INTERN PROJECT REPORT

DS_INT_SANDESH / Surprise Housing Case Study

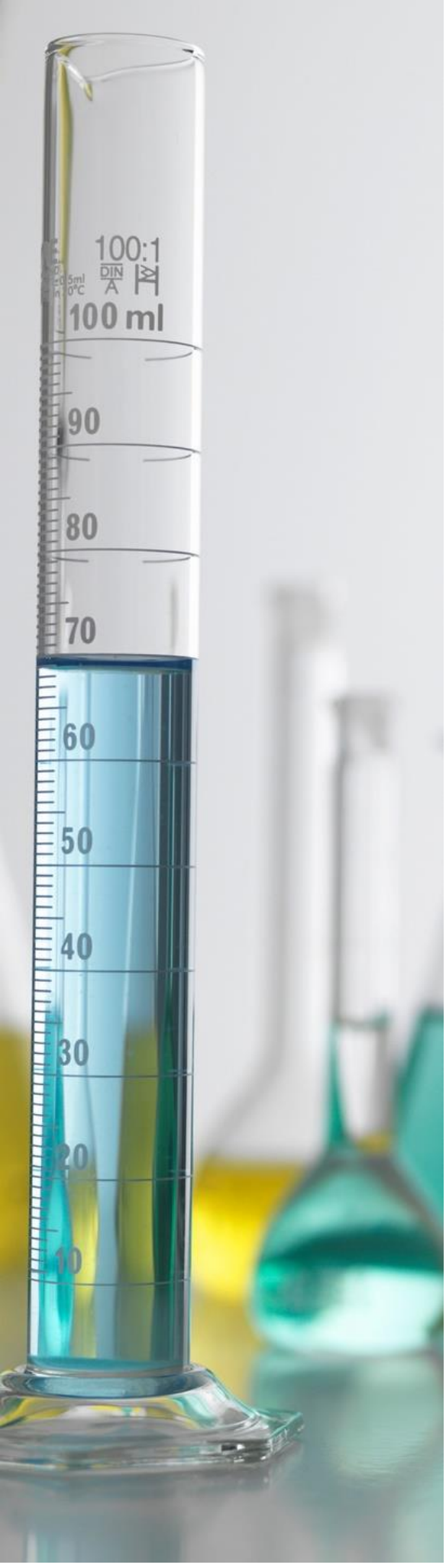
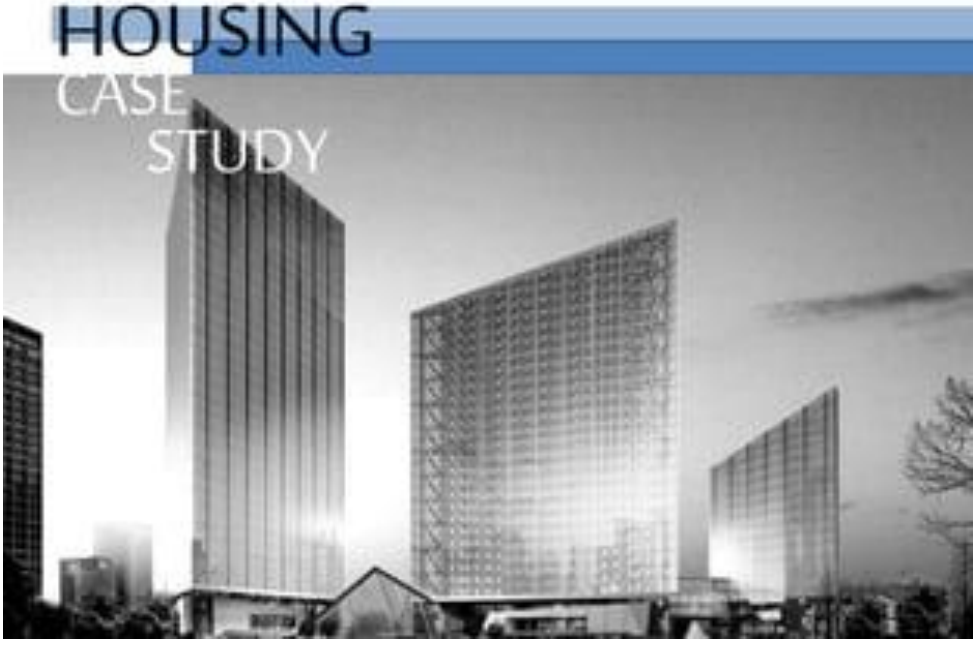


TABLE OF CONTENTS

ABSTRACT

The "Surprise Housing" case study examines a unique residential development project aimed at addressing the critical issue of affordable housing in rapidly growing urban areas. This case study delves into the innovative strategies employed in the design, funding, and implementation of the Surprise Housing initiative. By analyzing the project's integration of sustainable practices, community involvement, and public-private partnerships, the study provides insights into its success factors and challenges. Key elements include the adaptation of modular construction techniques, use of green technologies, and the establishment of a robust stakeholder engagement process. The findings highlight how such a comprehensive approach can serve as a model for similar initiatives, offering valuable lessons for policymakers, developers, and community leaders in their efforts to create sustainable and inclusive housing solutions. The case study also identifies areas for improvement and future research to enhance the effectiveness of affordable housing projects.

Housing Case Study



OBJECTIVE

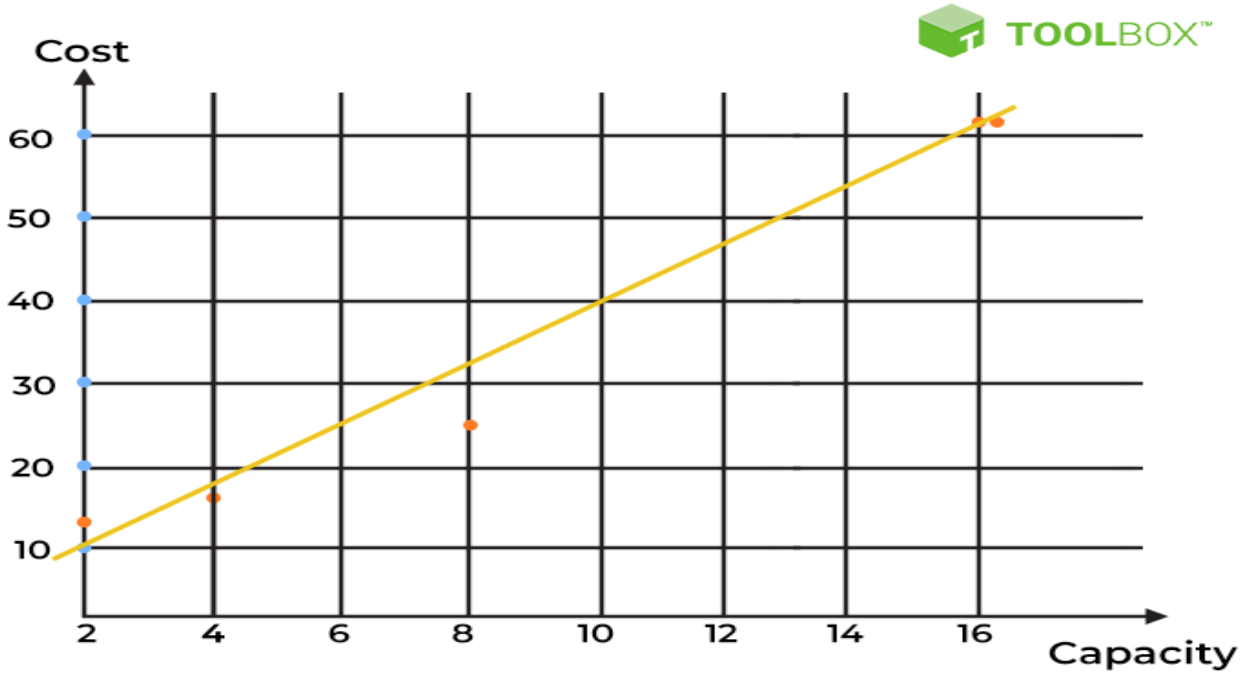
Objective

- **Determine the Influence of Key Variables:** Assess how different independent variables—such as construction costs, design features, and stakeholder engagement—affect dependent variables such as affordability, resident satisfaction, and sustainability.

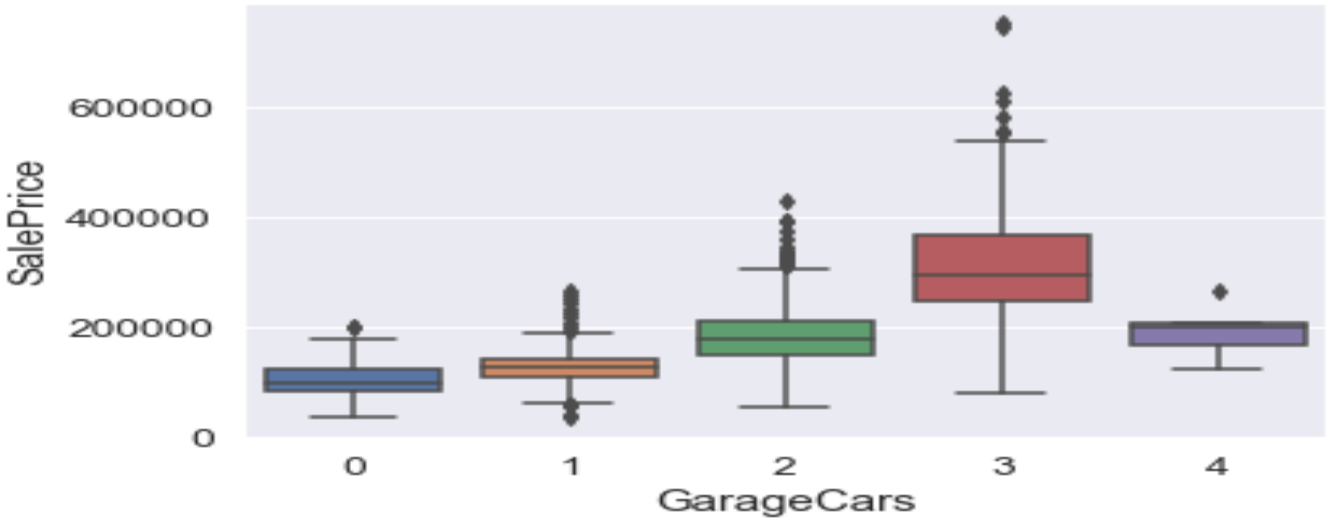
- **Identify Significant Predictors:** Identify which factors are statistically significant predictors of project success metrics, including affordability and resident satisfaction, and quantify their effects.

- **Analyze Relationships between Variables:** Explore the relationships between independent variables and project outcomes to understand how they interact and contribute to overall project effectiveness.

- **Evaluate Model Fit and Robustness:** Assess the goodness-of-fit of the regression models to ensure they accurately represent the relationships between variables and provide reliable results.



METHODOLOGY



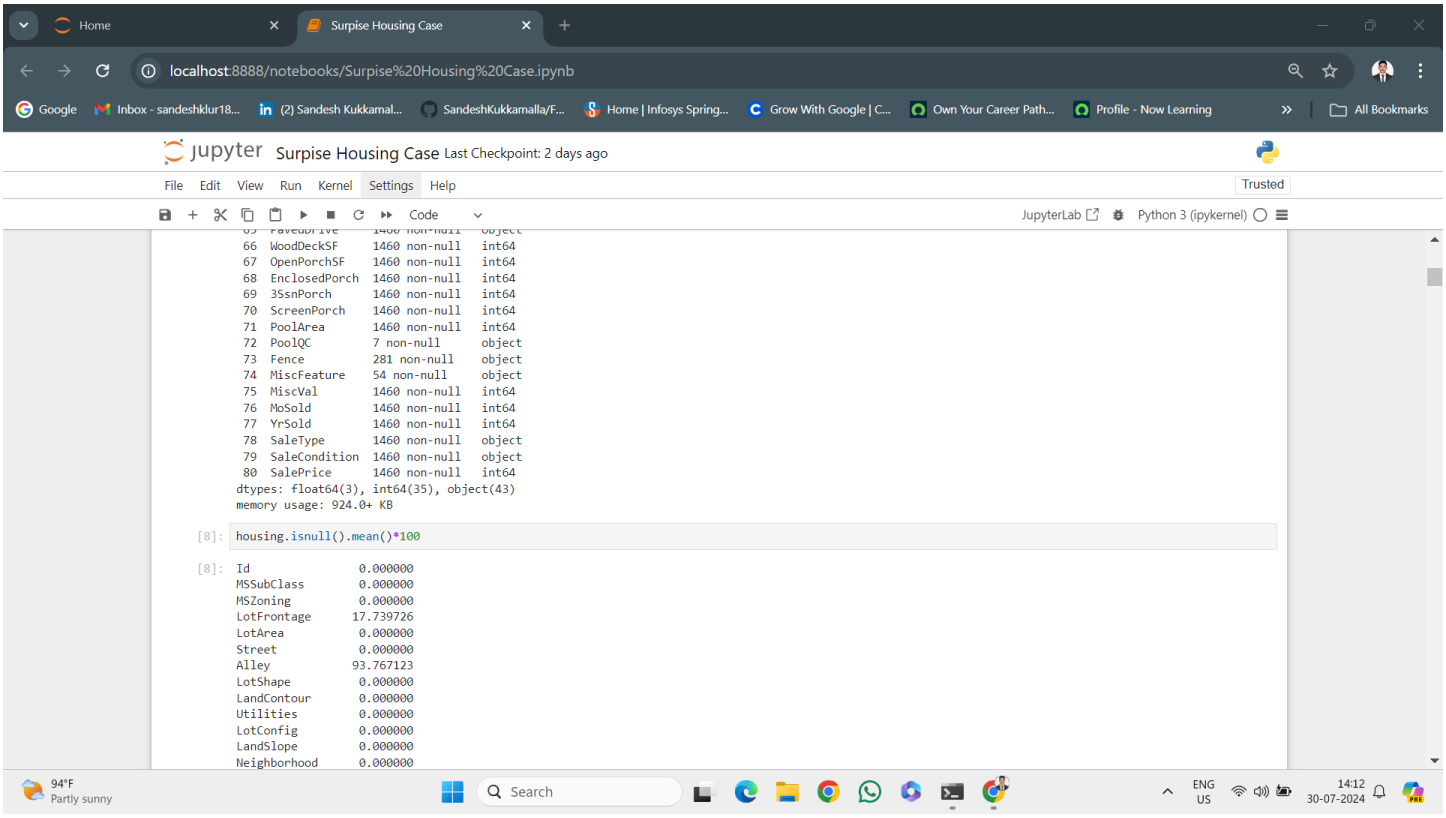
Data Preparation:

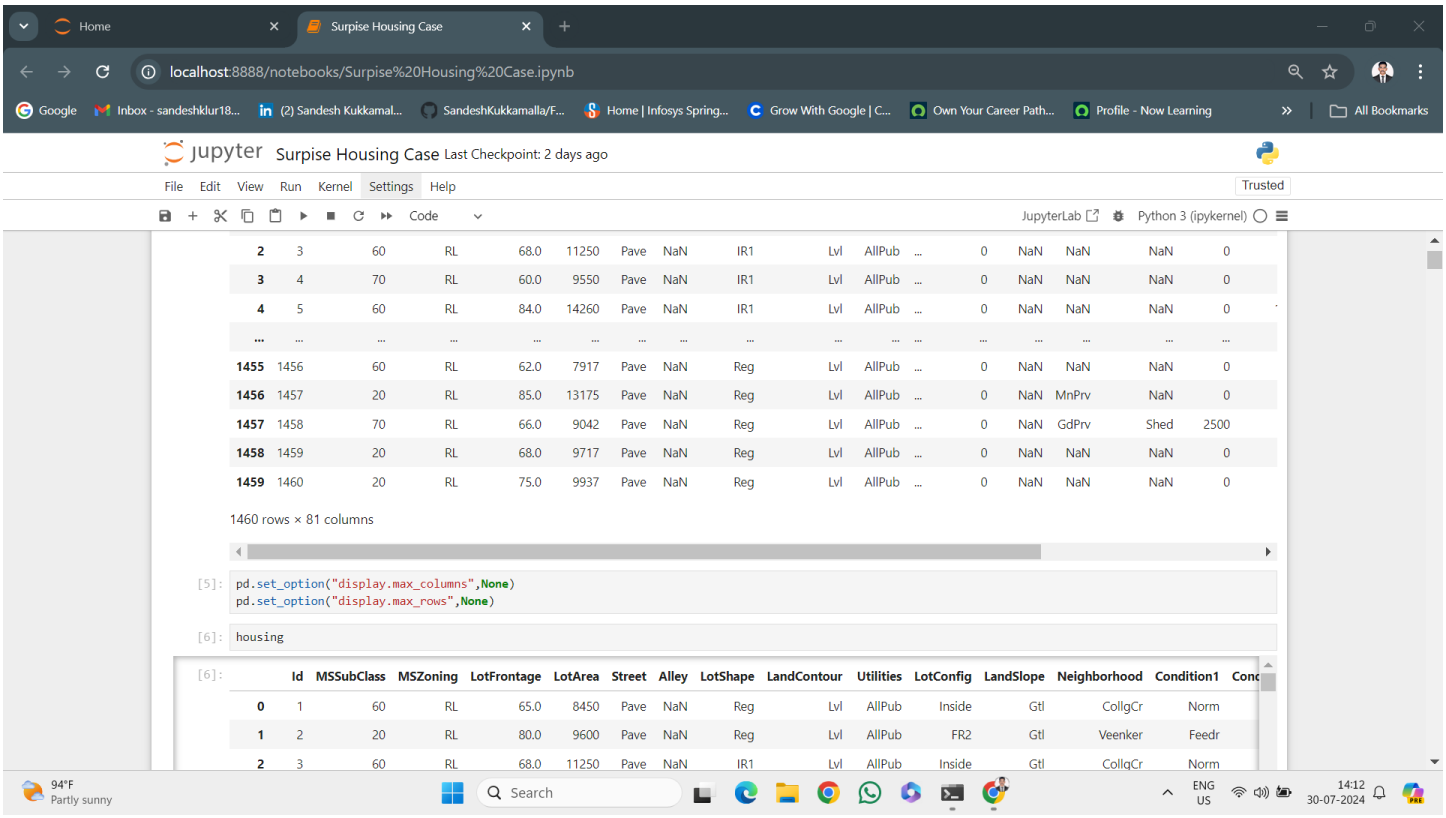
- **Data Cleaning:** Review and clean data to handle missing values, outliers, and inconsistencies.
- **Data Transformation:** Standardize variables and perform necessary transformations (e.g., log transformations for skewed data) to prepare for regression analysis.

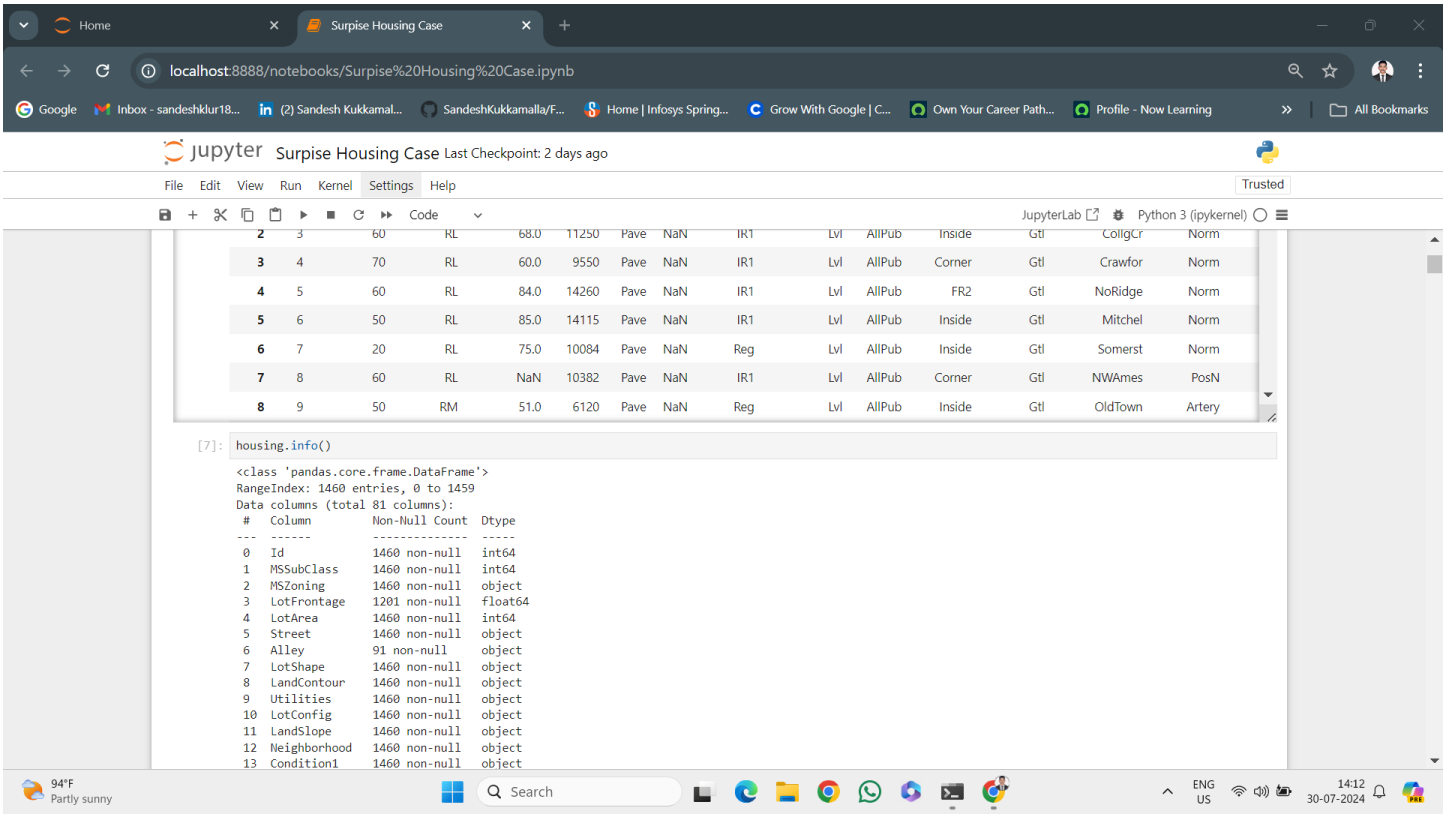
. Regression Analysis:

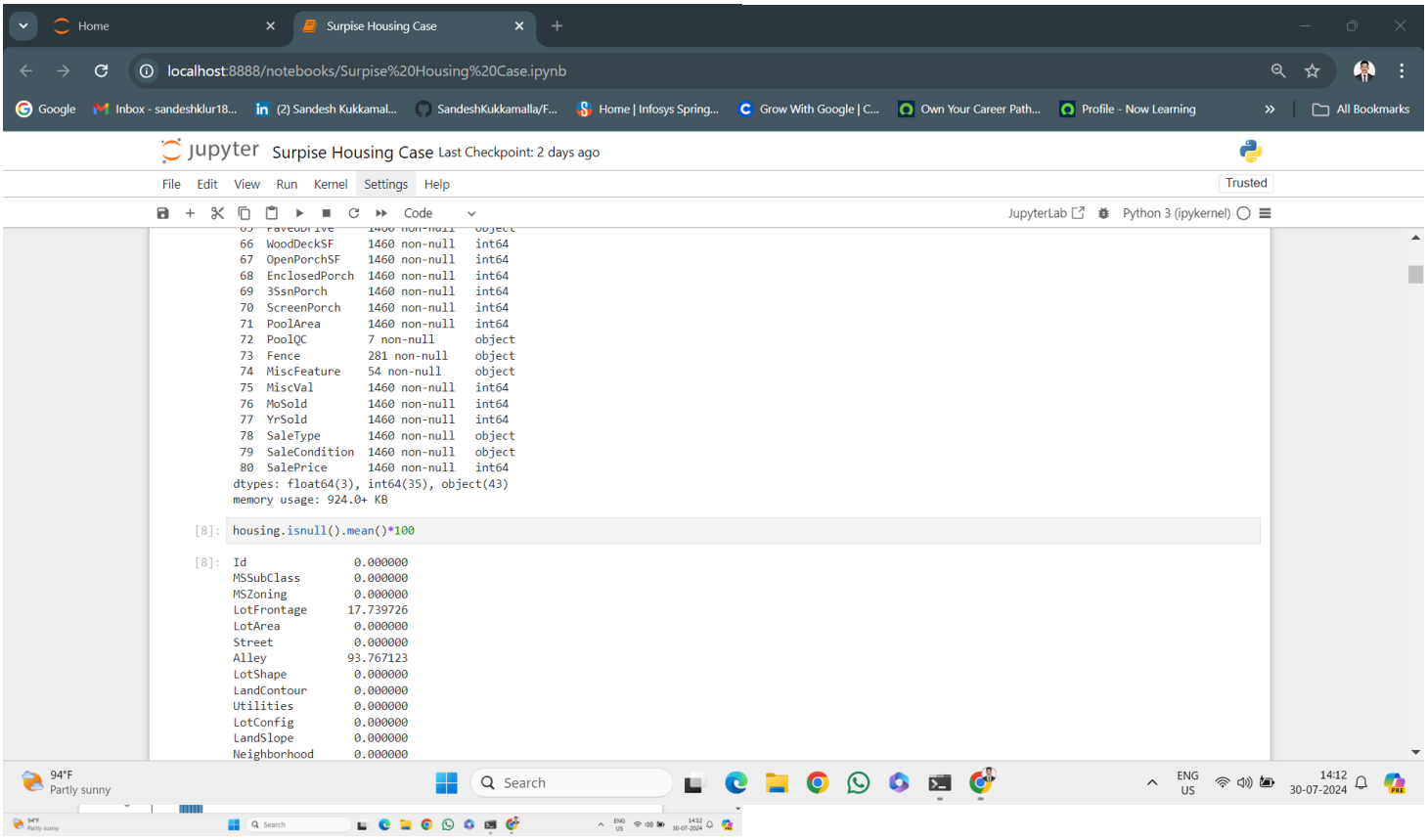
- **Model Selection:**
  - **Multiple Linear Regression:** Use multiple linear regression to analyze the relationship between independent variables (e.g., construction costs, design features) and dependent variables (e.g., affordability, resident satisfaction).
  - **Logistic Regression:** Apply logistic regression if analyzing categorical outcomes, such as the likelihood of meeting sustainability criteria.

