

real-state-insights

October 17, 2024

```
[72]: import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
```

```
[30]: data = pd.read_csv("Downloads/flats.csv")
```

```
[45]: flats_df = data
```

```
[47]: flats_df.head()
```

```
[47]:
```

| | property_name \ | link | society \ |
|---|---------------------------------|---|------------------------|
| 0 | 2 BHK Flat in Krishna Colony | https://www.99acres.com/2-bhk-bedroom-apartmen... | maa bhagwati residency |
| 1 | 2 BHK Flat in Ashok Vihar | https://www.99acres.com/2-bhk-bedroom-apartmen... | Apna Enclave |
| 2 | 2 BHK Flat in Sohna | https://www.99acres.com/2-bhk-bedroom-apartmen... | Tulsiani Easy in Homes |
| 3 | 2 BHK Flat in Sector 61 Gurgaon | https://www.99acres.com/2-bhk-bedroom-apartmen... | Smart World Orchard |
| 4 | 2 BHK Flat in Sector 92 Gurgaon | https://www.99acres.com/2-bhk-bedroom-apartmen... | Parkwood Westend |

| | price | rate | areaWithType \ |
|---|------------|---------------|--|
| 0 | 45 Lac | 5,000/sq.ft. | Carpet area: 900 (83.61 sq.m.) |
| 1 | 50 Lac | 7,692/sq.ft. | Carpet area: 650 (60.39 sq.m.) |
| 2 | 40 Lac | 6,722/sq.ft. | Carpet area: 595 (55.28 sq.m.) |
| 3 | 1.47 Crore | 12,250/sq.ft. | Carpet area: 1200 (111.48 sq.m.) |
| 4 | 70 Lac | 5,204/sq.ft. | Super Built up area 1345(124.95 sq.m.) |

| | bedRoom | bathroom | balcony | additionalRoom \ |
|---|------------|-------------|-------------|------------------|
| 0 | 2 Bedrooms | 2 Bathrooms | 1 Balcony | NaN |
| 1 | 2 Bedrooms | 2 Bathrooms | 1 Balcony | NaN |
| 2 | 2 Bedrooms | 2 Bathrooms | 3 Balconies | NaN |
| 3 | 2 Bedrooms | 2 Bathrooms | 2 Balconies | Study Room |
| 4 | 2 Bedrooms | 2 Bathrooms | 3 Balconies | Study Room |

| | address | floorNum | facing | \ |
|---|-------------------------------------|----------|--------------|------|
| 0 | Krishna Colony, Gurgaon, Haryana | 4th | of 4 Floors | West |
| 1 | 46b, Ashok Vihar, Gurgaon, Haryana | 1st | of 3 Floors | West |
| 2 | Sohna, Gurgaon, Haryana | 12nd | of 14 Floors | NaN |
| 3 | Sector 61 Gurgaon, Gurgaon, Haryana | 2nd | of 4 Floors | NaN |
| 4 | Sector 92 Gurgaon, Gurgaon, Haryana | 5th | of 8 Floors | NaN |

| | agePossession | nearbyLocations | \ |
|---|--------------------|---|---|
| 0 | 1 to 5 Year Old | ['Chintapurni Mandir', 'State bank ATM', 'Pear... | |
| 1 | 10+ Year Old | ['Chintapurni Mandir', 'Sheetla Mata Mandir', ... | |
| 2 | 0 to 1 Year Old | ['Huda City Metro', 'Golf Course extn road', '... | |
| 3 | Dec-23 | ['Sector 55-56 Metro station', 'Bestech Centra... | |
| 4 | Under Construction | ['Yadav Clinic', 'Bangali Clinic', 'Dr. J. S. ... | |

| | description | \ |
|---|---|---|
| 0 | So with lift.Maa bhagwati residency is one of ... | |
| 1 | Property situated on main road, railway statio... | |
| 2 | This property is 15 km away from badshapur, gu... | |
| 3 | Near to metro station of sector 56 and opposit... | |
| 4 | We are the proud owners of this 2 bhk alongwit... | |

| | furnishDetails | \ |
|---|---|-----|
| 0 | ['3 Fan', '4 Light', '1 Wardrobe', 'No AC', 'N... | |
| 1 | ['3 Wardrobe', '4 Fan', '1 Exhaust Fan', '1 Ge... | |
| 2 | | NaN |
| 3 | | NaN |
| 4 | | [] |

| | features | \ |
|---|---|-----|
| 0 | ['Feng Shui / Vaastu Compliant', 'Security / F... | |
| 1 | ['Security / Fire Alarm', 'Maintenance Staff',... | |
| 2 | ['Power Back-up', 'Feng Shui / Vaastu Complian... | |
| 3 | ['Security / Fire Alarm', 'Private Garden / Te... | |
| 4 | | NaN |

| | rating | property_id |
|---|---|---------------|
| 0 | ['Environment4 out of 5', 'Safety4 out of 5', ... | C68850746 |
| 1 | ['Environment4 out of 5', 'Safety4 out of 5', ... | H68850564 |
| 2 | ['Environment4 out of 5', 'Safety4 out of 5', ... | J68850120 |
| 3 | | NaN S68849476 |
| 4 | ['Environment5 out of 5', 'Safety3 out of 5', ... | L47956793 |

```
[49]: flats_df.tail()
```

```
[49]:
3012    3 BHK Flat in Sector 86 Gurgaon
```

3013 5 BHK Flat in Sector 48 Gurgaon
 3014 3 BHK Flat in Sector 108 Gurgaon
 3015 3 BHK Flat in DLF Phase 3
 3016 4 BHK Flat in Sector 54 Gurgaon

link \

3012 <https://www.99acres.com/3-bhk-bedroom-apartmen...>
 3013 <https://www.99acres.com/5-bhk-bedroom-apartmen...>
 3014 <https://www.99acres.com/3-bhk-bedroom-apartmen...>
 3015 <https://www.99acres.com/3-bhk-bedroom-apartmen...>
 3016 <https://www.99acres.com/4-bhk-bedroom-apartmen...>

| | society | price | rate \ |
|------|--------------------------|------------|---------------|
| 3012 | Ansal Heights 86 | 1.05 Crore | 5,541/sq.ft. |
| 3013 | Parsvnath Green Ville3.9 | 3.3 Crore | 9,984/sq.ft. |
| 3014 | Raheja Vedaanta3.6 | 95 Lac | 5,214/sq.ft. |
| 3015 | Ambience Lagoon3.9 | 5.8 Crore | 12,500/sq.ft. |
| 3016 | DLF The Crest3.6 | 11 Crore | 35,222/sq.ft. |

| | areaWithType | bedRoom \ |
|------|---|------------|
| 3012 | Super Built up area 1895(176.05 sq.m.) | 3 Bedrooms |
| 3013 | Super Built up area 3905(362.79 sq.m.)Built Up... | 5 Bedrooms |
| 3014 | Super Built up area 1822(169.27 sq.m.) | 3 Bedrooms |
| 3015 | Built Up area: 3700 (343.74 sq.m.) | 3 Bedrooms |
| 3016 | Super Built up area 3123(290.14 sq.m.) | 4 Bedrooms |

| | bathroom | balcony | additionalRoom \ |
|------|-------------|--------------|---|
| 3012 | 3 Bathrooms | 3 Balconies | Servant Room |
| 3013 | 5 Bathrooms | 3+ Balconies | Servant Room |
| 3014 | 3 Bathrooms | 3 Balconies | Others |
| 3015 | 4 Bathrooms | 3+ Balconies | Pooja Room,Study Room,Servant Room,Others |
| 3016 | 6 Bathrooms | 3 Balconies | Servant Room |

| | address | floorNum \ |
|------|--|------------------|
| 3012 | Tower C, Sector 86 Gurgaon, Gurgaon, Haryana | 9th of 13 Floors |
| 3013 | Sector 48 Gurgaon, Gurgaon, Haryana | 4th of 5 Floors |
| 3014 | Sector 108 Gurgaon, Gurgaon, Haryana | 3rd of 22 Floors |
| 3015 | Gurgaon, DLF Phase 3, Gurgaon, Haryana | 9th of 9 Floors |
| 3016 | Sector 54 Gurgaon, Gurgaon, Haryana | 7th of 36 Floors |

| | facing | agePossession \ |
|------|------------|--------------------|
| 3012 | North-East | Under Construction |
| 3013 | NaN | 10+ Year Old |
| 3014 | NaN | 1 to 5 Year Old |
| 3015 | North-East | 10+ Year Old |
| 3016 | NaN | 1 to 5 Year Old |

```

nearbyLocations \
3012 ['IRIS Broadway Mall', 'Delhi Jaipur Expresswa...
3013 ['Sri Radhe Krishna Temple', 'Icici bank ATM',...
3014 NaN
3015 ['Micromax moulisari avenue metro station', 'In...
3016 ['Sector 53-54 Metro Station', 'Ardee Mall', '...

description \
3012 Residential apartment for sell.Located in sect...
3013 5 bhk duplex penthouse in low rise building.Av...
3014 3 bedroom flat with full woodwork. Ready to mo...
3015 Luxury condominium complex located on delhi gu...
3016 Club & pool facing\nVrv air conditioning

furnishDetails \
3012 []
3013 NaN
3014 []
3015 ['1 Water Purifier', '10 Fan', '1 Fridge', '1 ...
3016 ['6 Fan', '1 Fridge', '1 Exhaust Fan', '5 Geys...

features \
3012 NaN
3013 ['Security / Fire Alarm', 'Private Garden / Te...
3014 ['Security / Fire Alarm', 'Feng Shui / Vaastu ...
3015 ['Security / Fire Alarm', 'Private Garden / Te...
3016 ['Security / Fire Alarm', 'Power Back-up', 'In...

rating property_id
3012 ['Safety4.5 out of 5', 'Lifestyle5 out of 5', ... D26586124
3013 ['Management4 out of 5', 'Green Area4 out of 5... J17123294
3014 ['Management3 out of 5', 'Green Area4 out of 5... A41215323
3015 ['Management5 out of 5', 'Green Area5 out of 5... J18888617
3016 ['Management4 out of 5', 'Green Area4 out of 5... V70296402

