

```
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
```

```
# Assuming your data is in a CSV file named 'property_data.csv'
df = pd.read_csv('property_data.csv')
```

```
df
```

		Place \
0	2 BHK Apartment for Sale in Sholinganallur Che...	
1	3 BHK Apartment for Sale in Sholinganallur Che...	
2	2 BHK Apartment for Sale in Kattupakkam Chennai	
3	5 BHK Apartment for Sale in Guindy Chennai	
4	3 BHK Builder Floor for Sale in Madambakkam Ch...	
..		...
501	4 BHK Villa for Sale in Karpagam Gardens Adyar...	
502	3 BHK Apartment for Sale in Velachery Chennai	
503	4 BHK Apartment for Sale in Korattur Chennai	
504	2 BHK Apartment for Sale in Thoraipakkam Chennai	
505	6 BHK House for Sale in Periyar Nagar West Pe...	

		Developer \
0		DRA D Elite
1		DRA D Elite
2		LML Arshiya
3	Puravankara Purva Somerset House	
4		N/a
..		...
501		Karpagam Gardens
502		Ayya Empire
503	Alliance Orchid Springss	
504		Grand Gems
505		N/a

	Summary	Price \
0	DRA D Elite is a project by DRA Group in Chenn...	₹32.5 Lac
1	DRA D Elite is the newest premium residential...	₹69 Lac
2	The area is filled with abundant open spaces a...	₹1.54 Cr
3	5 BHK Ultra Luxury Home with BluenNex, Automat...	₹48.5 Lac
4	3 bhk a brand new resediential flats in Madamb...	₹80.8 Lac
..		...
501	N/a	₹5.92 Cr
502	North facing, fully vasthu complaint, UDS 748 ...	₹1.31 Cr
503	Corner flat. Only 1 common wall. Bright sunlig...	₹1.70 Cr
504	Presenting an attractive 2 BHK flat for Resale...	₹72.5 Lac
505	Discover the epitome of spacious and luxurious...	₹2.80 Cr

	Sqft
0	₹6,983 per sqft

```

1    ₹7,300 per sqft
2    ₹6,400 per sqft
3    ₹5,607 per sqft
4    ₹7,638 per sqft
..
501    N/a
502    ₹9,562 per sqft
503    ₹7,509 per sqft
504    ₹7,104 per sqft
505    ₹6,829 per sqft

```

```
[506 rows x 5 columns]
```

```
import pandas as pd
```

```
# Assuming df is your dataframe
```

```
# Define a function to convert 'Price' values
```

```
def convert_price(price_str):
    try:
        # Extract numeric part
        numeric_part = float(''.join(filter(str.isdigit, price_str)))

        # Extract denomination part
        denomination = ''.join(filter(str.isalpha, price_str)).lower()

        # Convert to INR
        if denomination == 'lac':
            return numeric_part * 100000 # 1 Lac = 100,000
        elif denomination == 'cr':
            return numeric_part * 10000000 # 1 Cr = 10,000,000
        else:
            return None
    except (ValueError, AttributeError):
        return None
```

```
# Apply the function to the 'Price' column
```

```
df['Numeric Price'] = df['Price'].apply(convert_price)
```

```
# Display the dataframe with the converted 'Numeric Price' column
```

```
print(df[['Price', 'Numeric Price']])
```

	Price	Numeric Price
0	₹32.5 Lac	3.250000e+07
1	₹69 Lac	6.900000e+06
2	₹1.54 Cr	1.540000e+09
3	₹48.5 Lac	4.850000e+07
4	₹80.8 Lac	8.080000e+07
..	..	..
501	₹5.92 Cr	5.920000e+09

502	₹1.31 Cr	1.310000e+09
503	₹1.70 Cr	1.700000e+09
504	₹72.5 Lac	7.250000e+07
505	₹2.80 Cr	2.800000e+09

