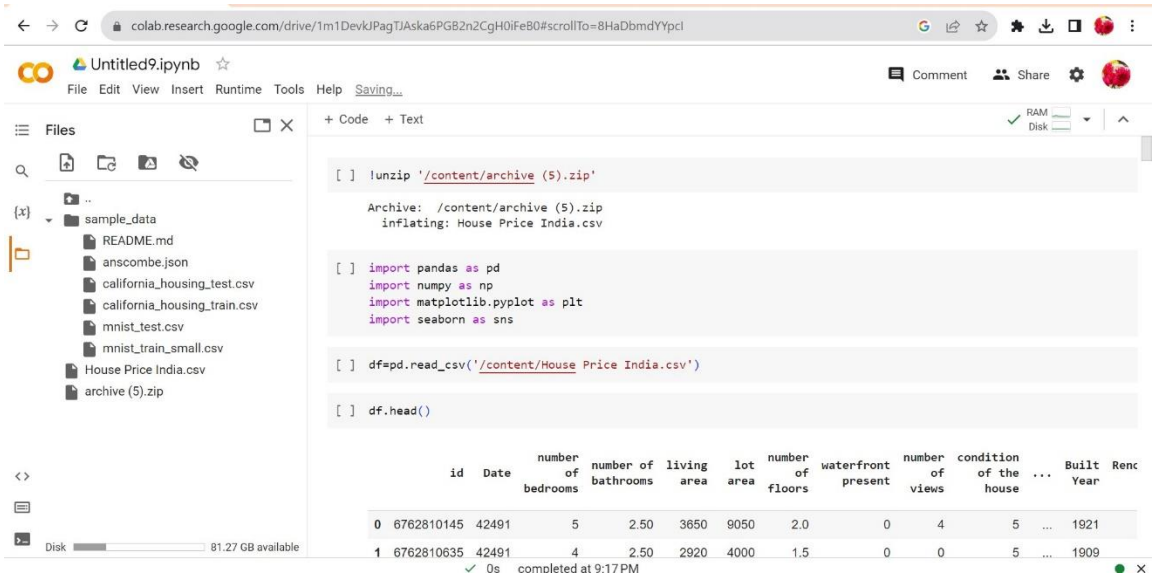
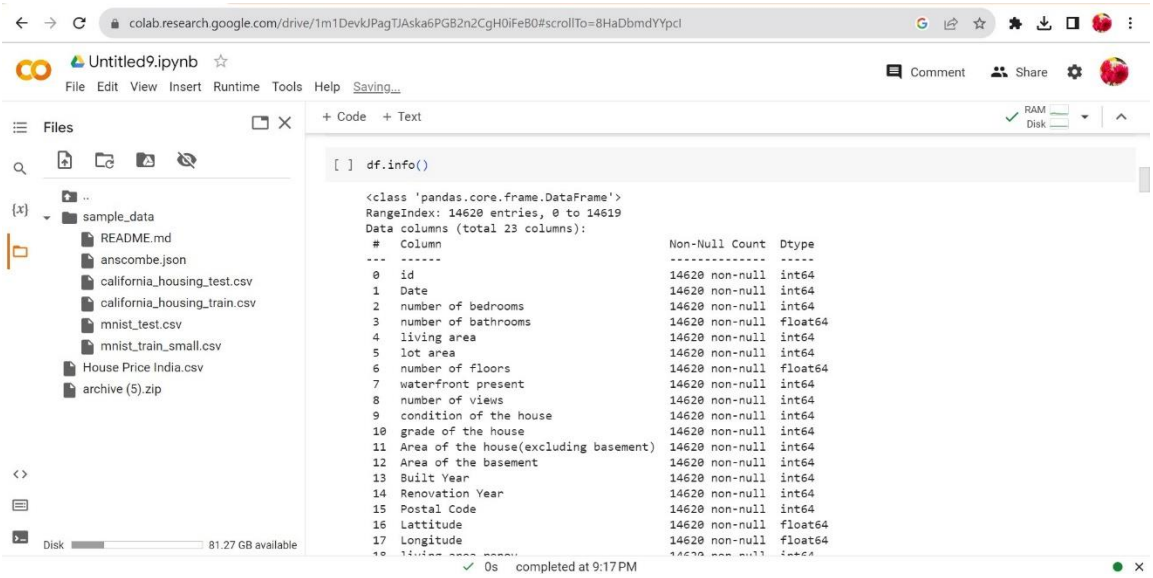