```

```
[51]: flats_df.columns
```

```
[51]: Index(['property_name', 'link', 'society', 'price', 'rate', 'areaWithType',
        'bedRoom', 'bathroom', 'balcony', 'additionalRoom', 'address',
        'floorNum', 'facing', 'agePossession', 'nearbyLocations', 'description',
        'furnishDetails', 'features', 'rating', 'property_id'],
        dtype='object')
```

```
[53]: flats_df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 3017 entries, 0 to 3016
```

Data columns (total 20 columns):

| # | Column | Non-Null Count | Dtype |
|----|-----------------|----------------|--------|
| 0 | property_name | 3017 non-null | object |
| 1 | link | 3017 non-null | object |
| 2 | society | 3016 non-null | object |
| 3 | price | 3007 non-null | object |
| 4 | rate | 3004 non-null | object |
| 5 | areaWithType | 3008 non-null | object |
| 6 | bedRoom | 3008 non-null | object |
| 7 | bathroom | 3008 non-null | object |
| 8 | balcony | 3008 non-null | object |
| 9 | additionalRoom | 1694 non-null | object |
| 10 | address | 3002 non-null | object |
| 11 | floorNum | 3006 non-null | object |
| 12 | facing | 2127 non-null | object |
| 13 | agePossession | 3007 non-null | object |
| 14 | nearbyLocations | 2913 non-null | object |
| 15 | description | 3008 non-null | object |
| 16 | furnishDetails | 2203 non-null | object |
| 17 | features | 2594 non-null | object |
| 18 | rating | 2676 non-null | object |
| 19 | property_id | 3008 non-null | object |

dtypes: object(20)

memory usage: 471.5+ KB

```
[55]: flats_df.describe()
```

```
[55]:
```

| | property_name \ | | link | society \ |
|--------|---------------------|--|---|----------------|
| count | 3017 | | 3017 | 3016 |
| unique | 404 | | 2880 | 638 |
| top | 2 BHK Flat in Sohna | | https://www.99acres.com/2-bhk-bedroom-apartmen... | SS The Leaf3.8 |
| freq | 126 | | 2 | 73 |

| | price | rate | areaWithType \ |
|--------|------------|---------------|---------------------------------------|
| count | 3007 | 3004 | 3008 |
| unique | 439 | 2130 | 1932 |
| top | 1.25 Crore | 10,000/sq.ft. | Super Built up area 1578(146.6 sq.m.) |
| freq | 79 | 19 | 17 |

| | bedRoom | bathroom | balcony additionalRoom \ |
|--|---------|----------|--------------------------|
|--|---------|----------|--------------------------|

| | | | | |
|--------|------------|-------------|-------------|--------------|
| count | 3008 | 3008 | 3008 | 1694 |
| unique | 6 | 7 | 5 | 49 |
| top | 3 Bedrooms | 2 Bathrooms | 3 Balconies | Servant Room |
| freq | 1439 | 1047 | 977 | 630 |

| | | | |
|--------|-------------------------|-----------------|------------|
| | address | floorNum | facing \ |
| count | 3002 | 3006 | 2127 |
| unique | 1287 | 550 | 8 |
| top | Sohna, Gurgaon, Haryana | 2nd of 4 Floors | North-East |
| freq | 77 | 74 | 506 |

| | | |
|--------|-----------------|---|
| | agePossession | nearbyLocations \ |
| count | 3007 | 2913 |
| unique | 48 | 599 |
| top | 1 to 5 Year Old | ['Airia Mall Sector 68', 'Southern Peripheral ... |
| freq | 1479 | 73 |

| | | |
|--------|---|------------------|
| | description | furnishDetails \ |
| count | 3008 | 2203 |
| unique | 2787 | 1367 |
| top | Best in class property available at sohna loca... | [] |
| freq | 7 | 367 |

| | |
|--------|---|
| | features \ |
| count | 2594 |
| unique | 1438 |
| top | ['Feng Shui / Vaastu Compliant', 'Security / F... |
| freq | 290 |

| | | |
|--------|---|-------------|
| | rating | property_id |
| count | 2676 | 3008 |
| unique | 341 | 2874 |
| top | ['Green Area5 out of 5', 'Construction4 out of... | I69037888 |
| freq | 128 | 2 |

```
[58]: flats_df.drop(columns=['link'], inplace=True)
```

```
[62]: flats_df.head()
```

```
[62]:
```

| | property_name | society | price \ |
|---|---------------------------------|------------------------|------------|
| 0 | 2 BHK Flat in Krishna Colony | maa bhagwati residency | 45 Lac |
| 1 | 2 BHK Flat in Ashok Vihar | Apna Enclave | 50 Lac |
| 2 | 2 BHK Flat in Sohna | Tulsiani Easy in Homes | 40 Lac |
| 3 | 2 BHK Flat in Sector 61 Gurgaon | Smart World Orchard | 1.47 Crore |
| 4 | 2 BHK Flat in Sector 92 Gurgaon | Parkwood Westend | 70 Lac |

| | rate | areaWithType | bedRoom \ |
|--|------|--------------|-----------|
|--|------|--------------|-----------|

| | | | |
|---|---------------|--|------------|
| 0 | 5,000/sq.ft. | Carpet area: 900 (83.61 sq.m.) | 2 Bedrooms |
| 1 | 7,692/sq.ft. | Carpet area: 650 (60.39 sq.m.) | 2 Bedrooms |
| 2 | 6,722/sq.ft. | Carpet area: 595 (55.28 sq.m.) | 2 Bedrooms |
| 3 | 12,250/sq.ft. | Carpet area: 1200 (111.48 sq.m.) | 2 Bedrooms |
| 4 | 5,204/sq.ft. | Super Built up area 1345(124.95 sq.m.) | 2 Bedrooms |

| | bathroom | balcony | additionalRoom | \ |
|---|-------------|-------------|----------------|-----|
| 0 | 2 Bathrooms | 1 Balcony | | NaN |
| 1 | 2 Bathrooms | 1 Balcony | | NaN |
| 2 | 2 Bathrooms | 3 Balconies | | NaN |
| 3 | 2 Bathrooms | 2 Balconies | Study Room | |
| 4 | 2 Bathrooms | 3 Balconies | Study Room | |

| | address | floorNum | facing | \ |
|---|-------------------------------------|----------|--------------|------|
| 0 | Krishna Colony, Gurgaon, Haryana | 4th | of 4 Floors | West |
| 1 | 46b, Ashok Vihar, Gurgaon, Haryana | 1st | of 3 Floors | West |
| 2 | Sohna, Gurgaon, Haryana | 12nd | of 14 Floors | NaN |
| 3 | Sector 61 Gurgaon, Gurgaon, Haryana | 2nd | of 4 Floors | NaN |
| 4 | Sector 92 Gurgaon, Gurgaon, Haryana | 5th | of 8 Floors | NaN |

| | agePossession | nearbyLocations | \ |
|---|--------------------|---|---|
| 0 | 1 to 5 Year Old | ['Chintapurni Mandir', 'State bank ATM', 'Pear... | |
| 1 | 10+ Year Old | ['Chintapurni Mandir', 'Sheetla Mata Mandir', ... | |
| 2 | 0 to 1 Year Old | ['Huda City Metro', 'Golf Course extn road', '... | |
| 3 | Dec-23 | ['Sector 55-56 Metro station', 'Bestech Centra... | |
| 4 | Under Construction | ['Yadav Clinic', 'Bangali Clinic', 'Dr. J. S. ... | |

| | description | \ |
|---|---|---|
| 0 | So with lift.Maa bhagwati residency is one of ... | |
| 1 | Property situated on main road, railway statio... | |
| 2 | This property is 15 km away from badshapur, gu... | |
| 3 | Near to metro station of sector 56 and opposit... | |
| 4 | We are the proud owners of this 2 bhk alongwit... | |

| | furnishDetails | \ |
|---|---|-----|
| 0 | ['3 Fan', '4 Light', '1 Wardrobe', 'No AC', 'N... | |
| 1 | ['3 Wardrobe', '4 Fan', '1 Exhaust Fan', '1 Ge... | |
| 2 | | NaN |
| 3 | | NaN |
| 4 | | [] |

| | features | \ |
|---|---|-----|
| 0 | ['Feng Shui / Vaastu Compliant', 'Security / F... | |
| 1 | ['Security / Fire Alarm', 'Maintenance Staff',... | |
| 2 | ['Power Back-up', 'Feng Shui / Vaastu Complian... | |
| 3 | ['Security / Fire Alarm', 'Private Garden / Te... | |
| 4 | | NaN |

| | | rating | property_id |
|---|---|--------|-------------|
| 0 | ['Environment4 out of 5', 'Safety4 out of 5', ... | | C68850746 |
| 1 | ['Environment4 out of 5', 'Safety4 out of 5', ... | | H68850564 |
| 2 | ['Environment4 out of 5', 'Safety4 out of 5', ... | | J68850120 |
| 3 | | NaN | S68849476 |
| 4 | ['Environment5 out of 5', 'Safety3 out of 5', ... | | L47956793 |

