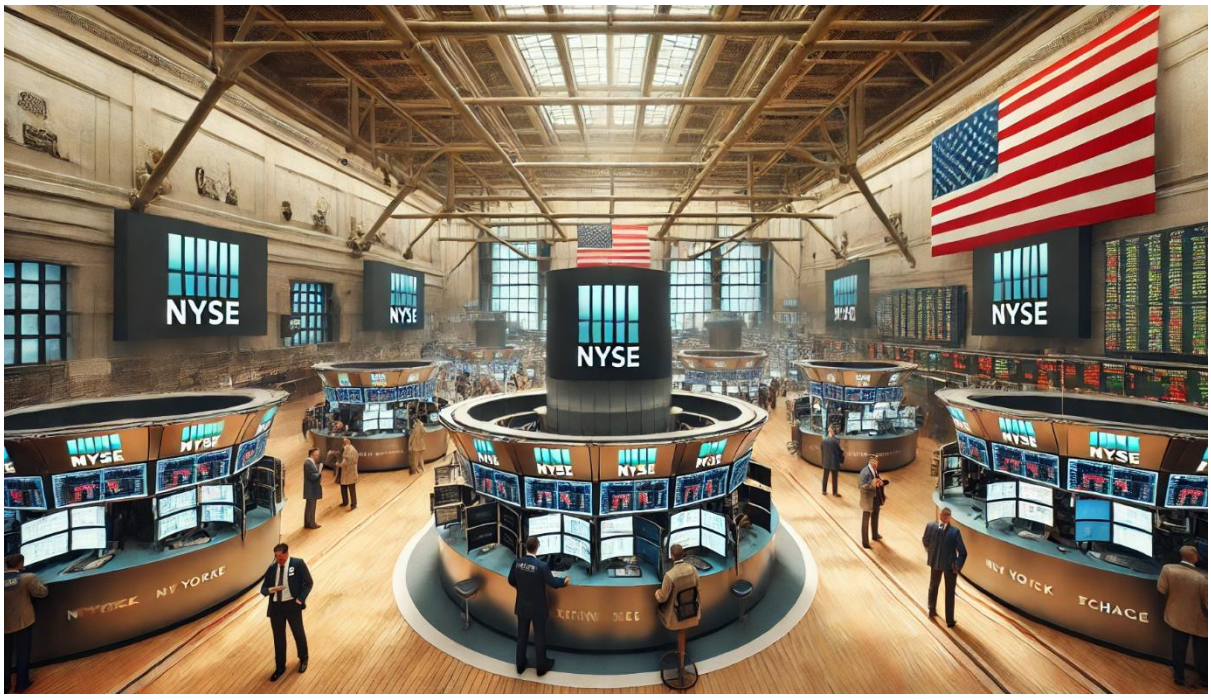
TECH CONATIVE – ASSIGNMENT

NAME: SANJEEV M ROLL NO : BE21B034

https://github.com/SanjeevM2004/TechConative_THA.git

1. *ChatGPT*

   PROMPT: "A busy and realistic scene inside the New York Stock Exchange (NYSE) trading floor. The room is filled with traders and people engaged in conversation, standing near circular desks (Focus a bit closer on one circular desk) filled with multiple monitors displaying stock market data. Large electronic displays with 'NYSE' logos are placed prominently throughout the space. American flags are hanging using a rod from the ceiling and there are 2 hanging flags on both sides and one hanging flag straight end of the picture. All flags are hung using white rods, and the floor has a wooden finish. The architecture of the building includes high ceilings with visible metal beams and a large glass window façade at the back, letting in natural light. Make the flags hanging using rods. And focus the table on the center with a perfect straight vertical view. No angled images. The environment is bustling with activity, capturing a financial and business atmosphere. Generate a "realistic image" based on the following environment envisioned. "



   https://www.freepik.com/pikaso/ai-image-generator

   PROMPT 2: "A busy and realistic scene inside the New York Stock Exchange NYSE trading floor. The room is filled with people engaged in conversation, standing near the center large circular desk. Focus a bit closer on one circular desk filled with multiple monitors displaying stock market data. Large electronic displays with 'NYSE' logos are placed prominently throughout the space. One American flag iis

