case as the correlation was 0.011 percent