[64]: # Clean price Column

```
def clean_price(price):
    if isinstance(price, str):
        price = price.strip()
        if 'Lac' in price:
            return float(price.replace(' Lac', '').replace(',', '')) * 100000
        elif 'Crore' in price:
            return float(price.replace(' Crore', '').replace(',', '')) * 10000000
        return 0

flats_df['price'] = flats_df['price'].apply(clean_price)
print(flats_df[['price']].head())
print(flats_df.info())
```

| | price |
|---|------------|
| 0 | 4500000.0 |
| 1 | 5000000.0 |
| 2 | 4000000.0 |
| 3 | 14700000.0 |
| 4 | 7000000.0 |

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 3017 entries, 0 to 3016
Data columns (total 19 columns):

| # | Column | Non-Null Count | Dtype |
|----|----------------|----------------|---------|
| 0 | property_name | 3017 non-null | object |
| 1 | society | 3016 non-null | object |
| 2 | price | 3017 non-null | float64 |
| 3 | rate | 3004 non-null | object |
| 4 | areaWithType | 3008 non-null | object |
| 5 | bedRoom | 3008 non-null | object |
| 6 | bathroom | 3008 non-null | object |
| 7 | balcony | 3008 non-null | object |
| 8 | additionalRoom | 1694 non-null | object |
| 9 | address | 3002 non-null | object |
| 10 | floorNum | 3006 non-null | object |
| 11 | facing | 2127 non-null | object |


```

12  agePossession      3007 non-null  object
13  nearbyLocations    2913 non-null  object
14  description        3008 non-null  object
15  furnishDetails     2203 non-null  object
16  features           2594 non-null  object
17  rating             2676 non-null  object
18  property_id        3008 non-null  object
dtypes: float64(1), object(18)
memory usage: 448.0+ KB
None

```

[78]: *# Q1 How many flats are listed in each society?*

```

society_counts = flats_df['society'].value_counts().head(10)
print(society_counts)

```

```

society
SS The Leaf3.8          73
Tulip Violet4.3         40
Shapoorji Pallonji Joyville Gurugram4.0    39
Signature Global Park4.0    36
Shree Vardhman Victoria3.8    35
Tulip Violet4.2         33
Emaar MGF Emerald Floors Premier3.8    32
Smart World Orchard      32
Smart World Gems         32
Paras Dews              31
Name: count, dtype: int64

```

[146]: *# Visualization*

```

plt.figure(figsize=(14, 8))
sns.barplot(x=society_counts.values, y=society_counts.index, palette='coolwarm')

plt.title('Top 10 Societies with Most Flats Listed', fontsize=20, weight='bold')
plt.xlabel('')
plt.ylabel('')

for index, value in enumerate(society_counts):
    plt.text(value, index, str(value), fontsize=12, color='black', va='center')

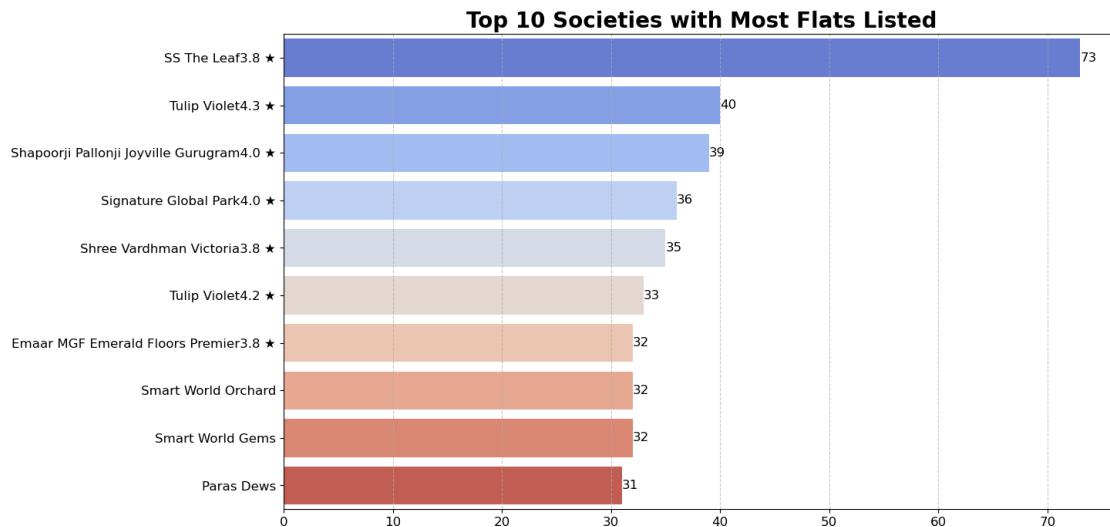
plt.xticks(fontsize=12)
plt.yticks(fontsize=12)
plt.grid(axis='x', linestyle='--', alpha=0.7)
plt.show()

```

C:\Users\Abc\AppData\Local\Temp\ipykernel_1100\1544059436.py:4: FutureWarning:

Passing `palette` without assigning `hue` is deprecated and will be removed in v0.14.0. Assign the `y` variable to `hue` and set `legend=False` for the same effect.

```
sns.barplot(x=society_counts.values, y=society_counts.index,
palette='coolwarm')
```



```
[97]: # Q2 What is the average price of 2 BHK flats?

average_price_2_bhk = flats_df[flats_df['bedRoom'] == '2 Bedrooms']['price'].
    ↪mean()
print("The average price of 2 BHK flats is: {:.2f}".
    ↪format(average_price_2_bhk))
```

The average price of 2 BHK flats is: 8848447.26

```
[ ]:
```

```
[128]: # Visualization

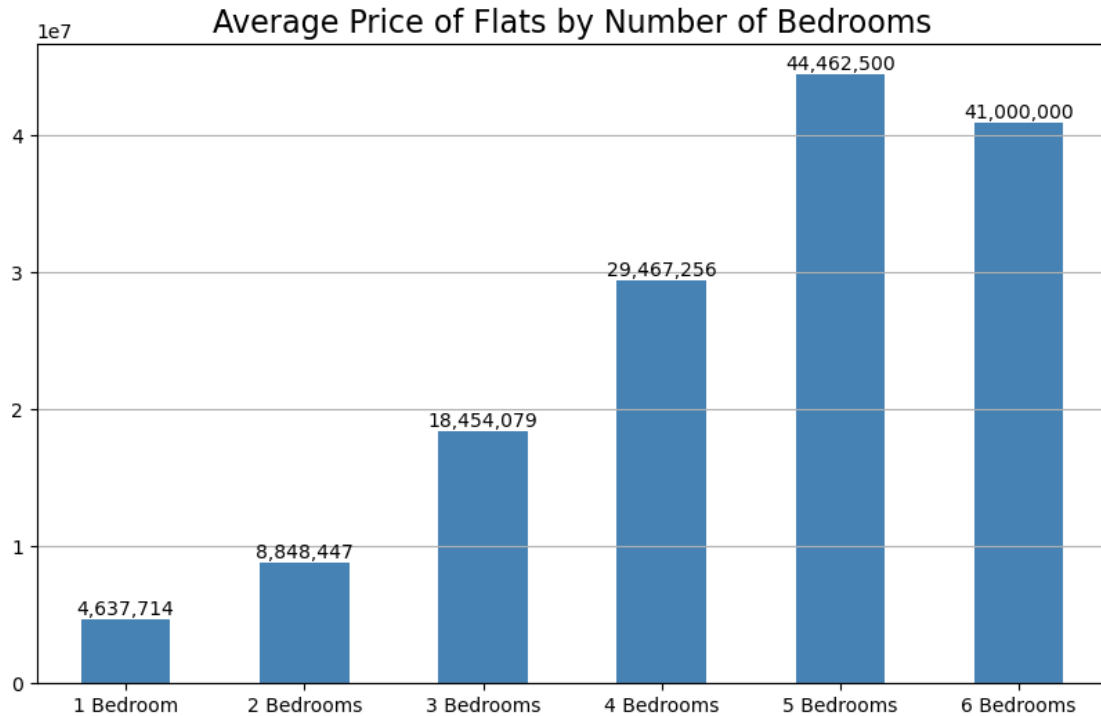
average_price_by_bedroom = flats_df.groupby('bedRoom')['price'].mean()
plt.figure(figsize=(10, 6))
bars = flats_df.groupby('bedRoom')['price'].mean().plot(kind='bar',
    ↪color='steelblue')

plt.title('Average Price of Flats by Number of Bedrooms', fontsize=16)
plt.xlabel('')
plt.ylabel('')
plt.xticks(rotation=0)
```

```
plt.grid(axis='y')

for bar in bars.patches:
    plt.text(bar.get_x() + bar.get_width() / 2, bar.get_height(), f'{bar.
        ↪get_height():,.0f}',
              ha='center', va='bottom')

plt.show()
```