[506 rows x 2 columns]

df

		Place \
0	2 BHK Apartment for Sale in Sholinganallur Che...	
1	3 BHK Apartment for Sale in Sholinganallur Che...	
2	2 BHK Apartment for Sale in Kattupakkam Chennai	
3	5 BHK Apartment for Sale in Guindy Chennai	
4	3 BHK Builder Floor for Sale in Madambakkam Ch...	
..	...	
501	4 BHK Villa for Sale in Karpagam Gardens Adyar...	
502	3 BHK Apartment for Sale in Velachery Chennai	
503	4 BHK Apartment for Sale in Korattur Chennai	
504	2 BHK Apartment for Sale in Thoraipakkam Chennai	
505	6 BHK House for Sale in Periyar Nagar West Pe...	

	Developer \
0	DRA D Elite
1	DRA D Elite
2	LML Arshiya
3	Puravankara Purva Somerset House
4	N/a
..	...
501	Karpagam Gardens
502	Ayya Empire
503	Alliance Orchid Springss
504	Grand Gems
505	N/a

	Summary	Price \
0	DRA D Elite is a project by DRA Group in Chenn...	₹32.5 Lac
1	DRA D Elite is the newest premium residential...	₹69 Lac
2	The area is filled with abundant open spaces a...	₹1.54 Cr
3	5 BHK Ultra Luxury Home with BluenNex, Automat...	₹48.5 Lac
4	3 bhk a brand new resediential flats in Madamb...	₹80.8 Lac
..	...	...
501	N/a	₹5.92 Cr
502	North facing, fully vasthu complaint, UDS 748 ...	₹1.31 Cr
503	Corner flat. Only 1 common wall. Bright sunlig...	₹1.70 Cr
504	Presenting an attractive 2 BHK flat for Resale...	₹72.5 Lac
505	Discover the epitome of spacious and luxurious...	₹2.80 Cr

	Sqft	Numeric Price
0	₹6,983 per sqft	3.250000e+07

1	₹7,300 per sqft	6.900000e+06
2	₹6,400 per sqft	1.540000e+09
3	₹5,607 per sqft	4.850000e+07
4	₹7,638 per sqft	8.080000e+07
...	...	...
501	N/a	5.920000e+09
502	₹9,562 per sqft	1.310000e+09
503	₹7,509 per sqft	1.700000e+09
504	₹7,104 per sqft	7.250000e+07
505	₹6,829 per sqft	2.800000e+09

[506 rows x 6 columns]

## EDA

```
# Display the first few rows of the dataset
print(df.head())
```

		Place \
0	2 BHK Apartment for Sale in Sholinganallur Che...	
1	3 BHK Apartment for Sale in Sholinganallur Che...	
2	2 BHK Apartment for Sale in Kattupakkam Chennai	
3	5 BHK Apartment for Sale in Guindy Chennai	
4	3 BHK Builder Floor for Sale in Madambakkam Ch...	

	Developer \
0	DRA D Elite
1	DRA D Elite
2	LML Arshiya
3	Puravankara Purva Somerset House
4	N/a

	Summary	Price \
0	DRA D Elite is a project by DRA Group in Chenn...	₹32.5 Lac
1	DRA D Elite is the newest premium residential...	₹69 Lac
2	The area is filled with abundant open spaces a...	₹1.54 Cr
3	5 BHK Ultra Luxury Home with BluenNex, Automat...	₹48.5 Lac
4	3 bhk a brand new resedential flats in Madamb...	₹80.8 Lac

	Sqft	Numeric Price
0	₹6,983 per sqft	3.250000e+07
1	₹7,300 per sqft	6.900000e+06
2	₹6,400 per sqft	1.540000e+09
3	₹5,607 per sqft	4.850000e+07
4	₹7,638 per sqft	8.080000e+07

```
# Get a summary of the dataset
print(df.info())
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 506 entries, 0 to 505
Data columns (total 6 columns):
#   Column                Non-Null Count  Dtype
---  -
0   Place                  506 non-null   object
1   Developer              506 non-null   object
2   Summary                506 non-null   object
3   Price                  506 non-null   object
4   Sqft                   506 non-null   object
5   Numeric Price          506 non-null   float64
dtypes: float64(1), object(5)
memory usage: 23.8+ KB
None
```