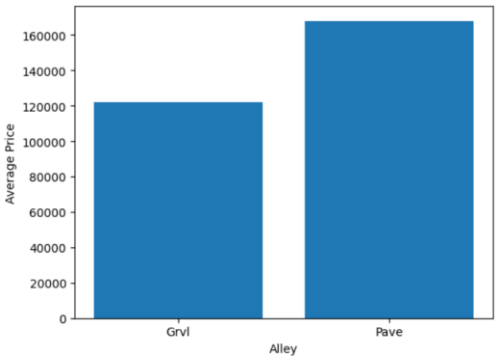
hanging using a rod from the ceiling on the left wall and One American flag on the right wall symmetrically and one flag hanging at the straight end of the picture. All flags are hung using white rods, and the floor has a wooden finish. The architecture of the building includes high ceilings with visible metal beams and a large glass window façade at the back, letting in natural light. Focus on the center with a perfect straight vertical view. The environment is bustling with activity, capturing a financial and business atmosphere. Generate a colorful 'realistic image' of "45 degree drone view" of the environment envisioned. 45 degree Drone view with people too”

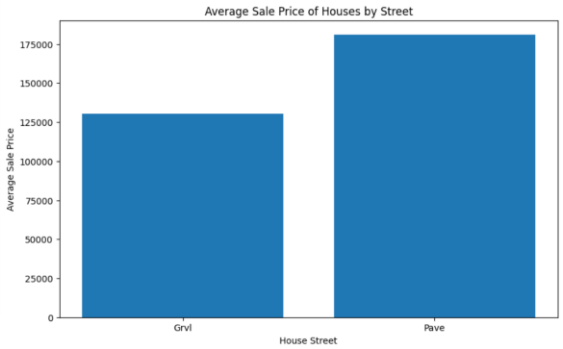
2. Link for the github repository:
https://github.com/SanjeevM2004/TechConative_THA.git

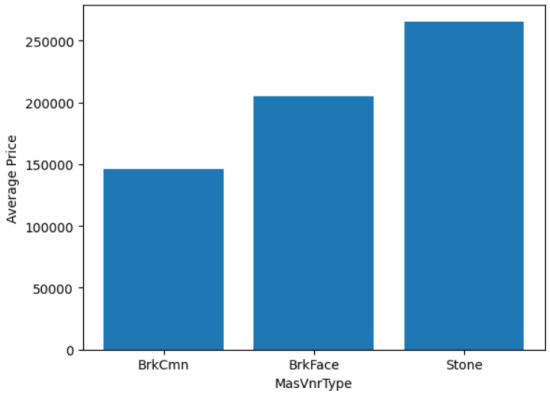
Analysis:
The data has 1460 rows and 81 columns

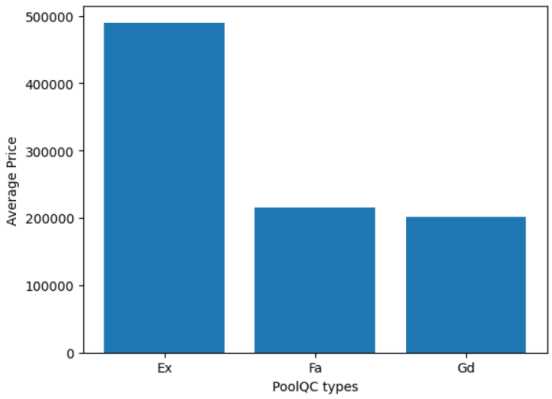
Columns with more than 100 NaN values:

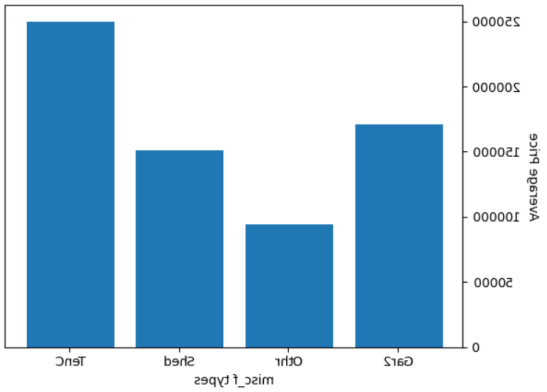
- Alley
- MasVnrType
- PoolQC
- Fence
- MiscFeature

