

Capstone Statistical Analysis Report

Venkatesh Krishnan

June 19, 2018

Interesting Facts:

Percentage of households with AGI under \$ 1M is approximately 83.99% and the corresponding home value range is \$ 11351 to \$ 3641564.

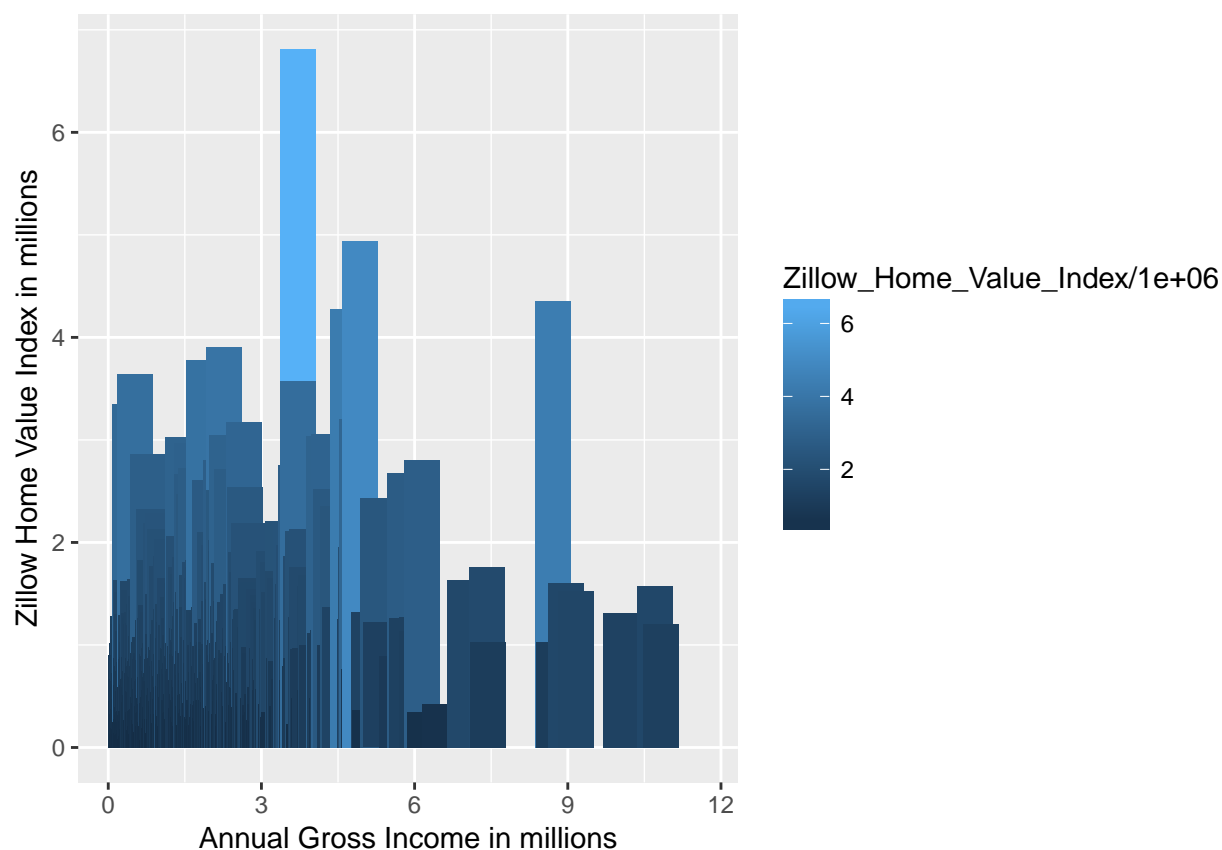
Percentage of households with AGI between \$ 1M and \$ 5M is approximately 15.85% and the corresponding home value range is \$ 64053 to \$ 6810939.

Percentage of households with AGI over \$ 5M is approximately 0.16% and the corresponding home value range is \$ 345803 to \$ 4350546.