*# Descriptive statistics*

```
print(df.describe())
```

	Numeric Price
count	5.060000e+02
mean	8.199994e+08
std	1.508098e+09
min	1.600000e+06
25%	7.500000e+06
50%	5.640000e+07
75%	1.260000e+09
max	1.380000e+10

*# Check for missing values*

```
print(df.isnull().sum())
```

Place	0
Developer	0
Summary	0
Price	0
Sqft	0
Numeric Price	0

dtype: int64

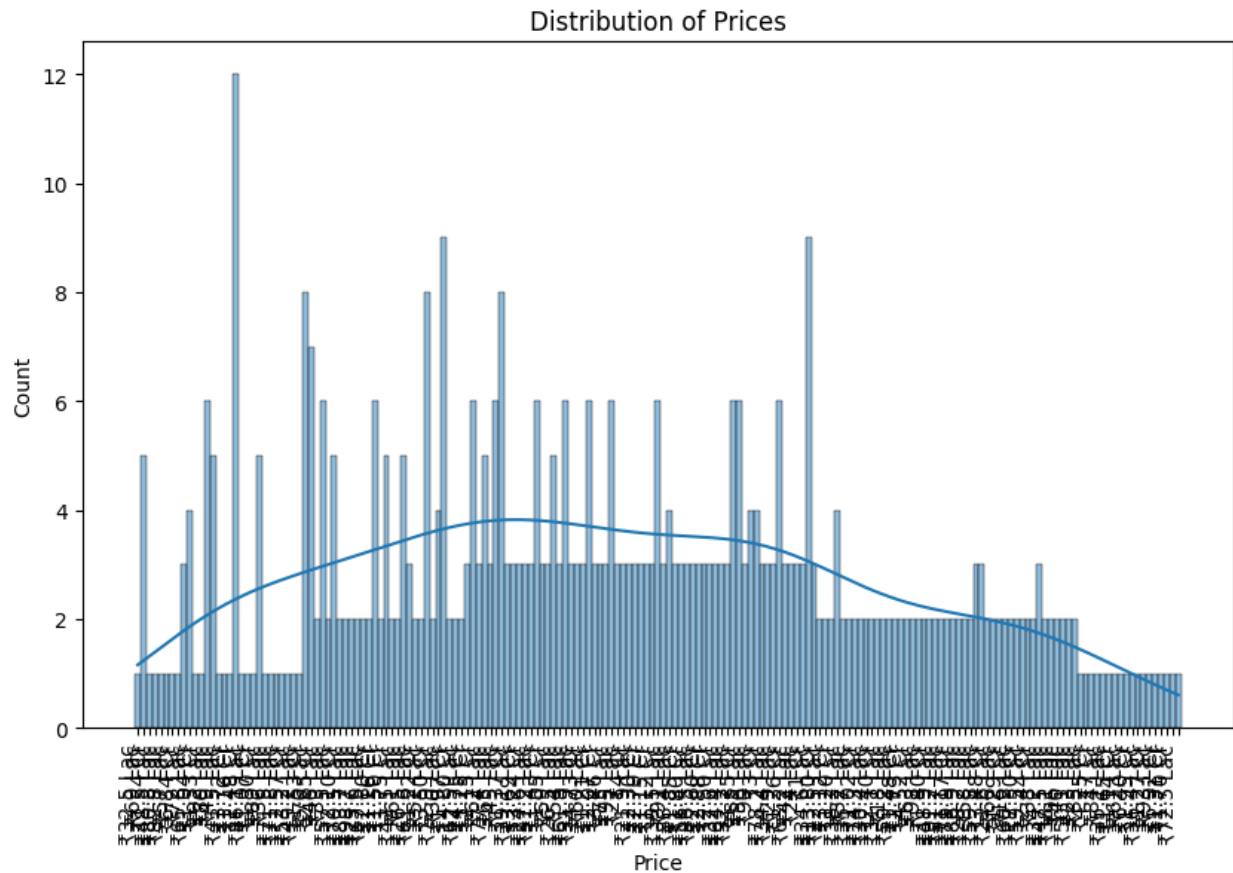
*# Distribution of Prices*

```
plt.figure(figsize=(10, 6))
sns.histplot(df['Price'], bins=30, kde=True)
plt.title('Distribution of Prices')
plt.xlabel('Price')
```