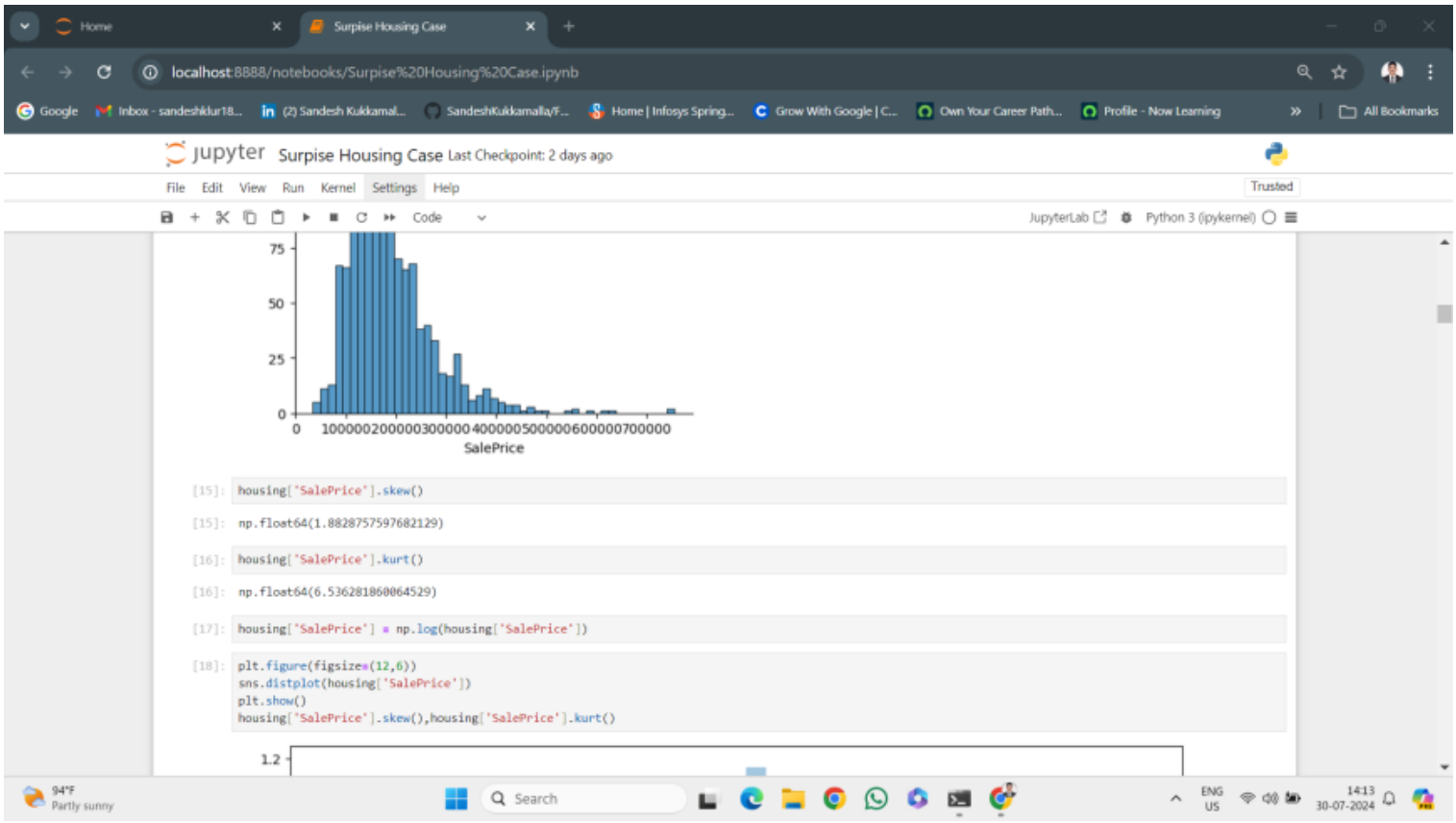
CODE

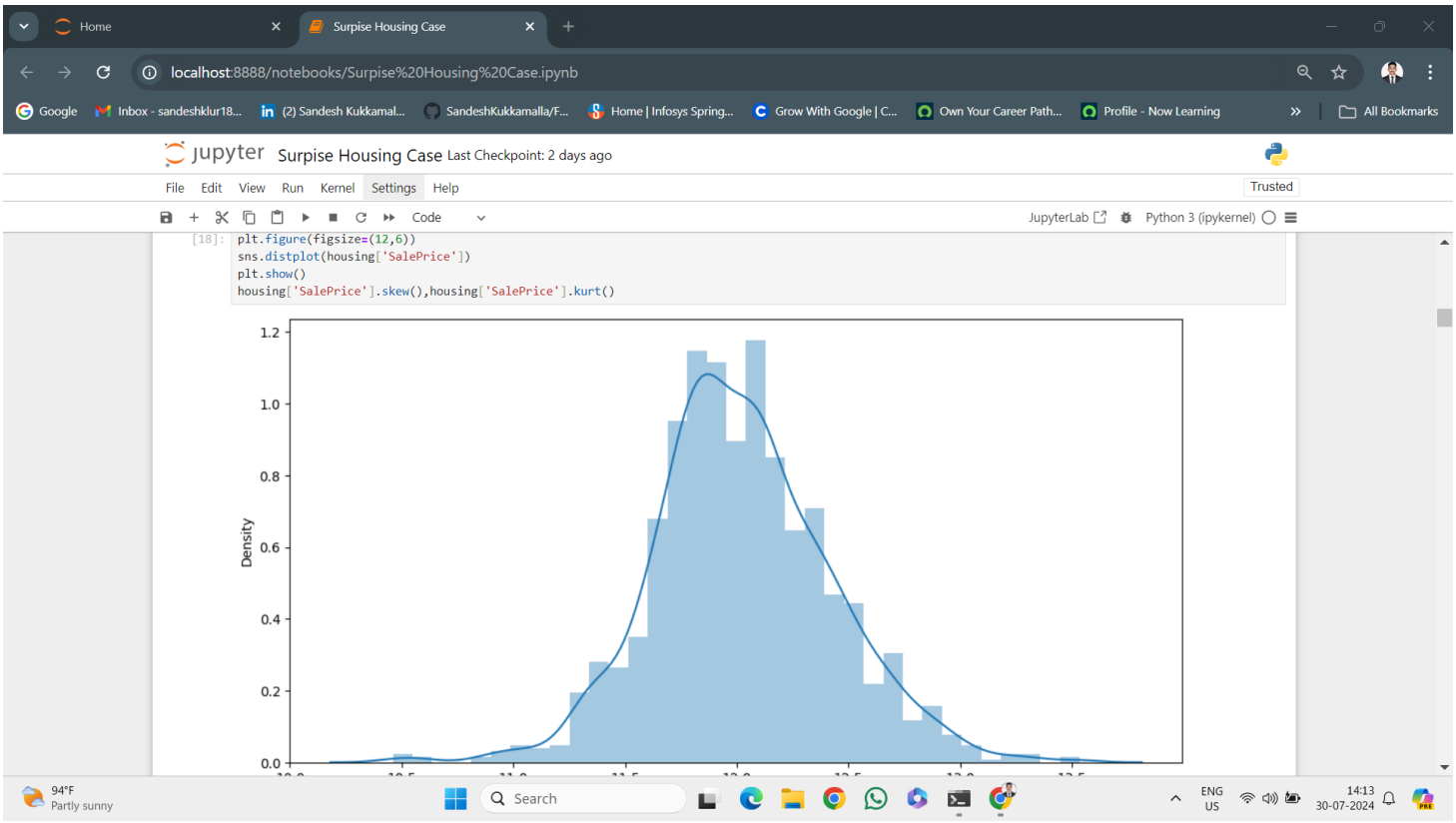

Home
Surpise Housing Case
localhost:8888/notebooks/Surpise%20Housing%20Case.ipynb
jupyter Surpise Housing Case Last Checkpoint: 2 days ago
File Edit View Run Kernel Settings Help
Trusted
Code
JupyterLab
Python 3 (ipykernel)
66 WoodDeckSF 1460 non-null int64
67 OpenPorchSF 1460 non-null int64
68 EnclosedPorch 1460 non-null int64
69 3SsnPorch 1460 non-null int64
70 ScreenPorch 1460 non-null int64
71 PoolArea 1460 non-null int64
72 PoolQC 7 non-null object
73 Fence 281 non-null object
74 MiscFeature 54 non-null object
75 MiscVal 1460 non-null int64
76 MoSold 1460 non-null int64
77 YrSold 1460 non-null int64
78 SaleType 1460 non-null object
79 SaleCondition 1460 non-null object
80 SalePrice 1460 non-null int64
dtypes: float64(3), int64(35), object(43)
memory usage: 924.0+ KB
[8]: housing.isnull().mean()*100
[8]: Id 0.000000
MSSubClass 0.000000
MSZoning 0.000000
LotFrontage 17.739726
LotArea 0.000000
Street 0.000000
Alley 93.767123
LotShape 0.000000
LandContour 0.000000
Utilities 0.000000
LotConfig 0.000000
LandSlope 0.000000
Neighborhood 0.000000
94°F
Partly sunny
Search
ENG
US
14:12
30-07-2024

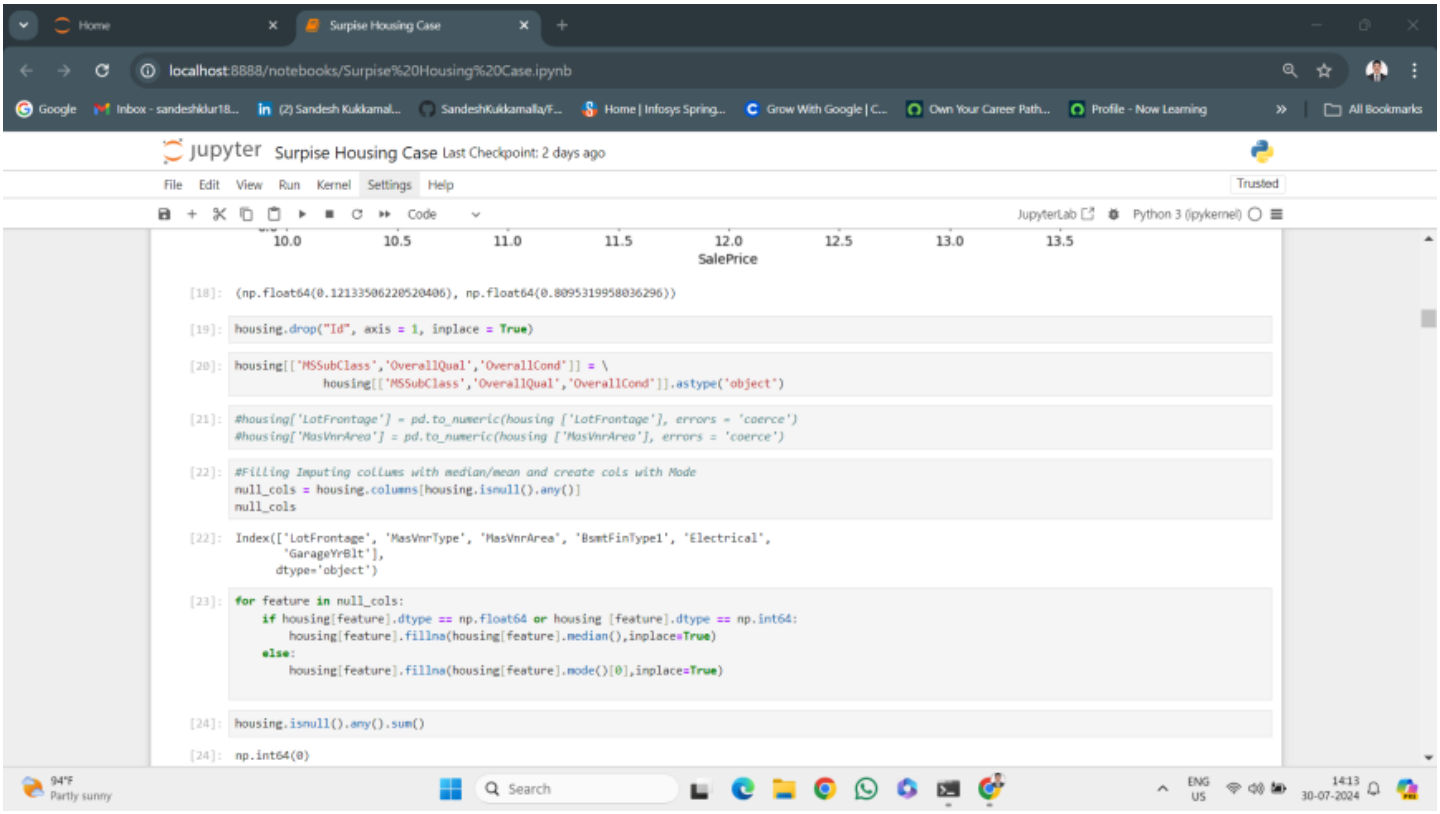

Home
Surpise Housing Case
localhost:8888/notebooks/Surpise%20Housing%20Case.ipynb
jupyter Surpise Housing Case Last Checkpoint: 2 days ago
File Edit View Run Kernel Settings Help
Trusted
Code
JupyterLab
Python 3 (ipykernel)
1460 rows × 81 columns
[5]: pd.set_option("display.max_columns",None)
pd.set_option("display.max_rows",None)
[6]: housing
[6]: Id MSSubClass MSZoning LotFrontage LotArea Street Alley LotShape LandContour Utilities LotConfig LandSlope Neighborhood Condition1
0 1 60 RL 65.0 8450 Pave NaN Reg Lvl AllPub Inside Gtl CollgCr Norm
1 2 20 RL 80.0 9600 Pave NaN Reg Lvl AllPub FR2 Gtl Veenker Feedr
2 3 60 RL 68.0 11250 Pave NaN IR1 Lvl AllPub Inside Gtl CollgCr Norm
94°F Partly sunny
Search
ENG US
14:12 30-07-2024


Home
Surpise Housing Case
localhost:8888/notebooks/Surpise%20Housing%20Case.ipynb
jupyter Surpise Housing Case Last Checkpoint: 2 days ago
File Edit View Run Kernel Settings Help
Trusted
Code
JupyterLab
Python 3 (ipykernel)
2 3 60 RL 68.0 11250 Pave NaN IR1 Lvl AllPub Inside Gtl CollgCr Norm
3 4 70 RL 60.0 9550 Pave NaN IR1 Lvl AllPub Corner Gtl Crawfor Norm
4 5 60 RL 84.0 14260 Pave NaN IR1 Lvl AllPub FR2 Gtl NoRidge Norm
5 6 50 RL 85.0 14115 Pave NaN IR1 Lvl AllPub Inside Gtl Mitchel Norm
6 7 20 RL 75.0 10084 Pave NaN Reg Lvl AllPub Inside Gtl Somerst Norm
7 8 60 RL NaN 10382 Pave NaN IR1 Lvl AllPub Corner Gtl NWAmes PosN
8 9 50 RM 51.0 6120 Pave NaN Reg Lvl AllPub Inside Gtl OldTown Artery
[7]: housing.info()
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 1460 entries, 0 to 1459
Data columns (total 81 columns):
Column Non-Null Count Dtype
0 Id 1460 non-null int64
1 MSSubClass 1460 non-null int64
2 MSZoning 1460 non-null object
3 LotFrontage 1201 non-null float64
4 LotArea 1460 non-null int64
5 Street 1460 non-null object
6 Alley 91 non-null object
7 LotShape 1460 non-null object
8 LandContour 1460 non-null object
9 Utilities 1460 non-null object
10 LotConfig 1460 non-null object
11 LandSlope 1460 non-null object
12 Neighborhood 1460 non-null object
13 Condition1 1460 non-null object
94°F Partly sunny
Search
ENG US
14:12 30-07-2024


Home
Surpise Housing Case
localhost:8888/notebooks/Surpise%20Housing%20Case.ipynb
jupyter Surpise Housing Case Last Checkpoint: 2 days ago
File Edit View Run Kernel Settings Help
Trusted
Code
JupyterLab
Python 3 (ipykernel)
66 WoodDeckSF 1460 non-null int64
67 OpenPorchSF 1460 non-null int64
68 EnclosedPorch 1460 non-null int64
69 3SsnPorch 1460 non-null int64
70 ScreenPorch 1460 non-null int64
71 PoolArea 1460 non-null int64
72 PoolQC 7 non-null object
73 Fence 281 non-null object
74 MiscFeature 54 non-null object
75 MiscVal 1460 non-null int64
76 MoSold 1460 non-null int64
77 YrSold 1460 non-null int64
78 SaleType 1460 non-null object
79 SaleCondition 1460 non-null object
80 SalePrice 1460 non-null int64
dtypes: float64(3), int64(35), object(43)
memory usage: 924.0+ KB
[8]: housing.isnull().mean()*100
[8]: Id 0.000000
MSSubClass 0.000000
MSZoning 0.000000
LotFrontage 17.739726
LotArea 0.000000
Street 0.000000
Alley 93.767123
LotShape 0.000000
LandContour 0.000000
Utilities 0.000000
LotConfig 0.000000
LandSlope 0.000000
Neighborhood 0.000000
94°F
Partly sunny
Search
ENG
US
14:12
30-07-2024


Home
Surpise Housing Case
localhost:8888/notebooks/Surpise%20Housing%20Case.ipynb
jupyter Surpise Housing Case Last Checkpoint: 2 days ago
File Edit View Run Kernel Settings Help
Trusted
Code
JupyterLab
Python 3 (ipykernel)
75
50
25
0
0 100000 200000 300000 400000 500000 600000 700000
SalePrice
[15]: housing['SalePrice'].skew()
[15]: np.float64(1.8828757597682129)
[16]: housing['SalePrice'].kurt()
[16]: np.float64(6.536281860064529)
[17]: housing['SalePrice'] = np.log(housing['SalePrice'])
[18]: plt.figure(figsize=(12,6))
sns.distplot(housing['SalePrice'])
plt.show()
housing['SalePrice'].skew(),housing['SalePrice'].kurt()
1.2
94°F
Partly sunny
Search
ENG
US
14:13
30-07-2024


Home
Surpise Housing Case
localhost:8888/notebooks/Surpise%20Housing%20Case.ipynb
jupyter Surpise Housing Case Last Checkpoint: 2 days ago
File Edit View Run Kernel Settings Help
Trusted
Code
JupyterLab
Python 3 (ipykernel)
[18]:
plt.figure(figsize=(12,6))
sns.distplot(housing['SalePrice'])
plt.show()
housing['SalePrice'].skew(),housing['SalePrice'].kurt()
Density
1.2
1.0
0.8
0.6
0.4
0.2
0.0
94°F
Partly sunny
Search
ENG
US
14:13
30-07-2024


Home
Surpise Housing Case
localhost:8888/notebooks/Surpise%20Housing%20Case.ipynb
jupyter Surpise Housing Case Last Checkpoint: 2 days ago
File Edit View Run Kernel Settings Help
Trusted
Code
JupyterLab
Python 3 (ipykernel)
10.0
10.5
11.0
11.5
12.0
12.5
13.0
13.5
SalePrice
[18]: (np.float64(0.12133506220520406), np.float64(0.8095319958036296))
[19]: housing.drop("Id", axis = 1, inplace = True)
[20]: housing[['MSSubClass','OverallQual','OverallCond']] = \
housing[['MSSubClass','OverallQual','OverallCond']].astype('object')
[21]: #housing['LotFrontage'] = pd.to_numeric(housing ['LotFrontage'], errors = 'coerce')
#housing['MasVnrArea'] = pd.to_numeric(housing ['MasVnrArea'], errors = 'coerce')
[22]: #Filling Imputing collums with median/mean and create cols with Mode
null_cols = housing.columns[housing.isnull().any()]
null_cols
[22]: Index(['LotFrontage', 'MasVnrType', 'MasVnrArea', 'BsmtFinType1', 'Electrical',
'GarageYrBlt'],
dtype='object')
[23]: for feature in null_cols:
if housing[feature].dtype == np.float64 or housing [feature].dtype == np.int64:
housing[feature].fillna(housing[feature].median(),inplace=True)
else:
housing[feature].fillna(housing[feature].mode()[0],inplace=True)
[24]: housing.isnull().any().sum()
[24]: np.int64(0)
94°F
Partly sunny
Search
ENG
US
14:13
30-07-2024

Home | Surpise Housing Case

localhost:8888/notebooks/Surpise%20Housing%20Case.ipynb

jupyter Surpise Housing Case Last Checkpoint: 2 days ago

File Edit View Run Kernel Settings Help — Trusted

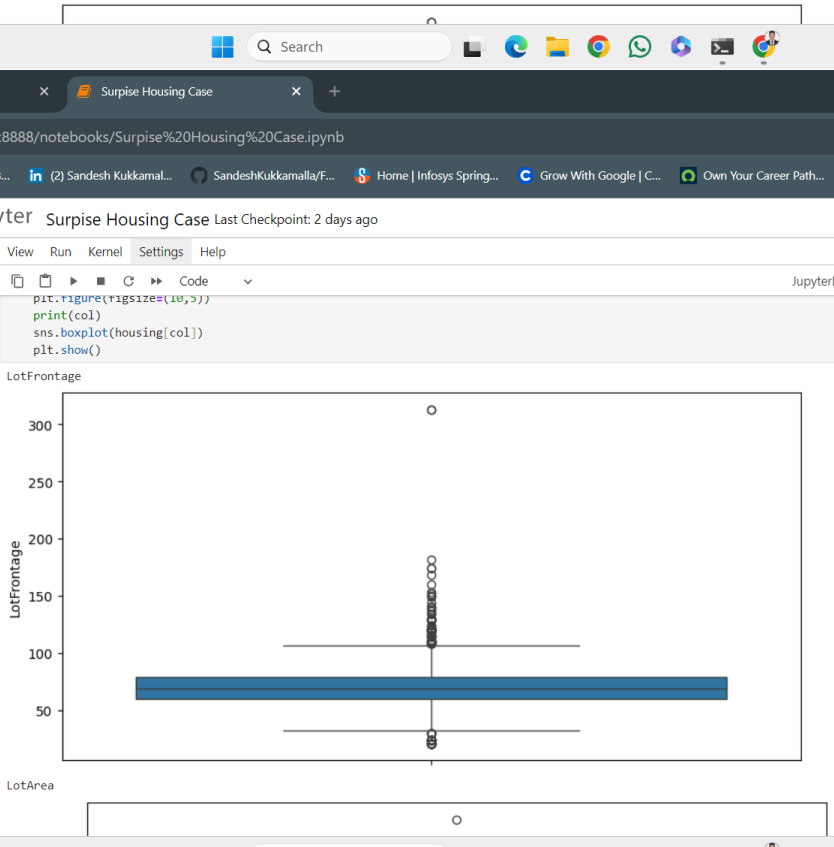
Code — JupyterLab — Python 3 (ipykernel)

[29]:
```
cat_cols = housing.select_dtypes(include = 'object').columns
num_cols = housing.select_dtypes(include = ['int64','float64']).columns
num_cols,cat_cols
```

[29]:
```
(Index(['LotFrontage', 'LotArea', 'YearBuilt', 'YearRemodAdd', 'MasVnrArea',
        'BsmtFinSF1', 'BsmtFinSF2', 'BsmtUnfSF', 'TotalBsmtSF', '1stFlrSF',
        '2ndFlrSF', 'LowQualFinSF', 'GrLivArea', 'BsmtFullBath', 'BsmtHalfBath',
        'FullBath', 'HalfBath', 'BedroomAbvGr', 'KitchenAbvGr', 'TotRmsAbvGrd',
        'Fireplaces', 'GarageYrBlt', 'GarageCars', 'GarageArea', 'WoodDeckSF',
        'OpenPorchSF', 'EnclosedPorch', '3SsnPorch', 'ScreenPorch', 'PoolArea',
        'MiscVal', 'MoSold', 'YrSold', 'SalePrice'],
       dtype='object'),
 Index(['MSSubClass', 'MSZoning', 'Street', 'Alley', 'LotShape', 'LandContour',
        'Utilities', 'LotConfig', 'LandSlope', 'Neighborhood', 'Condition1',
        'Condition2', 'BldgType', 'HouseStyle', 'OverallQual', 'OverallCond',
        'RoofStyle', 'RoofMatl', 'Exterior1st', 'Exterior2nd', 'MasVnrType',
        'ExterQual', 'ExterCond', 'Foundation', 'BsmtQual', 'BsmtCond',
        'BsmtExposure', 'BsmtFinType1', 'BsmtFinType2', 'Heating', 'HeatingQC',
        'CentralAir', 'Electrical', 'KitchenQual', 'Functional', 'FireplaceQu',
        'GarageType', 'GarageFinish', 'GarageQual', 'GarageCond', 'PavedDrive',
        'PoolQC', 'Fence', 'MiscFeature', 'SaleType', 'SaleCondition'],
       dtype='object'))
```

