Peer-graded Assignment: Submit your Project and Review Others

Reviews 2 left to complete

It looks like this is your first peer-graded assignment. Learn more

Links or files from other learners aren't verified for security by Coursera. If a file appears suspicious, use your preferred antivirus software before opening.

House Sale

KB by Karm Bhatt
December 28, 2022

Like Flag this submission

PROMPT

Question 1) Display the data types of each column using the attribute *dtypes*, then take a screenshot and submit it, include your code in the image.

```
In [15]: df.dtypes

Out[15]: id                int64
         date             object
         price           float64
         bedrooms          int64
         bathrooms       float64
         sqft_living       int64
         sqft_lot          int64
         floors          float64
         waterfront        int64
         view              int64
         condition         int64
         grade             int64
         sqft_above        int64
         sqft_basement     int64
         yr_built          int64
         yr_renovated      int64
         zipcode           int64
         lat             float64
         long            float64
         sqft_living15     int64
         sqft_lot15        int64
         dtype: object
```

RUBRIC

1 a) Does the assignment use the data attribute **dtypes** to get the data type and display the following image

```
Unnamed: 0         int64
id                 int64
date              object
price            float64
bedrooms         float64
bathrooms        float64
sqft_living        int64
sqft_lot           int64
floors           float64
waterfront         int64
view               int64
condition          int64
grade              int64
sqft_above         int64
sqft_basement      int64
yr_built           int64
yr_renovated       int64
zipcode            int64
lat              float64
long             float64
sqft_living15      int64
sqft_lot15         int64
dtype: object
```

- ○ 0 pts
  No
- ◉ 1 pt
  Yes

PROMPT

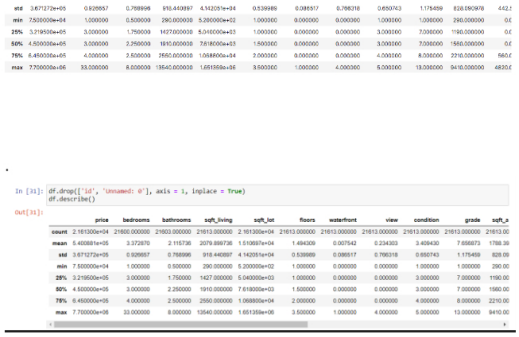
Question 2) Drop the columns *"id"* and *"Unnamed: 0"* from axis 1 using the method *drop()*, then use the method *describe()* to obtain a statistical summary of the data. Take a screenshot and submit it, make sure the *inplace* parameter is set to *True*. Your output should look like this:

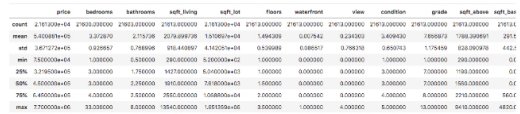
	price	bedrooms	bathrooms	sqft_living	sqft_lot	floors	waterfront	view	condition	grade	sqft_above	sqft_bas
count	2.161300e+04	21600.000000	21603.000000	21613.000000	2.161300e+04	21613.000000	21613.000000	21613.000000	21613.000000	21613.000000	21613.000000	21613.
mean	5.400881e+05	3.372870	2.115736	2079.899736	1.510697e+04	1.494309	0.007542	0.234303	3.409430	7.656873	1788.390691	291.

RUBRIC

Question 2) Does the assignment drop the columns *"id"* and *"Unnamed: 0"*

use the method **describe()** to obtain a statistical summary of the dataframe and produce the result. Note the missing collumns