[110]: *# Q3 Count the number of flats available in each bedroom category.*

```
bedroom_counts = flats_df['bedRoom'].value_counts()
print(bedroom_counts)
```

```
bedRoom
3 Bedrooms    1439
2 Bedrooms    948
4 Bedrooms    481
1 Bedroom     105
5 Bedrooms     32
6 Bedrooms      3
Name: count, dtype: int64
```

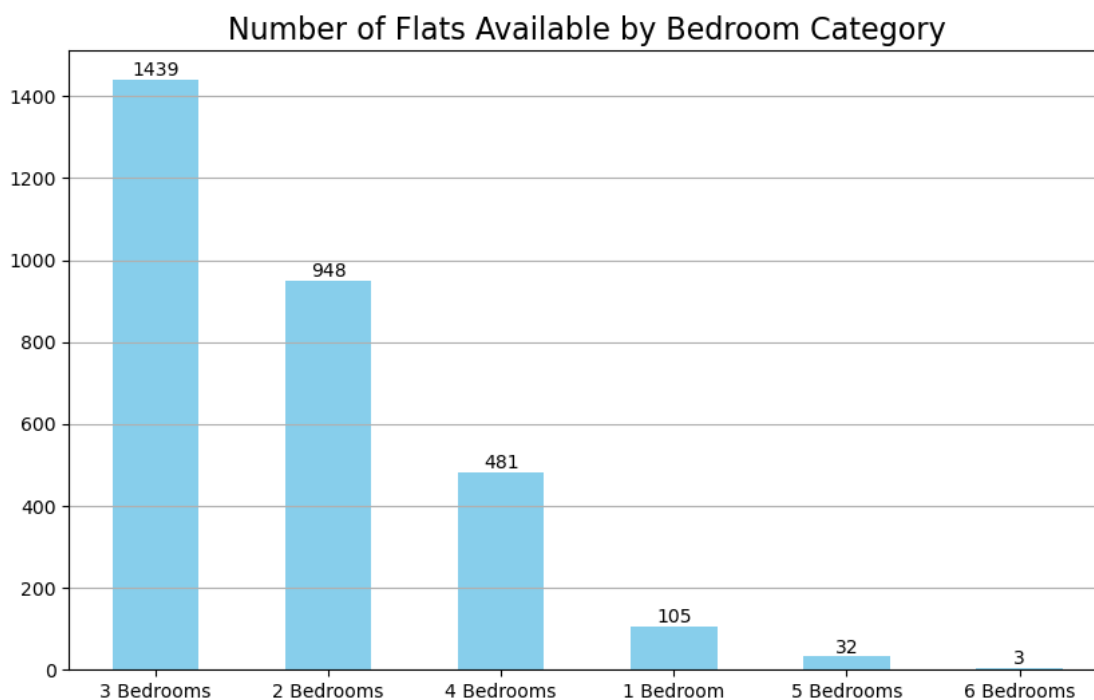
```
[144]: # Visualization

plt.figure(figsize=(10, 6))
bedroom_counts.plot(kind='bar', color='skyblue')

plt.title('Number of Flats Available by Bedroom Category', fontsize=16)
plt.xlabel('')
plt.ylabel('')
plt.xticks(rotation=0)
plt.grid(axis='y')

for index, value in enumerate(bedroom_counts):
    plt.text(index, value, str(value), ha='center', va='bottom')

plt.show()
```



```
[120]: # Q4 What is the distribution of flats based on the number of bathrooms?

bathroom_counts = flats_df['bathroom'].value_counts()
print(bathroom_counts)
```

```
bathroom
2 Bathrooms    1047
3 Bathrooms     991
4 Bathrooms     638
```

```
5 Bathrooms      170
1 Bathroom        115
6 Bathrooms       42
7 Bathrooms        5
Name: count, dtype: int64
```

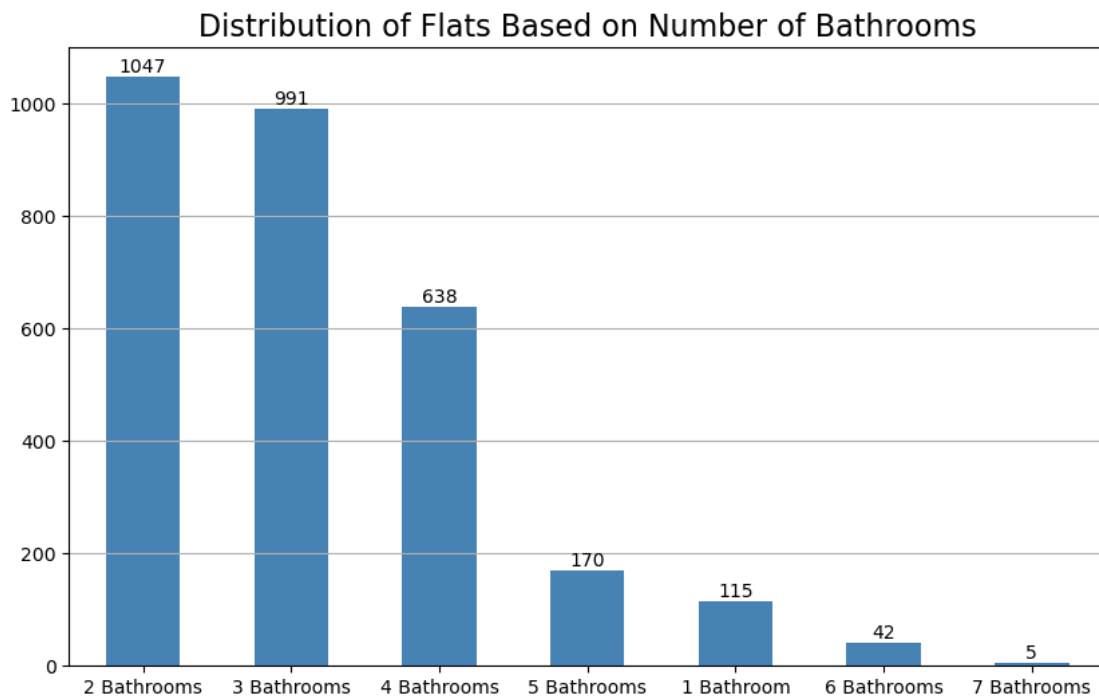
```
[142]: # Visualization

plt.figure(figsize=(10, 6))
bathroom_counts.plot(kind='bar', color='steelblue')

plt.title('Distribution of Flats Based on Number of Bathrooms', fontsize=16)
plt.xlabel('')
plt.ylabel('')
plt.xticks(rotation=0)
plt.grid(axis='y')

for index, value in enumerate(bathroom_counts):
    plt.text(index, value, str(value), ha='center', va='bottom')

plt.show()
```



```
[130]: flats_df.head()
```

[130]:

| | property_name | society | price \ |
|---|---------------------------------|------------------------|------------|
| 0 | 2 BHK Flat in Krishna Colony | maa bhagwati residency | 4500000.0 |
| 1 | 2 BHK Flat in Ashok Vihar | Apna Enclave | 5000000.0 |
| 2 | 2 BHK Flat in Sohna | Tulsiani Easy in Homes | 4000000.0 |
| 3 | 2 BHK Flat in Sector 61 Gurgaon | Smart World Orchard | 14700000.0 |
| 4 | 2 BHK Flat in Sector 92 Gurgaon | Parkwood Westend | 7000000.0 |

| | rate | areaWithType | bedRoom \ |
|---|---------------|--|------------|
| 0 | 5,000/sq.ft. | Carpet area: 900 (83.61 sq.m.) | 2 Bedrooms |
| 1 | 7,692/sq.ft. | Carpet area: 650 (60.39 sq.m.) | 2 Bedrooms |
| 2 | 6,722/sq.ft. | Carpet area: 595 (55.28 sq.m.) | 2 Bedrooms |
| 3 | 12,250/sq.ft. | Carpet area: 1200 (111.48 sq.m.) | 2 Bedrooms |
| 4 | 5,204/sq.ft. | Super Built up area 1345(124.95 sq.m.) | 2 Bedrooms |

| | bathroom | balcony | additionalRoom \ |
|---|-------------|-------------|------------------|
| 0 | 2 Bathrooms | 1 Balcony | NaN |
| 1 | 2 Bathrooms | 1 Balcony | NaN |
| 2 | 2 Bathrooms | 3 Balconies | NaN |
| 3 | 2 Bathrooms | 2 Balconies | Study Room |
| 4 | 2 Bathrooms | 3 Balconies | Study Room |

| | address | floorNum | facing \ |
|---|-------------------------------------|-------------------|----------|
| 0 | Krishna Colony, Gurgaon, Haryana | 4th of 4 Floors | West |
| 1 | 46b, Ashok Vihar, Gurgaon, Haryana | 1st of 3 Floors | West |
| 2 | Sohna, Gurgaon, Haryana | 12nd of 14 Floors | NaN |
| 3 | Sector 61 Gurgaon, Gurgaon, Haryana | 2nd of 4 Floors | NaN |
| 4 | Sector 92 Gurgaon, Gurgaon, Haryana | 5th of 8 Floors | NaN |

| | agePossession | nearbyLocations \ |
|---|--------------------|---|
| 0 | 1 to 5 Year Old | ['Chintapurni Mandir', 'State bank ATM', 'Pear... |
| 1 | 10+ Year Old | ['Chintapurni Mandir', 'Sheetla Mata Mandir', ... |
| 2 | 0 to 1 Year Old | ['Huda City Metro', 'Golf Course extn road', '... |
| 3 | Dec-23 | ['Sector 55-56 Metro station', 'Bestech Centra... |
| 4 | Under Construction | ['Yadav Clinic', 'Bangali Clinic', 'Dr. J. S. ... |

| | description \ |
|---|---|
| 0 | So with lift.Maa bhagwati residency is one of ... |
| 1 | Property situated on main road, railway statio... |
| 2 | This property is 15 km away from badshapur, gu... |
| 3 | Near to metro station of sector 56 and opposit... |
| 4 | We are the proud owners of this 2 bhk alongwit... |

| | furnishDetails \ |
|---|---|
| 0 | ['3 Fan', '4 Light', '1 Wardrobe', 'No AC', 'N... |
| 1 | ['3 Wardrobe', '4 Fan', '1 Exhaust Fan', '1 Ge... |
| 2 | NaN |
| 3 | NaN |