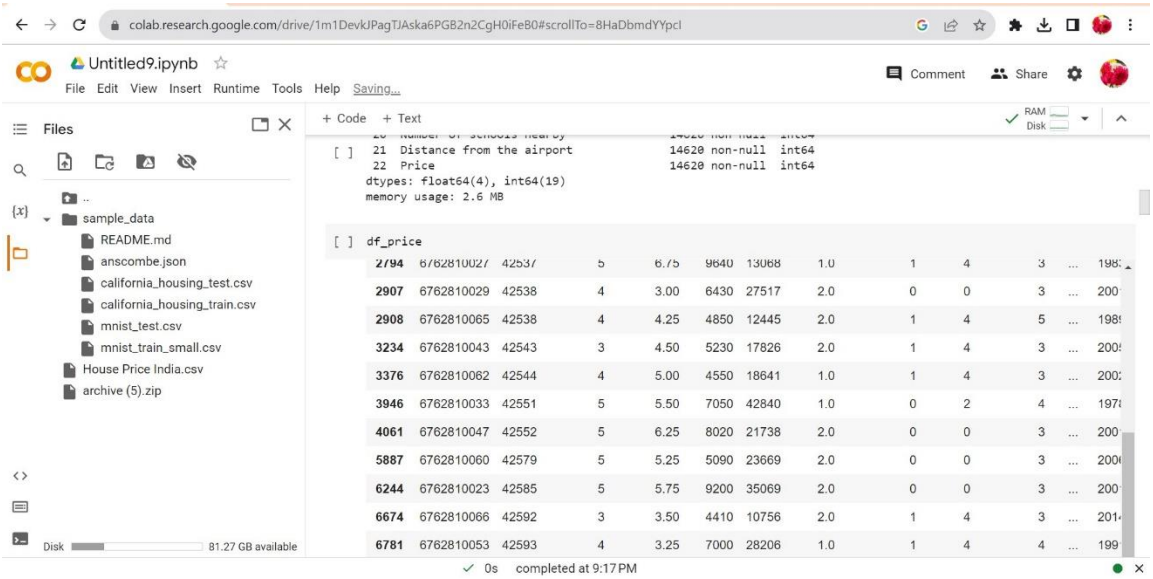
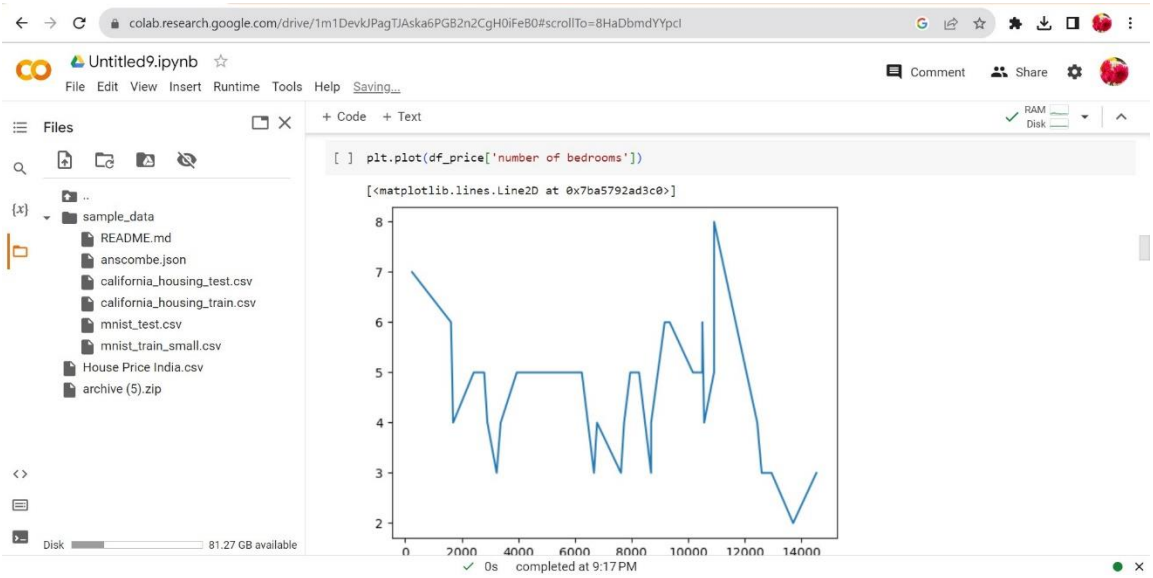
Assignment – 3