[38]:
```
for col in num_cols:
    plt.figure(figsize=(10,5))
    print(col)
    sns.boxplot(housing[col])
    plt.show()
```

LotFrontage



LotArea

Olympic Games Today's events — Search — ENG US — 14:13 30-07-2024

Home | Surpise Housing Case

localhost:8888/notebooks/Surpise%20Housing%20Case.ipynb

jupyter Surpise Housing Case Last Checkpoint: 2 days ago

File Edit View Run Kernel Settings Help — Trusted

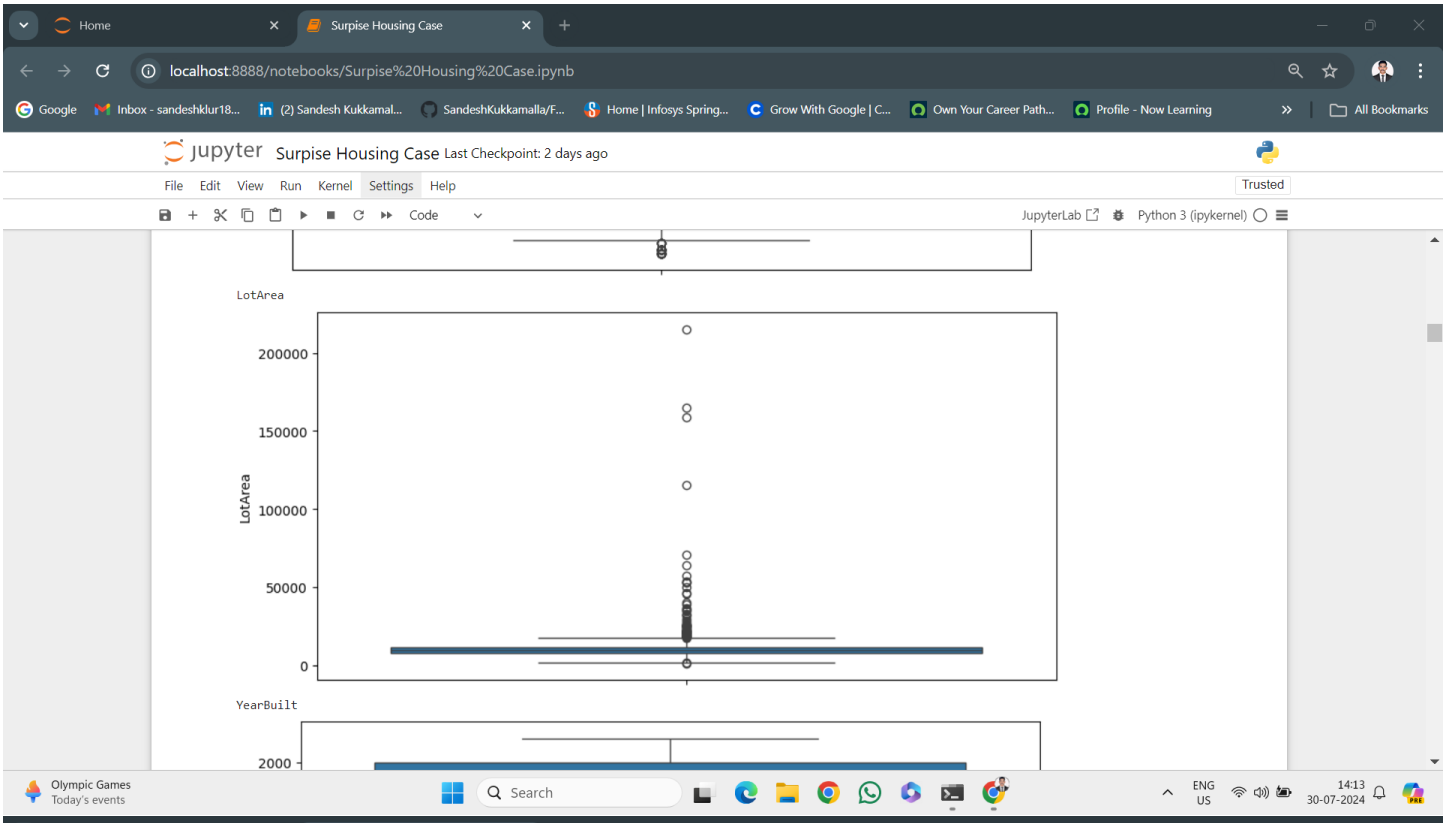
Code — JupyterLab — Python 3 (ipykernel)

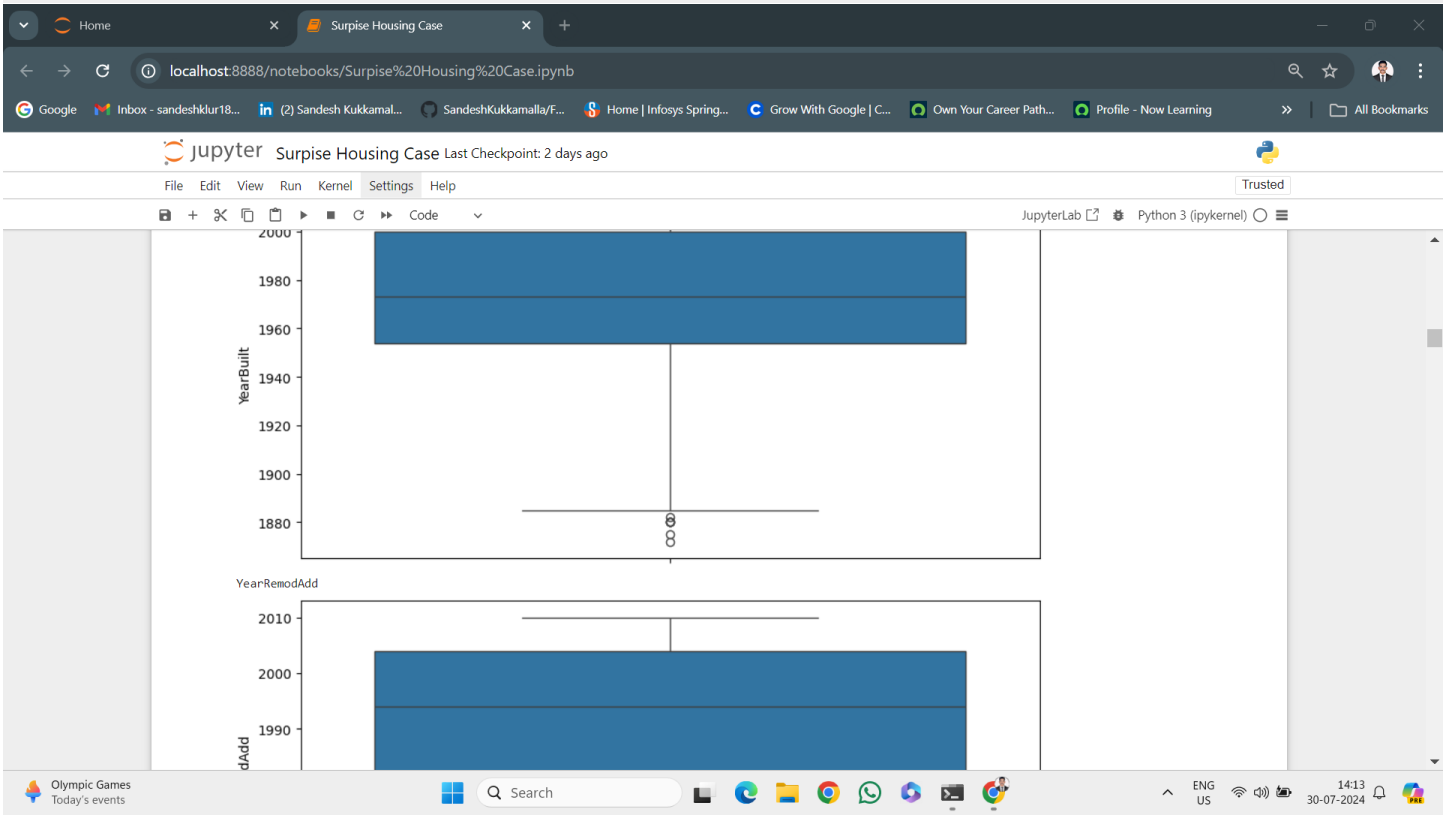
```
    plt.figure(figsize=(10,5))
    print(col)
    sns.boxplot(housing[col])
    plt.show()
```

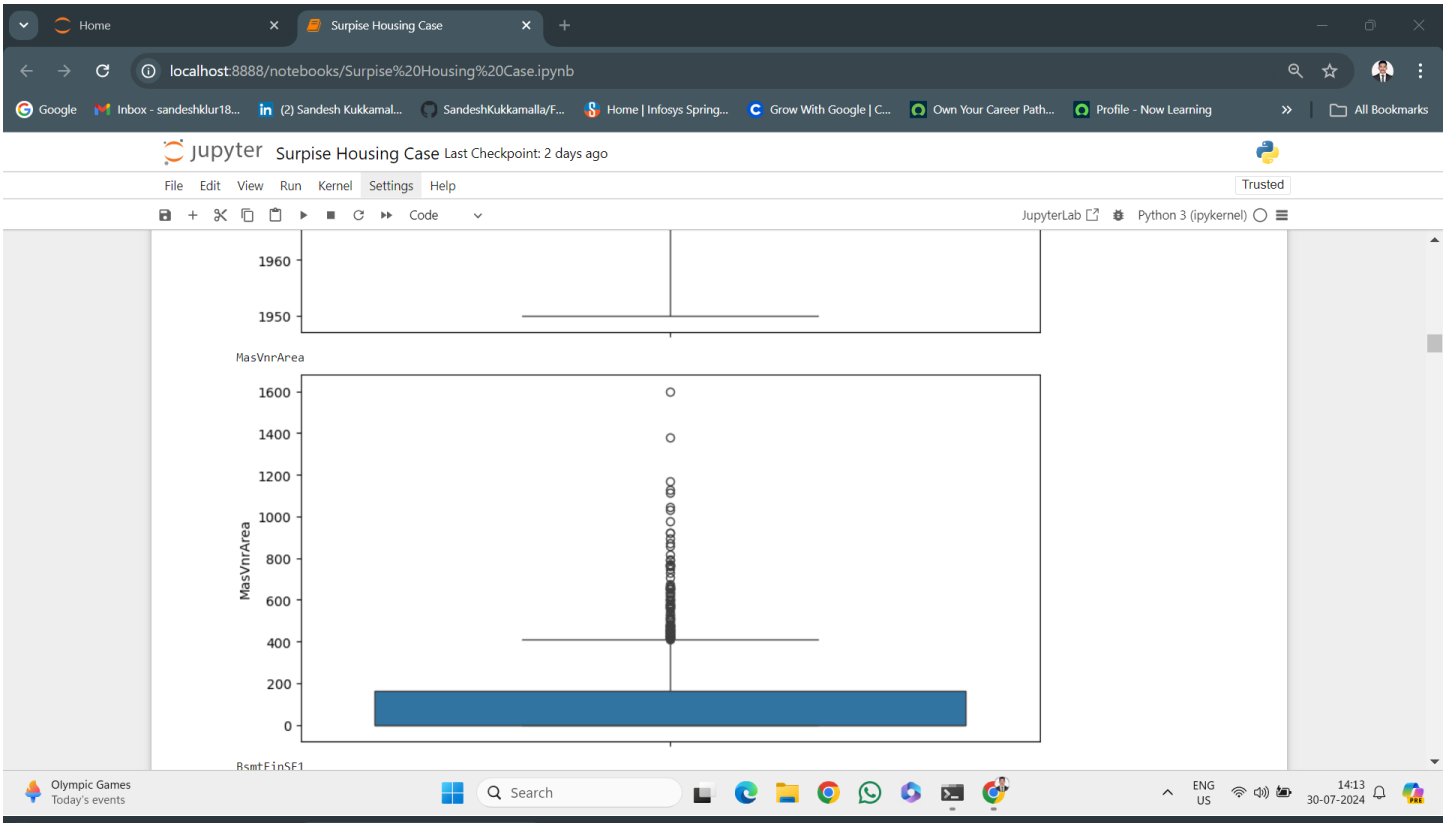
LotFrontage

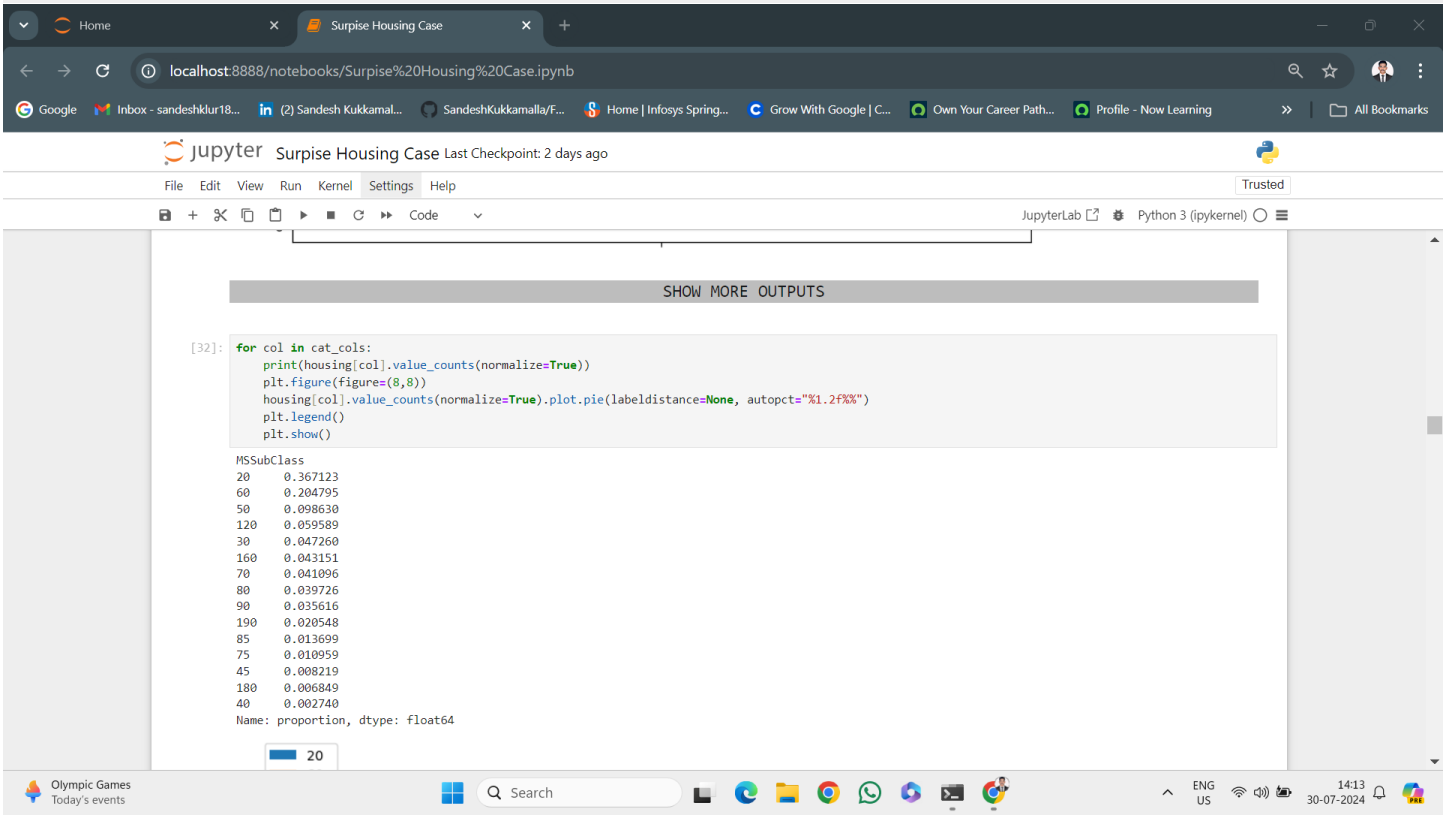
LotArea

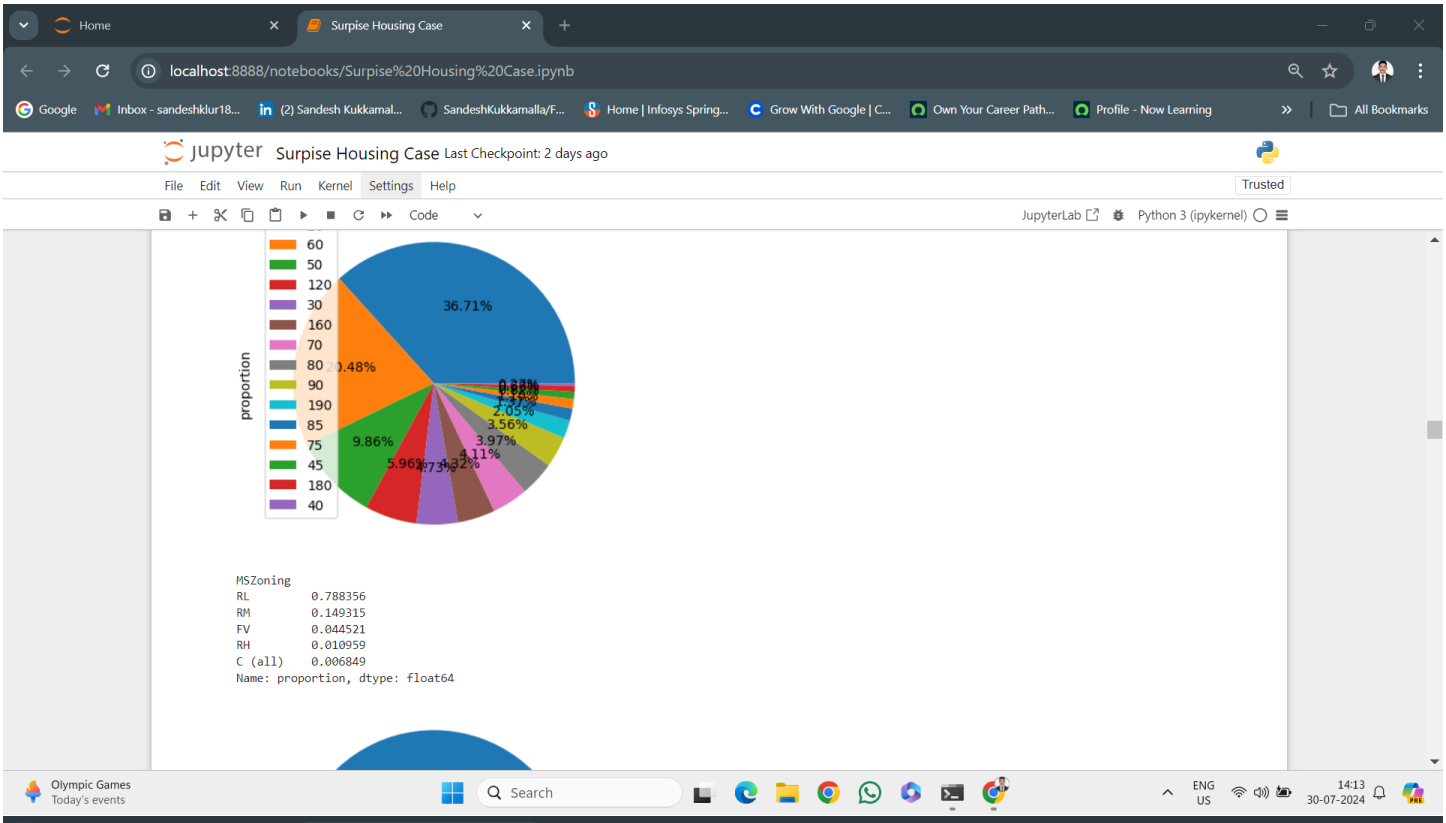
Olympic Games Today's events — Search — ENG US — 14:13 30-07-2024

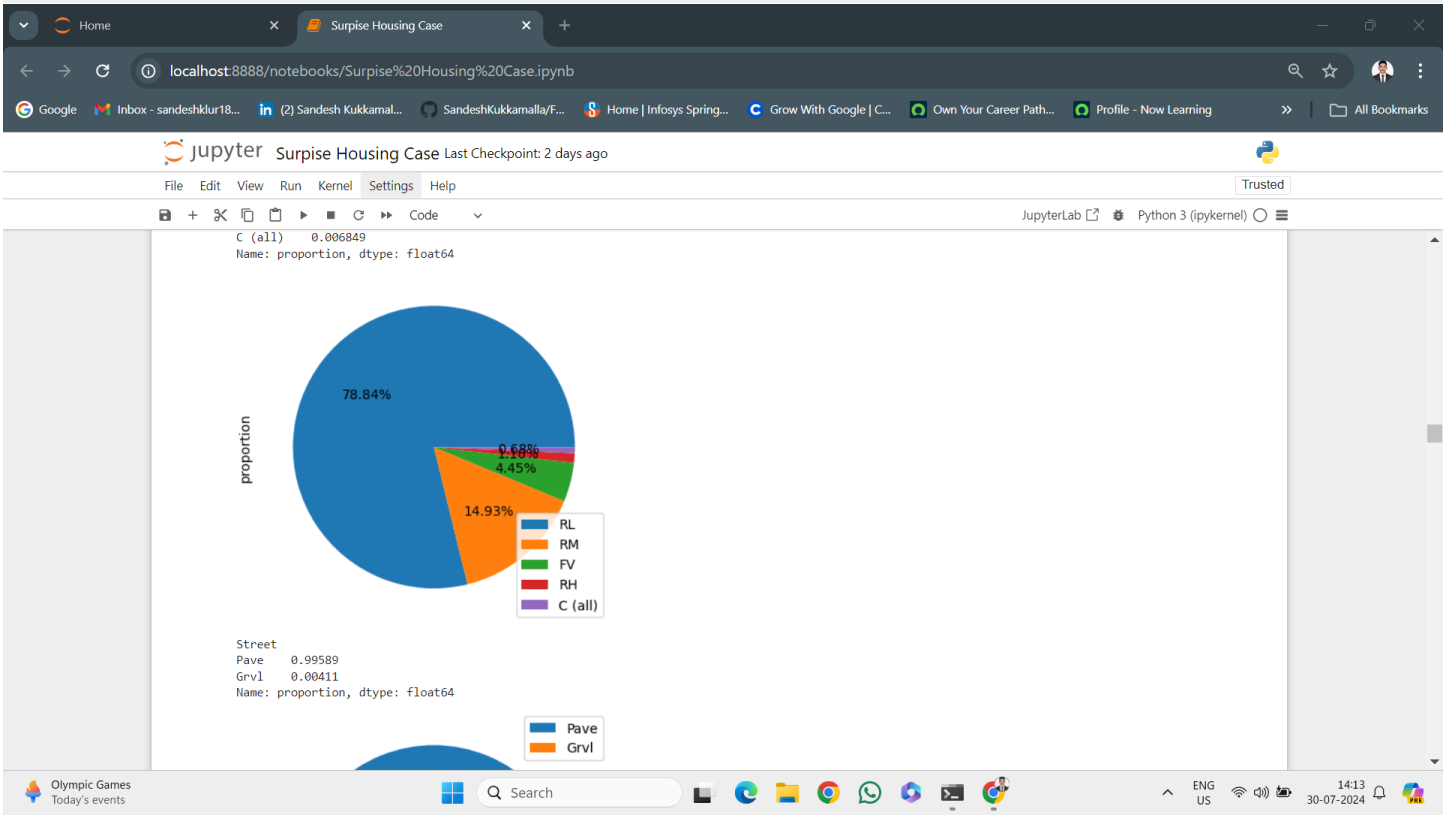

Home
Surpise Housing Case
localhost:8888/notebooks/Surpise%20Housing%20Case.ipynb
jupyter Surpise Housing Case Last Checkpoint: 2 days ago
File Edit View Run Kernel Settings Help
Trusted
Code
JupyterLab
Python 3 (ipykernel)
LotArea
200000
150000
100000
50000
0
YearBuilt
2000

