

- Files
- sample\_data
- House Price India.csv.zip

```
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
dataset=pd.read_csv('/content/House Price India.csv.zip')
#printing first 5 records of the dataset
print(dataset.head(5))
```

	id	Date	number of bedrooms	number of bathrooms	Living area
0	6762810145	42491	5	2.50	3650
1	6762810635	42491	4	2.50	2920
2	6762810998	42491	5	2.75	2910
3	6762812605	42491	4	2.50	3310
4	6762812919	42491	3	2.00	2710
	lot area	number of floors	waterfront	present	number of views
0	9050	2.0	0	0	4
1	4000	1.5	0	0	0
2	9480	1.5	0	0	0
3	42998	2.0	0	0	0
4	4500	1.5	0	0	0
	condition of the house	Built Year	Renovation Year	Postal Code	
0	5	1921	0	122003	
1	5	1909	0	122004	
2	3	1939	0	122004	
3	3	2001	0	122005	
4	4	1929	0	122006	



Microphone off

Set up voice access to continue

Download

Not now

House Price Prediction.ipynb x linked in - Yahoo India Search P x (19) My Network - Grow | Link x Kondakamaria-Sunya/SunyaH x +

colab.research.google.com/drive/1W6G2P6\_ziKQOIqJ0NUSvHCVPVlV9u1#scrollto=fjz--YVWgkNl

# House Price Prediction.ipynb

File Edit View Insert Runtime Tools Help All changes saved

Files + Code + Text

sample\_data  
House Price India.csv.zip

[2] dataset.shape  
(14620, 23)

#Data processing

```
obj=(dataset.dtypes=='object')  
obj_cols=list(obj[obj].index)  
print("Categorical Variables:",len(obj_cols))  
int=(dataset.dtypes=='int')  
num_cols=list(int[int].index)  
print("Numerical Variables:",len(num_cols))  
fl=(dataset.dtypes=='float')  
fl_cols=list(fl[fl].index)  
print("Float variables:",len(fl_cols))
```

Categorical Variables: 0

Numerical Variables: 19

Float variables: 4

[6] plt.figure(figsize=(12, 6))  
sns.heatmap(dataset.corr(),

```
cmap = 'BrBG',  
fmt = '.2f',  
linewidths = 2,  
annot = True)
```

FOR - CZE  
Game score

Search

ENG  
IN

14:23  
19-06-2024



Microphone off

1 Set up voice access to continue

Download

Not now



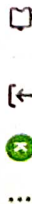
House Price Prediction.ipynb

linked in - Yahoo India Search

(19) My Network - Grow | Link

Kondakamala Surya/SuryaH

colab.research.google.com/drive/1tWG2P6\_zizKQOIq10NISvHCvPLvJ9uT#scrollto=fjz--YVWgKAI



House Price Prediction.ipynb

File Edit View Insert Runtime Tools Help All changes saved

Comment

Share

Files



+ Code + Text

RAM 1424

Disk 19-06-2024



3s

plt.figure(figsize=(12, 6))  
sns.heatmap(dataset.corr(),  
cmap = 'BrBG',  
fmt = '.2f',  
linewidths = 2,  
annot = True)