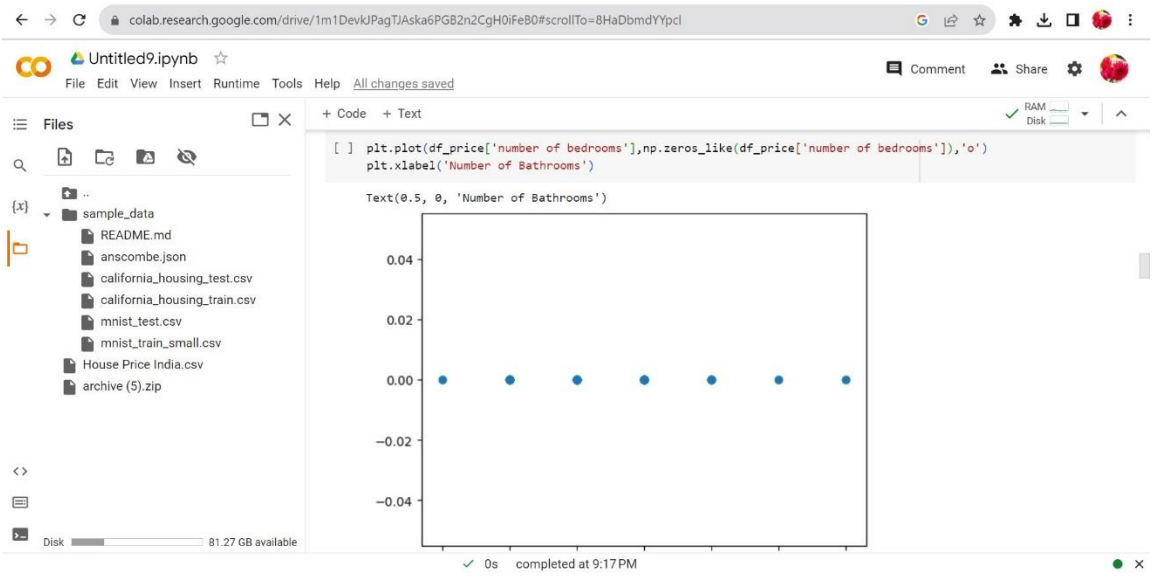
House price India

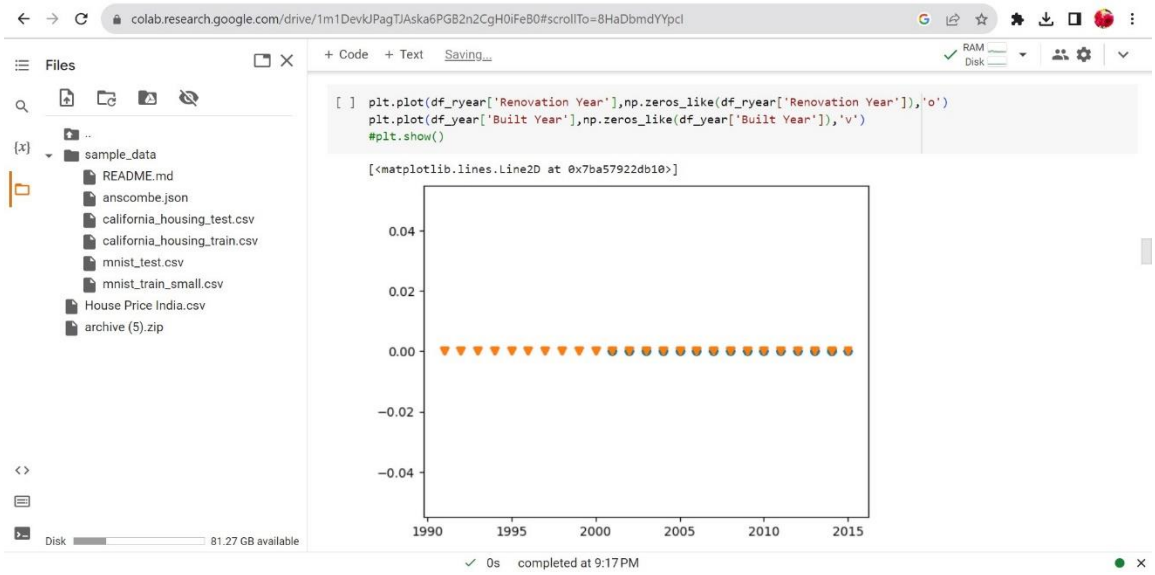


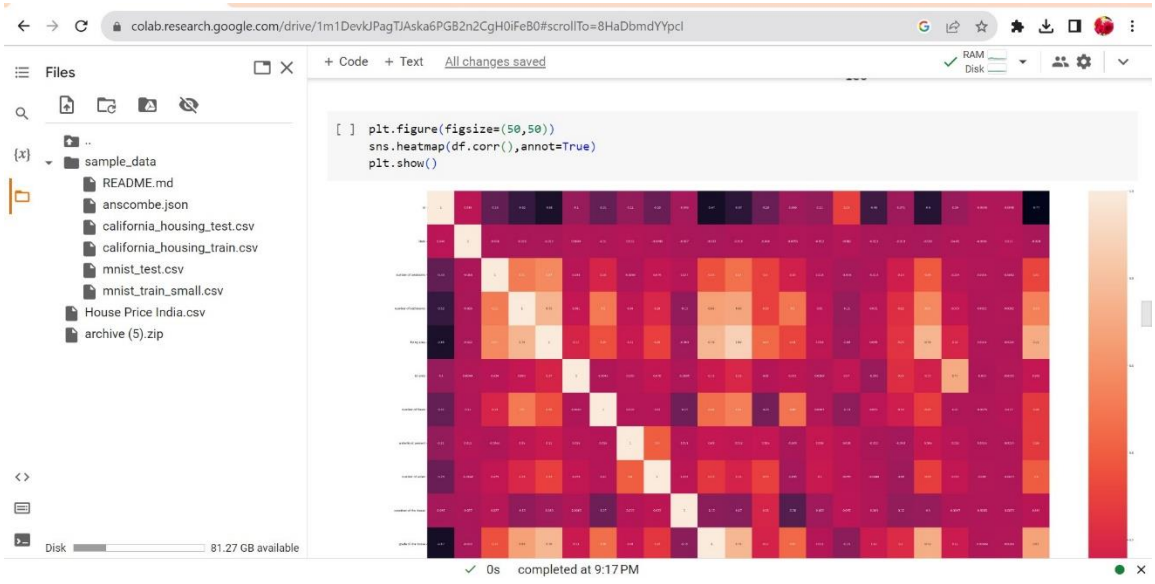

colab.research.google.com/drive/1m1DevkJPagTJAska6PGB2n2CgH0iFeB0#scrollTo=8HaDbmdYYpcI
Untitled9.ipynb
File Edit View Insert Runtime Tools Help Saving...
Comment
Share
Files
+ Code + Text
RAM
Disk
sample_data
README.md
anscombe.json
california_housing_test.csv
california_housing_train.csv
mnist_test.csv
mnist_train_small.csv
House Price India.csv
archive (5).zip
Disk 81.27 GB available
[ ] df.info()
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 14620 entries, 0 to 14619
Data columns (total 23 columns):
Column Non-Null Count Dtype
--- ------ -------------- -----
0 id 14620 non-null int64
1 Date 14620 non-null int64
2 number of bedrooms 14620 non-null int64
3 number of bathrooms 14620 non-null float64
4 living area 14620 non-null int64
5 lot area 14620 non-null int64
6 number of floors 14620 non-null float64
7 waterfront present 14620 non-null int64
8 number of views 14620 non-null int64
9 condition of the house 14620 non-null int64
10 grade of the house 14620 non-null int64
11 Area of the house(excluding basement) 14620 non-null int64
12 Area of the basement 14620 non-null int64
13 Built Year 14620 non-null int64
14 Renovation Year 14620 non-null int64
15 Postal Code 14620 non-null int64