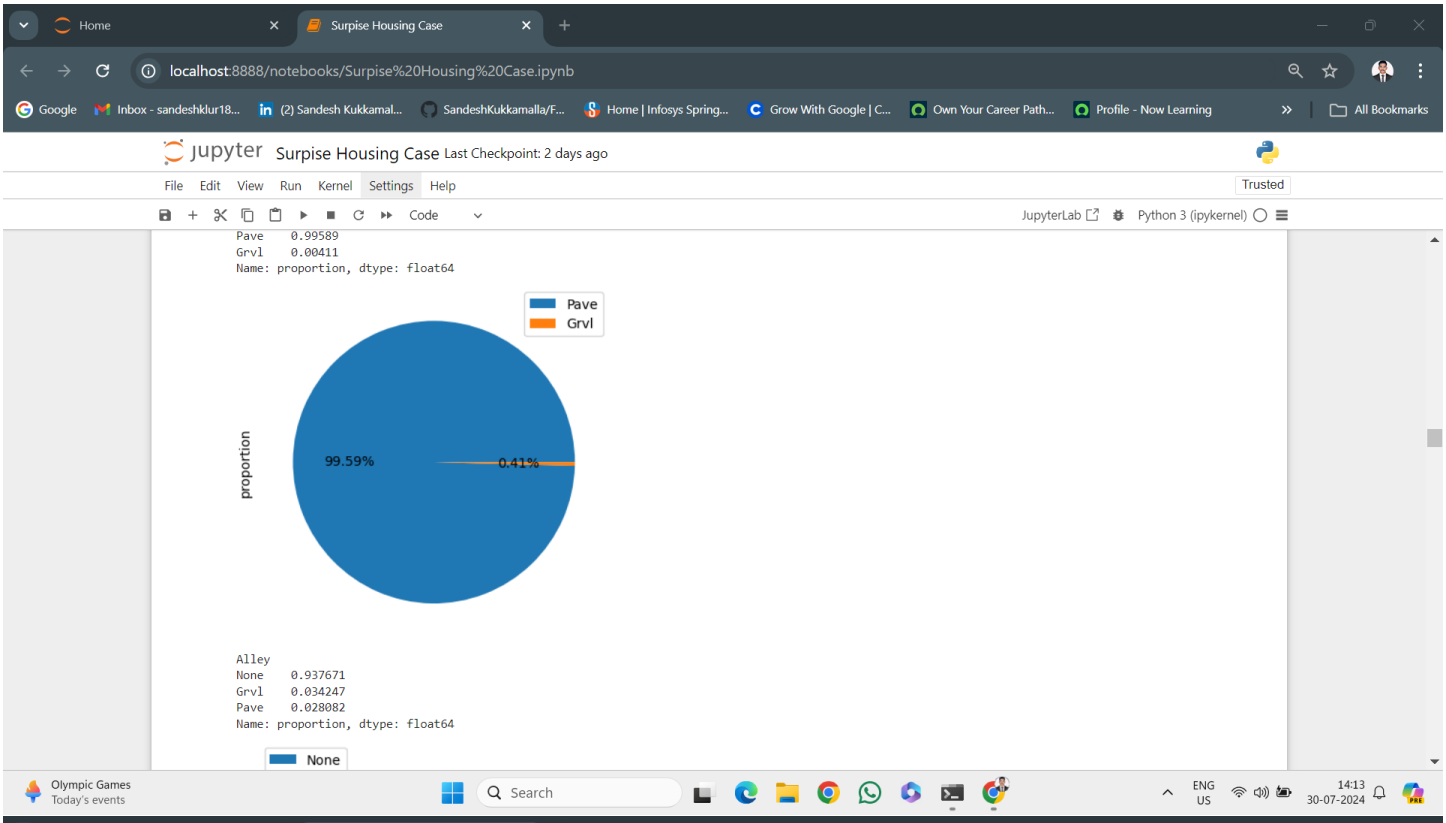

Home
Surpise Housing Case
localhost:8888/notebooks/Surpise%20Housing%20Case.ipynb
jupyter Surpise Housing Case Last Checkpoint: 2 days ago
File Edit View Run Kernel Settings Help
Trusted
Code
JupyterLab
Python 3 (ipykernel)
2000
1980
1960
1940
1920
1900
1880
YearBuilt
YearRemodAdd
2010
2000
1990

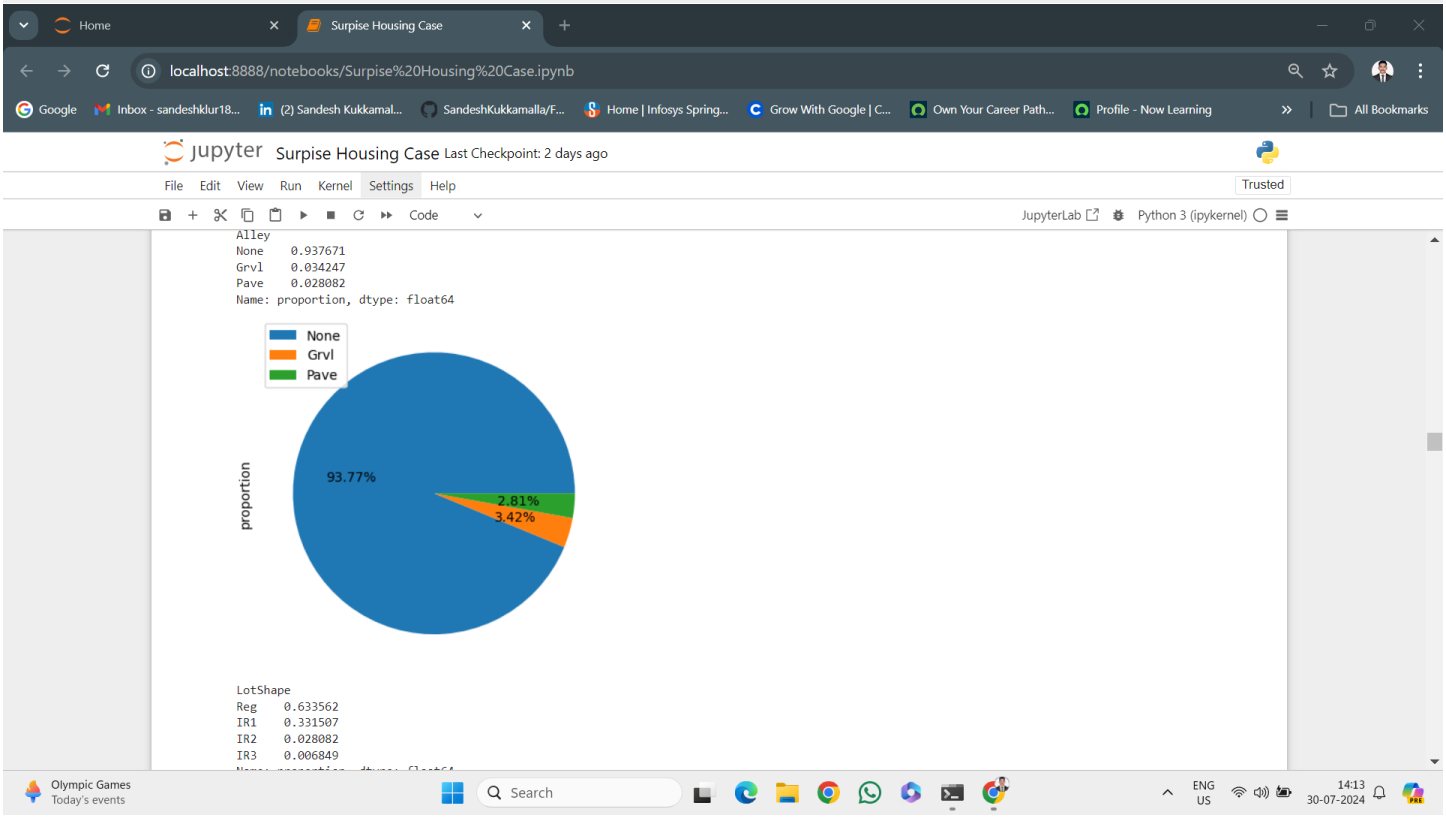

Home
Surpise Housing Case
localhost:8888/notebooks/Surpise%20Housing%20Case.ipynb
jupyter Surpise Housing Case Last Checkpoint: 2 days ago
File Edit View Run Kernel Settings Help
Trusted
Code
JupyterLab
Python 3 (ipykernel)
1960
1950
MasVnrArea
1600
1400
1200
1000
800
600
400
200
0
MasVnrArea
BsmtFinSF1
Search
ENG US
14:13 30-07-2024

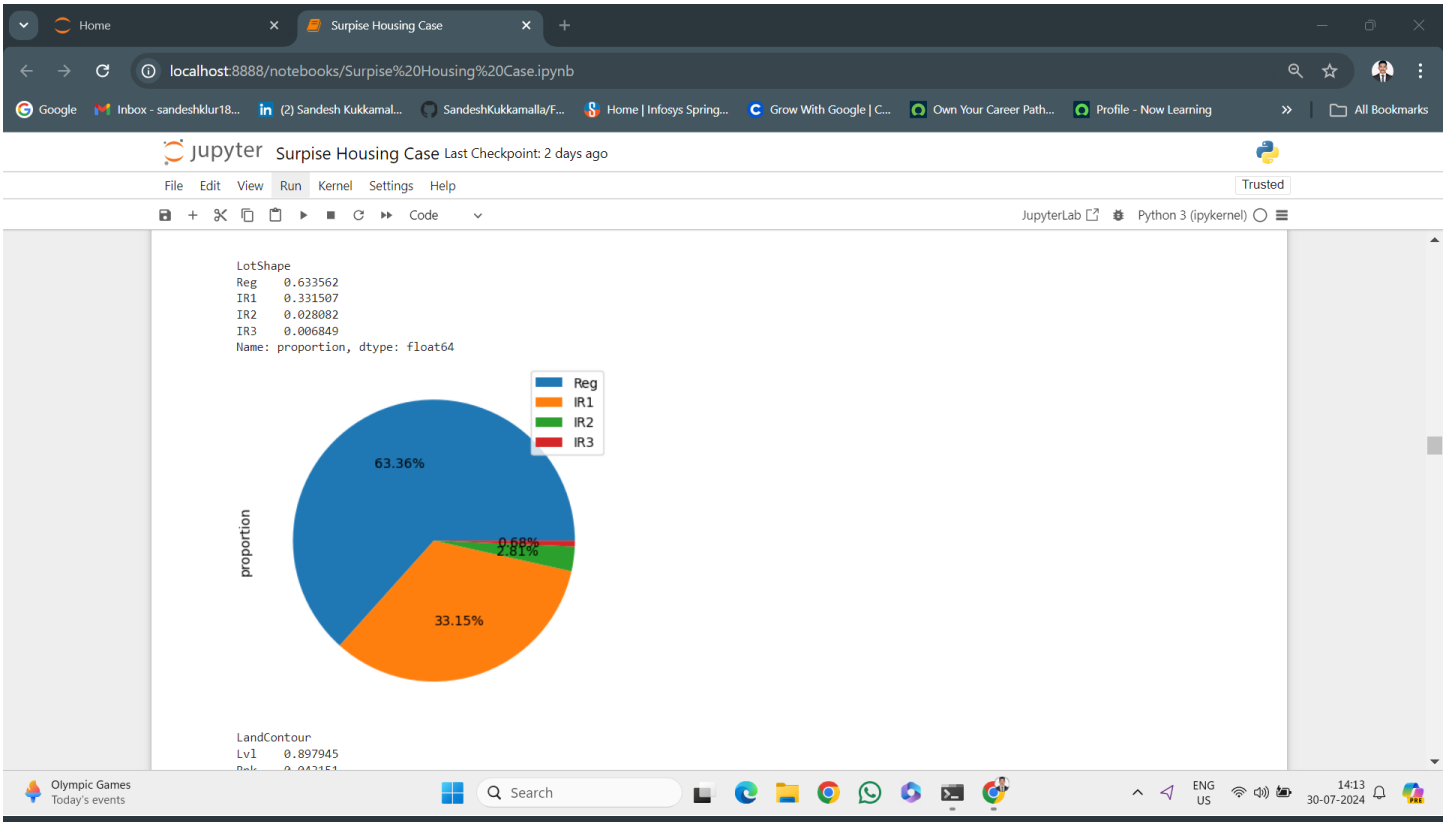

Home
Surpise Housing Case
localhost:8888/notebooks/Surpise%20Housing%20Case.ipynb
jupyter Surpise Housing Case Last Checkpoint: 2 days ago
File Edit View Run Kernel Settings Help
Trusted
Code
JupyterLab
Python 3 (ipykernel)
SHOW MORE OUTPUTS
[32]:
for col in cat_cols:
print(housing[col].value_counts(normalize=True))
plt.figure(figure=(8,8))
housing[col].value_counts(normalize=True).plot.pie(labeldistance=None, autopct="%1.2f%%")
plt.legend()
plt.show()
MSSubClass
20 0.367123
60 0.204795
50 0.098630
120 0.059589
30 0.047260
160 0.043151
70 0.041096
80 0.039726
90 0.035616
190 0.020548
85 0.013699
75 0.010959
45 0.008219
180 0.006849
40 0.002740
Name: proportion, dtype: float64
20
Search
ENG US
14:13 30-07-2024

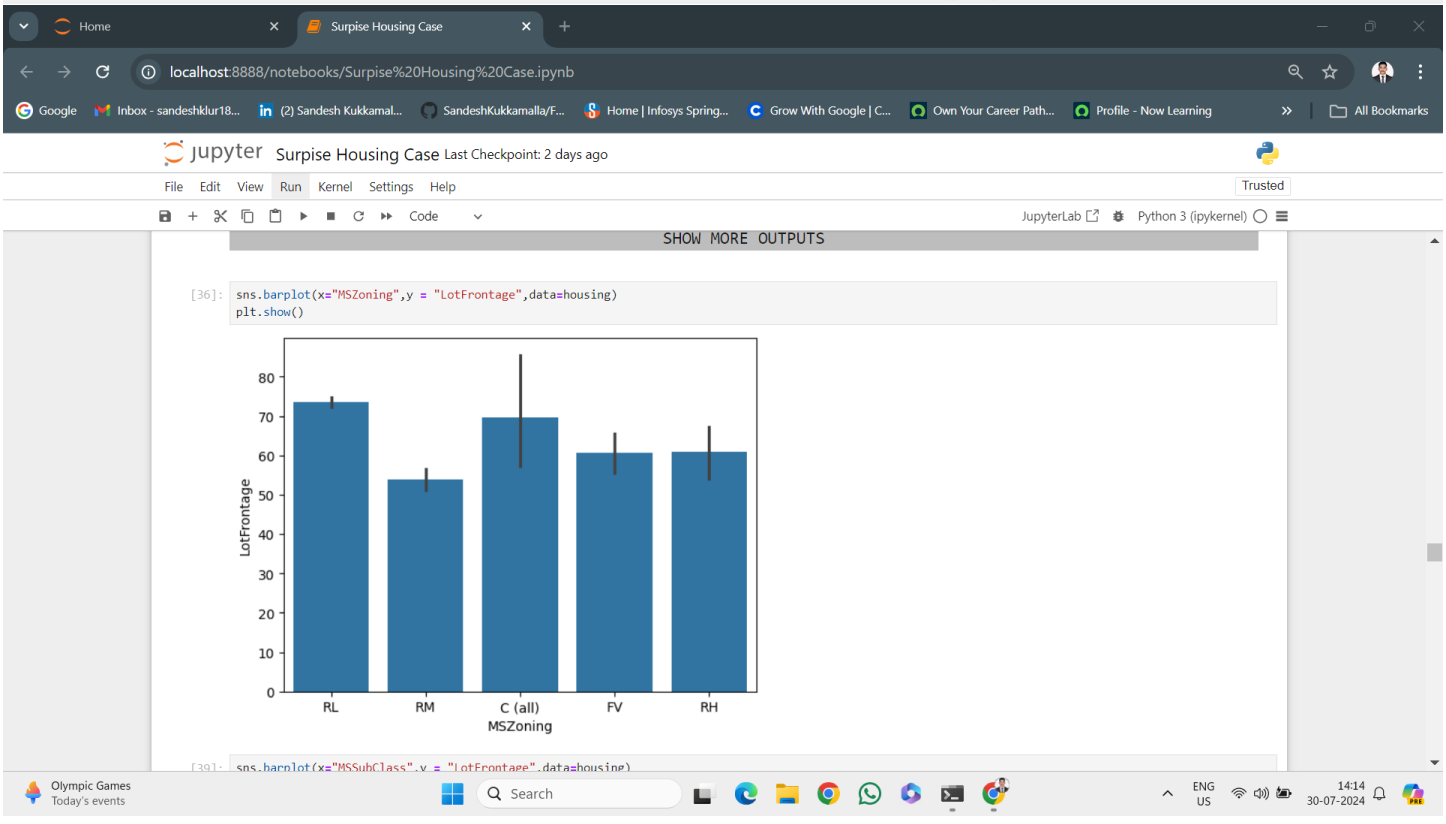

Home
Surpise Housing Case
localhost:8888/notebooks/Surpise%20Housing%20Case.ipynb
jupyter Surpise Housing Case Last Checkpoint: 2 days ago
File Edit View Run Kernel Settings Help
Trusted
Code
JupyterLab
Python 3 (ipykernel)
60
50
120
30
160
70
80
90
190
85
75
45
180
40
proportion
36.71%
9.86%
5.96%
4.11%
3.97%
3.56%
2.05%
MSZoning
RL 0.788356
RM 0.149315
FV 0.044521
RH 0.010959
C (all) 0.006849
Name: proportion, dtype: float64

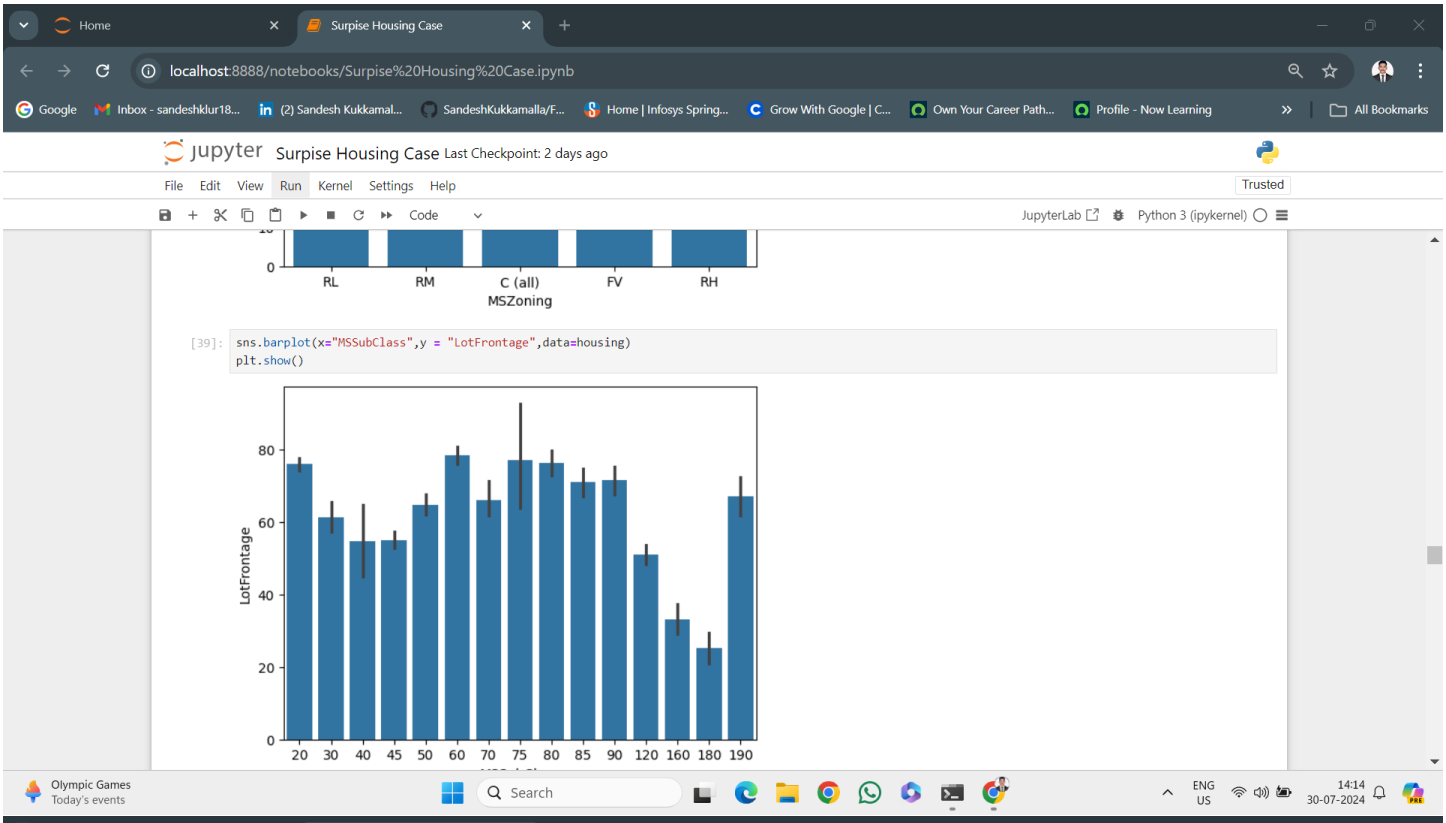

Home
Surpise Housing Case
localhost:8888/notebooks/Surpise%20Housing%20Case.ipynb
jupyter Surpise Housing Case Last Checkpoint: 2 days ago
File Edit View Run Kernel Settings Help
Trusted
Code
JupyterLab
Python 3 (ipykernel)
C (all) 0.006849
Name: proportion, dtype: float64
proportion
78.84%
14.93%
4.45%
RL
RM
FV
RH
C (all)
Street
Pave 0.99589
Grvl 0.00411
Name: proportion, dtype: float64
Pave
Grvl

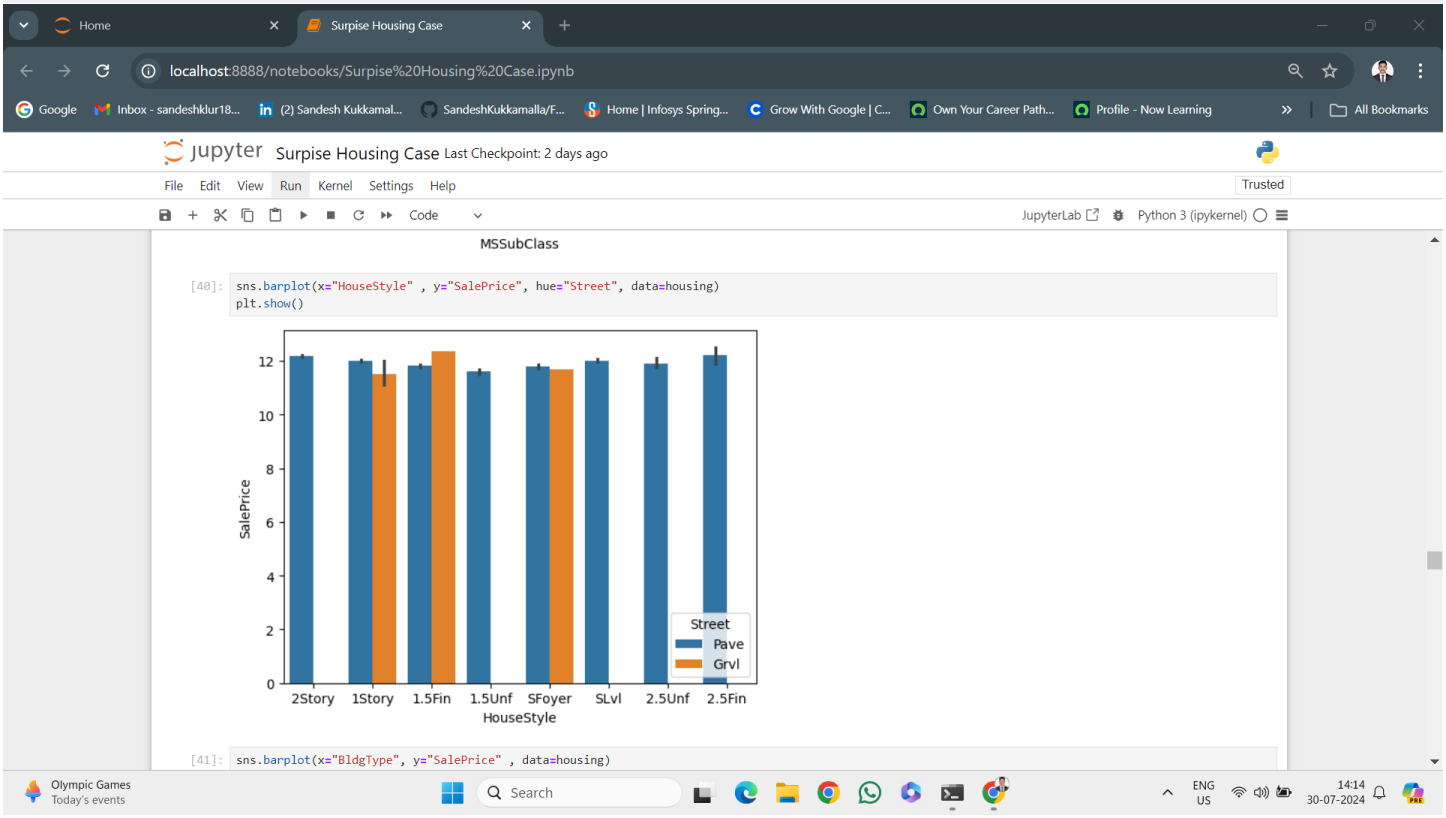

Home
Surpise Housing Case
localhost:8888/notebooks/Surpise%20Housing%20Case.ipynb
jupyter Surpise Housing Case Last Checkpoint: 2 days ago
File Edit View Run Kernel Settings Help
Trusted
Code
JupyterLab
Python 3 (ipykernel)
Pave 0.99589
Grvl 0.00411
Name: proportion, dtype: float64
Pave
Grvl
proportion
99.59%
0.41%
Alley
None 0.937671
Grvl 0.034247
Pave 0.028082
Name: proportion, dtype: float64
None
Olympic Games
Today's events
Search
ENG
US
14:13
30-07-2024