4

[]

```

                                features \
0 ['Feng Shui / Vaastu Compliant', 'Security / F...
1 ['Security / Fire Alarm', 'Maintenance Staff',...
2 ['Power Back-up', 'Feng Shui / Vaastu Complian...
3 ['Security / Fire Alarm', 'Private Garden / Te...
4                                     NaN

                                rating property_id
0 ['Environment4 out of 5', 'Safety4 out of 5', ... C68850746
1 ['Environment4 out of 5', 'Safety4 out of 5', ... H68850564
2 ['Environment4 out of 5', 'Safety4 out of 5', ... J68850120
3                                     NaN S68849476
4 ['Environment5 out of 5', 'Safety3 out of 5', ... L47956793

```

```

[132]: # Clean Area Column
flats_df['area'] = flats_df['areaWithType'].str.extract('(\d+\.? \d*)')
flats_df['area_type'] = flats_df['areaWithType'].str.extract('([A-Za-z ]+)')

flats_df['area'] = flats_df['area'].astype(float)
flats_df.head()

```

```

<>:2: SyntaxWarning: invalid escape sequence '\d'
<>:2: SyntaxWarning: invalid escape sequence '\d'
C:\Users\Abc\AppData\Local\Temp\ipykernel_1100\2579475786.py:2: SyntaxWarning:
invalid escape sequence '\d'
    flats_df['area'] = flats_df['areaWithType'].str.extract('(\d+\.? \d*)')

```

[136]:

```

[136]:
                                property_name          society      price \
0      2 BHK Flat in Krishna Colony  maa bhagwati residency  4500000.0
1      2 BHK Flat in Ashok Vihar      Apna Enclave          5000000.0
2      2 BHK Flat in Sohna           Tulsiani Easy in Homes  4000000.0
3  2 BHK Flat in Sector 61 Gurgaon    Smart World Orchard  14700000.0
4  2 BHK Flat in Sector 92 Gurgaon    Parkwood Westend    7000000.0

                                rate          areaWithType      bedRoom \
0      5,000/sq.ft.      Carpet area: 900 (83.61 sq.m.)  2 Bedrooms
1      7,692/sq.ft.      Carpet area: 650 (60.39 sq.m.)  2 Bedrooms
2      6,722/sq.ft.      Carpet area: 595 (55.28 sq.m.)  2 Bedrooms
3     12,250/sq.ft.      Carpet area: 1200 (111.48 sq.m.)  2 Bedrooms
4      5,204/sq.ft.  Super Built up area 1345(124.95 sq.m.)  2 Bedrooms

                                bathroom      balcony additionalRoom \
0      2 Bathrooms      1 Balcony      NaN

```

| | | | |
|---|-------------|-------------|------------|
| 1 | 2 Bathrooms | 1 Balcony | NaN |
| 2 | 2 Bathrooms | 3 Balconies | NaN |
| 3 | 2 Bathrooms | 2 Balconies | Study Room |
| 4 | 2 Bathrooms | 3 Balconies | Study Room |

| | address | ... facing | agePossession | \ |
|---|-------------------------------------|------------|--------------------|---|
| 0 | Krishna Colony, Gurgaon, Haryana | ... West | 1 to 5 Year Old | |
| 1 | 46b, Ashok Vihar, Gurgaon, Haryana | ... West | 10+ Year Old | |
| 2 | Sohna, Gurgaon, Haryana | ... NaN | 0 to 1 Year Old | |
| 3 | Sector 61 Gurgaon, Gurgaon, Haryana | ... NaN | Dec-23 | |
| 4 | Sector 92 Gurgaon, Gurgaon, Haryana | ... NaN | Under Construction | |

| | nearbyLocations | \ |
|---|---|---|
| 0 | ['Chintapurni Mandir', 'State bank ATM', 'Pear... | |
| 1 | ['Chintapurni Mandir', 'Sheetla Mata Mandir', ... | |
| 2 | ['Huda City Metro', 'Golf Course extn road', '... | |
| 3 | ['Sector 55-56 Metro station', 'Bestech Centra... | |
| 4 | ['Yadav Clinic', 'Bangali Clinic', 'Dr. J. S. ... | |

| | description | \ |
|---|---|---|
| 0 | So with lift.Maa bhagwati residency is one of ... | |
| 1 | Property situated on main road, railway statio... | |
| 2 | This property is 15 km away from badshapur, gu... | |
| 3 | Near to metro station of sector 56 and opposit... | |
| 4 | We are the proud owners of this 2 bhk alongwit... | |

| | furnishDetails | \ |
|---|---|---|
| 0 | ['3 Fan', '4 Light', '1 Wardrobe', 'No AC', 'N... | |
| 1 | ['3 Wardrobe', '4 Fan', '1 Exhaust Fan', '1 Ge... | |
| 2 | NaN | |
| 3 | NaN | |
| 4 | [] | |

| | features | \ |
|---|---|---|
| 0 | ['Feng Shui / Vaastu Compliant', 'Security / F... | |
| 1 | ['Security / Fire Alarm', 'Maintenance Staff',... | |
| 2 | ['Power Back-up', 'Feng Shui / Vaastu Complian... | |
| 3 | ['Security / Fire Alarm', 'Private Garden / Te... | |
| 4 | NaN | |

| | rating | property_id | area | \ |
|---|---|-------------|--------|---|
| 0 | ['Environment4 out of 5', 'Safety4 out of 5', ... | C68850746 | 900.0 | |
| 1 | ['Environment4 out of 5', 'Safety4 out of 5', ... | H68850564 | 650.0 | |
| 2 | ['Environment4 out of 5', 'Safety4 out of 5', ... | J68850120 | 595.0 | |
| 3 | NaN | S68849476 | 1200.0 | |
| 4 | ['Environment5 out of 5', 'Safety3 out of 5', ... | L47956793 | 1345.0 | |


```

        area_type
0      Carpet area
1      Carpet area
2      Carpet area
3      Carpet area
4  Super Built up area

```

```
[5 rows x 21 columns]
```

```
[138]: # Q5 Determine the average area of flats for each type of facing (e.g., North, ↵
        ↪South).
```

```

average_area_by_facing = flats_df.groupby('facing')['area'].mean()
print(average_area_by_facing)

```

```

facing
East      1930.989858
North     3931.697781
North-East 3162.792510
North-West 2413.443025
South     1921.257291
South-East 2025.120000
South-West 6500.352074
West      1870.265464
Name: area, dtype: float64

```

```
[160]: # Visualization
```

```

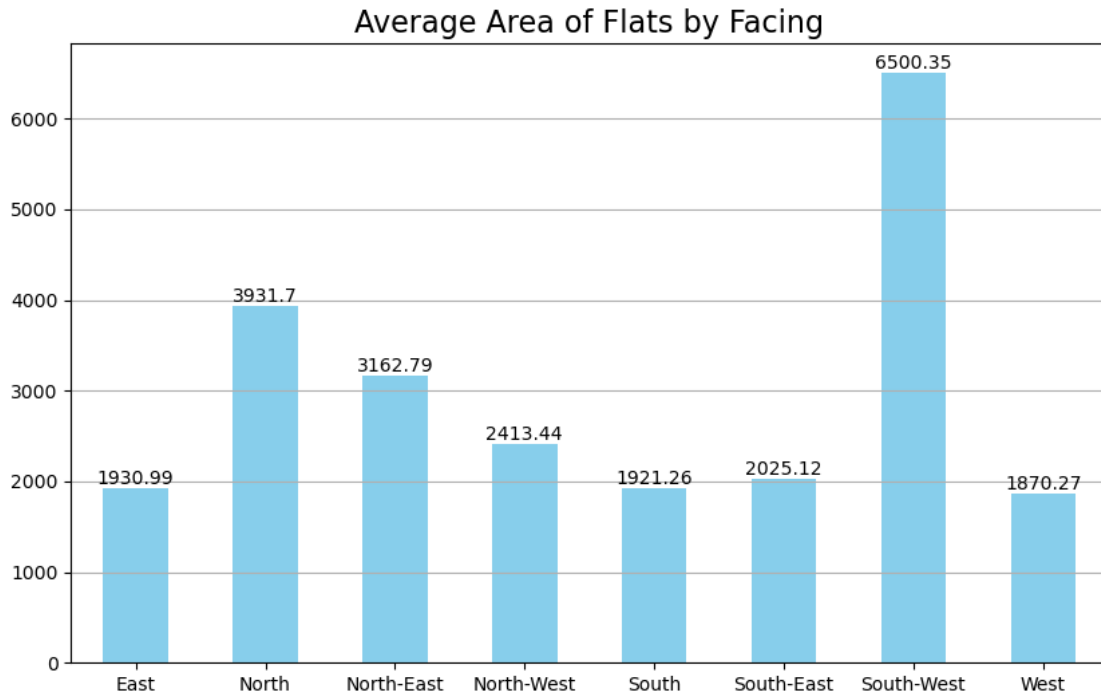
plt.figure(figsize=(10, 6))
average_area_by_facing.plot(kind='bar', color='skyblue')

plt.title('Average Area of Flats by Facing', fontsize=16)
plt.xlabel('')
plt.ylabel('')
plt.xticks(rotation=0)
plt.grid(axis='y')

for index, value in enumerate(average_area_by_facing):
    plt.text(index, value, round(value, 2), ha='center', va='bottom')

plt.show()