sample\_data

House Price India.csv.zip

< Axes: >

```
id - 1.00 0.05 0.33 0.52 0.65 0.100 0.31 0.110 0.29 0.050 0.67 0.57 0.29 0.070 1.10 0.29 0.44 0.07 0.60 0.09 0.00 0.07
Date - 0.05 1.00 0.02 0.03 0.02 0.00 0.01 0.01 0.00 0.03 0.03 0.02 0.02 0.01 0.01 0.02 0.02 0.02 0.03 0.00 0.01 0.03
number of bedrooms - 0.33 0.02 1.00 0.51 0.57 0.03 0.18 0.01 0.08 0.03 0.35 0.47 0.30 0.15 0.02 0.04 0.01 0.14 0.39 0.03 0.00 0.01 0.31
number of bathrooms - 0.52 0.03 0.51 1.00 0.75 0.08 0.50 0.06 0.18 0.13 0.66 0.68 0.29 0.50 0.05 0.11 0.03 0.22 0.57 0.08 0.00 0.01 0.53
living area - 0.65 0.02 0.57 0.75 1.00 0.17 0.35 0.11 0.29 0.06 0.76 0.88 0.44 0.31 0.06 0.08 0.05 0.24 0.76 0.18 0.00 0.00 0.71
lot area - 0.100 0.00 0.03 0.08 0.17 1.00 0.00 0.03 0.08 0.01 0.11 0.18 0.02 0.05 0.01 0.07 0.09 0.22 0.15 0.71 0.01 0.00 0.08
number of floors - 0.31 0.01 0.18 0.50 0.35 0.00 1.00 0.02 0.02 0.27 0.46 0.53 0.24 0.48 0.01 0.13 0.05 0.13 0.29 0.01 0.01 0.02 0.26
waterfront present - 0.11 0.01 0.01 0.06 0.11 0.03 0.02 1.00 0.40 0.02 0.08 0.07 0.09 0.02 0.09 0.04 0.02 0.05 0.09 0.03 0.00 0.00 0.26
number of views - 0.29 0.00 0.08 0.18 0.29 0.08 0.02 0.40 1.00 0.05 0.25 0.16 0.29 0.06 0.10 0.04 0.00 0.08 0.28 0.07 0.01 0.00 0.40
condition of the house - 0.05 0.03 0.03 0.13 0.06 0.01 0.27 0.02 0.05 1.00 0.15 0.17 0.18 0.33 0.06 0.05 0.00 0.12 0.10 0.00 0.01 0.00 0.04
grade of the house - 0.67 0.03 0.35 0.66 0.76 0.11 0.46 0.08 0.25 0.15 1.00 0.76 0.17 0.44 0.01 0.15 0.12 0.20 0.77 0.12 0.00 0.00 0.67
Area of the house(excluding basement) - 0.57 0.02 0.47 0.68 0.88 0.18 0.53 0.07 0.16 0.17 0.76 1.00 0.05 0.42 0.03 0.08 0.00 0.35 0.74 0.19 0.00 0.00 0.62
Area of the basement - 0.29 0.02 0.30 0.29 0.44 0.02 0.24 0.09 0.29 0.18 0.17 0.05 1.00 0.14 0.08 0.01 0.11 0.15 0.20 0.01 0.01 0.00 0.33
Built Year - 0.07 0.01 0.15 0.50 0.31 0.05 0.48 0.02 0.06 0.33 0.44 0.42 0.14 1.00 0.23 0.06 0.14 0.41 0.33 0.07 0.00 0.00 0.05
Renovation Year - 0.11 0.01 0.02 0.05 0.06 0.01 0.01 0.09 0.10 0.06 0.01 0.03 0.08 0.23 1.00 0.02 0.03 0.08 0.00 0.01 0.00 0.01 0.13
Postal Code - 0.29 0.02 0.04 0.11 0.08 0.07 0.13 0.04 0.04 0.05 0.15 0.08 0.01 0.06 0.02 1.00 0.31 0.10 0.11 0.08 0.01 0.01 0.12
Latitude - 0.44 0.02 0.01 0.03 0.05 0.09 0.05 0.02 0.00 0.00 0.12 0.00 0.11 0.14 0.03 0.31 1.00 0.13 0.05 0.09 0.01 0.01 0.30
Longitude - 0.07 0.02 0.14 0.22 0.24 0.22 0.13 0.05 0.08 0.12 0.20 0.35 0.15 0.41 0.08 0.10 0.13 1.00 0.34 0.26 0.01 0.00 0.02
living_area_renov - 0.60 0.03 0.39 0.57 0.76 0.15 0.29 0.09 0.28 0.10 0.72 0.74 0.20 0.33 0.00 0.11 0.05 0.34 1.00 0.19 0.00 0.01 0.53
```

POR - CZE  
Game score

Search





```
[16] #Replacing SalePrice empty values with their mean values to make the data distribution symmetric.
dataset['Price'] = dataset['Price'].fillna(
dataset['Price'].mean())
```

```
#Drop records with null values (as the empty records are very less).
```

```
new_dataset = dataset.dropna()
new_dataset
```

	Date	number of bedrooms	number of bathrooms	living area	lot area	number of floors	waterfront present	number of views	condition of the house	grade of the house	Built Year	Renovation Year	Postal Code	lattice
0	42491	5	2.50	3650	9050	2.0	0	4	5	10	1921	0	122003	52.8
1	42491	4	2.50	2920	4000	1.5	0	0	5	8	1909	0	122004	52.8
2	42491	5	2.75	2910	9480	1.5	0	0	3	8	1939	0	122004	52.8
3	42491	4	2.50	3310	42998	2.0	0	0	3	9	2001	0	122005	52.9
4	42491	3	2.00	2710	4500	1.5	0	0	4	8	1929	0	122006	52.9



Microphone off

1 Set up voice access to continue

Download

Not now

House Price Prediction.ipynb

linked in - Yahoo India Search

(19) My Network - Grow | Link

Kondakamarla Surya/Surya H

colab.research.google.com/drive/1nWG2P6\_zizKQOIq0NlSwHCVPlw/9ut#scrollTo=Fjz--YVWgKAI



House Price Prediction.ipynb

File Edit View Insert Runtime Tools Help All changes saved

Files



+ Code + Text

RAM  
Disk

Comment

Share

sample\_data

House Price India.csv.zip

#Checking features which have null values in the new dataframe (if there are still any).

new\_dataset.isnull().sum()

```
Date 0
number of bedrooms 0
number of bathrooms 0
living area 0
lot area 0
number of floors 0
waterfront present 0
number of views 0
condition of the house 0
grade of the house 0
Area of the house(excluding basement) 0
Area of the basement 0
Built Year 0
Renovation Year 0
Postal Code 0
Latitude 0
Longitude 0
living_area_renov 0
lot_area_renov 0
Number of schools nearby 0
Distance from the airport 0
Price
dtype: int64
```

32°C  
Haze

Search



14:24  
19-06-2024