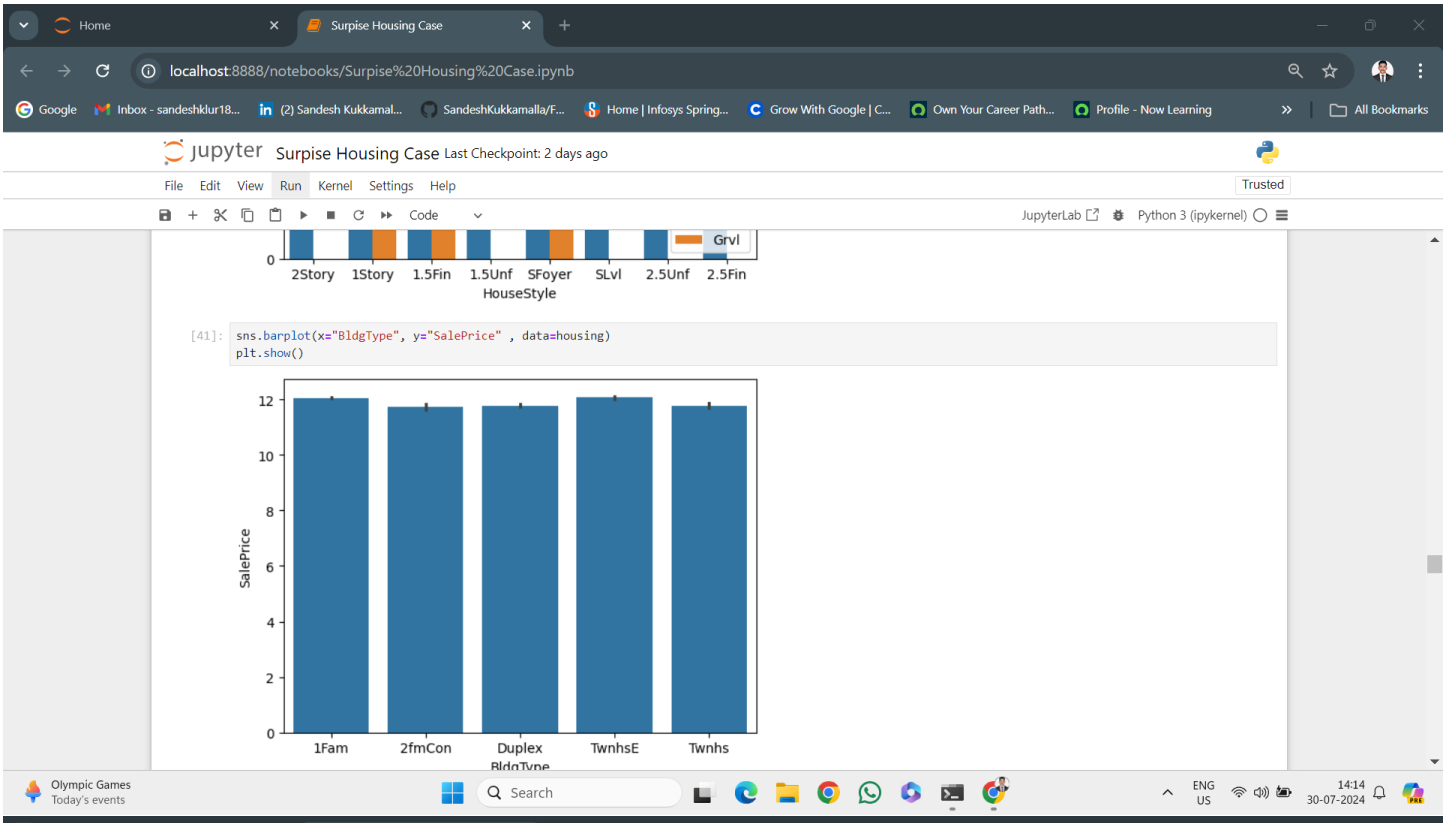

Home
Surpise Housing Case
localhost:8888/notebooks/Surpise%20Housing%20Case.ipynb
jupyter Surpise Housing Case Last Checkpoint: 2 days ago
File Edit View Run Kernel Settings Help
Trusted
Code
JupyterLab
Python 3 (ipykernel)
Alley
None 0.937671
Grvl 0.034247
Pave 0.028082
Name: proportion, dtype: float64
None
Grvl
Pave
proportion
93.77%
2.81%
3.42%
LotShape
Reg 0.633562
IR1 0.331507
IR2 0.028082
IR3 0.006849
Olympic Games
Today's events
Search
ENG
US
14:13
30-07-2024

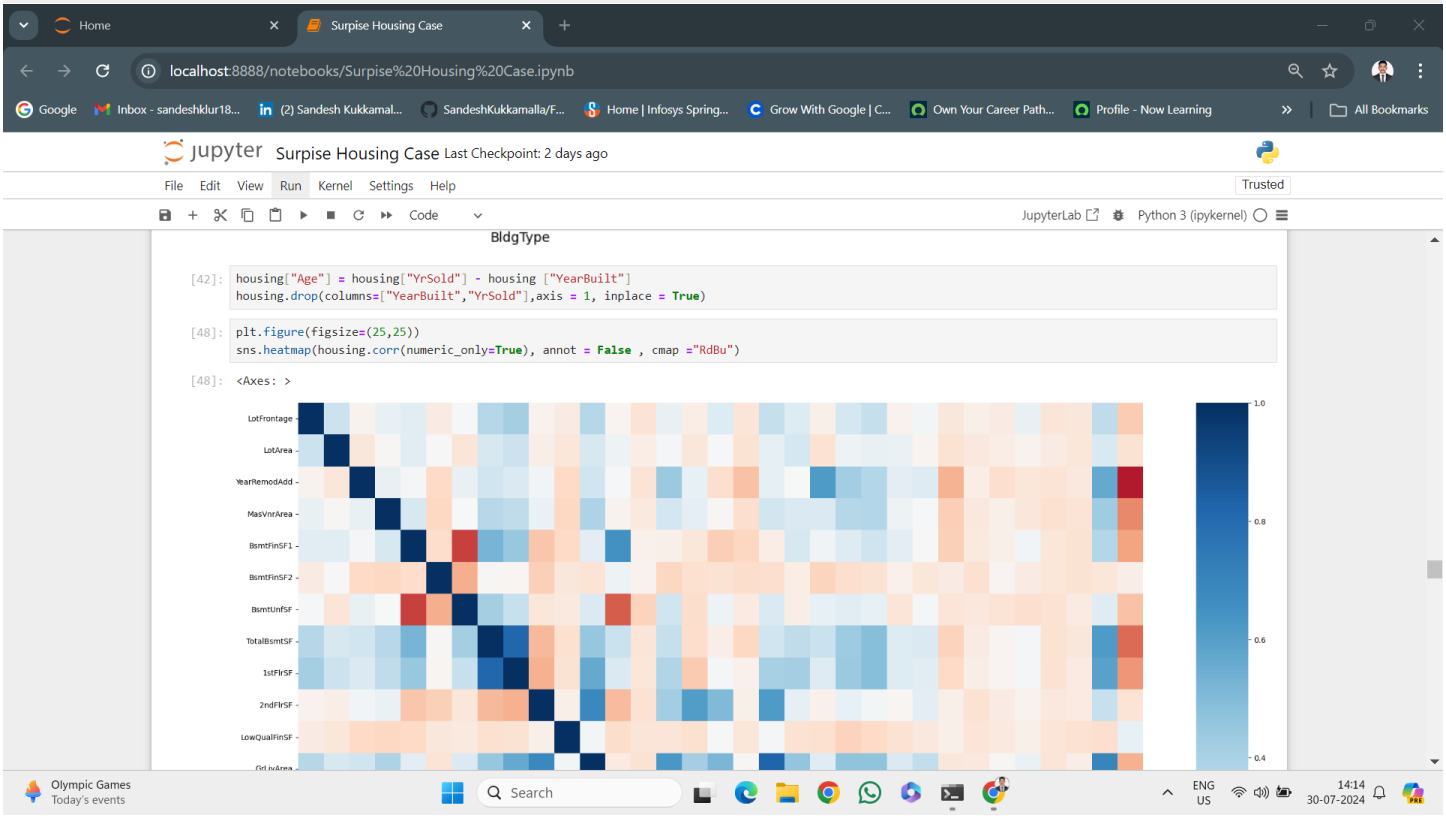

Home
Surpise Housing Case
localhost:8888/notebooks/Surpise%20Housing%20Case.ipynb
jupyter Surpise Housing Case Last Checkpoint: 2 days ago
File Edit View Run Kernel Settings Help
Trusted
Code
JupyterLab
Python 3 (ipykernel)
LotShape
Reg 0.633562
IR1 0.331507
IR2 0.028082
IR3 0.006849
Name: proportion, dtype: float64
Reg
IR1
IR2
IR3
63.36%
33.15%
2.81%
0.68%
proportion
LandContour
Lvl 0.897945

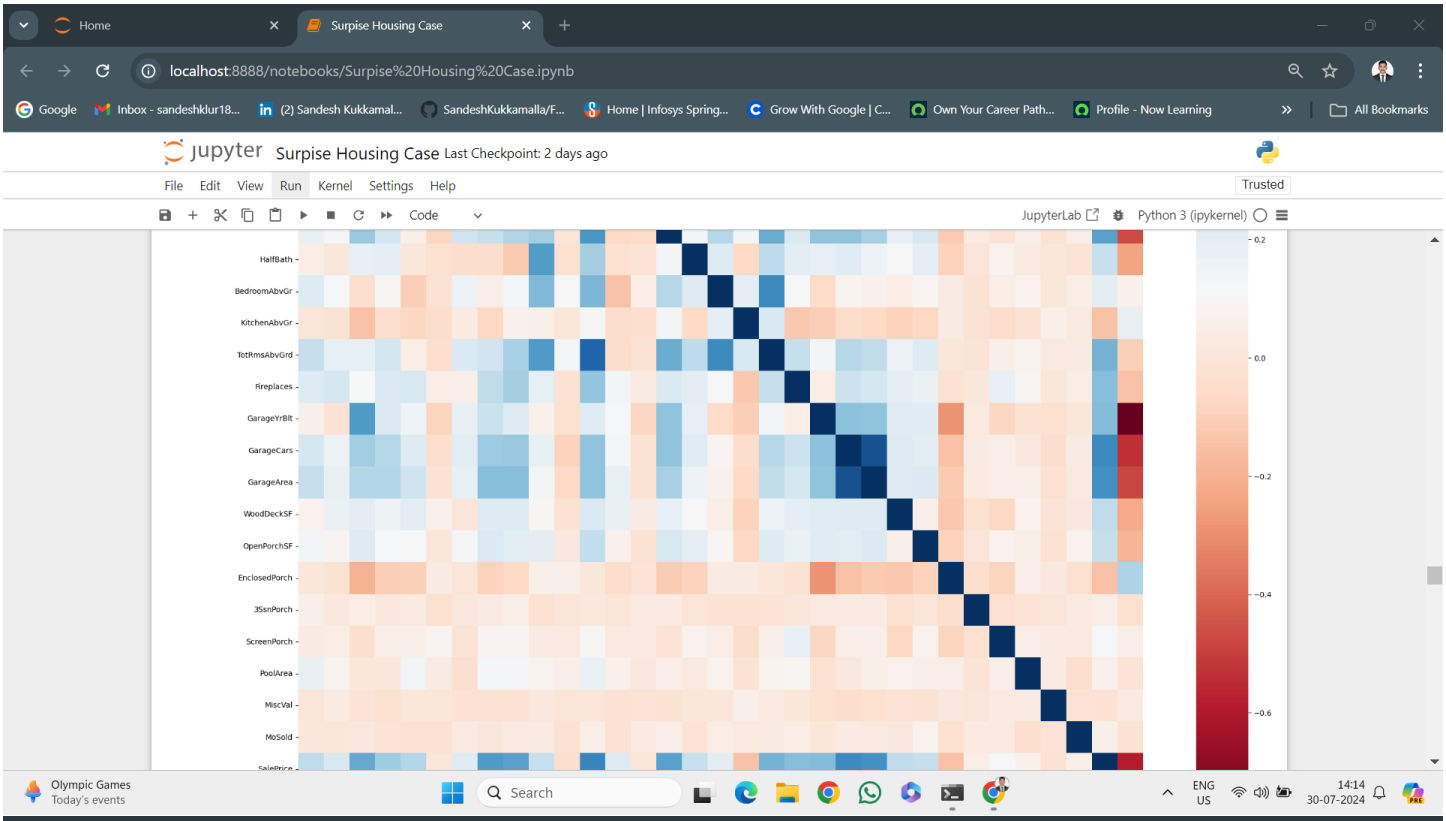

Home
Surpise Housing Case
localhost:8888/notebooks/Surpise%20Housing%20Case.ipynb
jupyter Surpise Housing Case Last Checkpoint: 2 days ago
File Edit View Run Kernel Settings Help
Trusted
Code
JupyterLab
Python 3 (ipykernel)
SHOW MORE OUTPUTS
[36]:
sns.barplot(x="MSZoning",y = "LotFrontage",data=housing)
plt.show()
LotFrontage
MSZoning
RL
RM
C (all)
FV
RH
0
10
20
30
40
50
60
70
80

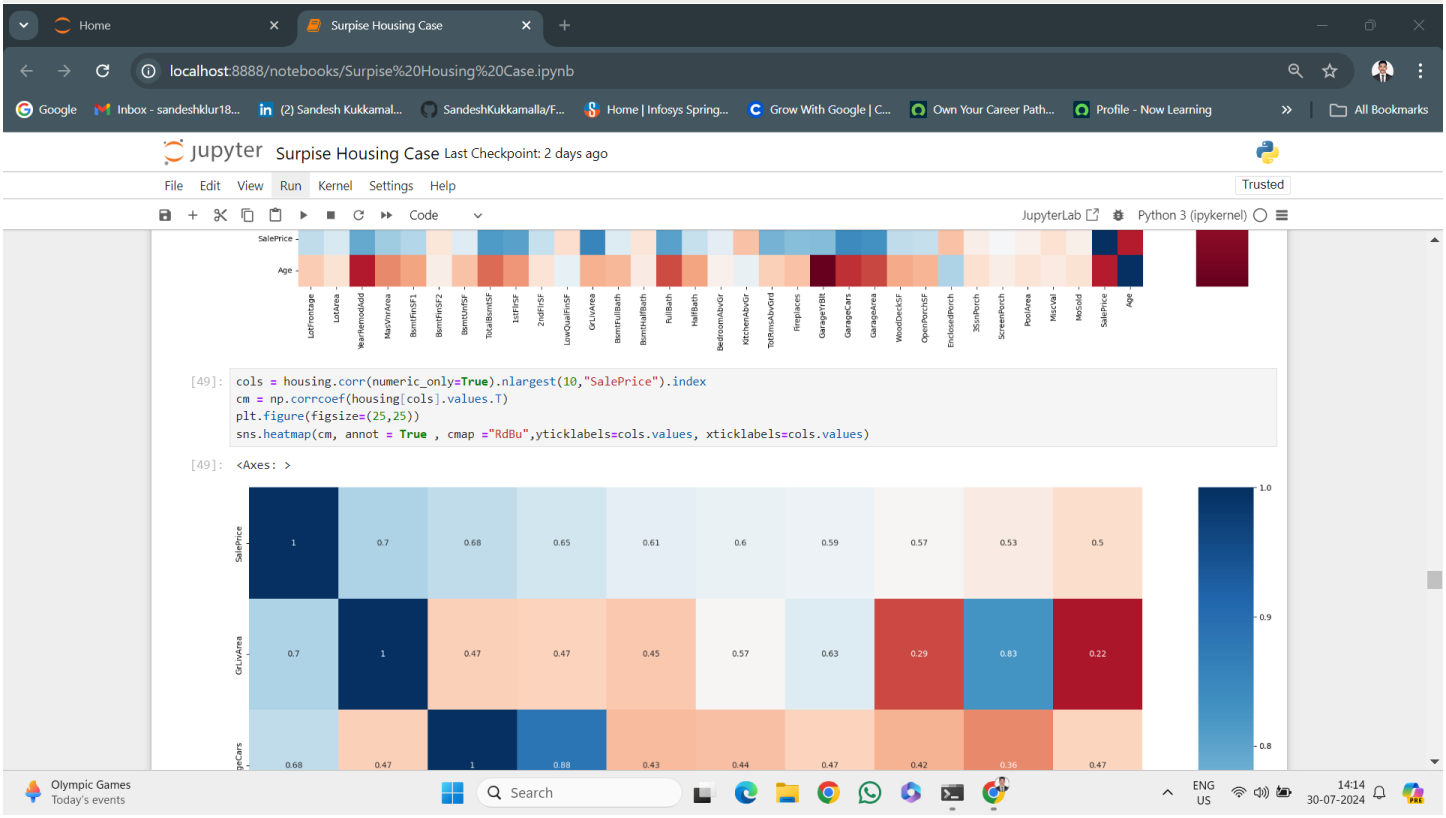

Home
Surpise Housing Case
localhost:8888/notebooks/Surpise%20Housing%20Case.ipynb
jupyter Surpise Housing Case Last Checkpoint: 2 days ago
File Edit View Run Kernel Settings Help
Trusted
Code
JupyterLab
Python 3 (ipykernel)
RL
RM
C (all)
FV
RH
MSZoning
[39]:
sns.barplot(x="MSSubClass",y = "LotFrontage",data=housing)
plt.show()
LotFrontage
20 30 40 45 50 60 70 75 80 85 90 120 160 180 190

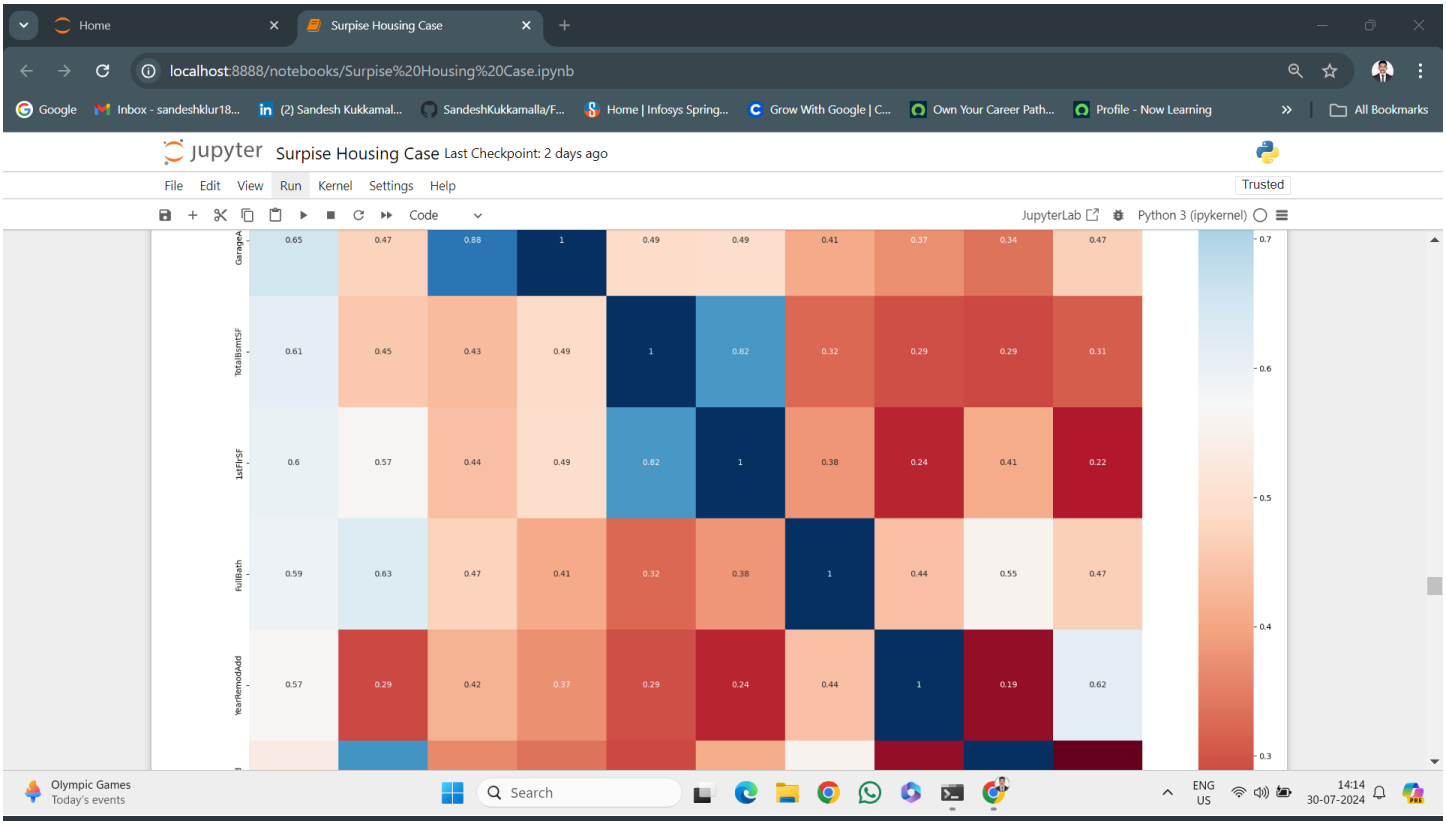

Home
Surpise Housing Case
localhost:8888/notebooks/Surpise%20Housing%20Case.ipynb
jupyter Surpise Housing Case Last Checkpoint: 2 days ago
File Edit View Run Kernel Settings Help
Trusted
Code
JupyterLab
Python 3 (ipykernel)
MSSubClass
[40]:
sns.barplot(x="HouseStyle" , y="SalePrice", hue="Street", data=housing)
plt.show()
SalePrice
Street
Pave
Grvl
2Story 1Story 1.5Fin 1.5Unf SFoyer SLvl 2.5Unf 2.5Fin
HouseStyle
[41]:
sns.barplot(x="BldgType", y="SalePrice" , data=housing)

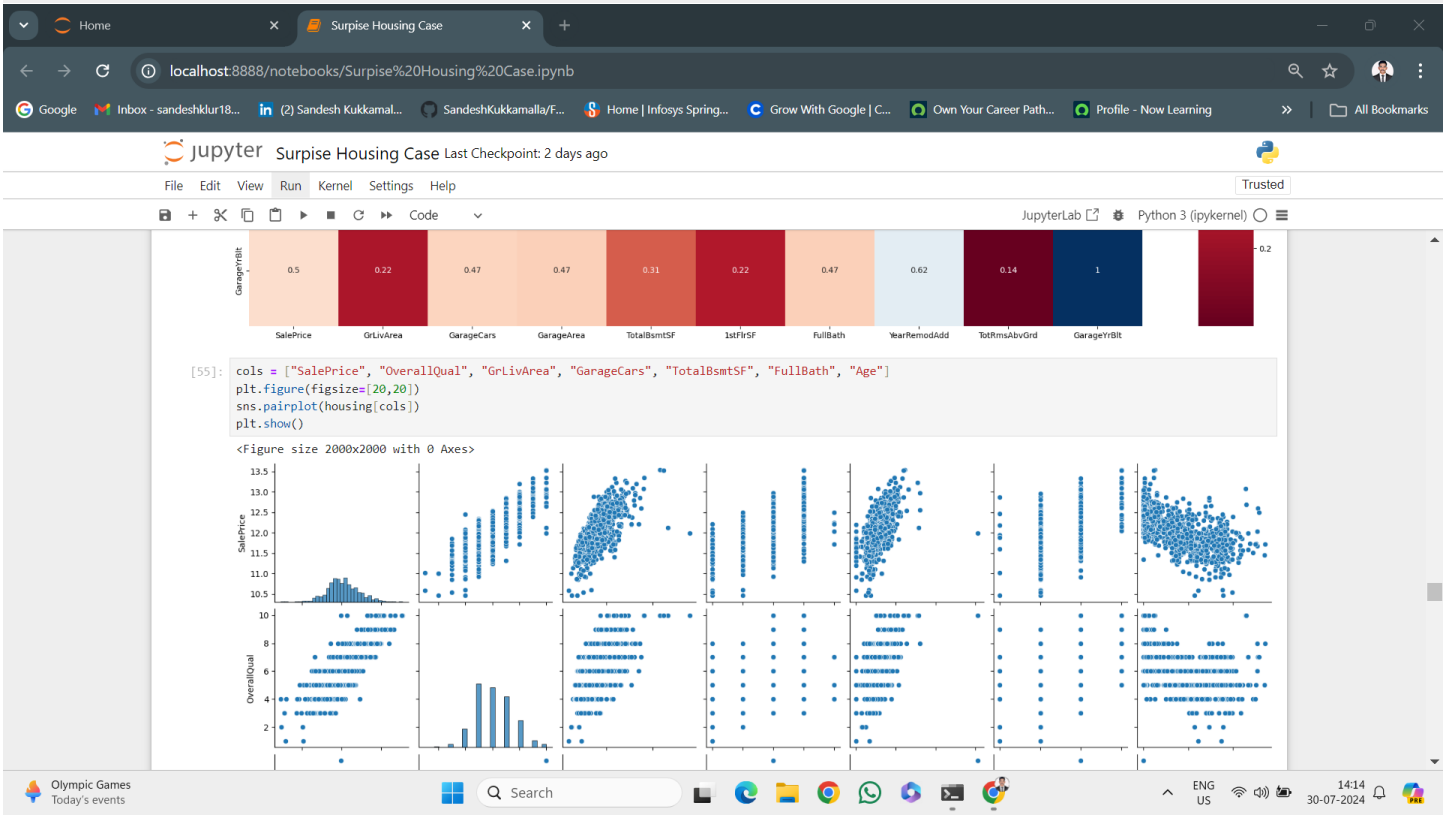

Home
Surpise Housing Case
localhost:8888/notebooks/Surpise%20Housing%20Case.ipynb
jupyter Surpise Housing Case Last Checkpoint: 2 days ago
File Edit View Run Kernel Settings Help
Trusted
Code
JupyterLab
Python 3 (ipykernel)
Grvl
0
2Story 1Story 1.5Fin 1.5Unf SFoyer SLvl 2.5Unf 2.5Fin
HouseStyle
[41]:
sns.barplot(x="BldgType", y="SalePrice" , data=housing)
plt.show()
SalePrice
0 2 4 6 8 10 12
1Fam 2fmCon Duplex TwnhsE Twnhs
BldgType