16 Lattitude 14620 non-null float64
17 Longitude 14620 non-null float64
0s completed at 9:17 PM

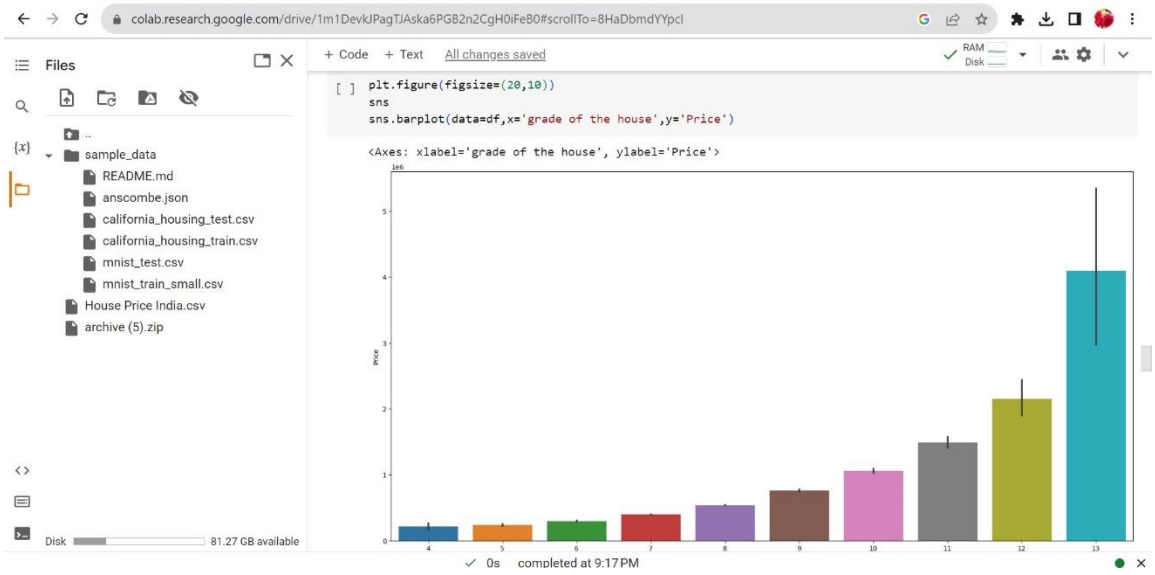

colab.research.google.com/drive/1m1DevkJPagTJAska6PGB2n2CgH0iFeB0#scrollTo=8HaDbmdYYpcI
Untitled9.ipynb
File Edit View Insert Runtime Tools Help Saving...
Comment
Share
Files
+ Code + Text
RAM
Disk
sample_data
README.md
anscombe.json
california_housing_test.csv
california_housing_train.csv
mnist_test.csv
mnist_train_small.csv
House Price India.csv
archive (5).zip
Disk 81.27 GB available
[ ] 21 Distance from the airport 14620 non-null int64
22 Price 14620 non-null int64
dtypes: float64(4), int64(19)
memory usage: 2.6 MB
[ ] df_price
2794 6762810027 42537 5 6.75 9640 13068 1.0 1 4 3 ... 198
2907 6762810029 42538 4 3.00 6430 27517 2.0 0 0 3 ... 200
2908 6762810065 42538 4 4.25 4850 12445 2.0 1 4 5 ... 198
3234 6762810043 42543 3 4.50 5230 17826 2.0 1 4 3 ... 200
3376 6762810062 42544 4 5.00 4550 18641 1.0 1 4 3 ... 200
3946 6762810033 42551 5 5.50 7050 42840 1.0 0 2 4 ... 197
4061 6762810047 42552 5 6.25 8020 21738 2.0 0 0 3 ... 200
5887 6762810060 42579 5 5.25 5090 23669 2.0 0 0 3 ... 200
6244 6762810023 42585 5 5.75 9200 35069 2.0 0 0 3 ... 200
6674 6762810066 42592 3 3.50 4410 10756 2.0 1 4 3 ... 201
6781 6762810053 42593 4 3.25 7000 28206 1.0 1 4 4 ... 199
0s completed at 9:17 PM