- 0 pts
  No
- 1 pt
  Partially complete
- 2 pts
  Yes

PROMPT

Question 3) use the method value_counts to count the number of houses with unique floor values, use the method .to_frame() to convert it to a dataframe. Your output should look like this :

	floors
1.0	10680
2.0	8241
1.5	1910
3.0	613
2.5	161
3.5	8

Y

```
In [45]: df.floors.value_counts().to_frame()
Out[45]:
```

	floors
1.0	10680
2.0	8241
1.5	1910
3.0	613
2.5	161
3.5	8

RUBRIC

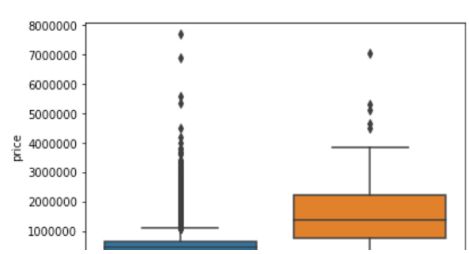
Question 3) does the assignment use the method **value_counts** to count the number of houses with unique floor values, then produce the following plot:

	floors
1.0	10680
2.0	8241
1.5	1910
3.0	613
2.5	161
3.5	8

- 0 pts
  No
- 0.5 pts
  They used the method **value_counts** on the correct column but did not produce the plot.
- 1 pt
  Yes

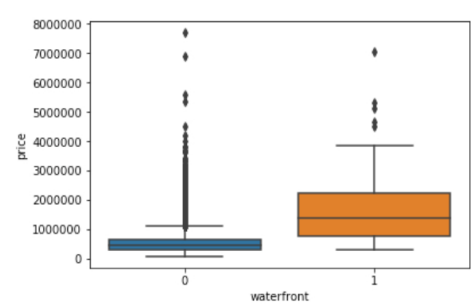
PROMPT

Question 4) use the function *boxplot* in the seaborn library to produce a plot that can be used to determine whether houses with a waterfront view or without a waterfront view have more price outliers. Your output should look like this with the code that produced it (the colors may be different) :



RUBRIC

Question 4) was the following plot produced:

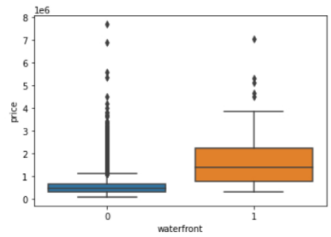


- 0 pts

.

```
sns.boxplot(x="waterfront", y="price", data=df)
```

```
<matplotlib.axes._subplots.AxesSubplot at 0x2cb2a432bb0>
```



.

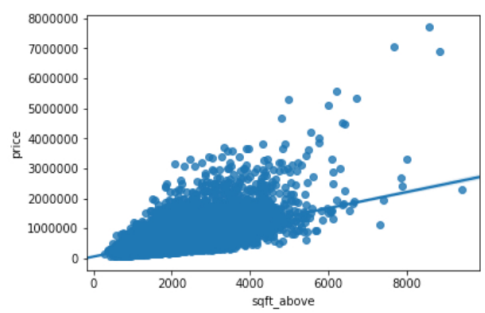
0 pts
No

1 pt
Yes

PROMPT

Question 5) Use the function *regplot* in the seaborn library to determine if the feature sqft_above is negatively or positively correlated with price. Take a screenshot of the plot and the code used to generate it.

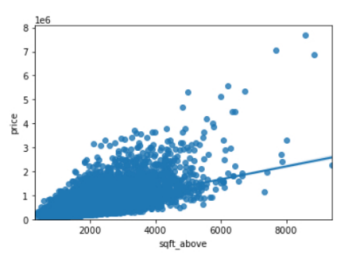
Your output should look like this with the code that produced it :



.

```
sns.regplot(x=df["sqft_above"],y=df["price"],data=df)
plt.ylim(0,)
```

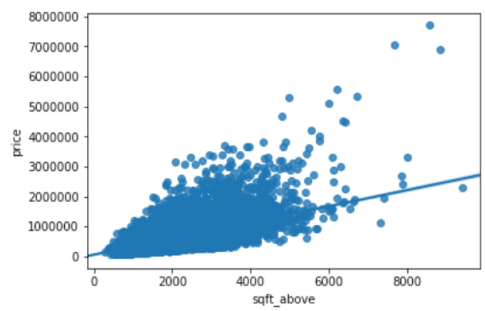
```
(0.0, 8081250.0)
```



.

RUBRIC

Question 5) Was the following plot produced?



0 pts
No

1 pt
Yes

PROMPT

Question 6) Fit a linear regression model to predict the price using the feature 'sqft_living' then calculate the R^2. Take a screenshot of your code and the value of the R^2.