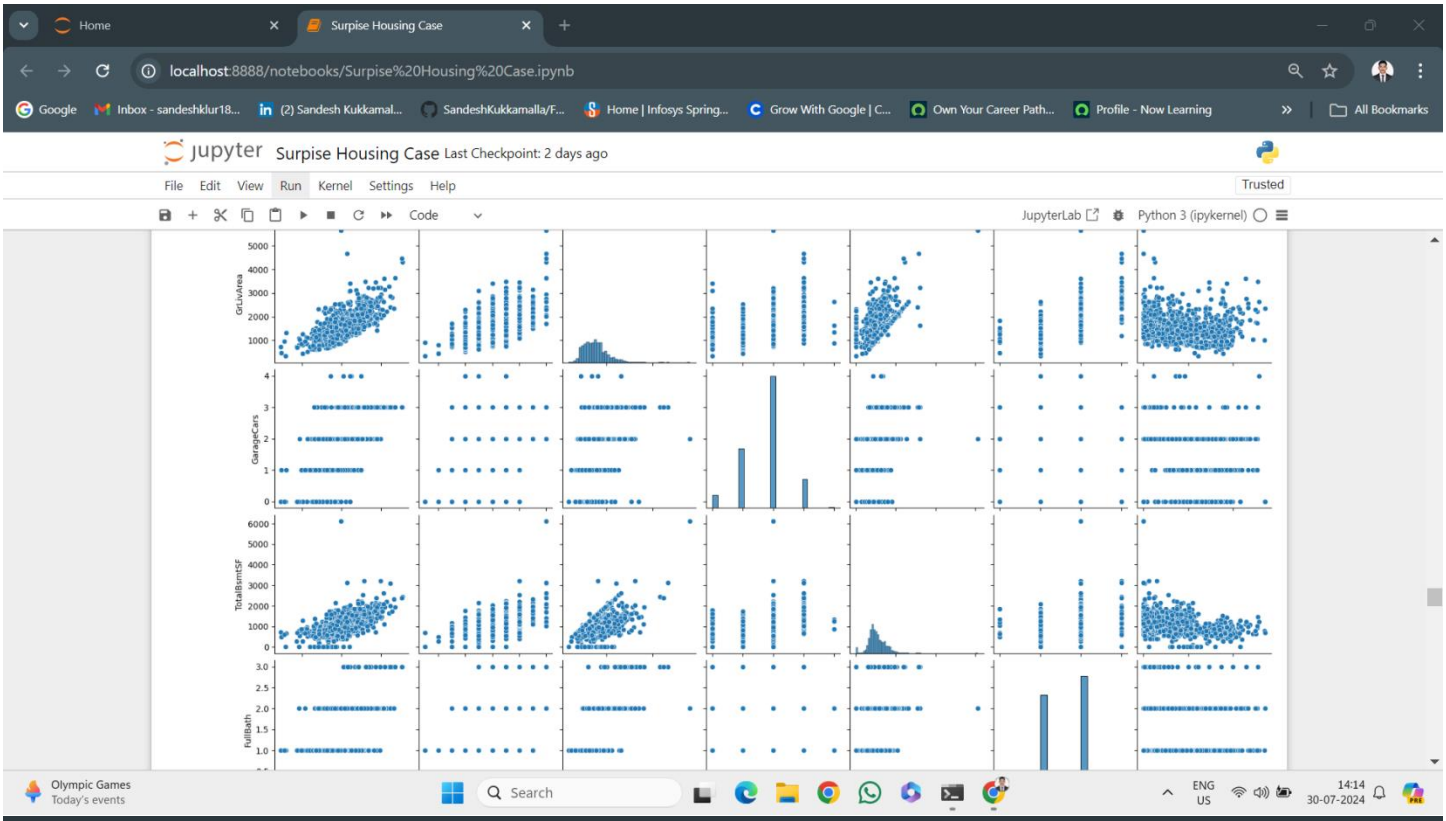

Home
Surpise Housing Case
localhost:8888/notebooks/Surpise%20Housing%20Case.ipynb
jupyter Surpise Housing Case Last Checkpoint: 2 days ago
File Edit View Run Kernel Settings Help
Trusted
Code
JupyterLab
Python 3 (ipykernel)
BldgType
[42]:
housing["Age"] = housing["YrSold"] - housing ["YearBuilt"]
housing.drop(columns=["YearBuilt","YrSold"],axis = 1, inplace = True)
[48]:
plt.figure(figsize=(25,25))
sns.heatmap(housing.corr(numeric_only=True), annot = False , cmap ="RdBu")
[48]:
<Axes: >
LotFrontage
LotArea
YearRemodAdd
MasVnrArea
BsmtFinSF1
BsmtFinSF2
BsmtUnfSF
TotalBsmtSF
1stFlrSF
2ndFlrSF
LowQualFinSF
1.0
0.8
0.6
0.4

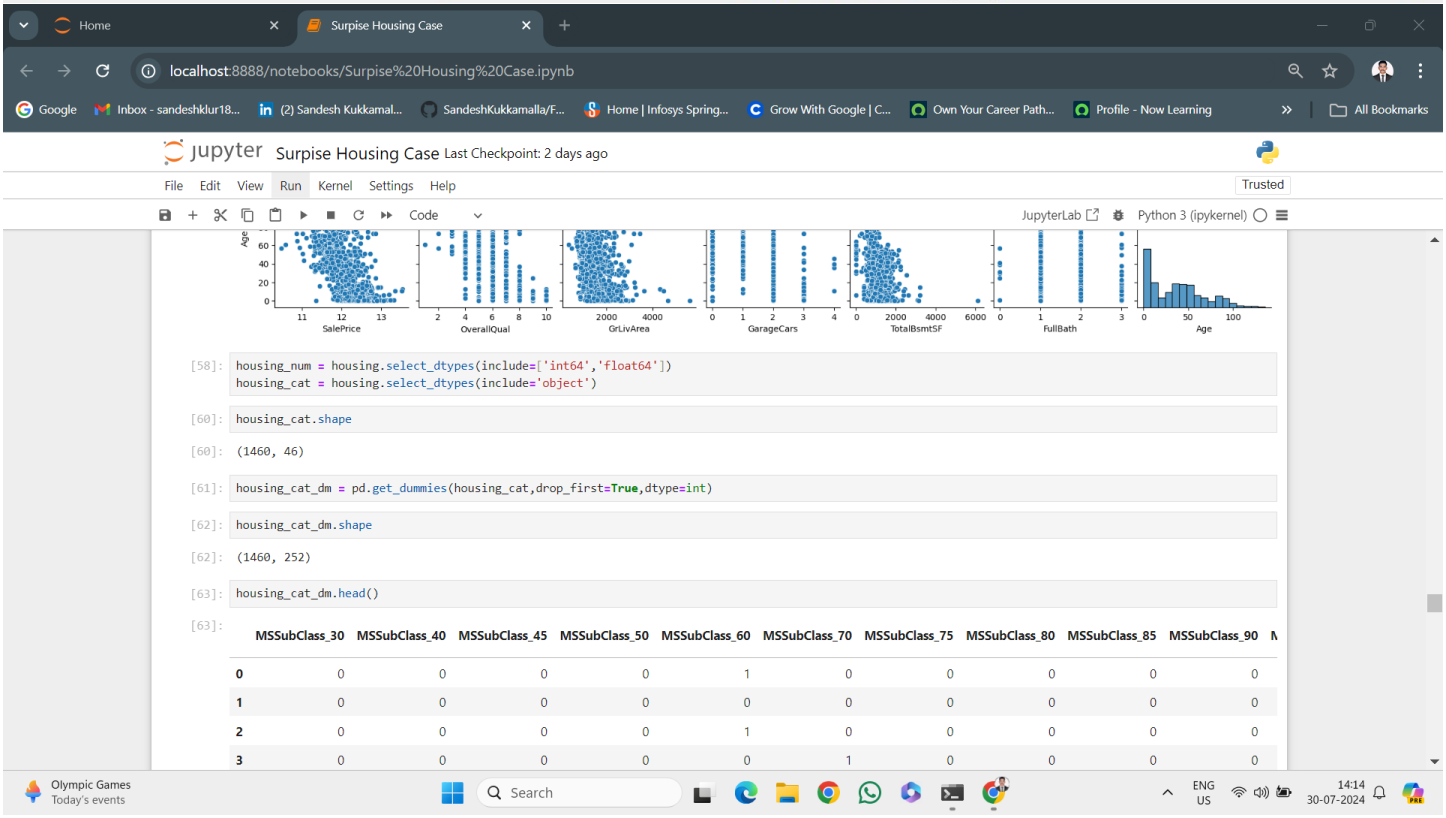

Home
Surpise Housing Case
localhost:8888/notebooks/Surpise%20Housing%20Case.ipynb
jupyter Surpise Housing Case Last Checkpoint: 2 days ago
File Edit View Run Kernel Settings Help
Trusted
Code
JupyterLab
Python 3 (ipykernel)
HalfBath
BedroomAbvGr
KitchenAbvGr
TotRmsAbvGrd
Fireplaces
GarageYrBlt
GarageCars
GarageArea
WoodDeckSF
OpenPorchSF
EnclosedPorch
3SsnPorch
ScreenPorch
PoolArea
MiscVal
MoSold
SalePrice
Olympic Games
Today's events
Search
ENG
US
14:14
30-07-2024


Home
Surpise Housing Case
localhost:8888/notebooks/Surpise%20Housing%20Case.ipynb
jupyter Surpise Housing Case Last Checkpoint: 2 days ago
File Edit View Run Kernel Settings Help
Trusted
Code
JupyterLab
Python 3 (ipykernel)
SalePrice
Age
LotFrontage
LotArea
YearRemodAdd
MasVnrArea
BsmtFinSF1
BsmtFinSF2
BsmtUnfSF
TotalBsmtSF
1stFlrSF
2ndFlrSF
LowQualFinSF
GrLivArea
BsmtFullBath
BsmtHalfBath
FullBath
HalfBath
BedroomAbvGr
KitchenAbvGr
TotRmsAbvGrd
Fireplaces
GarageYrBlt
GarageCars
GarageArea
WoodDeckSF
OpenPorchSF
EnclosedPorch
3SsnPorch
ScreenPorch
PoolArea
MiscVal
MoSold
SalePrice
Age
[49]:
cols = housing.corr(numeric_only=True).nlargest(10,"SalePrice").index
cm = np.corrcoef(housing[cols].values.T)
plt.figure(figsize=(25,25))
sns.heatmap(cm, annot = True , cmap ="RdBu",yticklabels=cols.values, xticklabels=cols.values)
[49]:
<Axes: >
SalePrice
GrLivArea
1 0.7 0.68 0.65 0.61 0.6 0.59 0.57 0.53 0.5
0.7 1 0.47 0.47 0.45 0.57 0.63 0.29 0.83 0.22
0.68 0.47 1 0.88 0.43 0.44 0.47 0.42 0.36 0.47
Olympic Games
Today's events
Search
ENG
US
14:14
30-07-2024


Home
Surpise Housing Case
localhost:8888/notebooks/Surpise%20Housing%20Case.ipynb
jupyter Surpise Housing Case Last Checkpoint: 2 days ago
File Edit View Run Kernel Settings Help
Trusted
Code
JupyterLab
Python 3 (ipykernel)
GarageA
TotalBsmtSF
1stFlrSF
FullBath
YearRemodAdd
0.65 0.47 0.88 1 0.49 0.49 0.41 0.37 0.34 0.47
0.61 0.45 0.43 0.49 1 0.82 0.32 0.29 0.29 0.31
0.6 0.57 0.44 0.49 0.82 1 0.38 0.24 0.41 0.22
0.59 0.63 0.47 0.41 0.32 0.38 1 0.44 0.55 0.47
0.57 0.29 0.42 0.37 0.29 0.24 0.44 1 0.19 0.62
0.7
0.6
0.5
0.4
0.3
Olympic Games
Today's events
Search
ENG
US
14:14
30-07-2024


Home
Surpise Housing Case
localhost:8888/notebooks/Surpise%20Housing%20Case.ipynb
jupyter Surpise Housing Case Last Checkpoint: 2 days ago
File Edit View Run Kernel Settings Help
Trusted
Code
JupyterLab
Python 3 (ipykernel)
GarageYrBlt
0.5 0.22 0.47 0.47 0.31 0.22 0.47 0.62 0.14 1
SalePrice
GrLivArea
GarageCars
GarageArea
TotalBsmtSF
1stFlrSF
FullBath
YearRemodAdd
TotRmsAbvGrd
GarageYrBlt
0.2
[55]:
cols = ["SalePrice", "OverallQual", "GrLivArea", "GarageCars", "TotalBsmtSF", "FullBath", "Age"]
plt.figure(figsize=[20,20])
sns.pairplot(housing[cols])
plt.show()
<Figure size 2000x2000 with 0 Axes>
SalePrice
OverallQual
Olympic Games
Today's events
Search
ENG
US
14:14
30-07-2024





[58]:
```
housing_num = housing.select_dtypes(include=['int64','float64'])
housing_cat = housing.select_dtypes(include='object')
```

[60]:
```
housing_cat.shape
```

[60]: (1460, 46)

[61]:
```
housing_cat_dm = pd.get_dummies(housing_cat,drop_first=True,dtype=int)
```

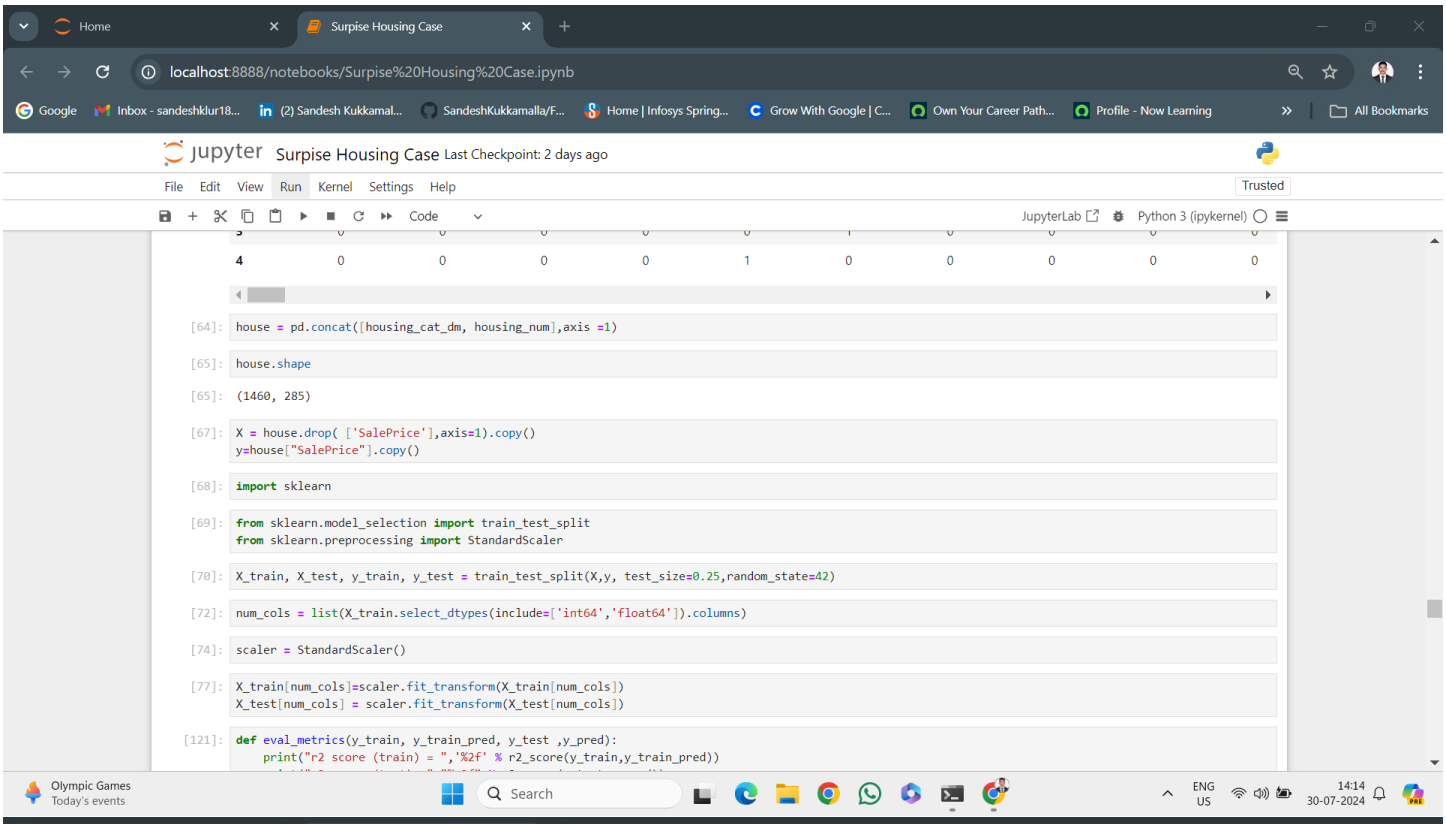
[62]:
```
housing_cat_dm.shape
```

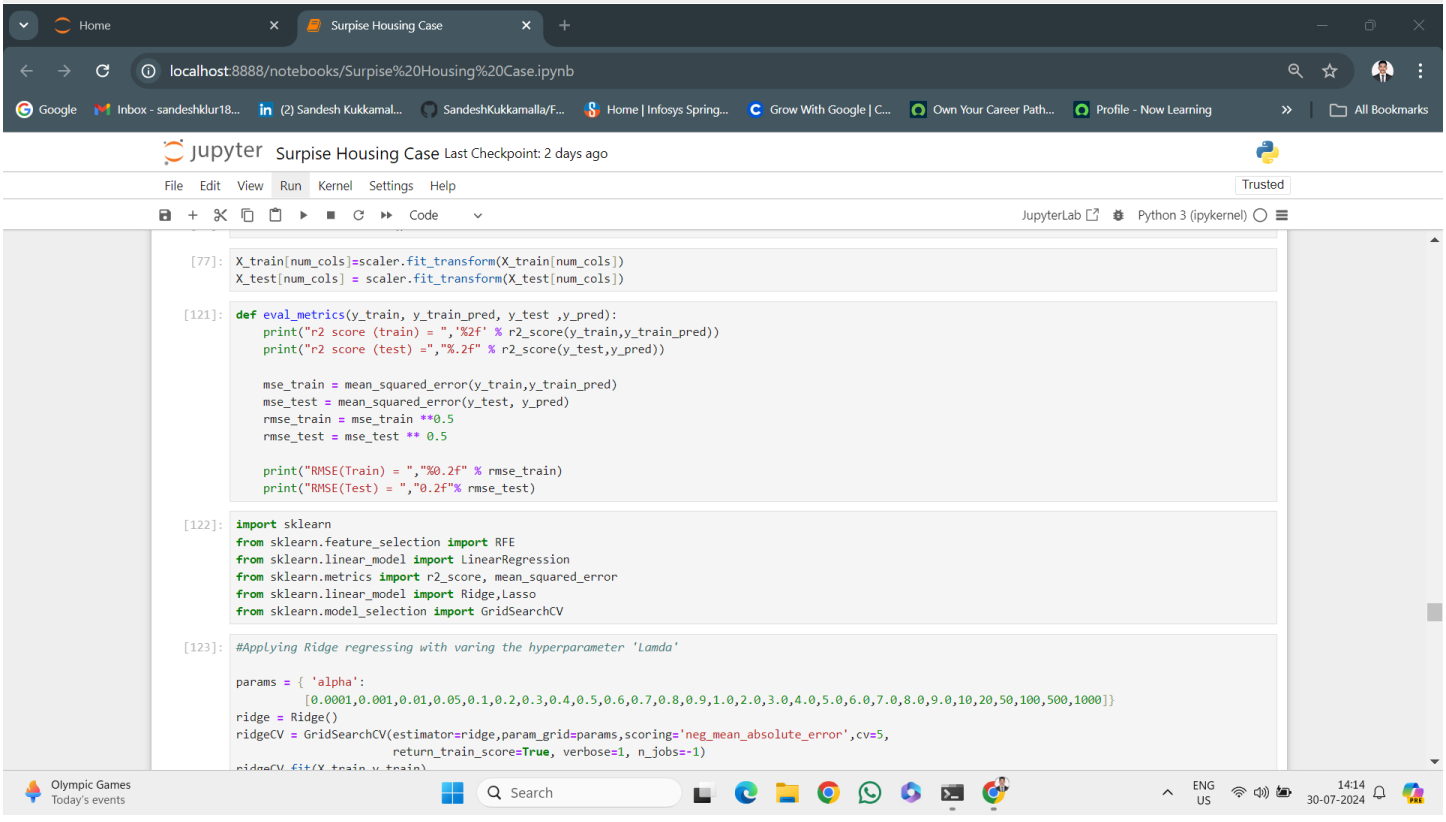
[62]: (1460, 252)

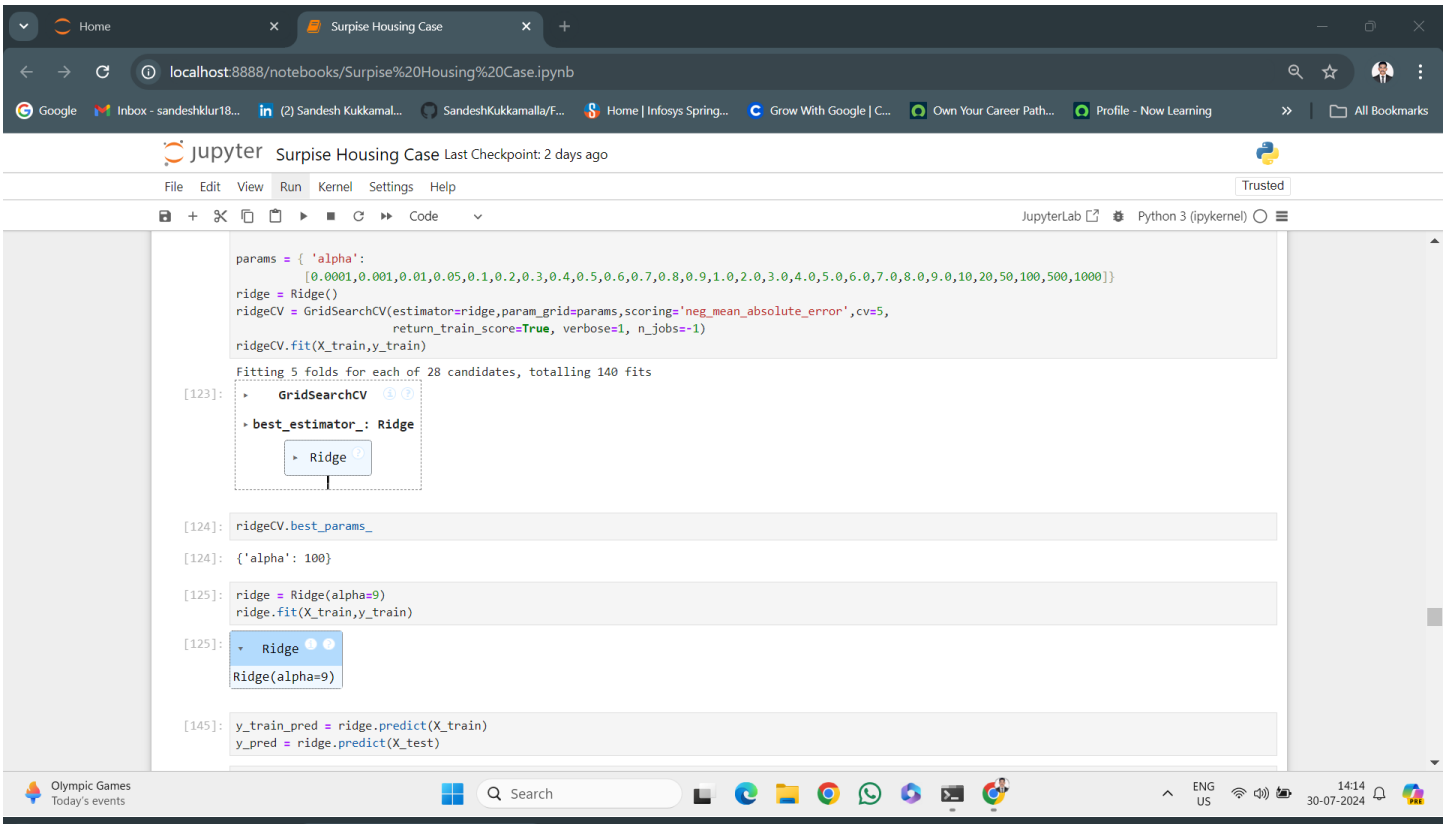
[63]:
```
housing_cat_dm.head()
```

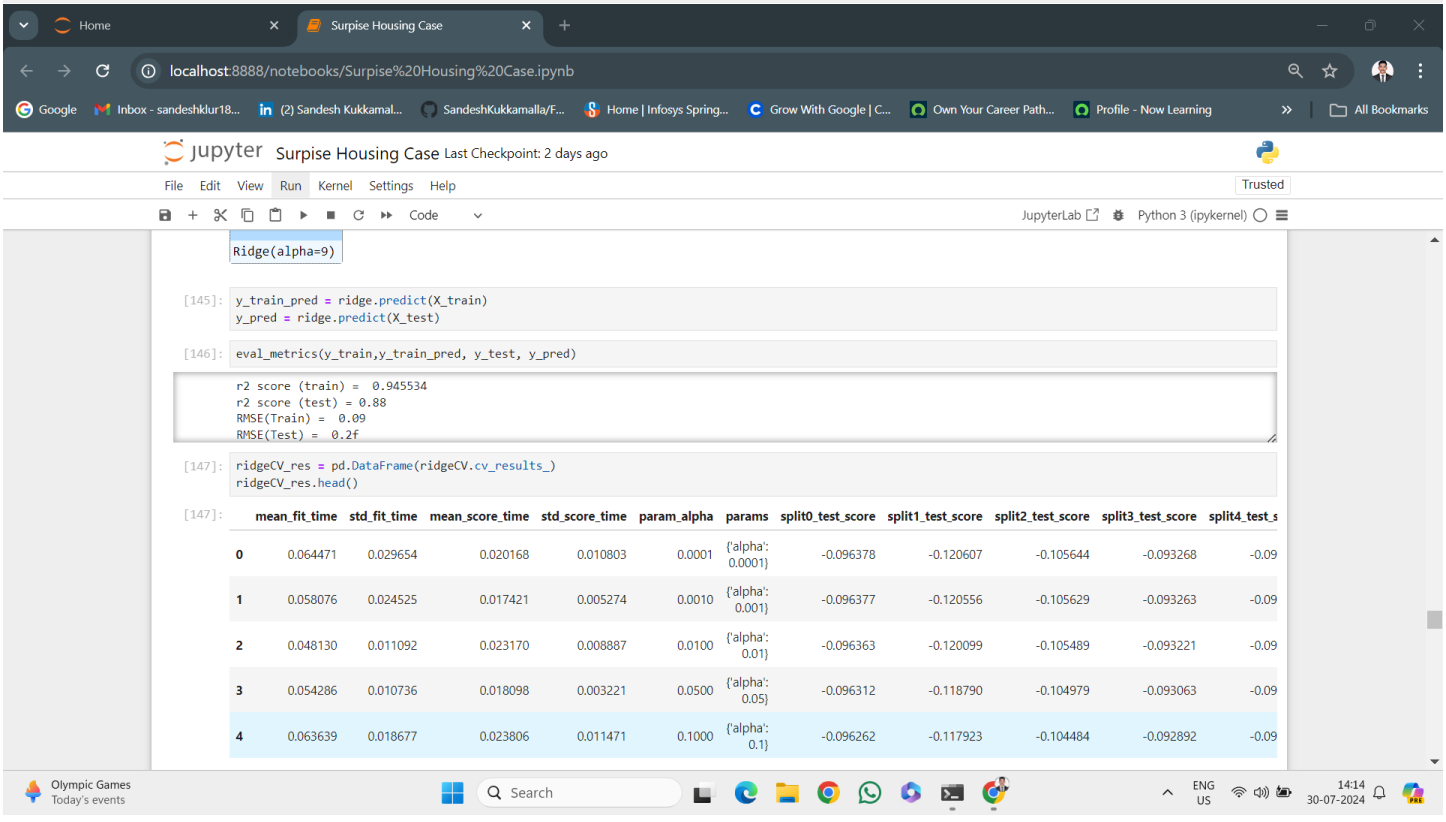
[63]:

