

Name: Sanika Jaiswal

Roll no : ET1-25

---

## 1. Find the average house price.

```
import pandas as pd
import numpy as np

data = pd.read_csv('house_prices.csv')

average_price = np.mean(data['SalePrice'])
print(f"Average House Price: {average_price}")
```

---

## 2. Find the maximum house price.

```
max_price = np.max(data['SalePrice'])
print(f"Maximum House Price: {max_price}")
```

---

## 3. Find the minimum house price.

```
min_price = np.min(data['SalePrice'])
print(f"Minimum House Price: {min_price}")
```

---

## 4. Find the standard deviation of house prices.

```
std_dev_price = np.std(data['SalePrice'])
print(f"Standard Deviation of Prices: {std_dev_price}")
```

---

## 5. Count how many houses have 4 or more bedrooms.

```
bedroom_4plus = np.sum(data['BedroomAbvGr'] >= 4)
print(f"Houses with 4+ bedrooms: {bedroom_4plus}")
```

---

## 6. Calculate the average living area (GrLivArea) of houses.

```
average_living_area = np.mean(data['GrLivArea'])
```

```
print(f"Average Living Area: {average_living_area}")
```

---

## 7. Find the correlation between Sale Price and Living Area.

```
correlation = data['SalePrice'].corr(data['GrLivArea'])  
print(f"Correlation between Price and Living Area: {correlation}")
```

---

## 8. Identify how many houses were built before 1980.

```
built_before_1980 = np.sum(data['YearBuilt'] < 1980)  
print(f"Houses built before 1980: {built_before_1980}")
```

---

## 9. Find the average garage size for houses with 2 or more garages.

```
avg_garage_area = np.mean(data[data['GarageCars'] >= 2]['GarageArea'])  
print(f"Average Garage Area (2+ cars): {avg_garage_area}")
```

---

## 10. Find top 5 houses with the largest lot area.

```
top5_lotarea = data.nlargest(5, 'LotArea')  
print(top5_lotarea[['Id', 'LotArea', 'SalePrice']])
```

---

## 11. Determine how many houses have no basement area.

```
no_basement = np.sum(data['TotalBsmtSF'] == 0)  
print(f"Houses without basement: {no_basement}")
```

---

## 12. Calculate the median house price.

```
median_price = np.median(data['SalePrice'])  
print(f"Median House Price: {median_price}")
```

---

## 13. Find the range (max-min) of Sale Prices.

```
price_range = np.ptp(data['SalePrice']) # ptp = Peak-to-Peak (max-min)
print(f"Price Range: {price_range}")
```

---

## 14. Find the number of houses priced above \$500,000.

```
above_500k = np.sum(data['SalePrice'] > 500000)
print(f"Houses priced above $500,000: {above_500k}")
```

---

## 15. Group houses by number of bedrooms and calculate average SalePrice.

```
avg_price_bedrooms = data.groupby('BedroomAbvGr')['SalePrice'].mean()
print(avg_price_bedrooms)
```

---

## 16. Find the house with the maximum number of garage spaces.

```
max_garage = data.loc[np.argmax(data['GarageCars'])]
print(max_garage[['Id', 'GarageCars', 'SalePrice']])
```

---

## 17. Find percentage of houses that have at least 2 full bathrooms.

```
percentage_2baths = (np.sum(data['FullBath'] >= 2) / len(data)) * 100
print(f"Percentage with 2+ Full Bathrooms: {percentage_2baths}%")
```

---

## 18. Calculate the mean year houses were built.

```
mean_year_built = np.mean(data['YearBuilt'])
print(f"Mean Year Built: {mean_year_built}")
```

---

19. Identify houses that have both large lot area (>10,000 sqft) and large living area (>2,000 sqft).

```
large_houses = data[(data['LotArea'] > 10000) & (data['GrLivArea'] > 2000)]  
print(large_houses[['Id', 'LotArea', 'GrLivArea', 'SalePrice']])
```

---

20. Find the average SalePrice for houses with no garage.

```
avg_price_no_garage = np.mean(data[data['GarageCars'] == 0]['SalePrice'])  
print(f"Average Price (No Garage): {avg_price_no_garage}")
```

---