```



[162]: *# Q6 Analyze the correlation between the area of flats and their prices.*

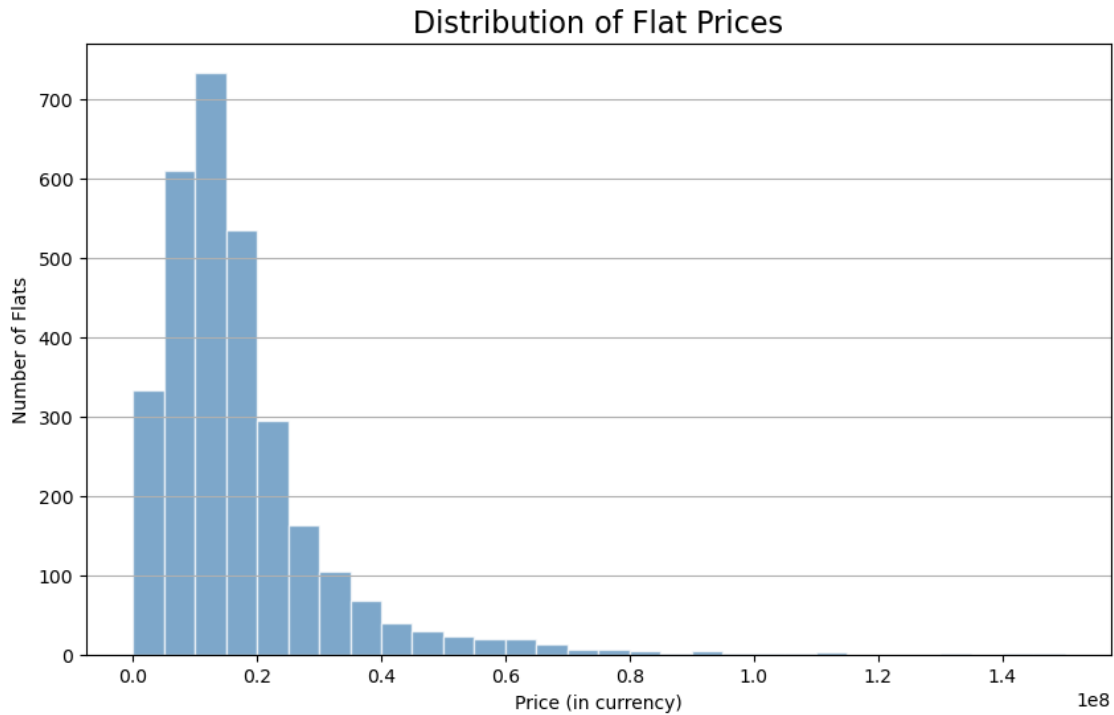
```
correlation = flats_df['area'].corr(flats_df['price'])
print("Correlation between area and price:", correlation)
```

Correlation between area and price: -0.00929493102328238

[168]: *# Q7 Create a histogram to visualize the distribution of flat prices.*

```
plt.figure(figsize=(10, 6))
plt.hist(flats_df['price'], bins=30, color='steelblue', edgecolor='white',
        alpha=0.7)
plt.title('Distribution of Flat Prices', fontsize=16)
plt.xlabel('Price (in currency)')
plt.ylabel('Number of Flats')
plt.grid(axis='y')

plt.show()
```



```
[212]: new = pd.read_csv("Downloads/houses.csv")
```

```
[217]: houses_df = new
```

```
[219]: houses_df.head()
```

```
[219]:
```

| | property_name \ |
|---|--|
| 0 | 5 Bedroom House for sale in Sector 70A Gurgaon |
| 1 | 5 Bedroom House for sale in Sector 21A Faridabad |
| 2 | 10 Bedroom House for sale in Sushant Lok Phase 1 |
| 3 | 21 Bedroom House for sale in Sector 54 Gurgaon |
| 4 | 12 Bedroom House for sale in Sushant Lok Phase 1 |

| | link | society \ |
|---|---|------------------|
| 0 | https://www.99acres.com/5-bhk-bedroom-independ... | Bptp Visionnaire |
| 1 | https://www.99acres.com/5-bhk-bedroom-independ... | NaN |
| 2 | https://www.99acres.com/10-bhk-bedroom-indepen... | NaN |
| 3 | https://www.99acres.com/21-bhk-bedroom-indepen... | NaN |
| 4 | https://www.99acres.com/12-bhk-bedroom-indepen... | NaN |

| | price | rate | area \ |
|---|------------|----------------|-----------------------|
| 0 | 5.25 Crore | 20,115/sq.ft. | (242 sq.m.) Plot Area |
| 1 | 5.7 Crore | 105,751/sq.ft. | (50 sq.m.) Plot Area |
| 2 | 2.1 Crore | 38,251/sq.ft. | (51 sq.m.) Plot Area |

| | | | |
|---|---------|---------------|-----------------------|
| 3 | 5 Crore | 43,066/sq.ft. | (108 sq.m.) Plot Area |
| 4 | 3 Crore | 53,763/sq.ft. | (52 sq.m.) Plot Area |

| | areaWithType | bedRoom | bathroom | balcony | ... | \ |
|---|-----------------------------|-------------|--------------|--------------|-----|---|
| 0 | Plot area 290(242.48 sq.m.) | 5 Bedrooms | 4 Bathrooms | 3+ Balconies | ... | |
| 1 | Plot area 539(50.07 sq.m.) | 5 Bedrooms | 4 Bathrooms | 2 Balconies | ... | |
| 2 | Plot area 61(51 sq.m.) | 10 Bedrooms | 10 Bathrooms | 3+ Balconies | ... | |
| 3 | Plot area 129(107.86 sq.m.) | 21 Bedrooms | 21 Bathrooms | 3+ Balconies | ... | |
| 4 | Plot area 62(51.84 sq.m.) | 12 Bedrooms | 12 Bathrooms | 3+ Balconies | ... | |

| | address | noOfFloor | facing | \ |
|---|---|-----------|------------|---|
| 0 | 29b, Sector 70A Gurgaon, Gurgaon, Haryana | 3 Floors | North-East | |
| 1 | Sector 21A Faridabad, Gurgaon, Haryana | 2 Floors | NaN | |
| 2 | Sushant Lok Phase 1, Gurgaon, Haryana | 5 Floors | West | |
| 3 | Sector 54 Gurgaon, Gurgaon, Haryana | 5 Floors | North | |
| 4 | 1228, Sushant Lok Phase 1, Gurgaon, Haryana | 5 Floors | West | |

| | agePossession | nearbyLocations | \ |
|---|------------------|---|---|
| 0 | 0 to 1 Year Old | ['Good Earth City Center 2', 'Kunskapsskolan I... | |
| 1 | 5 to 10 Year Old | NaN | |
| 2 | 0 to 1 Year Old | ['Sector 42-43 metro station', 'Huda city cent... | |
| 3 | 1 to 5 Year Old | ['Sector 53-54 metro station', 'Sector 54 chow... | |
| 4 | Within 6 months | ['Sector 42-43 metro station', 'Huda city cent... | |

| | description | \ |
|---|---|---|
| 0 | Do you wish to buy an independent house in sec... | |
| 1 | Hi, we have an independent house/villa availab... | |
| 2 | Monthly rental income is rs1,40,000/- Best opt... | |
| 3 | 129 sq yd plot size. 5 floors built up , fully... | |
| 4 | Best for investment purpose. Fully furnished b... | |

| | furnishDetails | \ |
|---|---|---|
| 0 | ['1 Wardrobe', '1 Fan', '1 Exhaust Fan', '1 Ge... | |
| 1 | ['1 Water Purifier', '5 Fan', '1 Exhaust Fan',... | |
| 2 | ['10 Bed', '3 Fan', '10 Geyser', '2 Light', 'N... | |
| 3 | ['1 Water Purifier', '21 Fan', '1 Fridge', '1 ... | |
| 4 | ['1 Water Purifier', '1 Fridge', '12 Fan', '1 ... | |

| | features | \ |
|---|---|---|
| 0 | ['Feng Shui / Vaastu Compliant', 'Private Gard... | |
| 1 | ['Private Garden / Terrace', 'Park', 'Visitor ... | |
| 2 | ['Maintenance Staff', 'Water Storage', 'Visito... | |
| 3 | ['Feng Shui / Vaastu Compliant', 'Private Gard... | |
| 4 | ['Maintenance Staff', 'Water Storage', 'Visito... | |

| | rating | property_id |
|---|---|-------------|
| 0 | ['Environment5 out of 5', 'Lifestyle4 out of 5... | K70037724 |

```

1                                     NaN    E69288322
2  ['Environment5 out of 5', 'Lifestyle5 out of 5...  F69536898
3  ['Environment4 out of 5', 'Lifestyle5 out of 5...  R69483164
4  ['Environment5 out of 5', 'Lifestyle5 out of 5...  M69381272

