PANDAS ASSIGNMENT QUIZ

How many houses have a waterfront?

Marked Answer :

163

Correct Answer :

163

| | MARKS OBTAINED ✔ 1 | TOTAL MARKS : 1 |

Which zip code has the costliest house?

Marked Answer :

98102

Correct Answer :

98102

| | MARKS OBTAINED ✔ 1 | TOTAL MARKS : 1 |

How many houses are having grade 10?

Marked Answer :

1134

Correct Answer :

1134

| | MARKS OBTAINED ✔ 1 | TOTAL MARKS : 1 |

How many null values are there in the dataset?

Marked Answer :

0

Correct Answer :

0

| | MARKS OBTAINED ✔ 1 | TOTAL MARKS : 1 |

Does this 9126100861 customer have a waterfront?

Marked Answer :

No

Correct Answer :

No,

| | MARKS OBTAINED ✔ 1 | TOTAL MARKS : 1 |

How many houses have 3 views?

Marked Answer :

510

Correct Answer :

510

| | MARKS OBTAINED ✔ 1 | TOTAL MARKS : 1 |

What is the lowest price of the house?

Marked Answer :

75000

Correct Answer :

75000

| | MARKS OBTAINED ✔ 1 | TOTAL MARKS : 1 |

Which zip code area has the cheapest house?

Marked Answer :

98022

Correct Answer :

98022

MARKS OBTAINED ✔ 1   TOTAL MARKS : 1

Which area has the biggest sqft_living?

Marked Answer :

98053

Correct Answer :

98053

MARKS OBTAINED ✔ 1   TOTAL MARKS : 1

In which year the costliest house has been built?

Marked Answer :

1910

Correct Answer :

1910

MARKS OBTAINED ✔ 1   TOTAL MARKS : 1

Which of the following data types can a Pandas Series have?

Marked Answer :

All of the Above

Correct Answer :

All of the Above

MARKS OBTAINED ✔ 1   TOTAL MARKS : 1

Which of the following is having a one-dimensional array?

Marked Answer :

Series data type

Correct Answer :

Series data type

MARKS OBTAINED ✔ 1   TOTAL MARKS : 1

What value will you get on the left side after printing a series format data?

Marked Answer :

Index

Correct Answer :

Index

MARKS OBTAINED ✔ 1   TOTAL MARKS : 1

Which of the following is considered as data in pandas?

Marked Answer :

All of the above

Correct Answer :

All of the above

MARKS OBTAINED ✔ 1   TOTAL MARKS : 1

In what format the keys will get converted into when we convert the dictionary data into data frame format?

Marked Answer :

Column

Correct Answer :

Column

MARKS OBTAINED ✔ 1   TOTAL MARKS : 1

Write a function to perform the following operation:
Increase the grade of the house by 1 if the sqft of the house is greater than 0 and less than equal to 400. If the total sqft of the house is greater than 400, increase the grade of the house by 2.
After creating the above function, which syntax will be using when using apply function to run the above created function in the dataset?

Marked Answer :

data.apply(your_function_name(arg1))

Correct Answer :

None of the above

MARKS OBTAINED ✖ 0   TOTAL MARKS : 1

How to check the duplicate values on id,grade and location?

Marked Answer :

df.duplicated(subset=['id','zipcode','grade'])

Correct Answer :

df[df.duplicated(['id','zipcode','grade'])]

MARKS OBTAINED ✖ 0   TOTAL MARKS : 1

Which of the following code will help to display the 3rd, 4th and 5th rows from the 6th to 9th columns of data frame data?

Marked Answer :

data.iloc[3:6,6:10]

Correct Answer :

data.iloc[3:6,6:10]

MARKS OBTAINED ✔ 1   TOTAL MARKS : 1

Which of the following syntax will display the last two records of df?
import pandas as pd
df = pd.DataFrame({'A':[34, 78, 54], 'B':[12, 67, 43]}, index=['r1', 'r2', 'r3'])

Marked Answer :

df.loc['r2':'r3']

Correct Answer :

df.loc['r2':'r3']

MARKS OBTAINED ✔ 1   TOTAL MARKS : 1

Which of the following is/are true about loc in pandas:

Marked Answer :

All of the above

Correct Answer :

All of the above

MARKS OBTAINED ✔ 1   TOTAL MARKS : 1

Change the date column in the format (yyyy/mm/dd) using the pandas to_datetime()
Function.

Marked Answer :

pd.to_datetime(data['date'], format='%Y-%m-%d', utc=False, dayfirst=True) & pd.to_datetime(data['date'])

Correct Answer :

pd.to_datetime(data['date'], format='%Y-%m-%d', utc=False, dayfirst=True) & pd.to_datetime(data['date'])

MARKS OBTAINED ✔ 1   TOTAL MARKS : 1

Create a separate data frame that satisfies the conditions below.
1. Houses built before 1980
2. Have more than 2 bedrooms
3. Have more than 2 floors.