*# Rotate x-axis labels by 45 degrees and add spaces between labels*

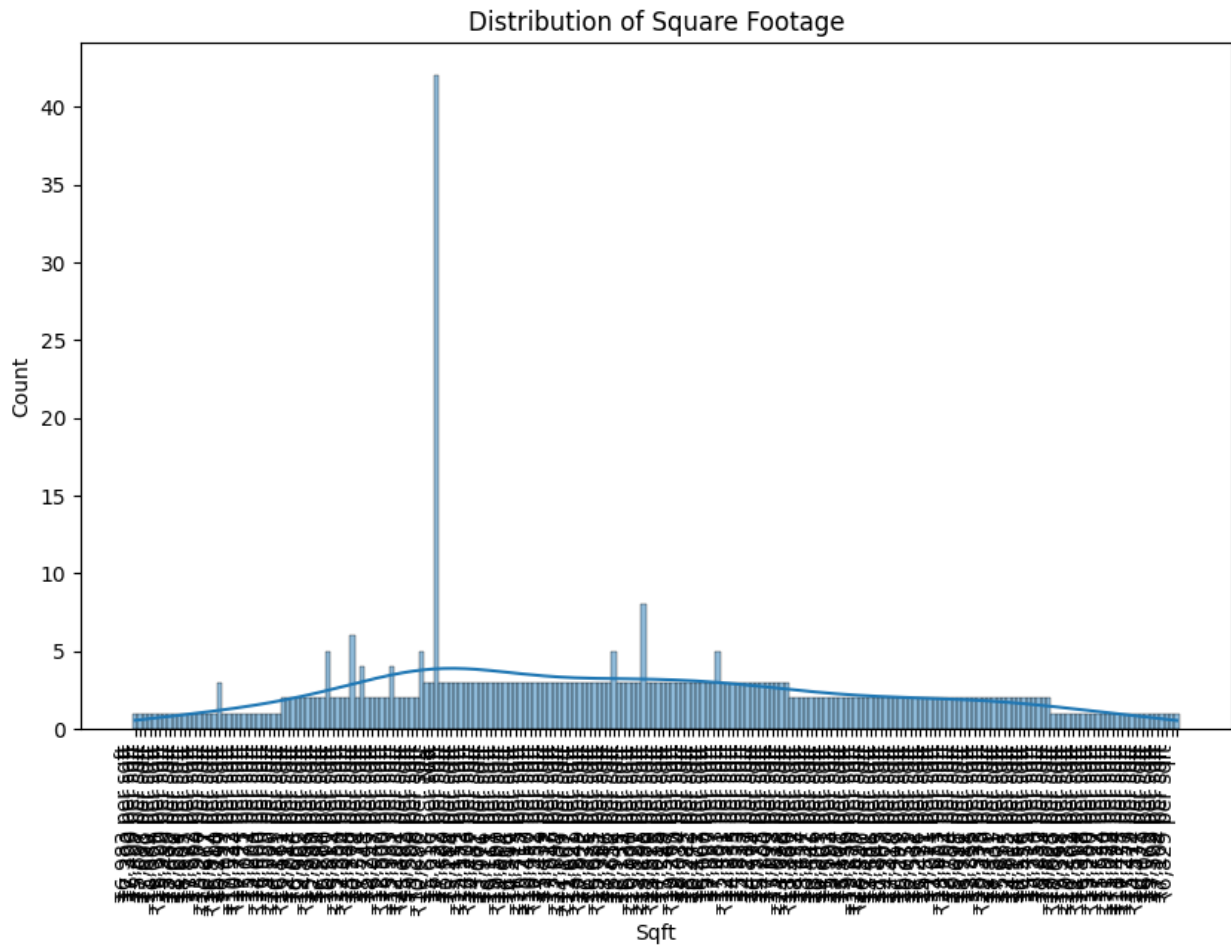
```
plt.xticks(rotation=90, ha='right')
```

```
plt.show()
```



```
# Distribution of Square Footage
plt.figure(figsize=(10, 6))
sns.histplot(df['Sqft'], bins=30, kde=True)
plt.title('Distribution of Square Footage')
plt.xlabel('Sqft')
# Rotate x-axis labels by 45 degrees and add spaces between labels
plt.xticks(rotation=90, ha='right')

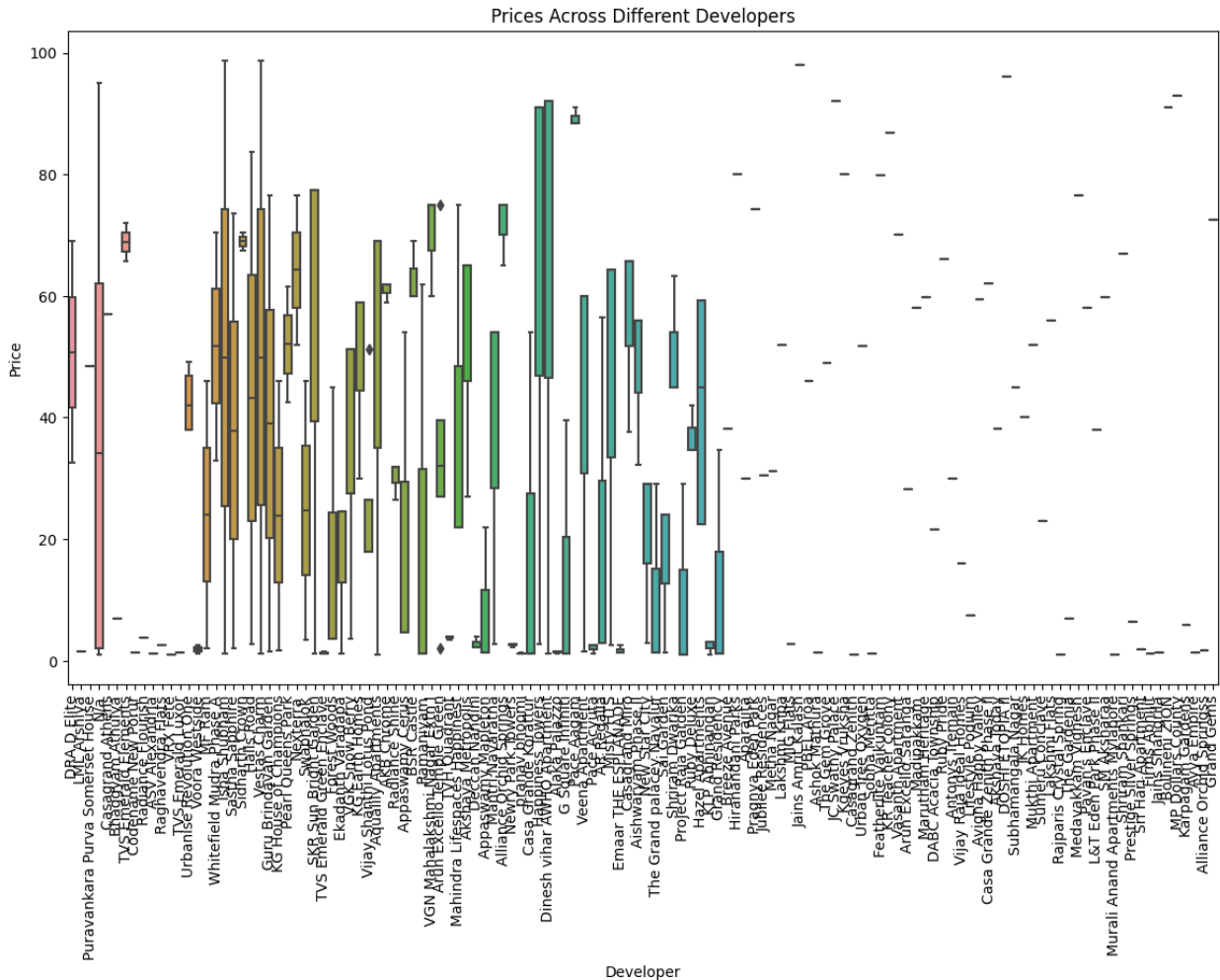
plt.show()
```



```
import pandas as pd