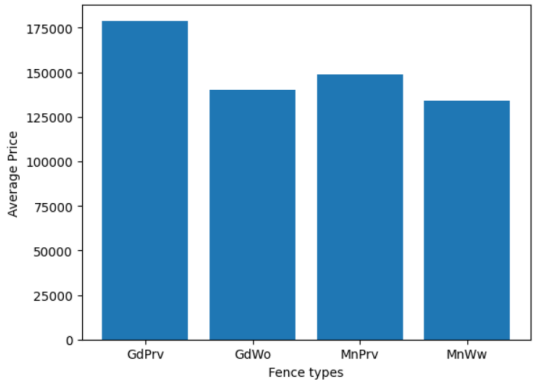










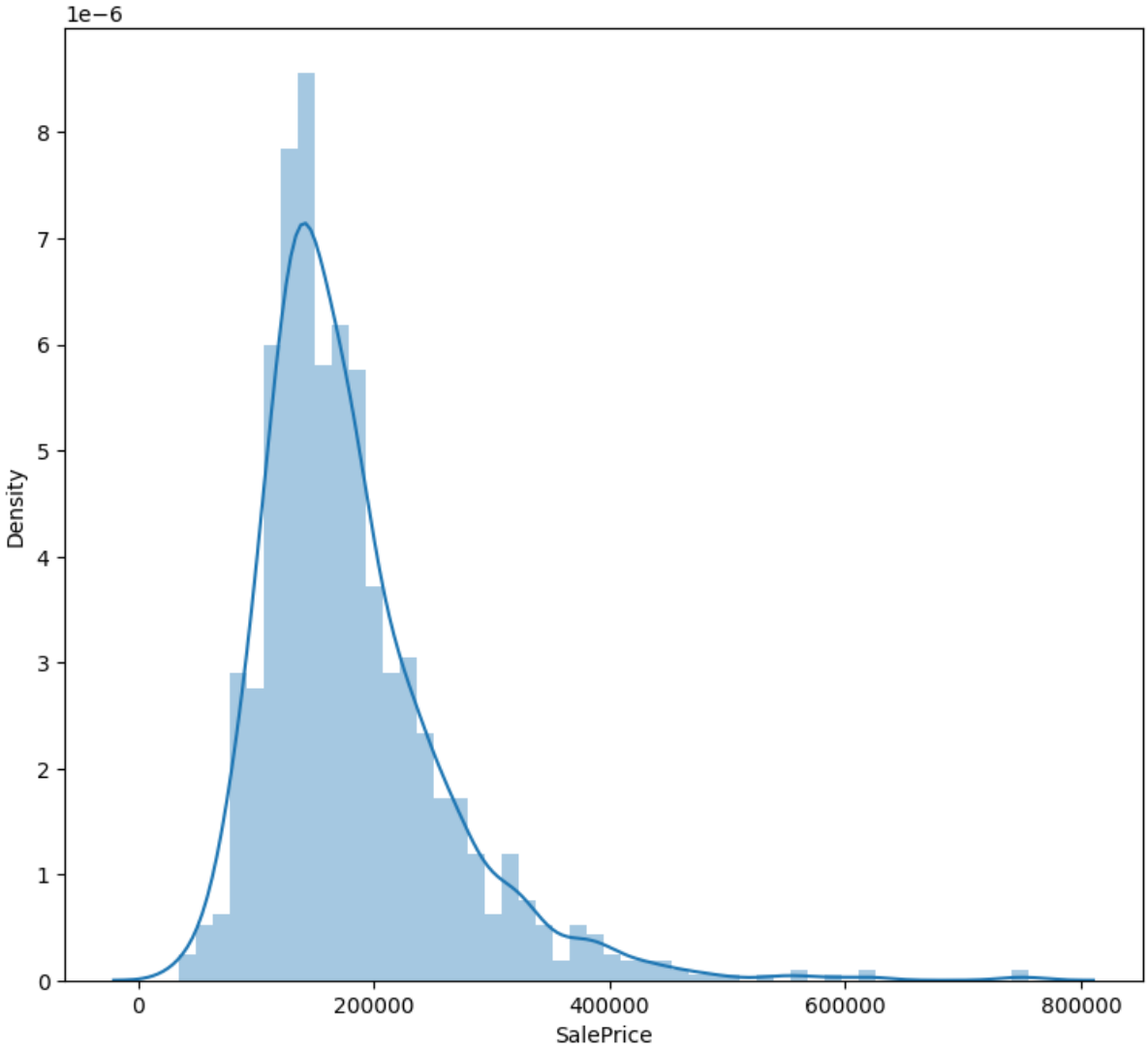


- The Alley information is given by Street without much nan value. atleast approximately. So we can delete Alley column
- We are not dropping MasVnrType as it provides useful info which is not compensated.
- Even PoolQC is useful info as the class Ex alone has shown drastic shoot up in price
- Misc Feature is useless as the shed shows hsoot up in price but its not important and since it adds different features, it adds high statistical bias
- We are not removing Fence as the classes of fence provide some information for classification. Also theoretically fencing adds extra cost.