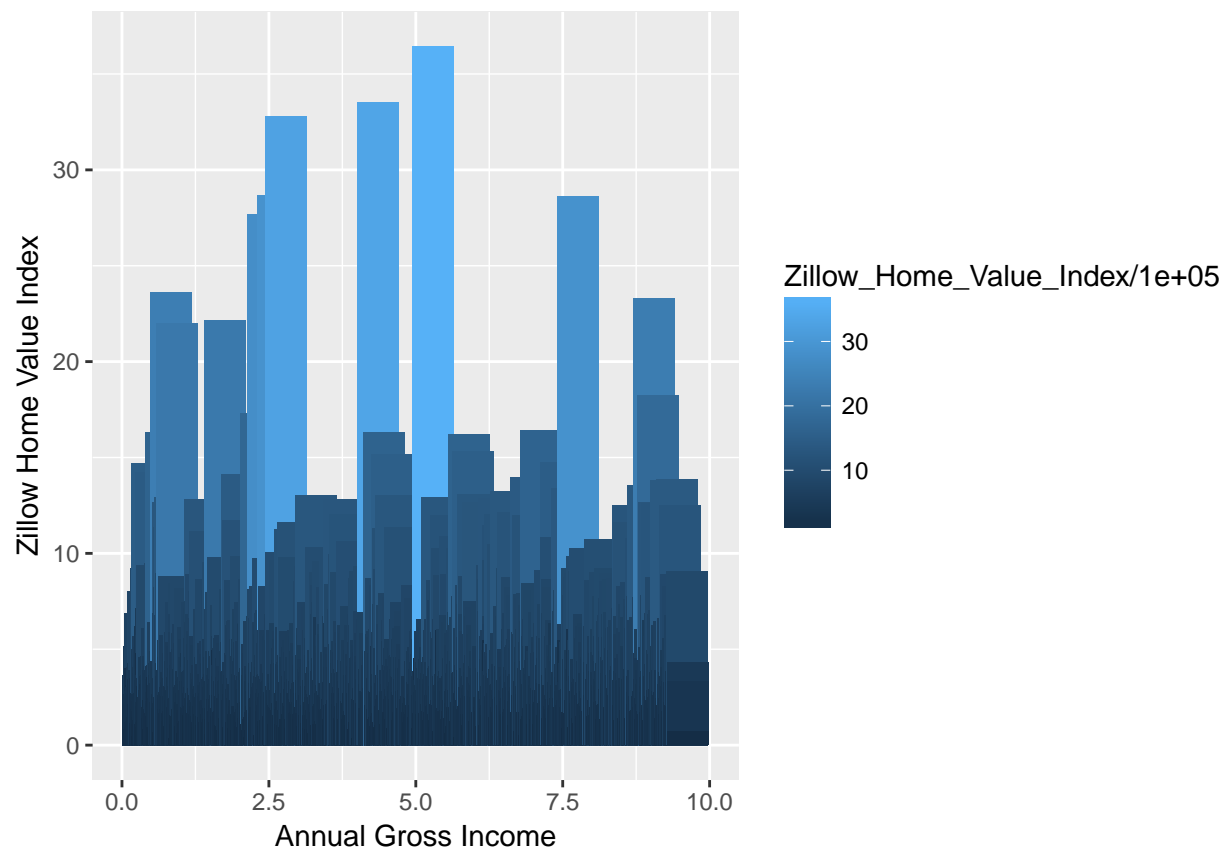
Trends:

There was a positive correlation of Household Incomes to Home Values.

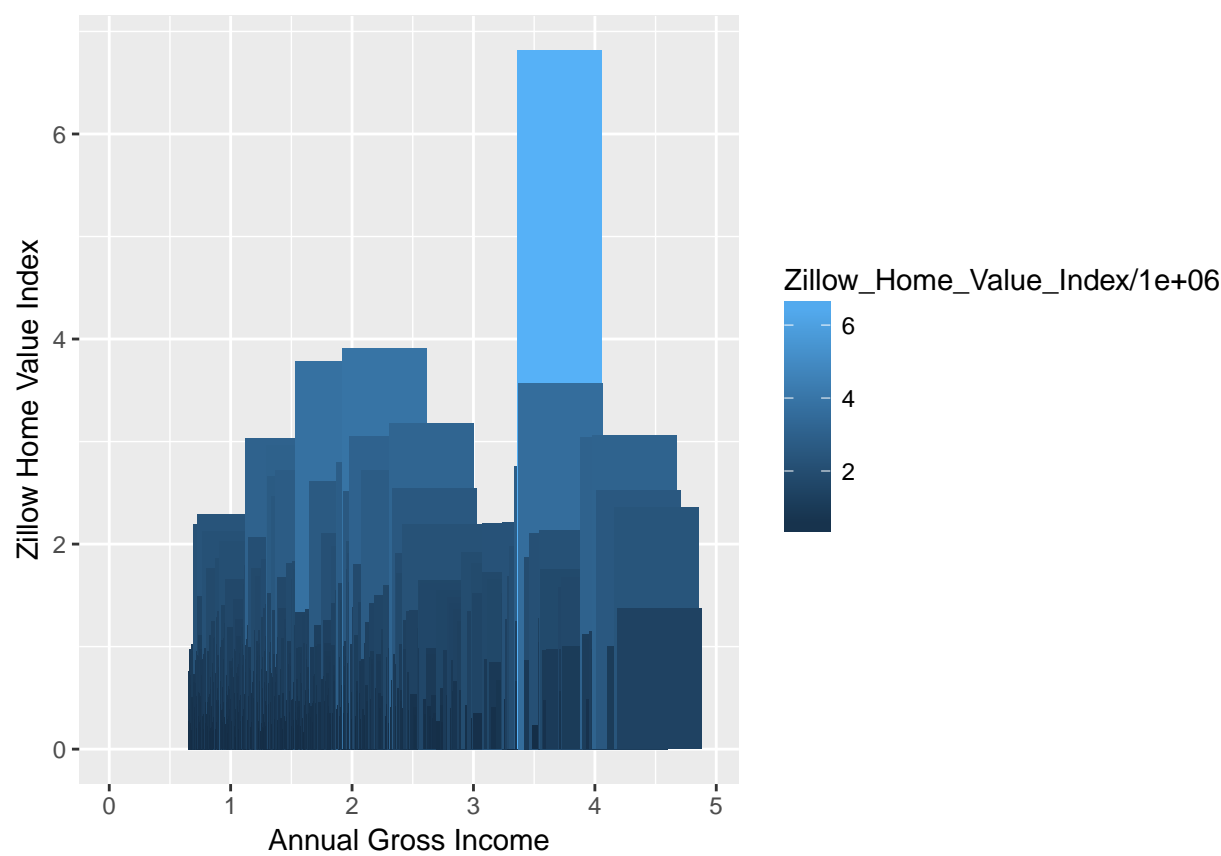
Bar Plot:



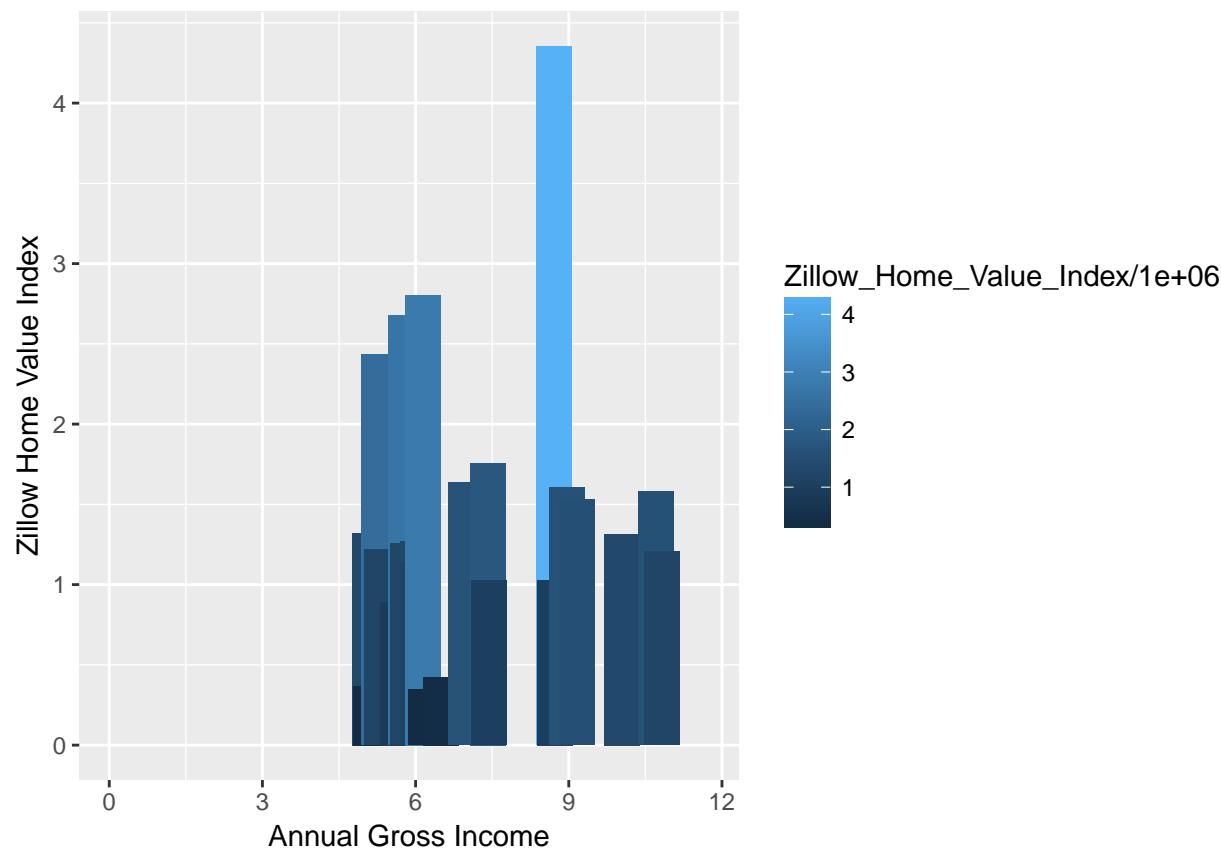
Household income less than a million



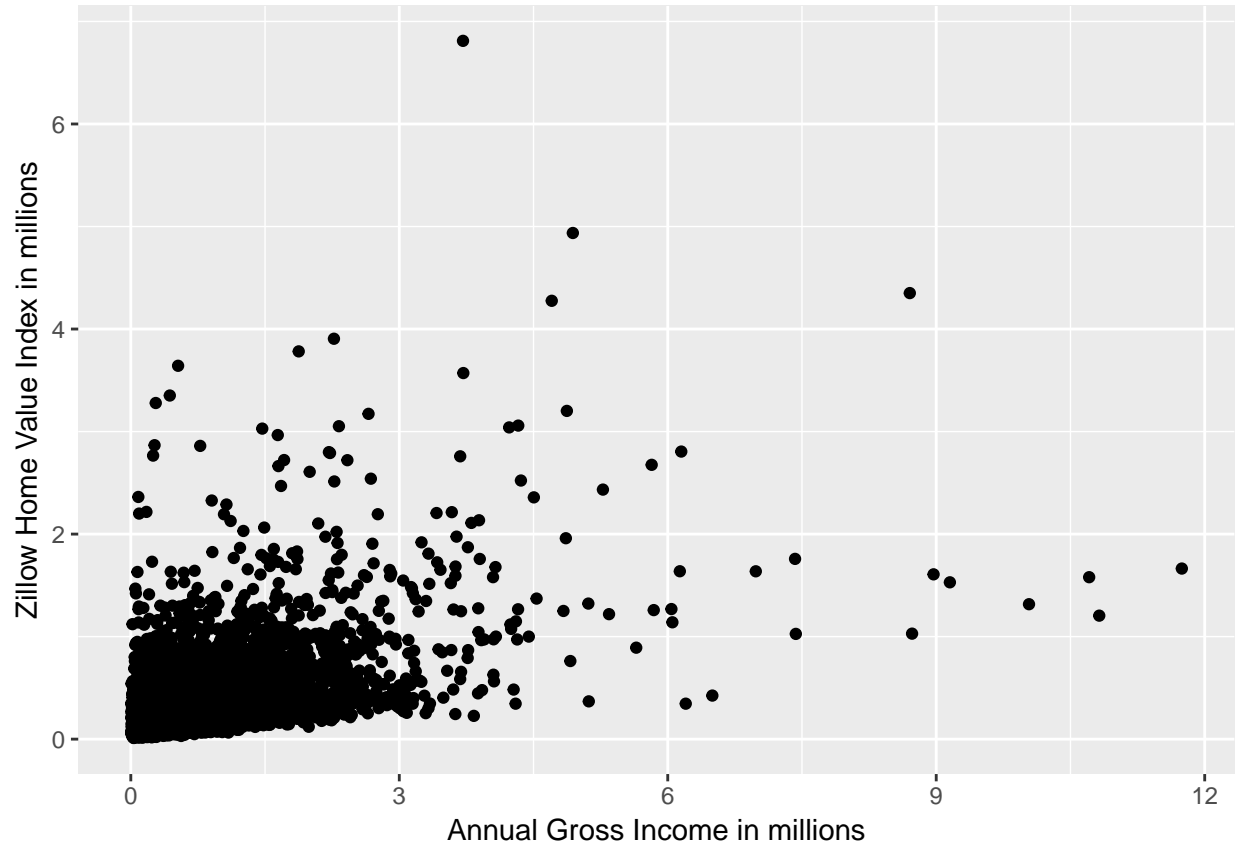
Household income between 1 & 5 million



Household income above 5 million



Scatter Plot:



Hypothesis for further Investigation:

We have only obtained a sample from approx. 14,000 zip codes of USA out of a total of approx. 42,000 zip codes in the country. So, further data is recommended to be obtained to extrapolate this analysis.

Leading Questions:

- 1) Can we analyze the data for each state and compare it with the country's data.
- 2) This is only the latest data from 2016. Can we also obtain historical data for the past 20 years or more to arrive at trend analysis. This will lead us to finding out the driving factors of a particular geography.
- 3) Currently what we see from Zillow is the actual price paid for a home and it's current value. What is not available now is the projection - 1 year from today, 3 years from today etc., This can provide insight for a customer into a particular market.
- 4) We can incorporate the crime data of a particular area / zip code that may throw some insight into home valuation.
- 5) Investments in technology, medicine, infrastructure, realstate (Commercial, Residential and Retail) can act as direct indicators for price prediction in an area.