```

[5 rows x 21 columns]

```
[221]: houses_df.tail()
```

```

[221]:                                     property_name \
1039  4 Bedroom House for sale in Sector 26 Gurgaon
1040          3 Bedroom House for sale in DLF Phase 2
1041  3 Bedroom House for sale in Sector 26 Gurgaon
1042  3 Bedroom House for sale in Sector 26 Gurgaon
1043          3 Bedroom House for sale in Bissar

                                     link \
1039  https://www.99acres.com/4-bhk-bedroom-independ...
1040  https://www.99acres.com/3-bhk-bedroom-independ...
1041  https://www.99acres.com/3-bhk-bedroom-independ...
1042  https://www.99acres.com/3-bhk-bedroom-independ...
1043  https://www.99acres.com/3-bhk-bedroom-independ...

                                     society      price      rate \
1039          DLF City Phase 1    5.5 Crore  30,556/sq.ft.
1040    DLF City Plots Phase 2    4.25 Crore  31,481/sq.ft.
1041          DLF City Phase 1    4.5 Crore  33,333/sq.ft.
1042          DLF City Phase 1    3.25 Crore  33,129/sq.ft.
1043  Silverglades Tarudhan Valley Golf Resort    3.6 Crore  14,815/sq.ft.

                                     area      areaWithType      bedRoom \
1039  (167 sq.m.) Plot Area  Plot area 200(167.23 sq.m.)  4 Bedrooms
1040  (125 sq.m.) Plot Area  Plot area 150(125.42 sq.m.)  3 Bedrooms
1041  (125 sq.m.) Plot Area  Plot area 150(125.42 sq.m.)  3 Bedrooms
1042  (91 sq.m.) Plot Area  Plot area 109(91.14 sq.m.)  3 Bedrooms
1043  (226 sq.m.) Plot Area  Plot area 270(225.75 sq.m.)  3 Bedrooms

      bathroom      balcony  ...      address \
1039  4 Bathrooms  3 Balconies  ...  Sector 26 Gurgaon, Gurgaon, Haryana
1040  2 Bathrooms  3 Balconies  ...      DLF Phase 2, Gurgaon, Haryana
1041  3 Bathrooms  2 Balconies  ...  Sector 26 Gurgaon, Gurgaon, Haryana
1042  3 Bathrooms  3 Balconies  ...  Sector 26 Gurgaon, Gurgaon, Haryana
1043  3 Bathrooms  2 Balconies  ...      Bissar, Gurgaon, Haryana

      noOfFloor      facing      agePossession \
1039  2 Floors  North-East  5 to 10 Year Old
1040  2 Floors      North      10+ Year Old

```

| | | | |
|------|----------|------|------------------|
| 1041 | 2 Floors | East | 5 to 10 Year Old |
| 1042 | 2 Floors | West | 10+ Year Old |
| 1043 | 2 Floors | NaN | 5 to 10 Year Old |

| | nearbyLocations \ |
|------|---|
| 1039 | ['Dlf phase 1 metro station', 'Sikandarpur met... |
| 1040 | ['Sahara Mall', 'Delhi Public School', 'Privat... |
| 1041 | ['Dlf phase 1 metro station', 'Sikandarpur met... |
| 1042 | ['Dlf phase 1 metro station', 'Sikandarpur met... |
| 1043 | NaN |

| | description \ |
|------|---|
| 1039 | Residential house, g block, dlf phase-1, gurga... |
| 1040 | Residential house, k block, dlf phase-2, gurga... |
| 1041 | Residential house, e block 150 sq.Yds , dlf ph... |
| 1042 | Residential house, e block, dlf phase-1, gurga... |
| 1043 | This 3 bhk house for sale in bissar is an idea... |

| | furnishDetails \ |
|------|---|
| 1039 | ['10 Light', 'No AC', 'No Bed', 'No Chimney', ... |
| 1040 | NaN |
| 1041 | ['6 Fan', '20 Light', 'No AC', 'No Bed', 'No C... |
| 1042 | ['5 Fan', '16 Light', 'No AC', 'No Bed', 'No C... |
| 1043 | ['1 Water Purifier', '10 Fan', '1 Fridge', '1 ... |

| | features \ |
|------|---|
| 1039 | ['Feng Shui / Vaastu Compliant', 'High Ceiling... |
| 1040 | ['High Ceiling Height', 'Maintenance Staff', '... |
| 1041 | ['Feng Shui / Vaastu Compliant', 'Private Gard... |
| 1042 | ['Feng Shui / Vaastu Compliant', 'Private Gard... |
| 1043 | ['Centrally Air Conditioned', 'Water purifier'... |

| | rating | property_id |
|------|---|-------------|
| 1039 | NaN | Q70281606 |
| 1040 | ['Environment5 out of 5', 'Safety5 out of 5', ... | B70281212 |
| 1041 | NaN | L70280308 |
| 1042 | NaN | J70280212 |
| 1043 | NaN | W70221444 |

[5 rows x 21 columns]

```
[223]: houses_df.columns
```

```
[223]: Index(['property_name', 'link', 'society', 'price', 'rate', 'area',
        'areaWithType', 'bedRoom', 'bathroom', 'balcony', 'additionalRoom',
        'address', 'noOfFloor', 'facing', 'agePossession', 'nearbyLocations',
        'description', 'furnishDetails', 'features', 'rating', 'property_id'],
```

```
dtype='object')
```

```
[225]: houses_df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 1044 entries, 0 to 1043
Data columns (total 21 columns):
#   Column                Non-Null Count  Dtype
---  -
0   property_name         1044 non-null   object
1   link                  1044 non-null   object
2   society               453 non-null    object
3   price                 968 non-null    object
4   rate                  1005 non-null   object
5   area                  1044 non-null   object
6   areaWithType          987 non-null    object
7   bedRoom               987 non-null    object
8   bathroom              987 non-null    object
9   balcony               987 non-null    object
10  additionalRoom         589 non-null    object
11  address                1031 non-null   object
12  noOfFloor              967 non-null    object
13  facing                 674 non-null    object
14  agePossession          987 non-null    object
15  nearbyLocations        913 non-null    object
16  description            1036 non-null   object
17  furnishDetails         743 non-null    object
18  features               674 non-null    object
19  rating                 907 non-null    object
20  property_id            1036 non-null   object
dtypes: object(21)
memory usage: 171.4+ KB
```

```
[227]: # Clean Price Column
```

```
def cleaned_price(price):
    if isinstance(price, str):
        price = price.strip()
        if 'Lac' in price:
            return float(price.replace(' Lac', '').replace(',', '')) * 100000
        elif 'Crore' in price:
            return float(price.replace(' Crore', '').replace(',', '')) * 10000000
    return 0

houses_df['price'] = houses_df['price'].apply(cleaned_price)
houses_df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 1044 entries, 0 to 1043
Data columns (total 21 columns):
#   Column                Non-Null Count  Dtype
---  -
0   property_name          1044 non-null   object
1   link                   1044 non-null   object
2   society                 453 non-null    object
3   price                   1044 non-null   float64
4   rate                   1005 non-null   object
5   area                   1044 non-null   object
6   areaWithType           987 non-null    object
7   bedRoom                 987 non-null    object
8   bathroom               987 non-null    object
9   balcony                987 non-null    object
10  additionalRoom          589 non-null    object
11  address                 1031 non-null   object
12  noOfFloor               967 non-null    object
13  facing                 674 non-null    object
14  agePossession           987 non-null    object
15  nearbyLocations        913 non-null    object
16  description             1036 non-null   object
17  furnishDetails          743 non-null    object
18  features                674 non-null    object
19  rating                 907 non-null    object
20  property_id            1036 non-null   object
dtypes: float64(1), object(20)
memory usage: 171.4+ KB
```

```
[231]: # Q1 What is the most common number of floors in the listed houses?
```

```
most_common_floors = houses_df['noOfFloor'].mode()[0]
print(most_common_floors)
```

3 Floors

```
[245]: # Q2 Calculate the average price of houses with more than 3 bedrooms.
```

```
average_price = houses_df[houses_df['bedRoom'] > 3]['price'].mean()
print(average_price)

average_price_by_bedroom = houses_df[houses_df['bedRoom'] > 3].
    ↳groupby('bedRoom')['price'].mean().head()
print(average_price_by_bedroom)
```

```
58821562.94200849
bedRoom
4.0      5.118146e+07
```



```
5.0    6.424264e+07
6.0    6.485366e+07
7.0    5.111250e+07
8.0    5.207632e+07
Name: price, dtype: float64
```