..

```
n [60]: x = df[['sqft_living']]
        y = df.price
```

RUBRIC

Question 6) Was a linear regression model fit and a R^2 of approximately 0.49285 calculated?

0 pts
No

```
lr = LinearRegression()
lr.fit(x, y)
lr.score(x, y)
```

Out[60]: 0.49285321790379316

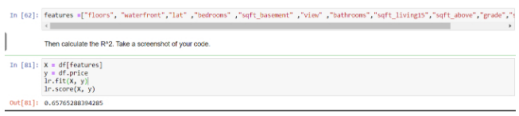
- 1 pt
  Partially complete
- 2 pts
  Yes

PROMPT

Question 7) Fit a linear regression model to predict the 'price' using the list of features:

- "floors"
- "waterfront"
- "lat"
- "bedrooms"
- "sqft_basement"
- "view"
- "bathrooms"
- "sqft_living15"
- "sqft_above"
- "grade"
- "sqft_living"

The calculate the R^2. Take a screenshot of your code and the value of the R^2.



RUBRIC

Question 7) Was a linear regression model fit and a R^2 of approximately 0.657695 calculated?

- 0 pts
  No
- 1 pt
  Partially complete
- 2 pts
  Yes

PROMPT

Question 8) Create a pipeline object that scales the data performs a polynomial transform and fits a linear regression model. Fit the object using the features in the question above, then fit the model and calculate the R^2. Take a screenshot of your code and the R^2.

There are some hints in the notebook

```
In [89]: X = df[features]
         y = df.price
         pipe = Pipeline(Input)
         pipe.fit(X, y)
         pipe.score(X, y)

Out[89]: 0.7473257069312564
```

RUBRIC

Question 8) Was an R^2 of approximately 0.75133 calculated?

- 0 pts
  No
- 1 pt
  Partially complete
- 2 pts
  Yes

PROMPT

Question 9) Create and fit a Ridge regression object using the training data, setting the regularization parameter to 0.1 and calculate the R^2 using the test data. Take a screenshot for your code and the R^2

```
from sklearn.linear_model import Ridge
```

```
rm = Ridge(alpha=0.1)
rm.fit(x_train, y_train)
```

RUBRIC

Question 9) Was the R^2 of approximately 0.647?

- 0 pts
  No
- 1 pt
  Partially complete

```
rm.score(x_test, y_test)
```

```
0.6478759163939112
```

.

2 pts
Yes

PROMPT

Question 10) Perform a second order polynomial transform on both the training data and testing data. Create and fit a Ridge regression object using the training data, setting the regularisation parameter to 0.1. Calculate the R^2 utilising the test data provided. Take a screenshot of your code and the R^2.

.

```
pr = PolynomialFeatures(degree = 2)
X_train_pr = pr.fit_transform(x_train)
X_test_pr = pr.fit_transform(x_test)

rr = Ridge(alpha = 0.1)
rr.fit(X_train_pr, y_train)
rr.score(X_test_pr, y_test)
```

```
0.7002744255607272
```

.

RUBRIC

Question 10) Was the R^2 of approximately 0.7?

- 0 pts
  No
- 1 pt
  Yes

PROMPT

Upload your notebook.

..

```
In [15]: df.dtypes

Out[15]: id                 int64
         date              object
         price            float64
         bedrooms           int64
         bathrooms        float64
         sqft_living        int64
         sqft_lot           int64
         floors           float64
         waterfront         int64
         view               int64
         condition          int64
         grade              int64
         sqft_above         int64
         sqft_basement      int64
         yr_built           int64
         yr_renovated       int64
         zipcode            int64
         lat              float64
         long             float64
         sqft_living15      int64
         sqft_lot15         int64
         dtype: object
```

.

RUBRIC

Did the user share their notebook?

- 0 pts
  No
- 3 pts
  Yes

Submit Review

Comments

Comments left for the learner are visible only to that learner and the person who left the comment.

RR

Share your thoughts...

Like Dislike Report an issue