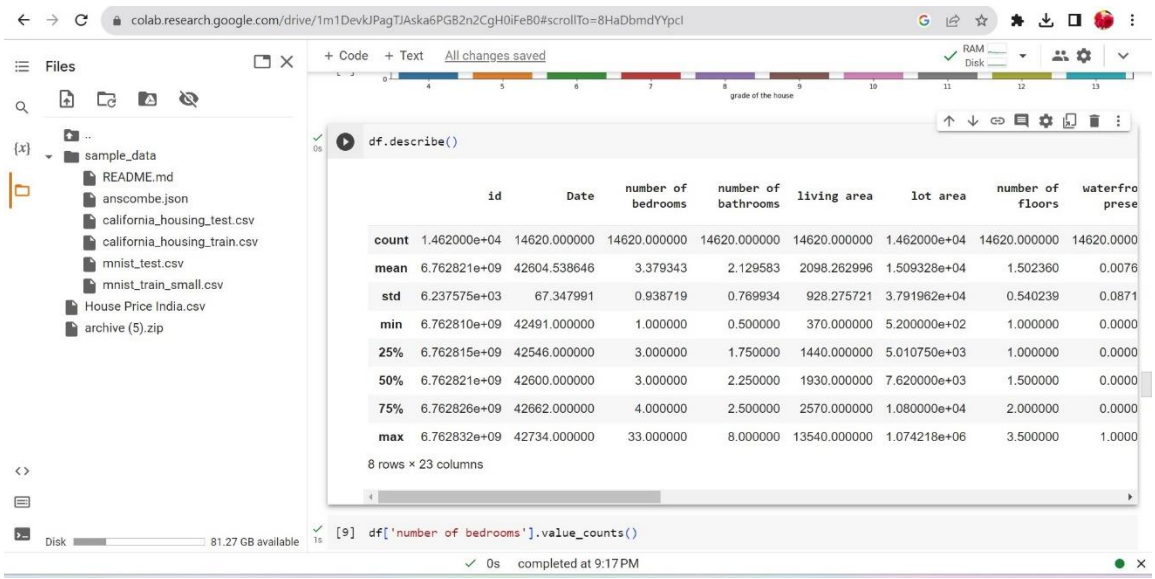

colab.research.google.com/drive/1m1DevkJPagTJAska6PGB2n2CgH0iFeB0#scrollTo=8HaDbmdYYpcI
Untitled9.ipynb
File Edit View Insert Runtime Tools Help Saving...
Comment
Share
Files
+ Code + Text
RAM
Disk
[ ] plt.plot(df_price['number of bedrooms'])
[<matplotlib.lines.Line2D at 0x7ba5792ad3c0>]
sample_data
README.md
anscombe.json
california_housing_test.csv
california_housing_train.csv
mnist_test.csv
mnist_train_small.csv
House Price India.csv
archive (5).zip
Disk
81.27 GB available
0s completed at 9:17 PM

