

Name - Ankit M. Chaudhari

PRN - 202401100034

Class - CS5-26

Dataset: Real Estate Listings

This document contains 20 problem statements and solutions using Numpy and Pandas based on a real-world housing dataset.

1. Problem - What is the total number of listings in the dataset?

Solution:

```
total_listings = df.shape[0]
```

2. Problem - How many unique cities are there in the dataset?

Solution:

```
unique_cities = df['city'].nunique()
```

3. Problem - What is the average price of all properties?

Solution:

```
avg_price = df['price'].mean()
```

4. Problem - What is the maximum number of bedrooms in any property?

Solution:

```
max_bedrooms = df['bedrooms'].max()
```

5. Problem - How many properties have more than 4 bedrooms?

Solution:

```
(df['bedrooms'] > 4).sum()
```

6. Problem - Find the average number of bathrooms in Seattle.

Solution:

```
avg_bathrooms_seattle = df[df['city'] == 'Seattle']['bathrooms'].mean()
```

7. Problem - How many properties have a waterfront view?

Solution:

```
(df['waterfront'] == 1).sum()
```

8. Problem - What is the average lot size of properties with 3 floors?

Solution:

```
avg_lot_3floors = df[df['floors'] == 3]['sqft_lot'].mean()
```

9. Problem - What is the oldest year a house was built?

Solution:

```
oldest_year = df['yr_built'].min()
```

10. Problem - How many properties have been renovated?

Solution:

```
(df['yr_renovated'] != 0).sum()
```

11. Problem - Which city has the highest average price?

Solution:

```
city_with_max_avg_price = df.groupby('city')['price'].mean().idxmax()
```

12. Problem - What is the most common condition rating?

Solution:

```
most_common_condition = df['condition'].mode()[0]
```

13. Problem - How many houses have a basement?

Solution:

```
(df['sqft_basement'] > 0).sum()
```

14. Problem - What is the average square footage of living area above ground?

Solution:

```
avg_sqft_above = df['sqft_above'].mean()
```

15. Problem - How many listings have more than 2000 sqft of living space?

Solution:

```
(df['sqft_living'] > 2000).sum()
```

16. Problem - What is the correlation between price and sqft_living?

Solution:

```
correlation = df['price'].corr(df['sqft_living'])
```

17. Problem - How many properties are located in the city of Kent?

Solution:

```
(df['city'] == 'Kent').sum()
```

18. Problem - What is the minimum price for properties with 2 bathrooms?

Solution:

```
min_price_2baths = df[df['bathrooms'] == 2]['price'].min()
```

19. Problem - Calculate the median price for homes with a view score of 4.

Solution:

```
median_price_view4 = df[df['view'] == 4]['price'].median()
```

20. Problem - List the top 5 cities with the most listings.

Solution:

```
top_cities = df['city'].value_counts().head(5)
```