	MSSubClass_30	MSSubClass_40	MSSubClass_45	MSSubClass_50	MSSubClass_60	MSSubClass_70	MSSubClass_75	MSSubClass_80	MSSubClass_85	MSSubClass_90	M
0	0	0	0	0	1	0	0	0	0	0	
1	0	0	0	0	0	0	0	0	0	0	
2	0	0	0	0	1	0	0	0	0	0	
3	0	0	0	0	0	1	0	0	0	0	


Home
Surpise Housing Case
localhost:8888/notebooks/Surpise%20Housing%20Case.ipynb
jupyter Surpise Housing Case Last Checkpoint: 2 days ago
File Edit View Run Kernel Settings Help
Trusted
Code
JupyterLab
Python 3 (ipykernel)
4 0 0 0 0 1 0 0 0 0 0
[64]: house = pd.concat([housing_cat_dm, housing_num],axis =1)
[65]: house.shape
[65]: (1460, 285)
[67]: X = house.drop(['SalePrice'],axis=1).copy()
y=house["SalePrice"].copy()
[68]: import sklearn
[69]: from sklearn.model_selection import train_test_split
from sklearn.preprocessing import StandardScaler
[70]: X_train, X_test, y_train, y_test = train_test_split(X,y, test_size=0.25,random_state=42)
[72]: num_cols = list(X_train.select_dtypes(include=['int64','float64']).columns)
[74]: scaler = StandardScaler()
[77]: X_train[num_cols]=scaler.fit_transform(X_train[num_cols])
X_test[num_cols] = scaler.fit_transform(X_test[num_cols])
[121]: def eval_metrics(y_train, y_train_pred, y_test ,y_pred):
print("r2 score (train) = ",'%2f' % r2_score(y_train,y_train_pred))
Olympic Games
Today's events
Search
ENG
US
14:14
30-07-2024


Home
Surpise Housing Case
localhost:8888/notebooks/Surpise%20Housing%20Case.ipynb
jupyter Surpise Housing Case Last Checkpoint: 2 days ago
File Edit View Run Kernel Settings Help
Trusted
Code
JupyterLab
Python 3 (ipykernel)
[77]: X_train[num_cols]=scaler.fit_transform(X_train[num_cols])
X_test[num_cols] = scaler.fit_transform(X_test[num_cols])
[121]: def eval_metrics(y_train, y_train_pred, y_test ,y_pred):
print("r2 score (train) = ",'%2f' % r2_score(y_train,y_train_pred))
print("r2 score (test) =","%.2f" % r2_score(y_test,y_pred))
mse_train = mean_squared_error(y_train,y_train_pred)
mse_test = mean_squared_error(y_test, y_pred)
rmse_train = mse_train **0.5
rmse_test = mse_test ** 0.5
print("RMSE(Train) = ","%0.2f" % rmse_train)
print("RMSE(Test) = ","0.2f"% rmse_test)
[122]: import sklearn
from sklearn.feature_selection import RFE
from sklearn.linear_model import LinearRegression
from sklearn.metrics import r2_score, mean_squared_error
from sklearn.linear_model import Ridge,Lasso
from sklearn.model_selection import GridSearchCV
[123]: #Applying Ridge regressing with varing the hyperparameter 'Lamda'
params = { 'alpha':
[0.0001,0.001,0.01,0.05,0.1,0.2,0.3,0.4,0.5,0.6,0.7,0.8,0.9,1.0,2.0,3.0,4.0,5.0,6.0,7.0,8.0,9.0,10,20,50,100,500,1000]}
ridge = Ridge()
ridgeCV = GridSearchCV(estimator=ridge,param_grid=params,scoring='neg_mean_absolute_error',cv=5,
return_train_score=True, verbose=1, n_jobs=-1)
Olympic Games
Today's events
Search
ENG
US
14:14
30-07-2024


Home
Surpise Housing Case
localhost:8888/notebooks/Surpise%20Housing%20Case.ipynb
jupyter Surpise Housing Case Last Checkpoint: 2 days ago
File Edit View Run Kernel Settings Help
Trusted
Code
JupyterLab
Python 3 (ipykernel)
params = { 'alpha':
[0.0001,0.001,0.01,0.05,0.1,0.2,0.3,0.4,0.5,0.6,0.7,0.8,0.9,1.0,2.0,3.0,4.0,5.0,6.0,7.0,8.0,9.0,10,20,50,100,500,1000]}
ridge = Ridge()
ridgeCV = GridSearchCV(estimator=ridge,param_grid=params,scoring='neg_mean_absolute_error',cv=5,
return_train_score=True, verbose=1, n_jobs=-1)
ridgeCV.fit(X_train,y_train)
Fitting 5 folds for each of 28 candidates, totalling 140 fits
[123]:
GridSearchCV
best_estimator_: Ridge
Ridge
[124]: ridgeCV.best_params_
[124]: {'alpha': 100}
[125]: ridge = Ridge(alpha=9)
ridge.fit(X_train,y_train)
[125]: Ridge
Ridge(alpha=9)
[145]: y_train_pred = ridge.predict(X_train)
y_pred = ridge.predict(X_test)
Olympic Games
Today's events
Search
ENG
US
14:14
30-07-2024


Home
Surpise Housing Case
localhost:8888/notebooks/Surpise%20Housing%20Case.ipynb
jupyter Surpise Housing Case Last Checkpoint: 2 days ago
File Edit View Run Kernel Settings Help
Trusted
Code
JupyterLab
Python 3 (ipykernel)
Ridge(alpha=9)
[145]: y_train_pred = ridge.predict(X_train)
y_pred = ridge.predict(X_test)
[146]: eval_metrics(y_train,y_train_pred, y_test, y_pred)
r2 score (train) = 0.945534
r2 score (test) = 0.88
RMSE(Train) = 0.09
RMSE(Test) = 0.2f
[147]: ridgeCV_res = pd.DataFrame(ridgeCV.cv_results_)
ridgeCV_res.head()
[147]:
mean_fit_time std_fit_time mean_score_time std_score_time param_alpha params split0_test_score split1_test_score split2_test_score split3_test_score split4_test_s
0 0.064471 0.029654 0.020168 0.010803 0.0001 {'alpha': 0.0001} -0.096378 -0.120607 -0.105644 -0.093268 -0.09
1 0.058076 0.024525 0.017421 0.005274 0.0010 {'alpha': 0.001} -0.096377 -0.120556 -0.105629 -0.093263 -0.09
2 0.048130 0.011092 0.023170 0.008887 0.0100 {'alpha': 0.01} -0.096363 -0.120099 -0.105489 -0.093221 -0.09
3 0.054286 0.010736 0.018098 0.003221 0.0500 {'alpha': 0.05} -0.096312 -0.118790 -0.104979 -0.093063 -0.09
4 0.063639 0.018677 0.023806 0.011471 0.1000 {'alpha': 0.1} -0.096262 -0.117923 -0.104484 -0.092892 -0.09
Olympic Games
Today's events
Search
ENG
US
14:14
30-07-2024

Home | Surpise Housing Case

localhost:8888/notebooks/Surpise%20Housing%20Case.ipynb

jupyter Surpise Housing Case Last Checkpoint: 2 days ago

File Edit View Run Kernel Settings Help — Trusted

Code — JupyterLab — Python 3 (ipykernel)

3	0.054286	0.010736	0.018098	0.003221	0.0500	{'alpha': 0.05}	-0.096312	-0.118790	-0.104979	-0.093063	-0.09
4	0.063639	0.018677	0.023806	0.011471	0.1000	{'alpha': 0.1}	-0.096262	-0.117923	-0.104484	-0.092892	-0.09

[148]:
```
plt.plot(ridgeCV_res['param_alpha'],ridgeCV_res['mean_train_score'],label ='train')
plt.plot(ridgeCV_res['param_alpha'],ridgeCV_res['mean_train_score'],label = 'test')
plt.xlabel('alpha')
plt.ylabel('R2_score')
plt.xscale('log')
plt.legend()
plt.show()
```

Olympic Games Today's events — Search — ENG US — 14:15 30-07-2024

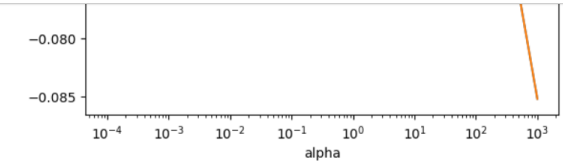
Home | Surpise Housing Case

localhost:8888/notebooks/Surpise%20Housing%20Case.ipynb

jupyter Surpise Housing Case Last Checkpoint: 2 days ago

File Edit View Run Kernel Settings Help — Trusted

Code — JupyterLab — Python 3 (ipykernel)



[149]:
```
params = { 'alpha':
           [0.0001,0.001,0.01,0.05,0.1,0.2,0.3,0.4,0.5,0.6,0.7,0.8,0.9,1.0,2.0,3.0,4.0,5.0,6.0,7.0,8.0,9.0,10,20,50,100,500,1000]}
lasso = Lasso()
lassoCV = GridSearchCV(estimator=ridge,param_grid=params,scoring='neg_mean_absolute_error',cv=5,
                       return_train_score=True, verbose=1, n_jobs=-1)
lassoCV.fit(X_train,y_train)
```

```
Fitting 5 folds for each of 28 candidates, totalling 140 fits
```

[149]:
```
▸ GridSearchCV
▸ best_estimator_: Ridge
    ▸ Ridge
```

[150]:
```
lassoCV.best_params_
```

[150]:
```
{'alpha': 100}
```

[151]:
```
lasso=Lasso(alpha=0.001)
```

Olympic Games Today's events — Search — ENG US — 14:15 30-07-2024

Home | Surpise Housing Case

localhost:8888/notebooks/Surpise%20Housing%20Case.ipynb

jupyter Surpise Housing Case Last Checkpoint: 2 days ago

File Edit View Run Kernel Settings Help — Trusted

Code — JupyterLab — Python 3 (ipykernel)

```
[150]: lassoCV.best_params_
```

```
[150]: {'alpha': 100}
```

```
[151]: lasso=Lasso(alpha=0.001)
       lasso.fit(X_train,y_train)
```

```
[151]: Lasso
       Lasso(alpha=0.001)
```

```
[157]: y_train_pred1 = lasso.predict(X_train)
       y_pred1 = lasso.predict(X_test)
```

```
[158]: eval_metrics(y_train,y_train_pred1, y_test, y_pred1)
```

```
r2 score (train) =  0.939695
r2 score (test) = 0.88
RMSE(Train) =  0.10
RMSE(Test) =  0.2f
```

```
[159]: lassoCV_res = pd.DataFrame(lassoCV.cv_results_)
       lassoCV_res.head()
```

[159]:

	mean_fit_time	std_fit_time	mean_score_time	std_score_time	param_alpha	params	split0_test_score	split1_test_score	split2_test_score	split3_test_score	split4_test_s
0	0.052941	0.010117	0.012133	0.004352	0.0001	{'alpha': 0.0001}	-0.096378	-0.120607	-0.105644	-0.093268	-0.09
1	0.043561	0.006814	0.015547	0.010085	0.0010	{'alpha': 0.001}	-0.096377	-0.120556	-0.105629	-0.093263	-0.09

Olympic Games Today's events — Search — ENG US — 14:15 30-07-2024

Home | Surpise Housing Case

localhost:8888/notebooks/Surpise%20Housing%20Case.ipynb

jupyter Surpise Housing Case Last Checkpoint: 2 days ago

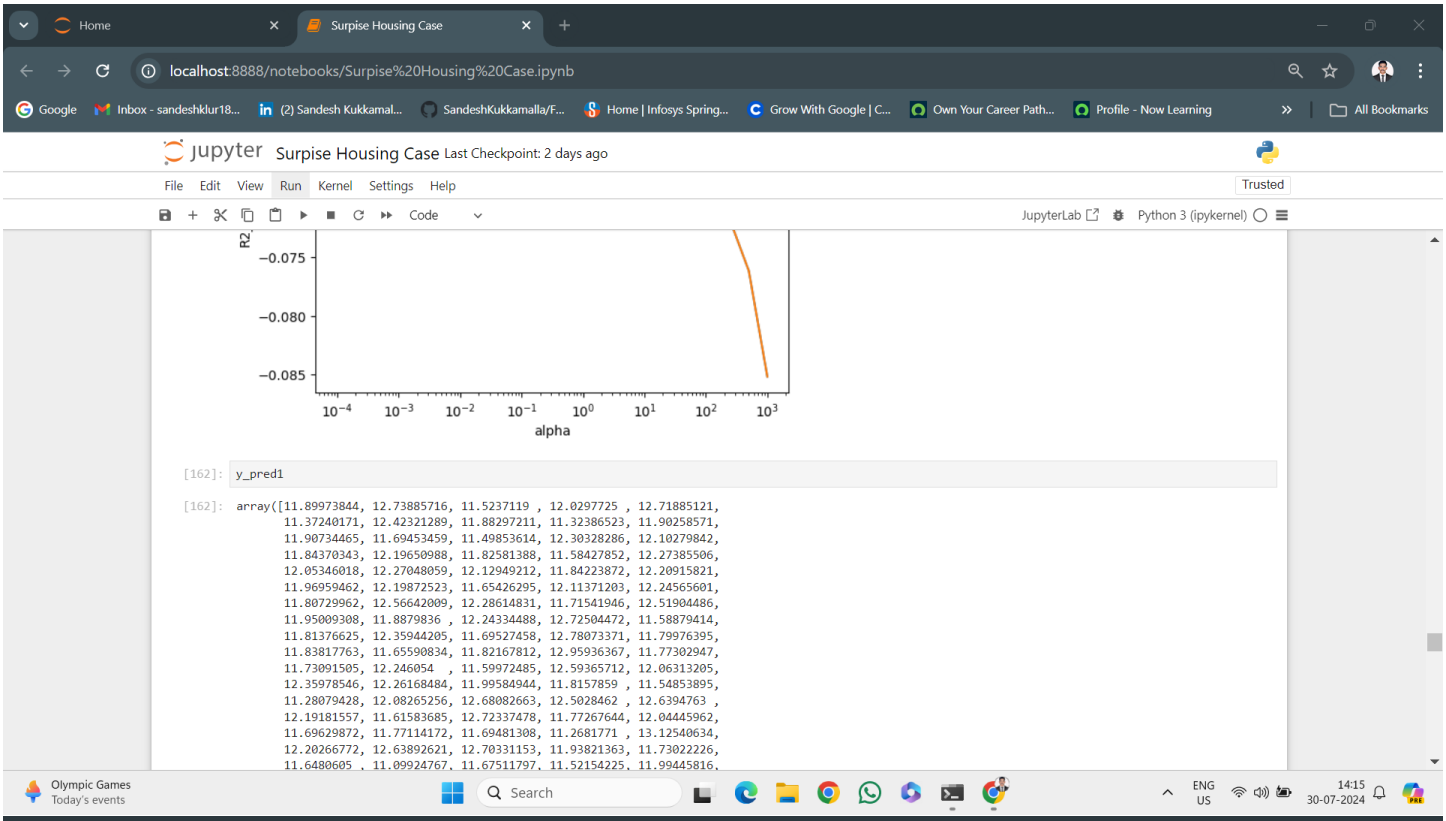
File Edit View Run Kernel Settings Help — Trusted

Code — JupyterLab — Python 3 (ipykernel)

						0.001}					
2	0.037662	0.005742	0.019127	0.007601	0.0100	{'alpha': 0.01}	-0.096363	-0.120099	-0.105489	-0.093221	-0.09
3	0.036825	0.005561	0.023140	0.017669	0.0500	{'alpha': 0.05}	-0.096312	-0.118790	-0.104979	-0.093063	-0.09
4	0.071927	0.040250	0.020768	0.015780	0.1000	{'alpha': 0.1}	-0.096262	-0.117923	-0.104484	-0.092892	-0.09

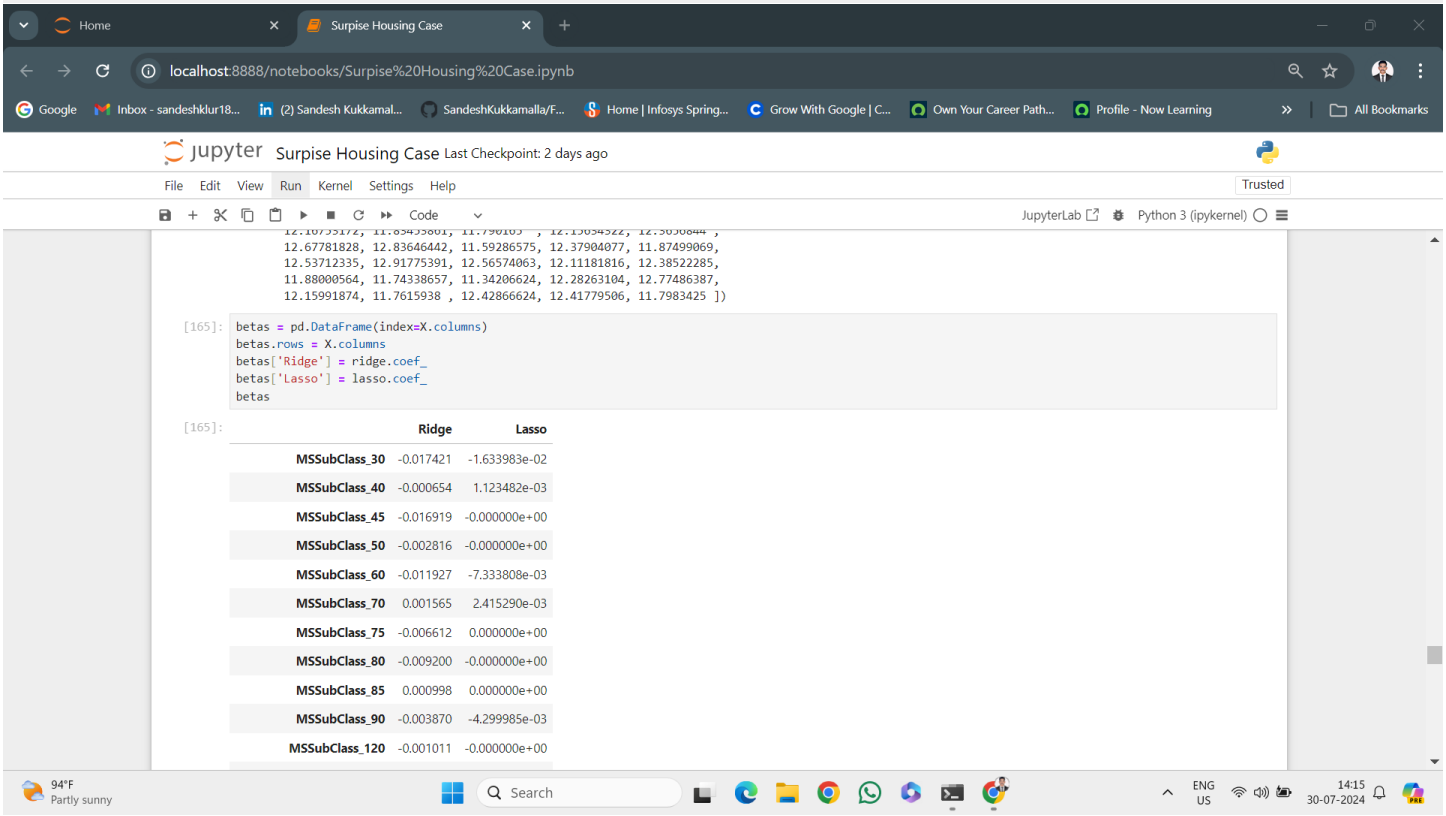
```
[160]: plt.plot(lassoCV_res['param_alpha'],lassoCV_res['mean_train_score'],label ='train')
       plt.plot(lassoCV_res['param_alpha'],lassoCV_res['mean_train_score'],label = 'test')
       plt.xlabel('alpha')
       plt.ylabel('R2_score')
       plt.xscale('log')
       plt.legend()
       plt.show()
```

Olympic Games Today's events — Search — ENG US — 14:15 30-07-2024



[162]:
```
y_pred1
```

[162]:
```
array([11.89973844, 12.73885716, 11.5237119 , 12.0297725 , 12.71885121,
       11.37240171, 12.42321289, 11.88297211, 11.32386523, 11.90258571,
       11.90734465, 11.69453459, 11.49853614, 12.30328286, 12.10279842,
       11.84370343, 12.19650988, 11.82581388, 11.58427852, 12.27385506,
       12.05346018, 12.27048059, 12.12949212, 11.84223872, 12.20915821,
       11.96959462, 12.19872523, 11.65426295, 12.11371203, 12.24565601,
       11.80729962, 12.56642009, 12.28614831, 11.71541946, 12.51904486,
       11.95009308, 11.8879836 , 12.24334488, 12.72504472, 11.58879414,
       11.81376625, 12.35944205, 11.69527458, 12.78073371, 11.79976395,
       11.83817763, 11.65590834, 11.82167812, 12.95936367, 11.77302947,
       11.73091505, 12.246054  , 11.59972485, 12.59365712, 12.06313205,
       12.35978546, 12.26168484, 11.99584944, 11.8157859 , 11.54853895,
       11.28079428, 12.08265256, 12.68082663, 12.5028462 , 12.6394763 ,
       12.19181557, 11.61583685, 12.72337478, 11.77267644, 12.04445962,
       11.69629872, 11.77114172, 11.69481308, 11.2681771 , 13.12540634,
       12.20266772, 12.63892621, 12.70331153, 11.93821363, 11.73022226,
       11.6480605 , 11.09924767, 11.67511797, 11.52154225, 11.99445816,
```