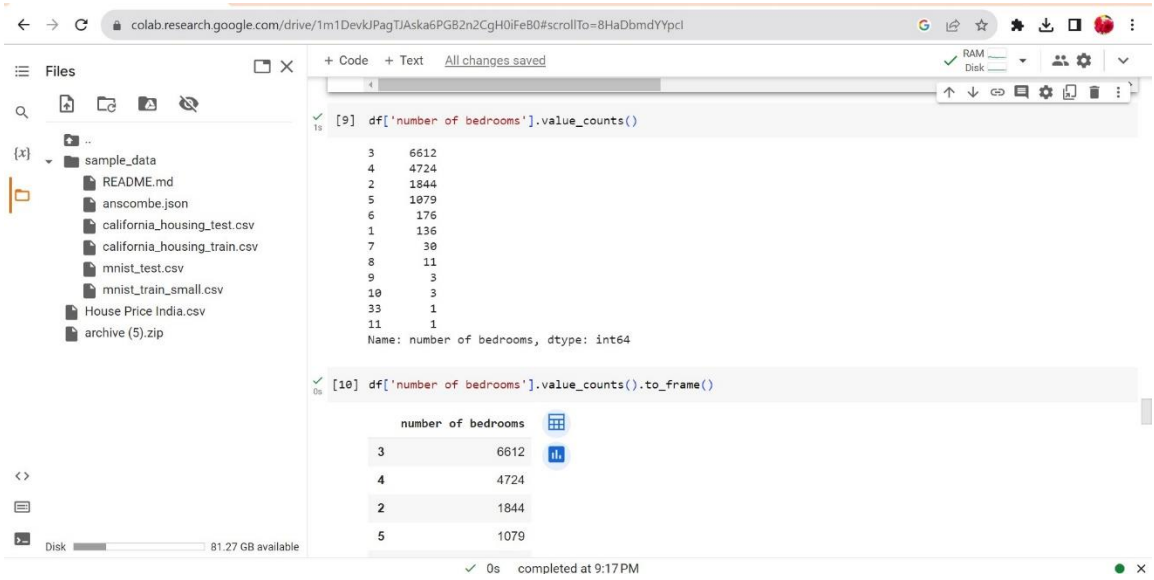

colab.research.google.com/drive/1m1DevkJPagTJAska6PGB2n2CgH0iFeB0#scrollTo=8HaDbmdYYpcI
Untitled9.ipynb
File Edit View Insert Runtime Tools Help All changes saved
Comment
Share
Files
+ Code + Text
RAM
Disk
[ ] plt.plot(df_price['number of bedrooms'],np.zeros_like(df_price['number of bedrooms']),'o')
plt.xlabel('Number of Bathrooms')
Text(0.5, 0, 'Number of Bathrooms')
sample_data
README.md
anscombe.json
california_housing_test.csv
california_housing_train.csv
mnist_test.csv
mnist_train_small.csv
House Price India.csv
archive (5).zip
Disk
81.27 GB available
0s completed at 9:17 PM

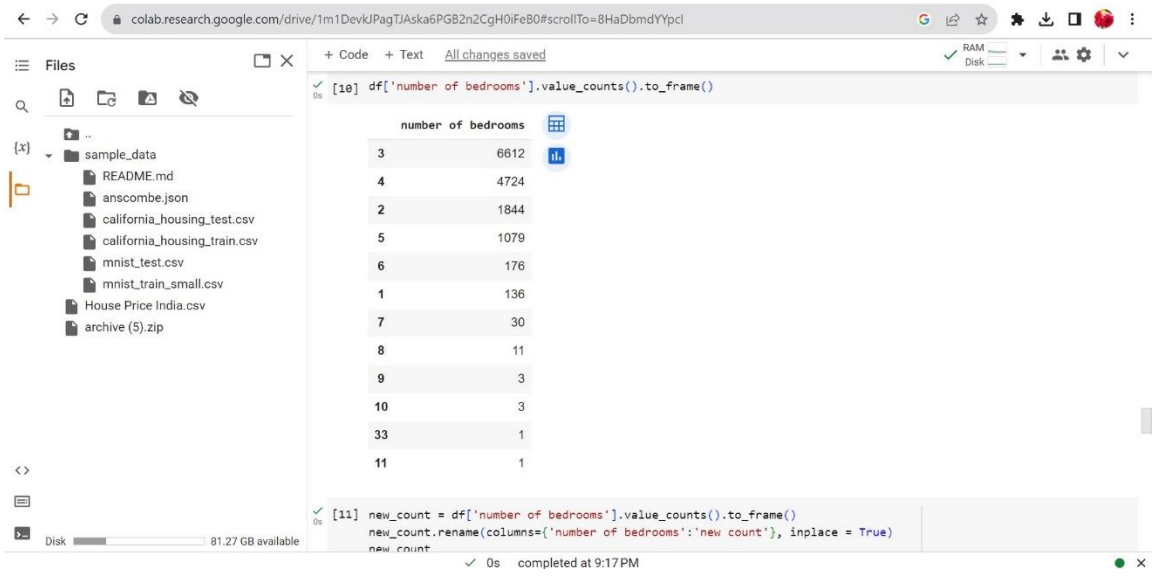

colab.research.google.com/drive/1m1DevkJPagTJAska6PGB2n2CgH0iFeB0#scrollTo=8HaDbmdYYpcI
Files
+ Code + Text Saving...
RAM
Disk
[ ] plt.plot(df_ryear['Renovation Year'],np.zeros_like(df_ryear['Renovation Year']),'o')
plt.plot(df_year['Built Year'],np.zeros_like(df_year['Built Year']),'v')
#plt.show()
[<matplotlib.lines.Line2D at 0x7ba57922db10>]
0.04
0.02
0.00
−0.02
−0.04
1990
1995
2000
2005
2010
2015
..
sample_data
README.md
anscombe.json
california_housing_test.csv
california_housing_train.csv
mnist_test.csv
mnist_train_small.csv
House Price India.csv
archive (5).zip
Disk
81.27 GB available
0s completed at 9:17 PM