MSSubClass MSZoning Street Alley LotShape LandContour Utilities LotConfig LandSlope Condition1 Condition2 BldgType HouseStyle OverallQual OverallCond RoofStyle RoofMatl Exterior1st Exterior2nd MasVnrType ExterQual ExterCond Foundation BsmtQual BsmtCond BsmtExposure BsmtFinType1 BsmtFinType2

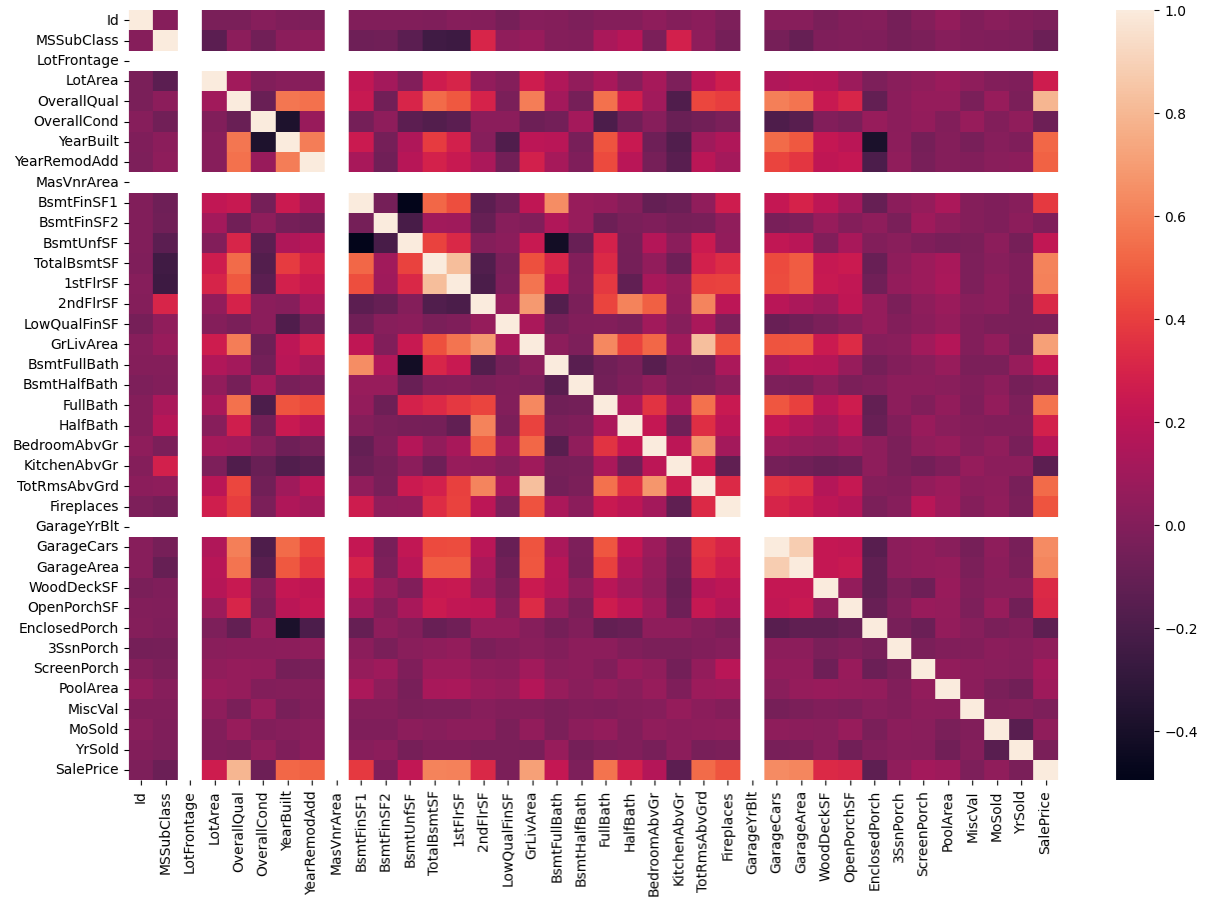
Heating HeatingQC CentralAir Electrical BsmtFullBath BsmtHalfBath FullBath HalfBath BedroomAbvGr KitchenAbvGr KitchenQual TotRmsAbvGrd Functional Fireplaces FireplaceQu GarageType GarageFinish GarageCars GarageQual GarageCond PavedDrive PoolArea PoolQC Fence MiscFeature MoSold YrSold SaleType SaleCondition