# Assuming df is your dataframe
# Remove non-numeric characters from 'Price' and convert to numeric
df['Price'] = pd.to_numeric(df['Price'].replace('[^\d.]', ''),
                             regex=True))

# Now, try the boxplot again
plt.figure(figsize=(14, 8))
sns.boxplot(x='Developer', y='Price', data=df)
plt.title('Prices Across Different Developers')
plt.xticks(rotation=90)
plt.show()
```



```
import matplotlib.pyplot as plt
import seaborn as sns

# Assuming df is your dataframe
# Remove non-numeric characters from 'Price' and convert to numeric
df['Price'] = pd.to_numeric(df['Price'].replace('[^\d.]', ''),
                             regex=True))

# Remove non-numeric characters from 'Sqft' and convert to numeric
df['Sqft'] = pd.to_numeric(df['Sqft'].replace('[^\d.]', ''),
                             regex=True))

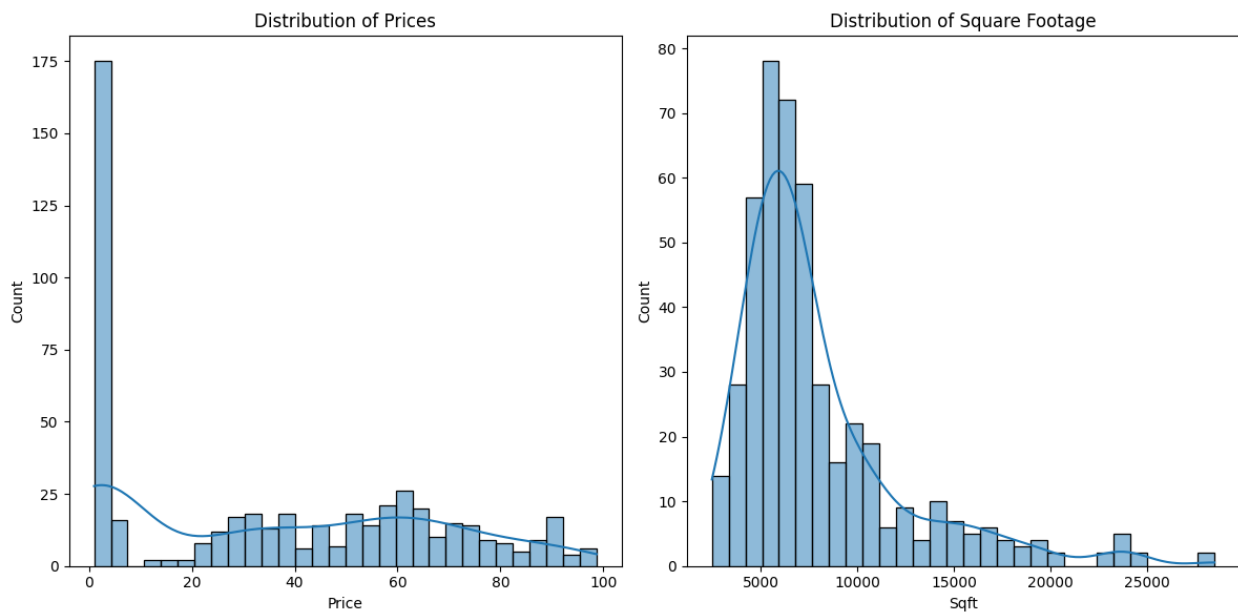
# Distribution of Prices
plt.figure(figsize=(12, 6))
plt.subplot(1, 2, 1)
sns.histplot(df['Price'], bins=30, kde=True)
plt.title('Distribution of Prices')
plt.xlabel('Price')

# Distribution of Square Footage
```