Marked Answer :

data.loc[(data["yr_built"] < 1980) & (data['floors'] > 2) & (data['bedrooms'] > 2)]

Correct Answer :

data.loc[(data["yr_built"] < 1980) & (data['floors'] > 2) & (data['bedrooms'] > 2)]

MARKS OBTAINED ✔ 1   TOTAL MARKS : 1

For a given nested list, convert the same into a dataframe.

sample_list = [['Carl', 22],

['Martha', 25],

['Calvin', 12],

['Stuart', 15]

]

The resulting dataframe must contain the column names as 'Name', and 'Age' with the

respective values from the sample_list.

Marked Answer :

pd.DataFrame(sample_list, columns=['Name', 'Age'])

Correct Answer :

pd.DataFrame(sample_list, columns=['Name', 'Age'])

MARKS OBTAINED ✔ 1   TOTAL MARKS : 1

For a given dictionary, convert the same into a dataframe.
sample_dict = {'Cristiano': ['Ronaldo','Man U', 801],
'Lionel': ['Messi','PSG', 758],
'Luis': ['Suarez','Atletico Madrid', 509],
'Robert': ['Lewandowski','Bayern Munich', 527],
'Zlatan': ['Ibrahimovic','AC Milan',553]
}

Marked Answer :

df1 = pd.DataFrame(sample_dict) df1 = df1.transpose() df1.reset_index(inplace = True) df1.columns = ['First Name','Last Name', 'Club', 'Goals']

Correct Answer :

df1 = pd.DataFrame(sample_dict) df1 = df1.transpose() df1.reset_index(inplace = True) df1.columns = ['First Name','Last Name', 'Club', 'Goals']

MARKS OBTAINED ✔ 1   TOTAL MARKS : 1

For a given tuple, convert the same into a dataframe.

sample_tuple = ([1, 'one', 3],

[2, 'two', 3],

[3, 'Three', 5],

[4, 'Four', 4],

[5, 'Five', 4])

Marked Answer :

pd.DataFrame(sample_tuple, columns=['Number', 'Number_text', 'txtlen'])

Correct Answer :

pd.DataFrame(sample_tuple, columns=['Number', 'Number_text', 'txtlen'])

MARKS OBTAINED ✔ 1   TOTAL MARKS : 1

Create a separate dataframe that contains houses ordered in ascending or descending
order of the prices of each house.

Marked Answer :

ascending = housing.sort_values('price', ascending=True) descending = housing.sort_values('price', ascending=False)

Correct Answer :

ascending = housing.sort_values('price', ascending=True) descending = housing.sort_values('price', ascending=False)

MARKS OBTAINED ✔ 1   TOTAL MARKS : 1

Calculate the mean and standard deviation of all the numerical values in the dataset. For
example – the mean for the bedrooms column is 3.370 and the standard deviation is 0.930.

Marked Answer :

housing.describe()

Correct Answer :

housing.describe()

MARKS OBTAINED ✔ 1   TOTAL MARKS : 1

Perform the following operations on the pandas dataframe.

1. Get the data starting from the indexes 25 to 35.

2. Get the price of the houses located at the longitude–122.045 latitude- 47.6168.

Marked Answer :

data[25:36], data.loc[(housing["lat"] == 47.6168) & (data["long"] == -122.045)]

Correct Answer :

data[25:36], data.loc[(housing["lat"] == 47.6168) & (data["long"] == -122.045)]

MARKS OBTAINED ✔ 1   TOTAL MARKS : 1

Create a new column with the floor area(sqft_living, sqft_lot, sqft_above, sqft_basement, all
combined in one column).

Marked Answer :

df['Floor Area'] = data['sqft_living'] + data['sqft_lot'] + data['sqft_basement'] + data['sqft_above']

Correct Answer :

df['Floor Area'] = data['sqft_living'] + data['sqft_lot'] + data['sqft_basement'] + data['sqft_above']

MARKS OBTAINED ✔ 1   TOTAL MARKS : 1

Perform the following operations on the dataframes given below.

A = pd.DataFrame([['Carl', 22],['Martha', 25],['Calvin', 12],['Stuart', 15]], columns=['Name', 'Age'])

B = pd.DataFrame([['Melvin', 25],['Martha', 34],['Lewis', 32],['Leo', 25]], columns=['Name', 'Age'])

1. Left Outer Join

2. Outer Join

3. Inner Join

4. Right Outer Join

Marked Answer :

inner = pd.merge(A, B, on='Age', how='inner') outer = pd.merge(A, B, on='Age', how='outer') left_outer = pd.merge(A, B, on='Age', how='left') right_outer = pd.merge(A, B, on='Age', how='right')

Correct Answer :

inner = pd.merge(A, B, on='Age', how='inner') outer = pd.merge(A, B, on='Age', how='outer') left_outer = pd.merge(A, B, on='Age', how='left') right_outer = pd.merge(A, B, on='Age', how='right')

| | MARKS OBTAINED ✔ 1 | TOTAL MARKS : 1 |

What will be the correlation between the columns sqft_living and sqft_above?

Marked Answer :

0.876597

Correct Answer :

0.876597

| | MARKS OBTAINED ✔ 1 | TOTAL MARKS : 1 |

If the correlation between the columns bathrooms and sqft_living is 0.754665, what all interpretations can be made about the two columns?

Marked Answer :

A positive correlation between the two columns

Correct Answer :

A positive correlation between the two columns

| | MARKS OBTAINED ✔ 1 | TOTAL MARKS : 1 |

Total Marks   30 / 32