The above are all the columns that have features rather than proportionate values or random numbers. All those have classes and they alone can sufficiently decide using Decision tree classification, etc.



Distribution of the SalesPrice is right skewed as it is higher at the left hand side with maximum house sold at the rate of 200000.

So the mean of the houses would be somewhere around 200000-400000 and the median would be at 200000.



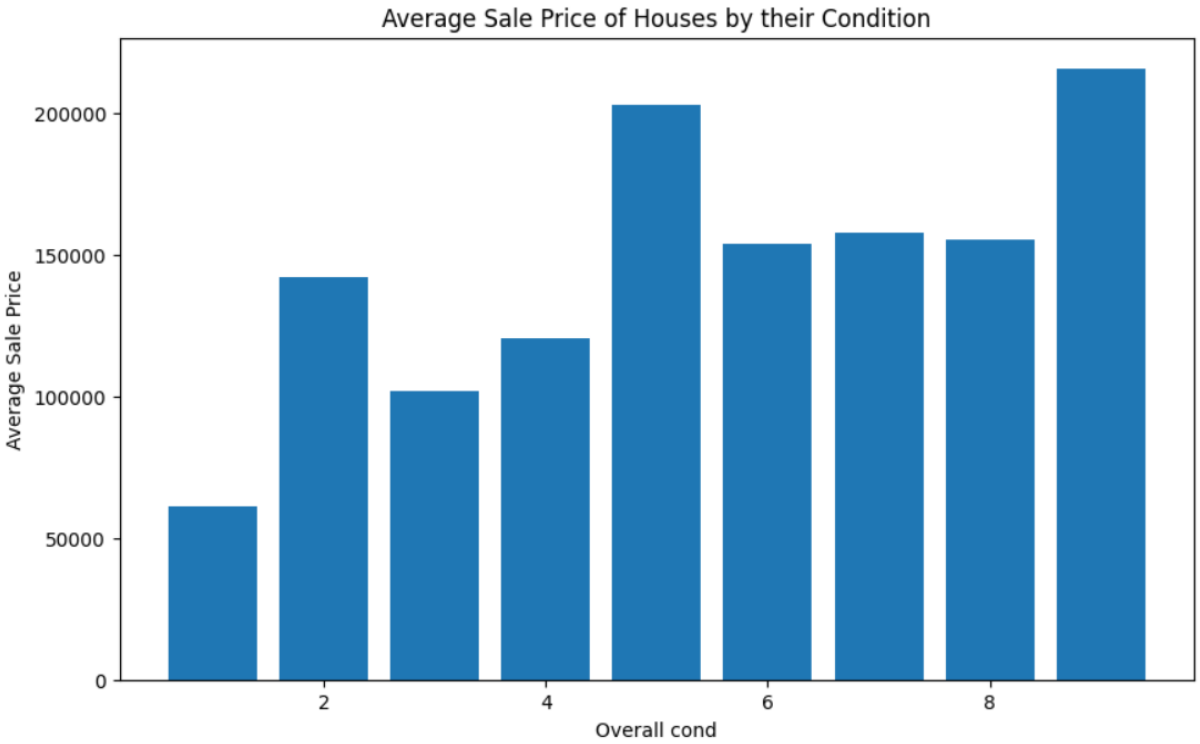
This is the heat Map which tells us about relation between different columns of data.

```
# Based on some understanding of Google information about House sale pricing and from
data['overall cond-qual']        = data['OverallQual']  + data['OverallCond']
data['YearRemodAdd-YearBuilt ']  = data['YearRemodAdd'] - data['YearBuilt']
data['BsmtBaths']                = data['BsmtFullBath'] + data['BsmtHalfBath']*0.5
data['AboveBaths']               = data['FullBath']     + data['HalfBath']*0.5
data['TotalBaths']               = data['BsmtBaths']    + data['AboveBaths']
data['OverallExteriorLivingSF'] = (data['WoodDeckSF'] +
                                   data['OpenPorchSF'] +
                                   data['EnclosedPorch'] +
                                   data['3SsnPorch'] +
                                   data['ScreenPorch'] +
                                   data['PoolArea'])
data['TotalAboveGradeLivingSF'] = data['GrLivArea'] + 0.5 * data['LowQualFinSF']
data['TotalBasementSF'] = data['BsmtFinSF2'] + data['BsmtUnfSF']
data['OverallUsbleSf'] = data['GrLivArea']+data['OverallExteriorLivingSF']
```