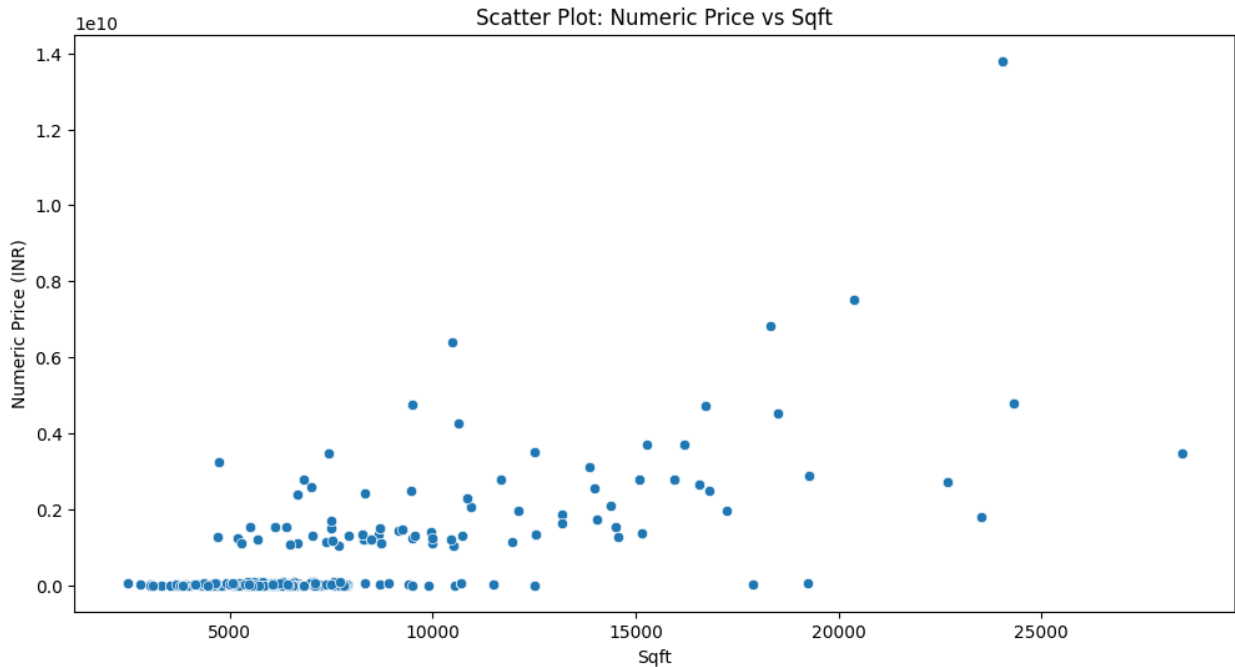
```
plt.subplot(1, 2, 2)
sns.histplot(df['Sqft'], bins=30, kde=True)
plt.title('Distribution of Square Footage')
plt.xlabel('Sqft')

plt.tight_layout()
plt.show()
```



```
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns

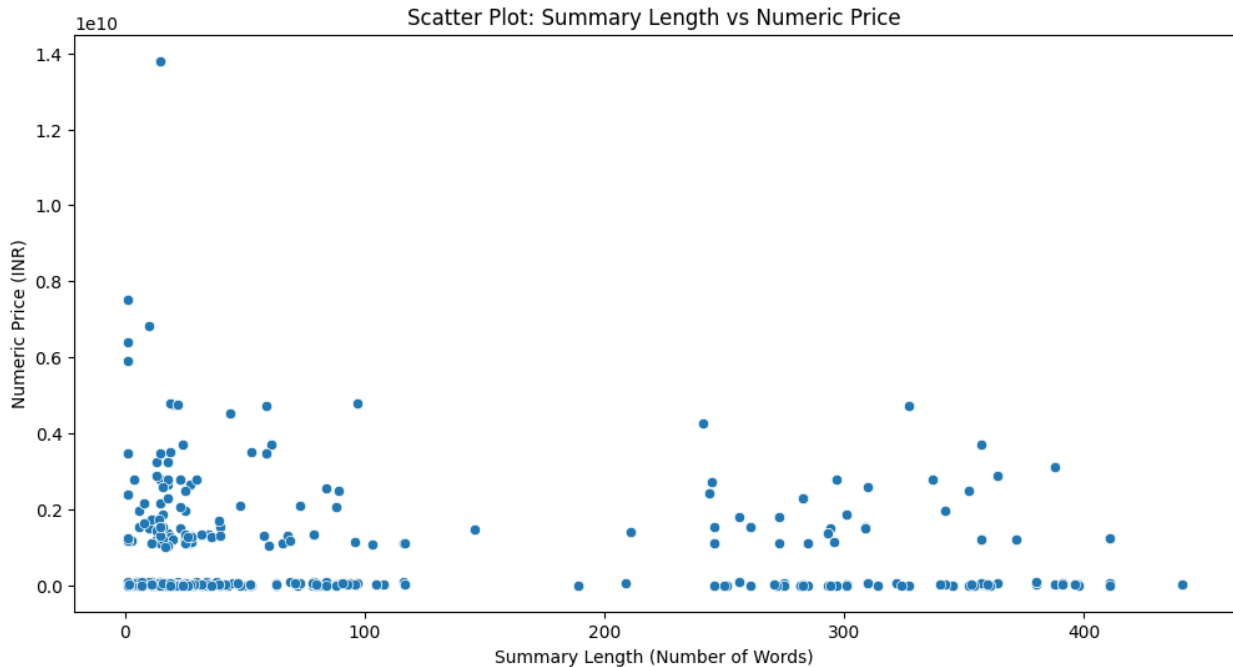
# Scatter plot: Numeric Price vs Sqft
plt.figure(figsize=(12, 6))
sns.scatterplot(x='Sqft', y='Numeric Price', data=df)
plt.title('Scatter Plot: Numeric Price vs Sqft')
plt.xlabel('Sqft')
plt.ylabel('Numeric Price (INR)')
plt.show()
```



```
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns

# Create a new column for summary lengths
df['Summary Length'] = df['Summary'].apply(lambda x:
len(str(x).split()))

# Scatter plot: Summary Length vs Numeric Price
plt.figure(figsize=(12, 6))
sns.scatterplot(x='Summary Length', y='Numeric Price', data=df)
plt.title('Scatter Plot: Summary Length vs Numeric Price')
plt.xlabel('Summary Length (Number of Words)')
plt.ylabel('Numeric Price (INR)')
plt.show()
```



```
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns

# Boxplot: Numeric Price by Developer
plt.figure(figsize=(14, 8))
sns.boxplot(x='Developer', y='Numeric Price', data=df)
plt.title('Boxplot: Numeric Price Across Different Developers')
plt.xlabel('Developer')
plt.ylabel('Numeric Price (INR)')
plt.xticks(rotation=45, ha='right') # Rotate x-axis labels for better readability
plt.show()
```

