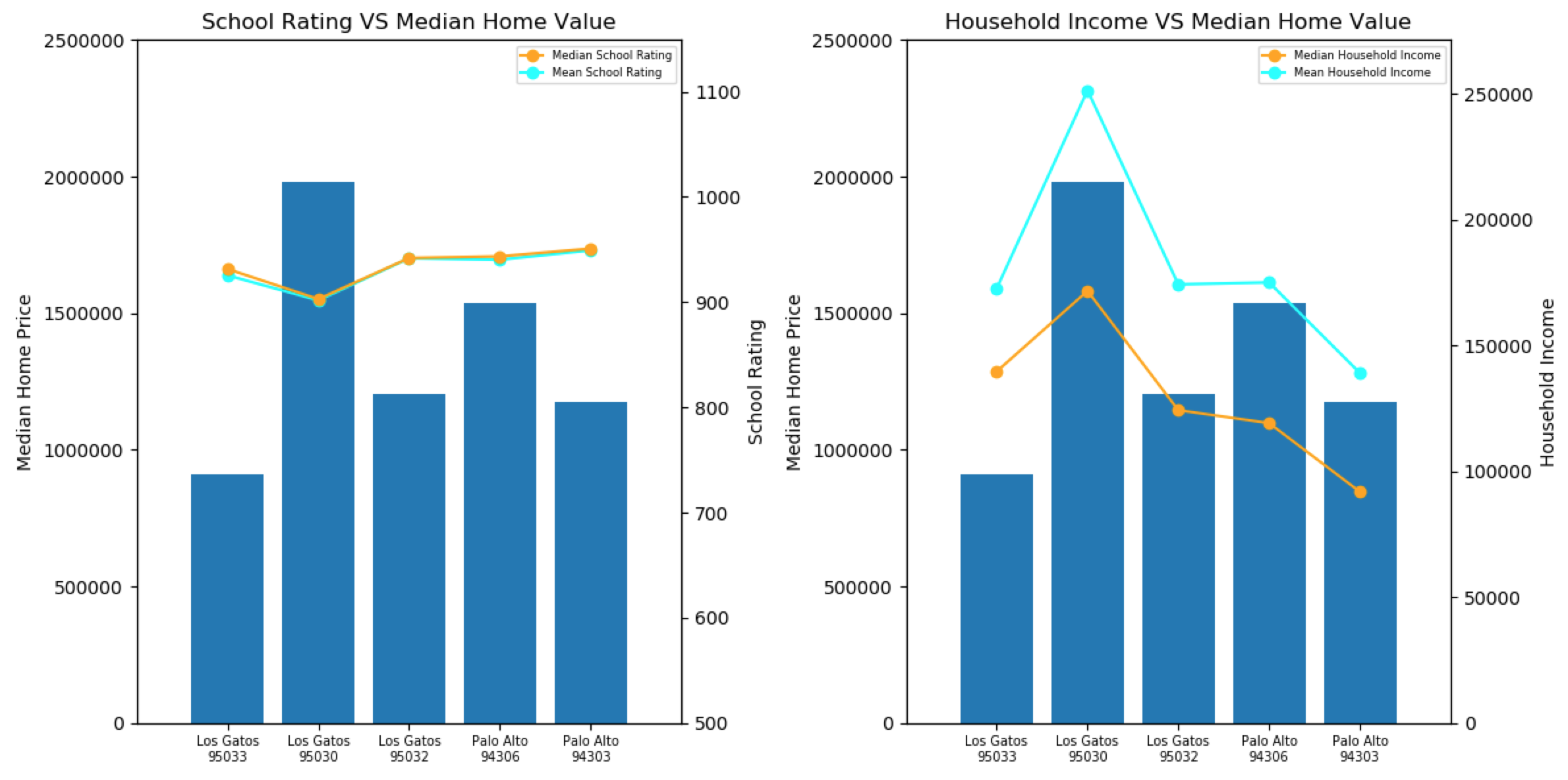
'''