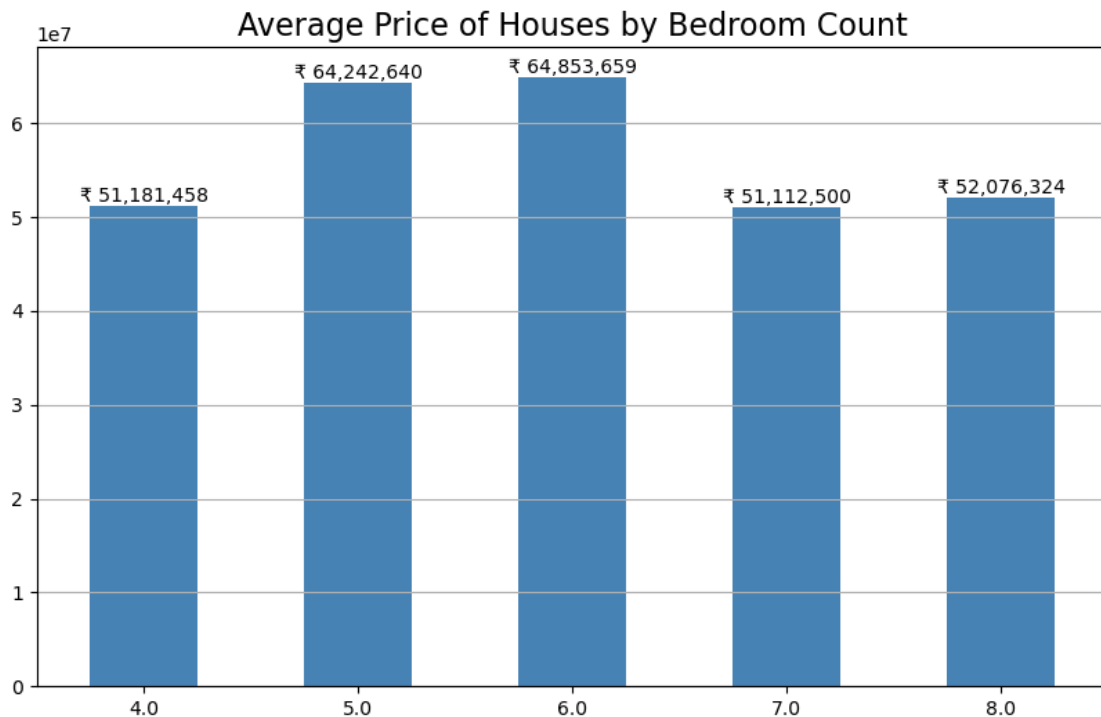
```
[255]: # Visualization

plt.figure(figsize=(10, 6))
average_price_by_bedroom.plot(kind='bar', color='steelblue')

plt.title('Average Price of Houses by Bedroom Count', fontsize=16)
plt.xlabel('')
plt.ylabel('')
plt.xticks(rotation=0)
plt.grid(axis='y')

for index, value in enumerate(average_price_by_bedroom):
    plt.text(index, value, f' ₹ {value:,.0f}', ha='center', va='bottom')

plt.show()
```



```
[259]: # Q3 Compare the average areas of furnished vs. unfurnished houses.
```

```
houses_df[['furnishDetails']]
```

```
[259]:
```

| | furnishDetails |
|------|---|
| 0 | ['1 Wardrobe', '1 Fan', '1 Exhaust Fan', '1 Ge... |
| 1 | ['1 Water Purifier', '5 Fan', '1 Exhaust Fan',... |
| 2 | ['10 Bed', '3 Fan', '10 Geyser', '2 Light', 'N... |
| 3 | ['1 Water Purifier', '21 Fan', '1 Fridge', '1 ... |
| 4 | ['1 Water Purifier', '1 Fridge', '12 Fan', '1 ... |
| ... | ... |
| 1039 | ['10 Light', 'No AC', 'No Bed', 'No Chimney', ... |
| 1040 | NaN |
| 1041 | ['6 Fan', '20 Light', 'No AC', 'No Bed', 'No C... |
| 1042 | ['5 Fan', '16 Light', 'No AC', 'No Bed', 'No C... |
| 1043 | ['1 Water Purifier', '10 Fan', '1 Fridge', '1 ... |

[1044 rows x 1 columns]

```
[265]: # Q3 Compare the average areas of furnished vs. unfurnished houses.
```

```
def determine_furnishing(furnish_details):
    if pd.isna(furnish_details):
        return 'Unfurnished'

    items = eval(furnish_details) # Convert string representation to list
    furnished_items = ['Wardrobe', 'Fan', 'Fridge', 'Bed', 'Geyser', 'Light', 'AC']

    # Check if any furnished item is present in the furnishDetails
    for item in items:
        if any(furnished_item in item for furnished_item in furnished_items):
            return 'Furnished'

    return 'Unfurnished'

houses_df['status'] = houses_df['furnishDetails'].apply(determine_furnishing)
print(houses_df[['furnishDetails', 'status']].head())
```

| | furnishDetails | status |
|---|---|-----------|
| 0 | ['1 Wardrobe', '1 Fan', '1 Exhaust Fan', '1 Ge... | Furnished |
| 1 | ['1 Water Purifier', '5 Fan', '1 Exhaust Fan',... | Furnished |
| 2 | ['10 Bed', '3 Fan', '10 Geyser', '2 Light', 'N... | Furnished |
| 3 | ['1 Water Purifier', '21 Fan', '1 Fridge', '1 ... | Furnished |
| 4 | ['1 Water Purifier', '1 Fridge', '12 Fan', '1 ... | Furnished |

```
[271]: # Q3 Compare the average areas of furnished vs. unfurnished houses.
```

```
# Clean Area Column
houses_df['area_value'] = houses_df['areaWithType'].str.extract('(\d+\.\d*)')
houses_df['area_type'] = houses_df['areaWithType'].str.extract('([A-Za-z ]+)')
houses_df['area_value'] = houses_df['area_value'].astype(float)
print(houses_df[['areaWithType', 'area_type', 'area_value']].head())
```

[279]: # Q3 Compare the average areas of furnished vs. unfurnished houses.

```
average_area_comparison = houses_df.groupby('status')['area_value'].mean()
print(average_area_comparison)
```

```
status
Furnished      959.917946
Unfurnished    918.916781
Name: area_value, dtype: float64
```

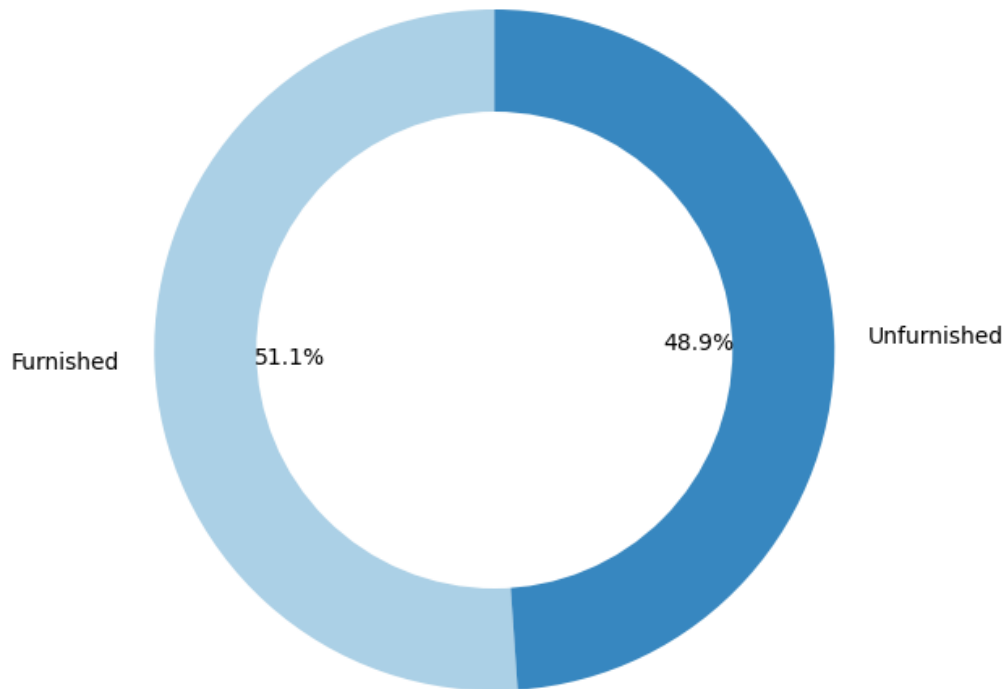
[285]: # Visualization

```
plt.figure(figsize=(8, 6))
plt.pie(average_area_comparison, labels=average_area_comparison.index,
        autopct='%1.1f%%', startangle=90, colors=sns.color_palette("Blues",
        len(average_area_comparison)))

centre_circle = plt.Circle((0, 0), 0.70, fc='white')
fig = plt.gcf()
fig.gca().add_artist(centre_circle)

plt.title('Average Area Comparison: Furnished vs Unfurnished Houses',
        fontsize=16)
plt.axis('equal')
plt.show()
```

Average Area Comparison: Furnished vs Unfurnished Houses



[293]: *# Key Performance Indicator (KPI)*

```
average_price_flats = flats_df['price'].mean()
average_price_houses = houses_df['price'].mean()
average_area_flats = flats_df['area'].mean()
average_area_houses = houses_df['area_value'].mean()
```

```
print("Flats Data KPIs:")
print(f"Average Price per Flat: {average_price_flats:,.2f}")
print(f"Average Area per Flat: {average_area_flats:.2f} sq.ft")
```

Flats Data KPIs:

Average Price per Flat: 0.00

Average Area per Flat: 2696.63 sq.ft

```
[295]: print("\nHouses Data KPIs:")
print(f"Average Price per House: {average_price_houses:,.2f}")
print(f"Average Area per House: {average_area_houses:.2f} sq.ft")
```

Houses Data KPIs:

Average Price per House: 45,224,675.29

Average Area per House: 946.62 sq.ft

[]: