→ MOUNT GOOGLE DRIVE

```
from google.colab import drive
drive.mount('/content/drive')
```

Drive already mounted at /content/drive; to attempt to forcibly remount, call drive.mount("/content/drive", force_remount=True).

▼ IMPORT REQUIRED PACKAGES

import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
%matplotlib inline
import seaborn as sns
import math

→ LOAD THE DATASET

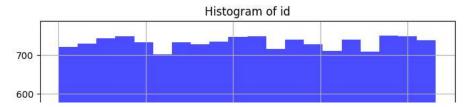
df=pd.read_csv("/content/drive/MyDrive/House Price India.csv")
df.head()

	id	Date	number of bedrooms	number of bathrooms	living area	lot area	number of floors	waterfront present	number of views
0	6762810145	42491	5	2.50	3650	9050	2.0	0	4
1	6762810635	42491	4	2.50	2920	4000	1.5	0	0
2	6762810998	42491	5	2.75	2910	9480	1.5	0	0
3	6762812605	42491	4	2.50	3310	42998	2.0	0	0
4	6762812919	42491	3	2.00	2710	4500	1.5	0	0

5 rows × 23 columns

→ UNIVARIATE ANALYSIS

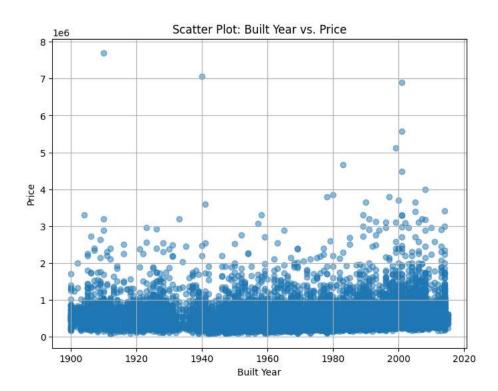
```
plt.figure(figsize=(8, 6))
column_name="id"
plt.hist(df[column_name], bins=20, color='blue', alpha=0.7)
plt.title("Histogram of " + column_name)
plt.xlabel(column_name)
plt.ylabel("Frequency")
plt.grid(True)
plt.show()
```



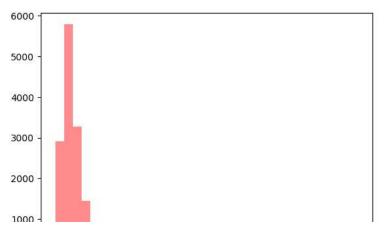
▼ BIVARIATE ANALYSIS

```
x_variable = 'Built Year'
y_variable = 'Price'

# Create a scatter plot
plt.figure(figsize=(8, 6))
plt.scatter(df[x_variable], df[y_variable], alpha=0.5)
plt.title("Scatter Plot: " + x_variable + " vs. " + y_variable)
plt.xlabel(x_variable)
plt.ylabel(y_variable)
plt.grid(True)
plt.show()
```



→ MULTIVARIATE ANALYSIS



→ DESCRIPTIVE STATISTICS

0 1 2 3 4 3 0 /

```
descriptive_stats = df.describe()
```

Display the descriptive statistics
print("Descriptive Statistics:")
print(descriptive_stats)

Std 6.237575e+03 67.347991 0.938719 0.7699		iptive_statist.	ics.)						
count 1.462000e+04 14620.000000 14620.000000 14620.000000 mean 6.762821e+09 42604.538646 3.379343 2.1295 std 6.237575e+08 67.347991 0.938719 0.7699 min 6.76281e+09 42546.000000 3.000000 1.7500 56% 6.76282e+09 42660.000000 3.000000 2.5500 75% 6.76283e+09 42600.000000 4.000000 2.5600 max 6.76283e+09 42734.000000 33.000000 8.0000 mean 2082.62996 1.509328e+04 1.502360 0.007661 std 928.275721 3.791962e+04 0.540239 0.087193 min 370.000000 5.200000e+02 1.000000 0.000000 25% 1440.000000 5.200000e+03 1.590000 0.000000 25% 1440.000000 1.000000 1.000000 0.000000 75% 2570.000000 1.000000 1.000000 1.000000 mean 9.233105 3.430506	Descri	ptive Statistic	s:						
mean 6.762821e+09 42604.538646 3.379343 2.1295 std 6.237575e+03 67.347991 0.938719 0.7669 min 6.76281be+09 42451.000000 1.000000 0.50000 0.50000 50% 6.762815e+09 42564.000000 3.000000 1.7500 50% 6.762826e+09 42662.000000 4.000000 2.5800 max 6.76282e+09 42662.000000 4.000000 2.5800 max 6.76282e+09 42662.000000 33.000000 8.0000 Living area 10t area number of floors waterfront present Logo 2000 1.4620.00000 3.000000 1.6020.00000 1.000000 std 928.75721 3.791962e+04 1.502360 0.007661 9.00000 std 928.75721 3.791962e+04 1.502360 0.007661 9.00000 Logo 2000000 7.000000 0.00000 0.00000 0.00000 Logo 2000000 7.000000 0.00000 0.00000		id	Date	number	of bedr	ooms	number of	bathroom	1S
Std 6.237575e+03 67.347991 0.938719 0.7699	count	1.462000e+04	14620.000000		14620.00	0000	146	20.00000	90
min 6.762810e+09 42491.000000 1.000000 0.5000 25% 6.762812e+09 42646.000000 3.000000 1.7500 25% 6.762812e+09 42660.000000 3.000000 2.5000 75% 6.762826e+09 42662.000000 4.000000 2.5000 max 6.762832e+09 42632.000000 4.000000 8.00000 living area	mean	6.762821e+09	42604.538646		3.37	9343		2.12958	33
1.7506	std		67.347991		0.93	8719		0.76993	
50% 6.762821e+09	min	6.762810e+09		1.00	0000	0.500000			
T5% 6.762826+09 42662.000000 4.000000 2.50000 8.00000 8.00000 8.00000 8.00000 1.0000000 1.0000000 1.0000000 1.0000000 1.0000000 1.0000000 1.0000000 1.0000000 1.0000000 1.0000000 1.0000000 1.0000000 1.0000000 0.00000000	25%	6.762815e+09		3.00	0000	0 1.750000			
Tiving area lot area number of floors waterfront present		6.762821e+09		3.00	0000				
Count 14620.00000	75%	6.762826e+09	42662.000000	4.00000					
count 14620.00000 1.4620.00000 14620.000000 14620.000000 mean 2098.262996 1.509328e+04 1.502360 0.007661 std 928.275721 3.791962e+04 0.549239 0.087193 min 370.000000 5.200000e+02 1.000000 0.000000 50% 1930.000000 7.620000e+03 1.500000 0.000000 75% 2570.000000 1.000000 0.000000 mx 13540.00000 1.000000 0.000000 max 13540.00000 1.000000 1.000000 mean 0.233105 3.430566 1970.926402 std 0.766259 0.664151 29.493625 min 0.000000 1.000000 1.990.000000 25% 0.000000 3.000000 1.975.000000 25% 0.000000 3.000000 1.997.000000 25% 0.000000 1.4620.000000 1.4620.000000 1.4620.000000 26% 0.000000 1.4620.000000 1.4620.000000 1.4620.000000	max	6.762832e+09	42734.000000		33.00	0000		8.00000	90
mean 2098.262996 1.509328e+04 1.502360 0.007661 std 928.275721 3.791962e+04 0.540239 0.087193 min 370.000000 5.200000e+02 1.000000 0.000000 25% 1440.000000 5.010750e+03 1.000000 0.000000 50% 1930.000000 1.080000e+04 2.000000 0.000000 75% 2570.000000 1.080000e+04 2.000000 0.000000 max 13540.000000 1.080000e+04 2.000000 1.000000 max 0.233105 3.430506 1970.926402 std 0.766259 0.664151 29.493625 min 0.00000 1.000000 1975.000000 25% 0.000000 3.000000 1975.000000 25% 0.000000 3.000000 1975.000000 25% 0.000000 3.000000 1975.000000 25% 0.000000 4.000000 1975.00000 25% 0.000000 122033.002244 52.792848 -114.79000		-							\
std 928.275721 3.791962e+04 0.540239 0.087193 min 370.000000 5.200000e+02 1.000000 0.000000 25% 1440.00000 5.01750e+03 1.000000 0.000000 50% 1930.000000 1.620000e+03 1.500000 0.000000 75% 2570.000000 1.080000e+04 2.000000 0.000000 max 13540.000000 1.074218e+06 3.500000 1.000000 mean 0.233105 3.430506 1970.926402 std 0.766259 0.664151 29.493625 min 0.000000 1.000000 1970.926402 25% 0.000000 3.000000 1970.900000 25% 0.000000 3.000000 1975.000000 75% 0.000000 3.000000 1975.000000 75% 0.000000 12033.062244 52.792848 -114.4000 8ean 90.924008 122003.00000 14620.000000 14620.000000 75% 0.000000 122033.000000 52.385900	count								
min 370.000000 5.200000e+02 1.000000 0.000000 25% 1440.000000 5.010750e+03 1.000000 0.000000 50% 1930.000000 1.080000e+04 2.000000 0.000000 75% 2570.000000 1.080000e+04 2.000000 0.000000 max 13540.000000 1.074218e+06 3.500000 1.000000 mean 0.233105 3.430506 1970.926402 std 0.766259 0.664151 29.493625 min 0.000000 1.000000 1900.00000 25% 0.000000 3.000000 1975.000000 75% 0.000000 4.000000 1975.000000 75% 0.000000 14620.000000 14620.00000 1997.000000 8enovation Year Postal Code Lattitude Longitude 14620.00000 std 416.216661 19.082418 1.37522 0.141326 min 0.000000 122033.062244 52.792848 -114.404007 50% 0.000000 1	mean			1.502360					
25% 1440.000000 5.010750e+03 1.0000000 0.0000000 50% 1930.0000000 1.080000e+04 2.0000000 0.0000000 max 13540.000000 1.074218e+06 3.5000000 1.0000000									
1930.000000 7.620000e+03 1.500000 0.000000				30000e+02 1.00					
Total									
max 13540.000000 1.074218e+06 3.500000 1.000000 count 14620.000000 14620.000000 Built Year \ count 14620.000000 14620.000000 14620.000000 mean 0.233105 3.430506 1970.926402 std 0.766259 0.664151 29.493625 min 0.000000 1.000000 1990.00000 25% 0.000000 3.000000 1997.000000 75% 0.000000 4.000000 1997.000000 max 4.000000 5.000000 1997.000000 mean 90.924008 122033.062244 52.792848 -114.404007 std 416.216661 19.082418 0.137522 0.141326 min 0.000000 122032.000000 52.385900 -114.709000 25% 0.000000 122032.000000 52.806400 -114.51900 max 2015.000000 122032.000000 52.806400<									