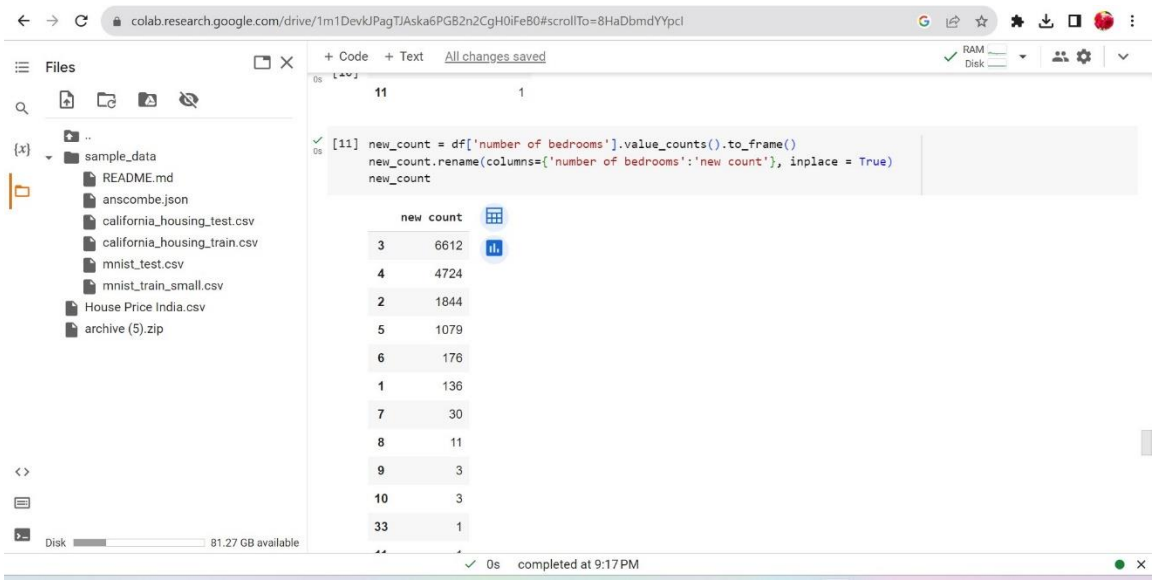

colab.research.google.com/drive/1m1DevkJPagTJAska6PGB2n2CgH0iFeB0#scrollTo=8HaDbmdYYpcI
Files
+ Code + Text All changes saved
RAM
Disk
[ ] plt.figure(figsize=(50,50))
sns.heatmap(df.corr(),annot=True)
plt.show()
..
sample_data
README.md
anscombe.json
california_housing_test.csv
california_housing_train.csv
mnist_test.csv
mnist_train_small.csv
House Price India.csv
archive (5).zip
Disk
81.27 GB available
0s completed at 9:17 PM

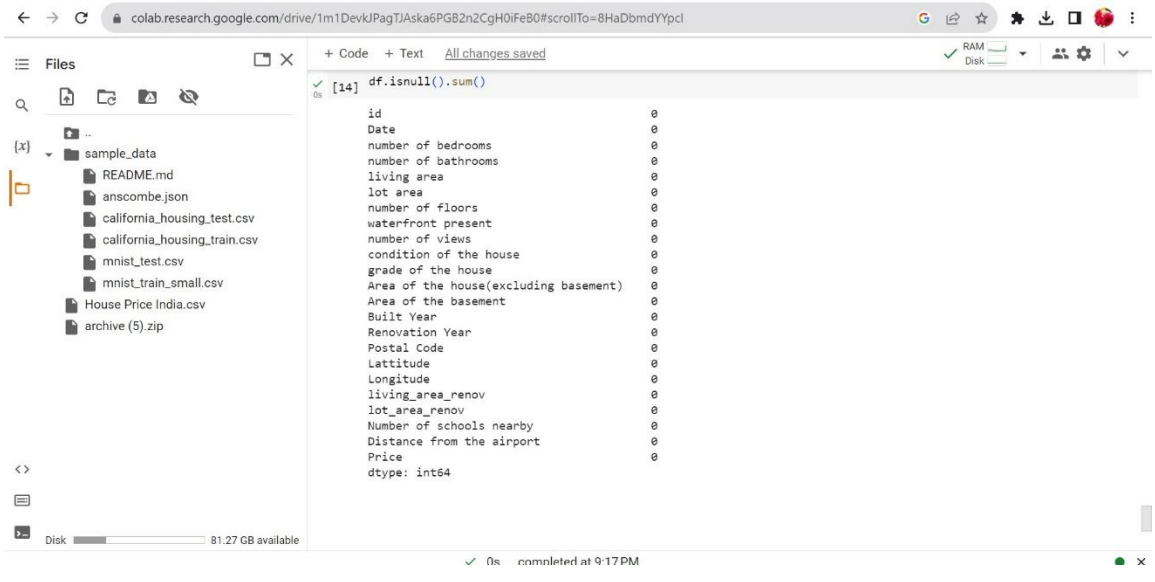

colab.research.google.com/drive/1m1DevkJPagTJAska6PGB2n2CgH0iFeB0#scrollTo=8HaDbmdYYpcI
Files
+ Code + Text All changes saved
RAM
Disk
[ ] plt.figure(figsize=(20,10))
sns
sns.barplot(data=df,x='grade of the house',y='Price')
<Axes: xlabel='grade of the house', ylabel='Price'>
1e6
Price
grade of the house
..
sample_data
README.md
anscombe.json
california_housing_test.csv
california_housing_train.csv
mnist_test.csv
mnist_train_small.csv
House Price India.csv
archive (5).zip
Disk
81.27 GB available
0s completed at 9:17 PM