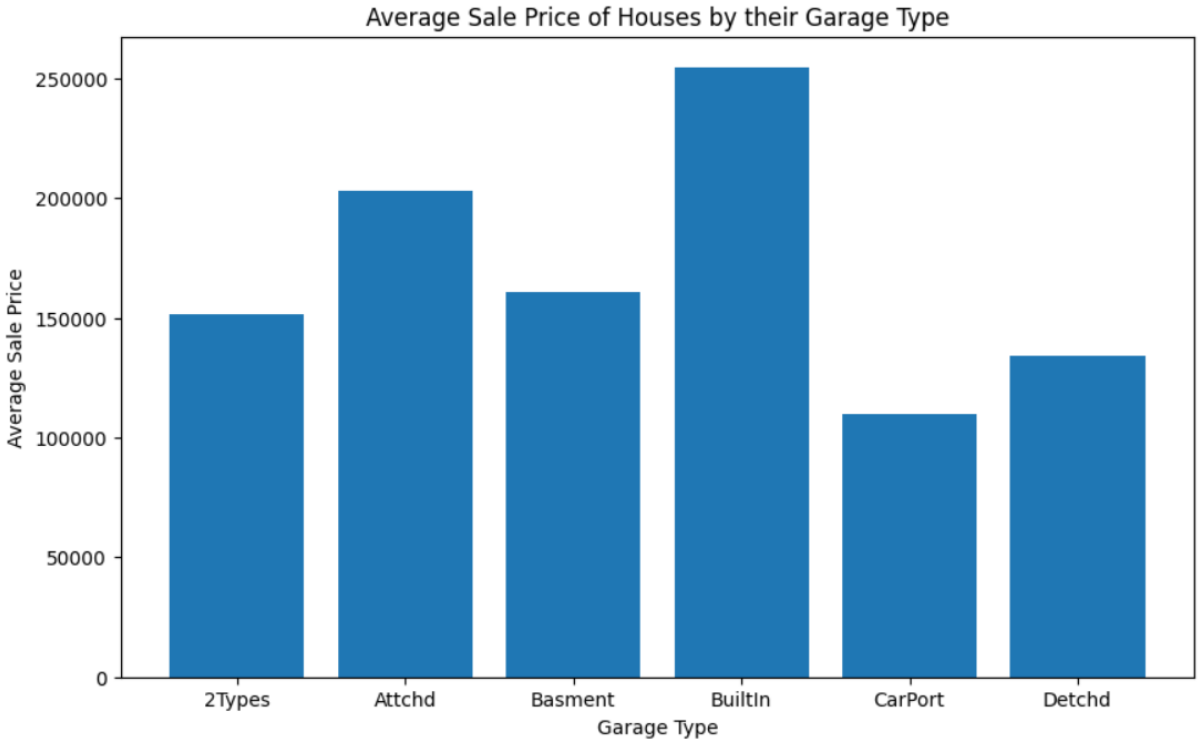
So these are some artificial variables I constructed to increase the r2 score which also increased y 0.06, which is 6%

The graph between Overall quality and SalePrice is given as

Which proves they both are proportional and go hand in hand.