number of views condition of the house Built Year count 14620.000000 14620.000000 14620.000000 mean 0.233105 3.430506 1970.926402 std 0.766259 0.664151 29.493625 min 0.000000 1.000000 1990.000000 25% 0.000000 3.000000 1951.00000 75% 0.000000 4.000000 1975.00000 max 4.000000 5.000000 1975.000000 max 4.000000 14620.00000 1975.000000 mean 90.924008 122033.062244 52.792848 -114.404007 std 416.216661 19.082418 0.137522 0.141326 min 0.000000 122017.000000 52.385900 -114.709000 50% 0.000000 122017.000000 52.386400 -114.519000 50% 0.000000 122048.000000 52.806400 -114.421000 75% 0.000000 122072.000000 53.007600 -113.505000 max									
count 14620.000000 14620.000000 14620.000000 mean 0.233105 3.430506 1970.926402 std 0.766259 0.664151 29.493625 min 0.000000 1.000000 1990.000000 25% 0.000000 3.000000 1951.000000 50% 0.000000 3.000000 1975.000000 75% 0.000000 4.000000 1997.000000 max 4.000000 14620.000000 14620.000000 14620.000000 max 4.000000 12033.062244 52.792848 -114.404007 std 416.216661 19.082418 0.137522 0.141326 min 0.000000 122033.062244 52.792848 -114.404007 std 416.216661 19.082418 0.137522 0.141326 min 0.000000 122033.000000 52.707600 -114.519000 25% 0.000000 122032.000000 52.806400 -114.4	max	13540.000000	1.074218e+06		3.5000	100	1.	000000	
mean 0.233105 3.430506 1970.926402 std 0.766259 0.664151 29.493625 min 0.000000 1.000000 1900.000000 25% 0.000000 3.000000 1951.000000 50% 0.000000 3.000000 1975.000000 75% 0.000000 4.000000 1977.000000 max 4.000000 14620.00000 1.000000 max 4.000000 14620.000000 14620.00000 mean 90.924008 122033.062244 52.792848 -114.404007 std 416.216661 19.082418 0.137522 0.141326 min 0.000000 122032.00000 52.385900 -114.799000 25% 0.000000 122032.00000 52.707600 -114.519000 50% 0.000000 122032.00000 52.988900 -114.315000 max 2015.00000 122072.00000 53.007600 -113.505000 max 2015.00000 14620.000000 1.000000 1.000000								\	
### Std									
min 0.000000 1.000000 1900.000000 25% 0.000000 3.000000 1951.000000 50% 0.000000 3.000000 1975.000000 75% 0.000000 4.000000 1997.000000 max 4.000000 5.000000 2015.000000 max 4.000000 14620.000000 14620.000000 14620.000000 mean 90.924008 122033.062244 52.792848 -114.404007 std 416.216661 19.082418 0.137522 0.141326 min 0.000000 122032.000000 52.385900 -114.709000 25% 0.000000 122032.000000 52.707600 -114.519000 50% 0.000000 122032.000000 52.806400 -114.421000 75% 0.000000 122048.000000 52.908900 -114.315000 max 2015.000000 14620.000000 53.007600 -113.505000 mean 1996.702257 12753.500068 2.01224									
25% 0.000000 3.000000 1951.000000 50% 0.000000 3.000000 1975.000000 75% 0.000000 4.000000 1997.000000 max 4.000000 5.000000 2015.000000 max 4.000000 5.000000 2015.000000 2015.000000 max 4.000000 14620.000000 14620.000000 14620.000000 14620.000000 14620.000000 14620.000000 mean 90.924008 122033.062244 52.792848 -114.404007 std 416.216661 19.082418 0.137522 0.141326 min 0.000000 122032.000000 52.385900 -114.799000 25% 0.000000 122017.000000 52.707600 -114.519000 50% 0.000000 122032.000000 52.806400 -114.519000 50% 0.000000 122032.000000 52.806400 -114.421000 75% 0.000000 122048.000000 52.908900 -114.315000 max 2015.000000 14620.000000 53.007600 -113.505000									
50% 0.000000 3.000000 1975.000000 75% 0.000000 4.000000 1997.000000 max 4.000000 5.000000 2015.000000 Renovation Year Postal Code Lattitude Longitude \ count 14620.000000 14620.000000 14620.000000 14620.000000 mean 90.924008 122033.062244 52.792848 -114.404007 std 416.216661 19.082418 0.137522 0.141326 min 0.000000 122003.000000 52.385900 -114.709000 25% 0.000000 122032.000000 52.707600 -114.519000 75% 0.000000 122048.000000 52.806400 -114.315000 max 2015.000000 122072.000000 53.007600 -113.505000 count 14620.000000 14620.000000 14620.00000 14620.000000 14620.00000 50% 1490.000000 5097.750000 1.000000 2.000000									
Renovation Year									
Renovation Year Postal Code Lattitude Longitude \ count 14620.000000 122003.000000 52.792848 -114.404007 std 416.216661 19.082418 0.137522 0.141326 min 0.000000 122003.000000 52.385900 -114.709000 25% 0.000000 122017.000000 52.707600 -114.519000 50% 0.000000 122032.000000 52.806400 -114.421000 75% 0.000000 122072.000000 52.908900 -114.315000 max 2015.000000 122072.000000 53.007600 -113.505000 1200000 122072.000000 53.007600 -113.505000 14620.000000 14620.000000 14620.000000 14620.000000 14620.000000 14620.000000 1.0000000 25% 1490.000000 651.000000 1.000000 1.000000 25% 1490.000000 5997.750000 1.000000 1.000000 5097.750000 1.000000 3.000000 max 6110.000000 560617.000000 3.000000 3.000000 max 6110.000000 560617.000000 3.000000 3.000000 max 6110.000000 560617.000000 3.000000 3.000000 max 619.000000 7.8000000 7.8000000+04									
Renovation Year									
count 14620.000000 14620.000000 14620.000000 14620.000000 14620.000000 mean 90.924008 122033.062244 52.792848 -114.404007 std 416.216661 19.082418 0.137522 0.141326 min 0.000000 122003.000000 52.385900 -114.799000 25% 0.000000 122017.000000 52.707600 -114.519000 50% 0.000000 122032.000000 52.806400 -114.421000 75% 0.000000 122048.000000 52.908900 -114.421000 max 2015.000000 122072.000000 53.007600 -113.505000 living_area_renov lot_area_renov Number of schools nearby \ count 14620.000000 14620.000000 14620.000000 14620.000000 min 460.000000 651.000000 1.000000 1.000000 50% 1850.000000 7620.000000 3.000000 3.000000 Distance from the airport Price Price									

