In this project I did major cleaning on the raw_house_dataset. In the beginning,I figured out the null values in every column and replaced it with the appropriate value. The first column I selected was lot_acres and I replaced all null values with the median of the lot_acres.I repeated this procedure with the rest of columns which include sqr_ft,taxes,year_built,HOA,garage,and bathrooms. The only column I changed null values to mode was for kitchen_features and this was because most houses stay consistent with the most popular kitchen arrangement.

- Used google cloud library to mount google drive
- Used pandas to load csv file
- Accounted null values using df.isnull().sum() command
- Converted hoa object to float
- Replaced null values using df.fillna() command
- Replaced null values using list comprehensions and assignment
- Used seaborn to plot boxplots and figure out outliers