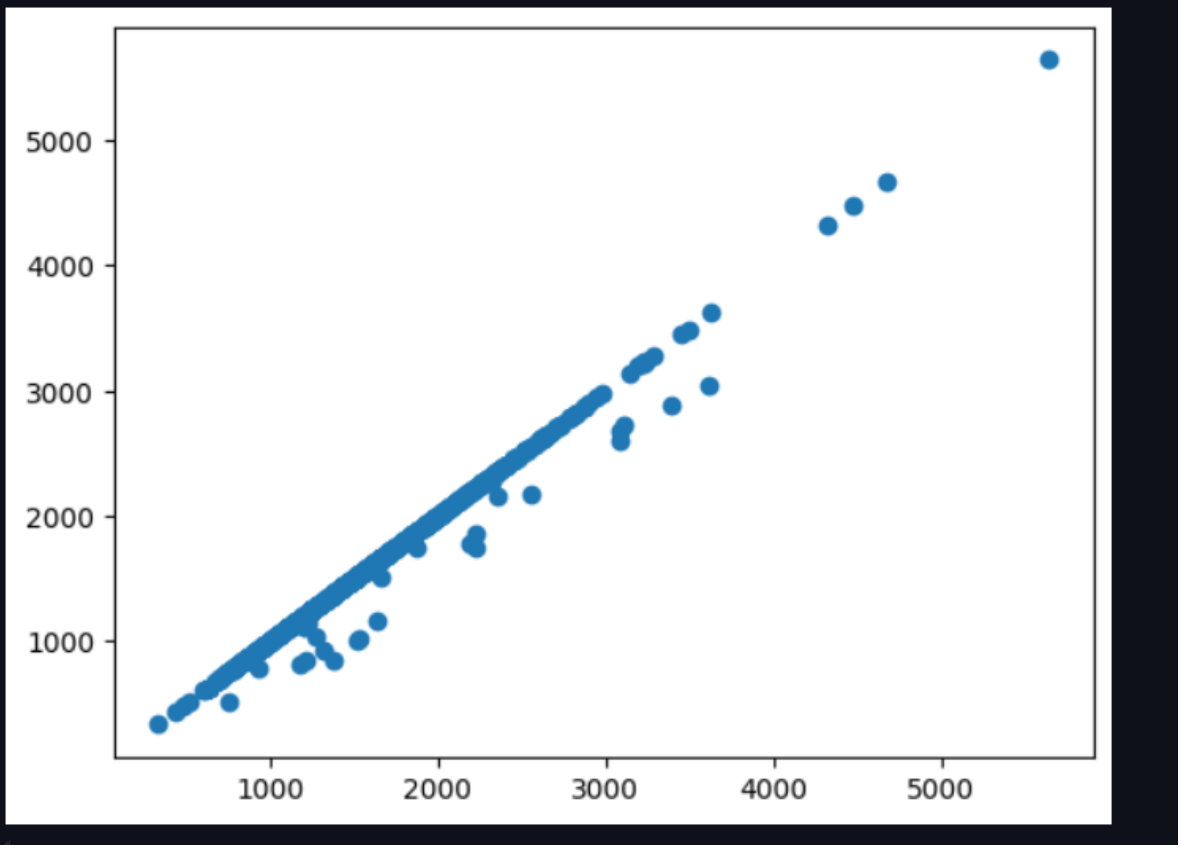
As u can see below The Garage types are plotted with respect to average price and we can clearly observe that my Average Sale Price is high for Built in garages



```
plt.scatter(data['GrLivArea'], data['1stFlrSF']+data['2ndFlrSF'])
```