Author: Anna Cho

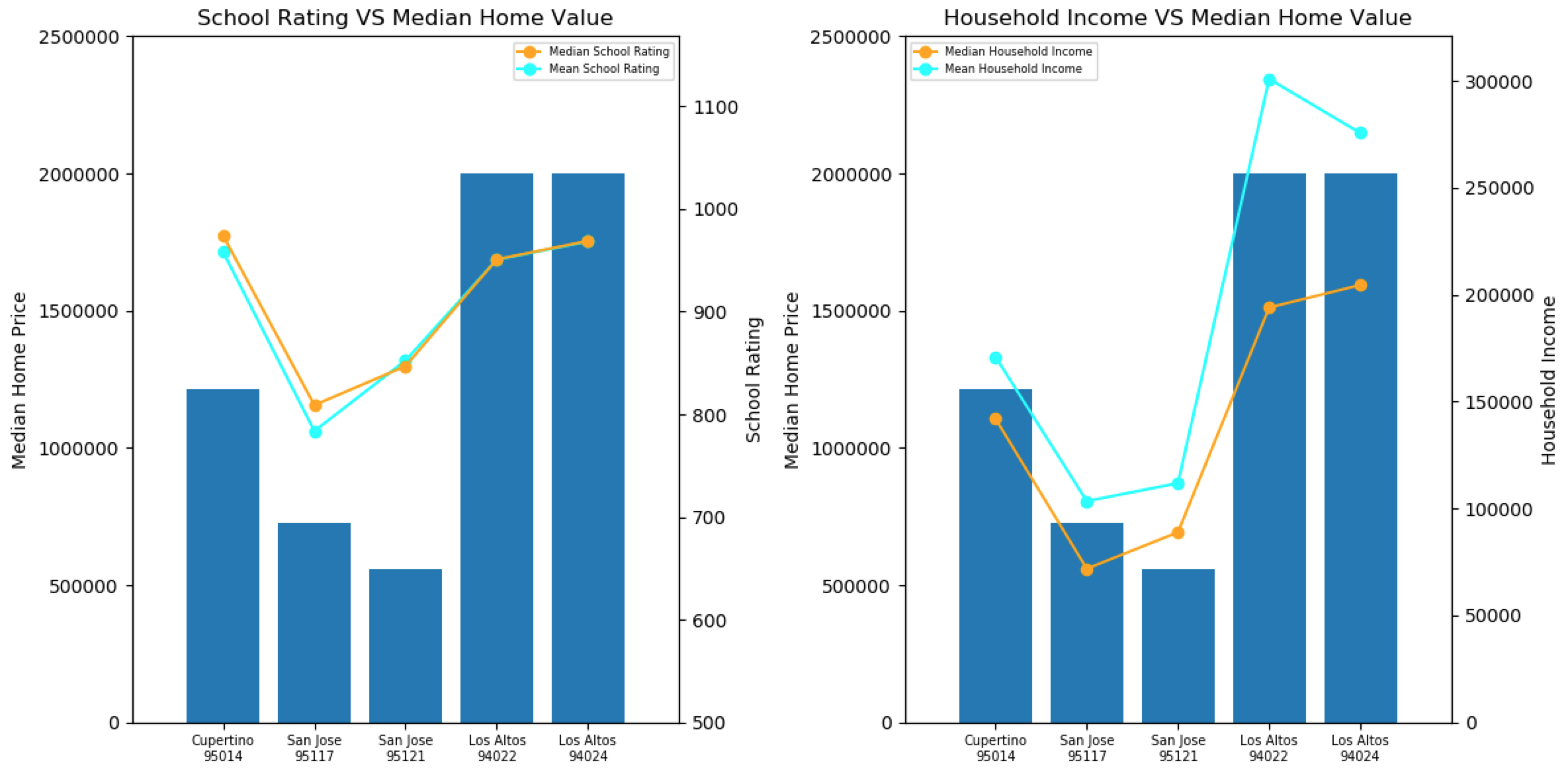
'''

The main goal of the data analysis in this project is to find correlations between median housing price in Santa Clara county and: (1) the median and mean household income, and (2) the median and mean public school rating among the user-chosen zip codes

The general trend we found within the data is that zip codes with high home prices have high public school ratings and high household incomes, but we see a substantial difference between mean and median household incomes.



The plot on the right side graph shows that the difference between mean and median household incomes is nearly $100,000 (Los Gatos 95030). In most cases, Santa Clara County shows a similar pattern: the mean household income is higher than the median household income. When the Census bureau measures the median household income, they only look at the income of those households that are in the 50th percentile. But, the Census’s mean household income calculation includes outliers making a small number of high income households to drive a high mean household income value. This suggests that in Santa Clara County, the income is distributed unequally.



Unlike the household income, the difference between mean and median values in public school rating is not significantly different. The zip codes with a high home price have a high public school rating. But, some zip codes, such as Cupertino 95014, have the higher school rating compare to zip codes with much higher median home price, Los Altos 94022 and 94024. This could be used as an indicator when a family with school age children is seeking to purchase house to maximize their purchasing power.