colab.research.google.com/drive/1m1DevkJPagTJAska6PGB2n2CgH0iFeB0#scrollTo=8HaDbmdYYpcI
Files
+ Code + Text All changes saved
RAM
Disk
grade of the house
df.describe()
id Date number of bedrooms number of bathrooms living area lot area number of floors waterfro prese
count 1.462000e+04 14620.000000 14620.000000 14620.000000 14620.000000 1.462000e+04 14620.000000 14620.0000
mean 6.762821e+09 42604.538646 3.379343 2.129583 2098.262996 1.509328e+04 1.502360 0.0076
std 6.237575e+03 67.347991 0.938719 0.769934 928.275721 3.791962e+04 0.540239 0.0871
min 6.762810e+09 42491.000000 1.000000 0.500000 370.000000 5.200000e+02 1.000000 0.0000
25% 6.762815e+09 42546.000000 3.000000 1.750000 1440.000000 5.010750e+03 1.000000 0.0000
50% 6.762821e+09 42600.000000 3.000000 2.250000 1930.000000 7.620000e+03 1.500000 0.0000
75% 6.762826e+09 42662.000000 4.000000 2.500000 2570.000000 1.080000e+04 2.000000 0.0000
max 6.762832e+09 42734.000000 33.000000 8.000000 13540.000000 1.074218e+06 3.500000 1.0000
8 rows × 23 columns
[9] df['number of bedrooms'].value_counts()
..
sample_data
README.md
anscombe.json
california_housing_test.csv
california_housing_train.csv
mnist_test.csv
mnist_train_small.csv
House Price India.csv
archive (5).zip
Disk
81.27 GB available
0s completed at 9:17 PM


colab.research.google.com/drive/1m1DevkJPagTJAska6PGB2n2CgH0iFeB0#scrollTo=8HaDbmdYYpcI
Files
+ Code + Text All changes saved
RAM
Disk
[9] df['number of bedrooms'].value_counts()
3 6612
4 4724
2 1844
5 1079
6 176
1 136
7 30
8 11
9 3
10 3
33 1
11 1
Name: number of bedrooms, dtype: int64
[10] df['number of bedrooms'].value_counts().to_frame()
number of bedrooms
3 6612
4 4724
2 1844
5 1079
sample_data
README.md
anscombe.json
california_housing_test.csv
california_housing_train.csv
mnist_test.csv
mnist_train_small.csv
House Price India.csv
archive (5).zip
Disk 81.27 GB available
0s completed at 9:17 PM


colab.research.google.com/drive/1m1DevkJPagTJAska6PGB2n2CgH0iFeB0#scrollTo=8HaDbmdYYpcI
Files
+ Code + Text All changes saved
RAM
Disk
[10] df['number of bedrooms'].value_counts().to_frame()
number of bedrooms
3 6612
4 4724
2 1844
5 1079
6 176
1 136
7 30
8 11
9 3
10 3
33 1
11 1
[11] new_count = df['number of bedrooms'].value_counts().to_frame()
new_count.rename(columns={'number of bedrooms':'new count'}, inplace = True)
new_count
sample_data
README.md
anscombe.json
california_housing_test.csv
california_housing_train.csv
mnist_test.csv
mnist_train_small.csv
House Price India.csv
archive (5).zip
Disk 81.27 GB available
0s completed at 9:17 PM


colab.research.google.com/drive/1m1DevkJPagTJAska6PGB2n2CgH0iFeB0#scrollTo=8HaDbmdYYpcI
Files
+ Code + Text All changes saved
RAM
Disk
11 1
[11] new_count = df['number of bedrooms'].value_counts().to_frame()
new_count.rename(columns={'number of bedrooms':'new count'}, inplace = True)
new_count
new count
3 6612
4 4724
2 1844
5 1079
6 176
1 136
7 30
8 11
9 3
10 3
33 1
sample_data
README.md
anscombe.json
california_housing_test.csv
california_housing_train.csv
mnist_test.csv
mnist_train_small.csv
House Price India.csv
archive (5).zip
Disk 81.27 GB available
0s completed at 9:17 PM


colab.research.google.com/drive/1m1DevkJPagTJAska6PGB2n2CgH0iFeB0#scrollTo=8HaDbmdYYpcI
Files
+ Code + Text All changes saved
RAM
Disk
[14] df.isnull().sum()
id 0
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
Lattitude 0
Longitude 0
living_area_renov 0
lot_area_renov 0
Number of schools nearby 0
Distance from the airport 0
Price 0
dtype: int64
sample_data
README.md
anscombe.json
california_housing_test.csv
california_housing_train.csv
mnist_test.csv
mnist_train_small.csv
House Price India.csv
archive (5).zip
Disk 81.27 GB available
0s completed at 9:17 PM