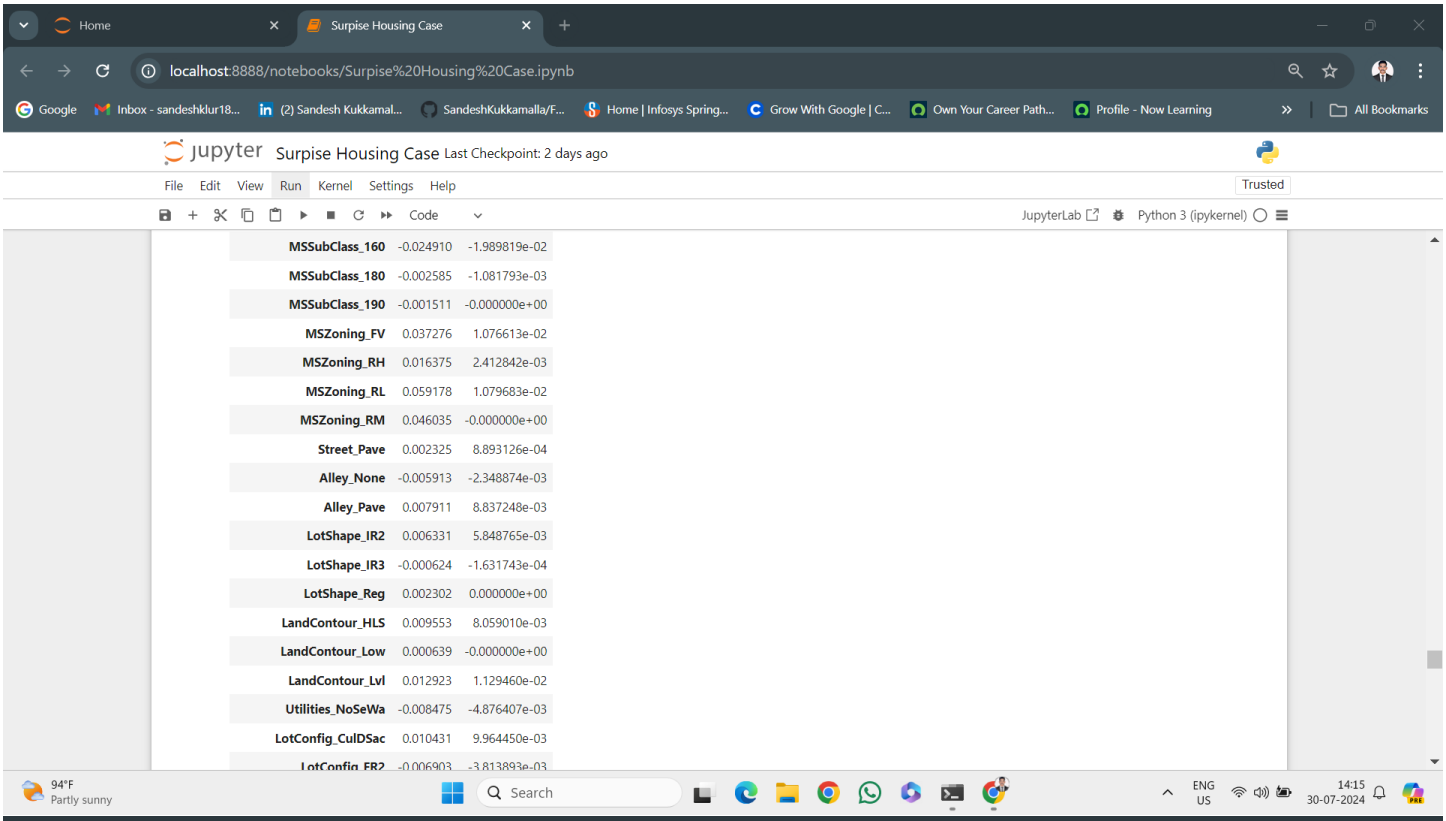



```
       12.67781828, 12.83646442, 11.59286575, 12.37904077, 11.87499069,
       12.53712335, 12.91775391, 12.56574063, 12.11181816, 12.38522285,
       11.88000564, 11.74338657, 11.34206624, 12.28263104, 12.77486387,
       12.15991874, 11.7615938 , 12.42866424, 12.41779506, 11.7983425 ])
```

[165]:
```
betas = pd.DataFrame(index=X.columns)
betas.rows = X.columns
betas['Ridge'] = ridge.coef_
betas['Lasso'] = lasso.coef_
betas
```

[165]:

	Ridge	Lasso
MSSubClass_30	-0.017421	-1.633983e-02
MSSubClass_40	-0.000654	1.123482e-03
MSSubClass_45	-0.016919	-0.000000e+00
MSSubClass_50	-0.002816	-0.000000e+00
MSSubClass_60	-0.011927	-7.333808e-03
MSSubClass_70	0.001565	2.415290e-03
MSSubClass_75	-0.006612	0.000000e+00
MSSubClass_80	-0.009200	-0.000000e+00
MSSubClass_85	0.000998	0.000000e+00
MSSubClass_90	-0.003870	-4.299985e-03
MSSubClass_120	-0.001011	-0.000000e+00



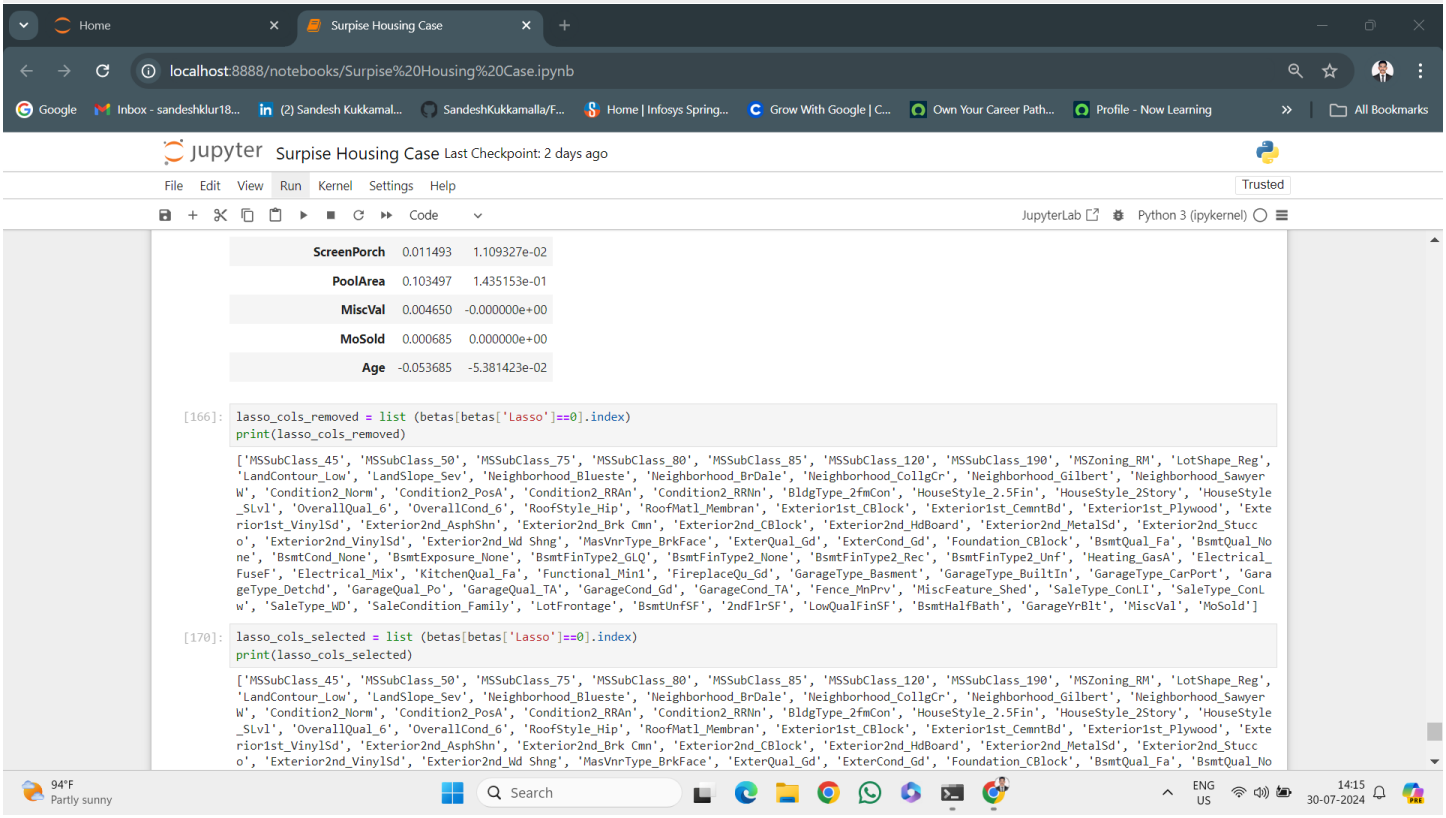
Home | Surpise Housing Case | localhost:8888/notebooks/Surpise%20Housing%20Case.ipynb

jupyter Surpise Housing Case Last Checkpoint: 2 days ago

File Edit View Run Kernel Settings Help | Trusted | Code | JupyterLab | Python 3 (ipykernel)

MSSubClass_160	-0.024910	-1.989819e-02
MSSubClass_180	-0.002585	-1.081793e-03
MSSubClass_190	-0.001511	-0.000000e+00
MSZoning_FV	0.037276	1.076613e-02
MSZoning_RH	0.016375	2.412842e-03
MSZoning_RL	0.059178	1.079683e-02
MSZoning_RM	0.046035	-0.000000e+00
Street_Pave	0.002325	8.893126e-04
Alley_None	-0.005913	-2.348874e-03
Alley_Pave	0.007911	8.837248e-03
LotShape_IR2	0.006331	5.848765e-03
LotShape_IR3	-0.000624	-1.631743e-04
LotShape_Reg	0.002302	0.000000e+00
LandContour_HLS	0.009553	8.059010e-03
LandContour_Low	0.000639	-0.000000e+00
LandContour_Lvl	0.012923	1.129460e-02
Utilities_NoSeWa	-0.008475	-4.876407e-03
LotConfig_CulDSac	0.010431	9.964450e-03
LotConfig_FR2	-0.006903	-3.813893e-03

94°F Partly sunny | Search | ENG US | 14:15 30-07-2024



Home | Surpise Housing Case | localhost:8888/notebooks/Surpise%20Housing%20Case.ipynb

jupyter Surpise Housing Case Last Checkpoint: 2 days ago

File Edit View Run Kernel Settings Help | Trusted | Code | JupyterLab | Python 3 (ipykernel)

ScreenPorch	0.011493	1.109327e-02
PoolArea	0.103497	1.435153e-01
MiscVal	0.004650	-0.000000e+00
MoSold	0.000685	0.000000e+00
Age	-0.053685	-5.381423e-02

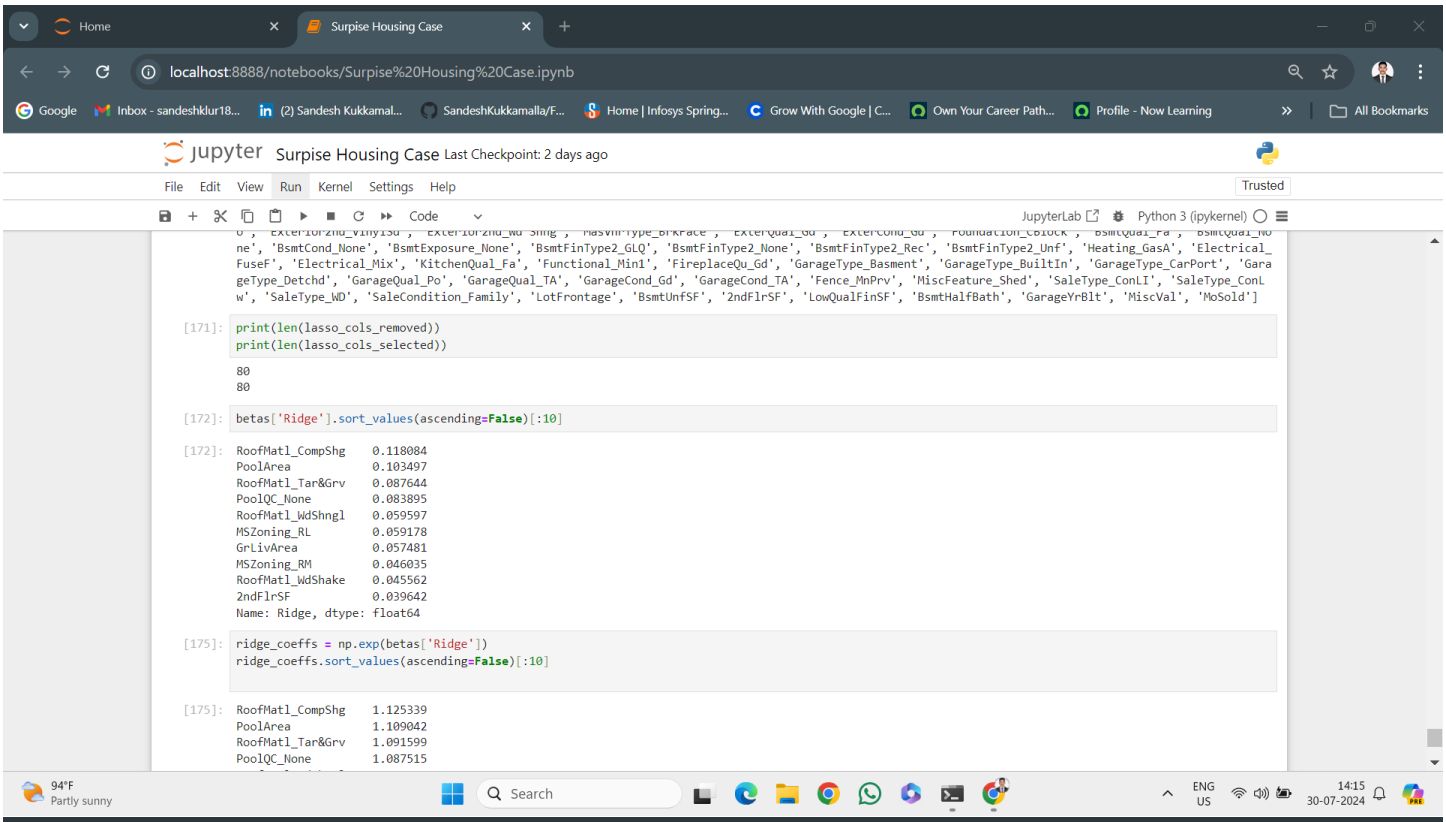
[166]:
```
lasso_cols_removed = list (betas[betas['Lasso']==0].index)
print(lasso_cols_removed)
```

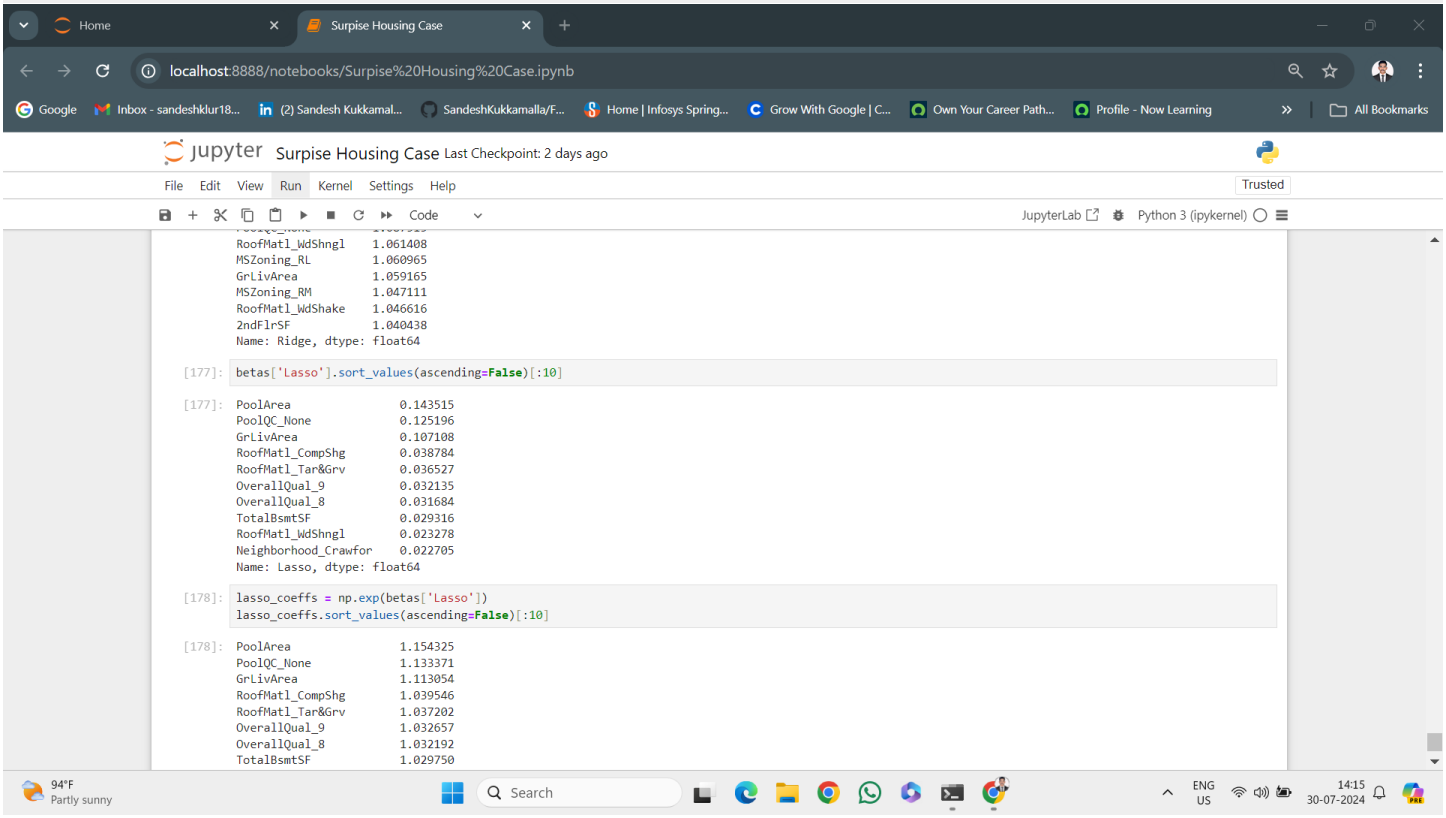
```
['MSSubClass_45', 'MSSubClass_50', 'MSSubClass_75', 'MSSubClass_80', 'MSSubClass_85', 'MSSubClass_120', 'MSSubClass_190', 'MSZoning_RM', 'LotShape_Reg',
'LandContour_Low', 'LandSlope_Sev', 'Neighborhood_Blueste', 'Neighborhood_BrDale', 'Neighborhood_CollgCr', 'Neighborhood_Gilbert', 'Neighborhood_Sawyer
W', 'Condition2_Norm', 'Condition2_PosA', 'Condition2_RRAn', 'Condition2_RRNn', 'BldgType_2fmCon', 'HouseStyle_2.5Fin', 'HouseStyle_2Story', 'HouseStyle
_SLvl', 'OverallQual_6', 'OverallCond_6', 'RoofStyle_Hip', 'RoofMatl_Membran', 'Exterior1st_CBlock', 'Exterior1st_CemntBd', 'Exterior1st_Plywood', 'Exte
rior1st_VinylSd', 'Exterior2nd_AsphShn', 'Exterior2nd_Brk Cmn', 'Exterior2nd_CBlock', 'Exterior2nd_HdBoard', 'Exterior2nd_MetalSd', 'Exterior2nd_Stucc
o', 'Exterior2nd_VinylSd', 'Exterior2nd_Wd Shng', 'MasVnrType_BrkFace', 'ExterQual_Gd', 'ExterCond_Gd', 'Foundation_CBlock', 'BsmtQual_Fa', 'BsmtQual_No
ne', 'BsmtCond_None', 'BsmtExposure_None', 'BsmtFinType2_GLQ', 'BsmtFinType2_None', 'BsmtFinType2_Rec', 'BsmtFinType2_Unf', 'Heating_GasA', 'Electrical_
FuseF', 'Electrical_Mix', 'KitchenQual_Fa', 'Functional_Min1', 'FireplaceQu_Gd', 'GarageType_Basment', 'GarageType_BuiltIn', 'GarageType_CarPort', 'Gara
geType_Detchd', 'GarageQual_Po', 'GarageQual_TA', 'GarageCond_Gd', 'GarageCond_TA', 'Fence_MnPrv', 'MiscFeature_Shed', 'SaleType_ConLI', 'SaleType_ConL
w', 'SaleType_WD', 'SaleCondition_Family', 'LotFrontage', 'BsmtUnfSF', '2ndFlrSF', 'LowQualFinSF', 'BsmtHalfBath', 'GarageYrBlt', 'MiscVal', 'MoSold']
```

[170]:
```
lasso_cols_selected = list (betas[betas['Lasso']==0].index)
print(lasso_cols_selected)
```

```
['MSSubClass_45', 'MSSubClass_50', 'MSSubClass_75', 'MSSubClass_80', 'MSSubClass_85', 'MSSubClass_120', 'MSSubClass_190', 'MSZoning_RM', 'LotShape_Reg',
'LandContour_Low', 'LandSlope_Sev', 'Neighborhood_Blueste', 'Neighborhood_BrDale', 'Neighborhood_CollgCr', 'Neighborhood_Gilbert', 'Neighborhood_Sawyer
W', 'Condition2_Norm', 'Condition2_PosA', 'Condition2_RRAn', 'Condition2_RRNn', 'BldgType_2fmCon', 'HouseStyle_2.5Fin', 'HouseStyle_2Story', 'HouseStyle
_SLvl', 'OverallQual_6', 'OverallCond_6', 'RoofStyle_Hip', 'RoofMatl_Membran', 'Exterior1st_CBlock', 'Exterior1st_CemntBd', 'Exterior1st_Plywood', 'Exte
rior1st_VinylSd', 'Exterior2nd_AsphShn', 'Exterior2nd_Brk Cmn', 'Exterior2nd_CBlock', 'Exterior2nd_HdBoard', 'Exterior2nd_MetalSd', 'Exterior2nd_Stucc
o', 'Exterior2nd_VinylSd', 'Exterior2nd_Wd Shng', 'MasVnrType_BrkFace', 'ExterQual_Gd', 'ExterCond_Gd', 'Foundation_CBlock', 'BsmtQual_Fa', 'BsmtQual_No
```

94°F Partly sunny | Search | ENG US | 14:15 30-07-2024


Home
Surpise Housing Case
localhost:8888/notebooks/Surpise%20Housing%20Case.ipynb
jupyter Surpise Housing Case Last Checkpoint: 2 days ago
File Edit View Run Kernel Settings Help
Trusted
Code
JupyterLab
Python 3 (ipykernel)
'BsmtCond_None', 'BsmtExposure_None', 'BsmtFinType2_GLQ', 'BsmtFinType2_None', 'BsmtFinType2_Rec', 'BsmtFinType2_Unf', 'Heating_GasA', 'Electrical_FuseF', 'Electrical_Mix', 'KitchenQual_Fa', 'Functional_Min1', 'FireplaceQu_Gd', 'GarageType_Basment', 'GarageType_BuiltIn', 'GarageType_CarPort', 'GarageType_Detchd', 'GarageQual_Po', 'GarageQual_TA', 'GarageCond_Gd', 'GarageCond_TA', 'Fence_MnPrv', 'MiscFeature_Shed', 'SaleType_ConLI', 'SaleType_ConLw', 'SaleType_WD', 'SaleCondition_Family', 'LotFrontage', 'BsmtUnfSF', '2ndFlrSF', 'LowQualFinSF', 'BsmtHalfBath', 'GarageYrBlt', 'MiscVal', 'MoSold']
[171]: print(len(lasso_cols_removed))
print(len(lasso_cols_selected))
80
80
[172]: betas['Ridge'].sort_values(ascending=False)[:10]
[172]: RoofMatl_CompShg 0.118084
PoolArea 0.103497
RoofMatl_Tar&Grv 0.087644
PoolQC_None 0.083895
RoofMatl_WdShngl 0.059597
MSZoning_RL 0.059178
GrLivArea 0.057481
MSZoning_RM 0.046035
RoofMatl_WdShake 0.045562
2ndFlrSF 0.039642
Name: Ridge, dtype: float64
[175]: ridge_coeffs = np.exp(betas['Ridge'])
ridge_coeffs.sort_values(ascending=False)[:10]
[175]: RoofMatl_CompShg 1.125339
PoolArea 1.109042
RoofMatl_Tar&Grv 1.091599
PoolQC_None 1.087515
94°F Partly sunny
Search
ENG US
14:15 30-07-2024


Home
Surpise Housing Case
localhost:8888/notebooks/Surpise%20Housing%20Case.ipynb
jupyter Surpise Housing Case Last Checkpoint: 2 days ago
File Edit View Run Kernel Settings Help
Trusted
Code
JupyterLab
Python 3 (ipykernel)
RoofMatl_WdShngl 1.061408
MSZoning_RL 1.060965
GrLivArea 1.059165
MSZoning_RM 1.047111
RoofMatl_WdShake 1.046616
2ndFlrSF 1.040438
Name: Ridge, dtype: float64
[177]: betas['Lasso'].sort_values(ascending=False)[:10]
[177]: PoolArea 0.143515
PoolQC_None 0.125196
GrLivArea 0.107108
RoofMatl_CompShg 0.038784
RoofMatl_Tar&Grv 0.036527
OverallQual_9 0.032135
OverallQual_8 0.031684
TotalBsmtSF 0.029316
RoofMatl_WdShngl 0.023278
Neighborhood_Crawfor 0.022705
Name: Lasso, dtype: float64
[178]: lasso_coeffs = np.exp(betas['Lasso'])
lasso_coeffs.sort_values(ascending=False)[:10]
[178]: PoolArea 1.154325
PoolQC_None 1.133371
GrLivArea 1.113054
RoofMatl_CompShg 1.039546
RoofMatl_Tar&Grv 1.037202
OverallQual_9 1.032657
OverallQual_8 1.032192
TotalBsmtSF 1.029750
94°F Partly sunny
Search
ENG US
14:15 30-07-2024

THANK YOU