57.000000 3.200000e+05

65.000000 4.500000e+05

50%

```
mean_value = df['Price'].mean()
print("Mean of 'your_column':", mean_value)

Mean of 'your_column': 538932.2183310534
```

Handle the Missing values.

```
# Check for missing values
missing_values = df.isnull().sum()
print("Missing Values:")
print(missing_values)
# Option 1: Remove rows with missing values
df_cleaned = df.dropna()
# Option 2: Fill missing values with a specific value (e.g., mean or median)
# Replace 'your_column' with the actual column name
mean_value = df['Price'].mean()
data_filled = df.fillna(mean_value)
# Option 3: Forward fill or backward fill missing values
data_ffill = df.ffill() # Forward fill missing values
data_bfill = df.bfill() # Backward fill missing values
# Option 4: Interpolate missing values
data_interpolated = df.interpolate()
     Missing Values:
     Date
     number of bedrooms
     number of bathrooms
     living area
     lot area
     number of floors
                                              0
     waterfront present
                                              0
     number of views
     condition of the house
     grade of the house
     Area of the house(excluding basement)
     Area of the basement
     Built Year
     Renovation Year
     Postal Code
     Lattitude
                                              0
     Longitude
     living_area_renov
                                              0
     lot_area_renov
     Number of schools nearby
     Distance from the airport
     Price
     dtype: int64
```