MIT ACDAMEY OF ENGINEERING

ESSENTIAL OF DATA SCIENCE

Topic : Houseprice Dataset

Name: Ashwini Arun Atole

Div : ET1

Roll no : ET1-68

PRN : 202401070117

pune_house_prices.csv

1.What is the average house price ?

```
Import numpy as np
prices = np.array([/* list of house prices */])
average_price  =  np.mean(prices)
print("Average house price:", average_price)

import pandas as pd
average_price = df['Price'].mean()
print("Average house price:", average_price)
```

2.What is the highest and lowest price ?

```
import numpy as np
prices = np.array([/* list of house prices */])
highest_price = np.max(prices)
 lowest_price = np.min(prices)
 print("Highest house price:", highest_price)
  print("Lowest house price:", lowest_price)
```

```
import pandas as pd
highest_price = df['Price'].max()
 lowest_price = df['Price'].min()
 print("Highest house price:", highest_price)
 print("Lowest house price:", lowest_price)
```

3.What is the minimum house price?

```
import numpy as np
 min_price = np.min(prices)
 print(min_price)
```

```
import pandas as pd
min_price = df['Price'].min()
 print(min_price)
```

4.How many houses have more than 3 bedrooms?

```
import numpy as np
bedrooms = df['Bedrooms'].values
count = np.sum(bedrooms > 3)
print(count)
```

5. What is the average number of bathrooms?

```
import numpy as np
bathrooms = df['Bathrooms'].value
average_bathrooms = np.mean(bathrooms)
print("Average number of bathrooms
(NumPy):", average_bathrooms)

import pandas as pd
average_bathrooms = df['Bathrooms'].mean()
print("Average number of bathrooms
(Pandas):", average_bathrooms)
```

6.What is the average square footage?

```
import numpy as np
 sqft = df['Sqft'].values
 average_sqft = np.mean(sqft)
 print("Average square footage
(NumPy):",average_sqft)

import pandas as pd
average_sqft = df['Sqft'].mean()
```

```
print("Average square footage (Pandas):",
average_sqft)
```

7.How many houses have a pool?

```
import numpy as np
pool = df['Pool'].values
count_with_pool = np.sum(pool == 'Yes')
print("Number of houses with a pool
(NumPy):", count_with_pool)

import pandas as pd
count_with_pool = df[df['Pool'] ==
'Yes'].shape[0]
print("Number of houses with a pool
(Pandas):", count_with_pool)
```

8. How many houses were built after 2010?

```
import numpy as np
year_built = df['YearBuilt'].values
 print("Number of houses built after 2010
(NumPy):", count_after_2010)

        import pandas as pd
```

```
count_after_2010 = df[df['YearBuilt'] > 2010].shape[0]
print("Number of houses built after 2010 (Pandas):", count_after_2010)
```

9.What is the oldest house year?

```
import numpy as np
year_built = df['YearBuilt'].values
oldest_year = np.min(year_built)
print("Oldest house year (NumPy):", oldest_year)

import pandas as pd
oldest_year = df['YearBuilt'].min()
print("Oldest house year (Pandas):", oldest_year)
```

10.What is the newest house year?

```
import numpy as np
year_built = df['YearBuilt'].values
newest_year = np.max(year_built)
print("Newest house year (NumPy):", newest_year)
```

```
import pandas as pd
newest_year = df['YearBuilt'].max()
print("Newest house year (Pandas):",
newest_year)
```

11.What is the median house price?

```
import numpy as np
prices = df['Price'].values
median_price = np.median(prices)
print("Median house price (NumPy):",
median_price)

import pandas as pd
median_price = df['Price'].median()
        print("Median house price (Pandas):",
median_price)
```

12.How many houses are smaller than 1000 sqft?

```
import numpy as np
sqft = df['Sqft'].values
small_houses = np.sum(sqft < 1000)
```

```
print("Number of houses smaller than 1000
sqft (NumPy):", small_houses)

import pandas as pd
small_houses = df[df['Sqft'] < 1000].shape[0]
print("Number of houses smaller than 1000
sqft (Pandas):", small_houses)
```

13.How many houses have 2 bathrooms exactly?

```
import numpy as np
bathrooms = df['Bathrooms'].values
two_bathrooms = np.sum(bathrooms == 2)
print("Number of houses with exactly 2
bathrooms (NumPy):", two_bathrooms)

import pandas as pd
two_bathrooms = df[df['Bathrooms'] ==
2].shape[0]
print("Number of houses with exactly 2
bathrooms (Pandas):", two_bathrooms)
```

14. How many houses are priced below 200,000?

```
import numpy as np
prices = df['Price'].values
cheap_houses = np.sum(prices < 200000)
print("Number of houses priced below 200,000
(NumPy):", cheap_houses)
```

15.What is the standard deviation of house prices?

```
import numpy as np
prices = df['Price'].values
std_price = np.std(prices)
print("Standard deviation of house prices
(NumPy):", std_price)

import pandas as pd
std_price = df['Price'].std()
print("Standard deviation of house prices
(Pandas):", std_price)
```

16.How many houses are in 'Kothrud' neighborhood?

```
import numpy as np
neighborhoods = df['Neighborhood'].values
kothrud_count = np.sum(neighborhoods ==
'Kothrud')
print("Number of houses in Kothrud
(NumPy):", kothrud_count)
```

```
import pandas as pd
kothrud_count = df[df['Neighborhood']
== 'Kothrud'].shape[0]
print("Number of houses in Kothrud
(Pandas):", kothrud_count)
```

17. What is the maximum square footage?

```
import numpy as np
sqft = df['Sqft'].values
max_sqft = np.max(sqft)
print("Maximum square footage (NumPy):",
max_sqft)
```

```
import pandas as pd
```

```
max_sqft = df['Sqft'].max()
print("Maximum square footage (Pandas):",
max_sqft)
```

18.Find the house with the highest price.

```
import numpy as np
prices = df['Price'].values
max_price_index = np.argmax(prices)
highest_price_house =
df.iloc[max_price_index]
print("House with the highest price
(NumPy):\n", highest_price_house)
```

```
import pandas as pd
highest_price_house = df[df['Price'] ==
df['Price'].max()]
print("House with the highest price
(Pandas):\n", highest_price_house)
```

19.Find the house with the lowest price.

```
import numpy as np
prices = df['Price'].values
min_price_index = np.argmin(prices)
lowest_price_house = df.iloc[min_price_index]
print("House with the lowest price
(NumPy):\n", lowest_price_house)

import pandas as pd
```

```
lowest_price_house = df[df['Price'] ==
df['Price'].min()]
print("House with the lowest price
(Pandas):\n", lowest_price_house)
```

20.How many houses have more than 2 bedrooms and 2 bathrooms?

```
import numpy as np
houses = np.array([[3, 2], [2, 1], [4, 3],
[3, 2], [5, 4], [1, 1], [3, 2]])
filtered_houses = houses[(houses[: ,
0] > 2) & (houses[: , 1] > 2)]
count = filtered_houses.shape[0]
print(count)


import pandas as pd
```

```
                data = {'bedrooms': [3, 2, 4, 3, 5, 1,
3], 'bathrooms': [2, 1, 3, 2, 4, 1, 2]}
                df = pd.DataFrame(data)
                filtered_df = df[(df['bedrooms'] > 2) &
  (df['bathrooms'] > 2)]
                count = filtered_df.shape[0]
                 print(count)
```