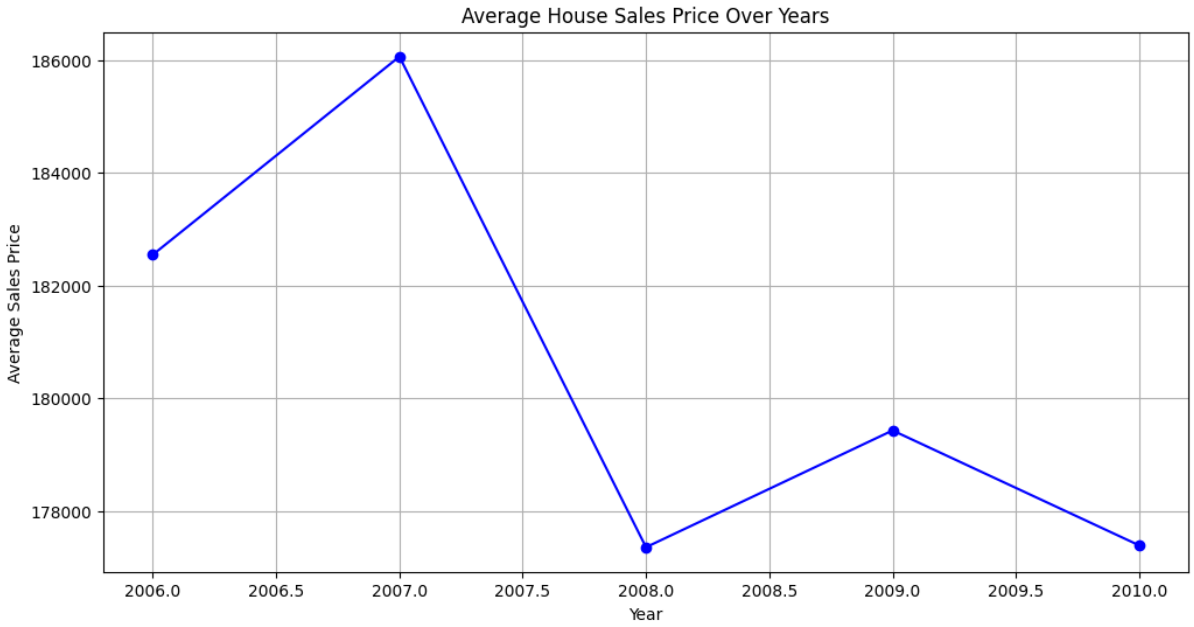
✓ 0.1s

```
<matplotlib.collections.PathCollection at 0x1fdd7da3190>
```



The GrLivArea is equal to the 2nd Flr SF in most cases.

I also did some time series analysis in this data.



We can see that the process have skyrocketed in 2007 compared to the other months

I also plotted boxplots for some important variables to find the boundaries and number of outliers. The code is available in:

Analysis_code

The Readme.md file is given below as link:

readme_file

The Output is given in the Readme file itself.

3.

Creative Idea 1:

I have an idea that could actually automate the process of music development. I want to fine tune an LLM model with music lyrics and their respective background music, the pitch frequency and the actual cuts and plays. So, my model must generate appropriate lyrics for the music theme generated. That is one of my creative ideas. That would revolutionalize the world of music development.

So, the basic dataset would contain databased on the audio tracks along with additional parameters such as emotion, storyline, prior and posterior track condition probabilities, etc which could help us generate appropriate lyrics for a background theme generated. But the dataset creation is very complex in this case as we need a track of the emotion of the track, the exact clip, etc which is not easy to derive. Based on my research in multimodal sections we have audio recognition pretrained models but there isn't one for the automatic development from music or themes. So, a new encoder/embedder is necessary in our case. So, such tasks really take time.

Creative idea 2:

This is using automated VR an AR technologies integrated with AI and spatial recognition as well situation analyser. One such implementation I started with was AMSA(Automatic Multimodal Situation Analyzer) similar to what Gemini had performed earlier to camera interactions. I want to make a software similar to E.D.I.T.H in Iron-Man where we have a model that analyses the surroundings and suggests appropriate measures. Also, we must be able to control our mobiles, Laptops, etc through it. The use of Laptops, mobiles all must be transferred to one simple device that changes the scenery in our eyes. This might sound similar to Apple vision pro but this is far higher. Apple vision pro is spoon-fed with existing data. But my glasses should optimize itself using the surroundings.

Creative Idea 3:

We all know that VR based Architectural modelling has been growing recently. I want to allow VR based processing that will allow a customer to have pluigin features in a VR and basically he can construct his own house virtually before asking suggestions. Of course, architectural designs are best usually.m Burt even after getting their designs of the house, imagine if we arew able to add some family touch to it automatically, then our AI should technically make each and every alteration possible. It would be great.

Although Ideas 1 and 3 are limited in their use case. But idea 2 can really make a change.