

dubai-real-estate-analysis

0.1 Data Science Project - Dubai Real Estate Goldmine, UAE Rental Market Data

- This dataset can be found on Kaggle : <https://www.kaggle.com/datasets/azharsaleem/real-estate-goldmine-dubai-uae-rental-market>

```
[1]: import pandas as pd
import numpy as np
import seaborn as sns
import matplotlib.pyplot as plt
import plotly.express as px
```

```
[2]: df = pd.read_csv("dubai_Rent.csv")
```

```
[3]: df.head(3)
```

```
[3]:
```

		Address	Rent	Beds	Baths	\
0		The Gate Tower 2, The Gate Tower, Shams Gate D...	124000	3	4	
1		Water's Edge, Yas Island, Abu Dhabi	140000	3	4	
2		Al Raha Lofts, Al Raha Beach, Abu Dhabi	99000	2	3	

	Type	Area_in_sqft	Rent_per_sqft	Rent_category	Frequency	\
0	Apartment	1785	69.467787	Medium	Yearly	
1	Apartment	1422	98.452883	Medium	Yearly	
2	Apartment	1314	75.342466	Medium	Yearly	

	Furnishing	Purpose	Posted_date	Age_of_listing_in_days	Location	\
0	Unfurnished	For Rent	2024-03-07	45	Al Reem Island	
1	Unfurnished	For Rent	2024-03-08	44	Yas Island	
2	Furnished	For Rent	2024-03-21	31	Al Raha Beach	

	City	Latitude	Longitude
0	Abu Dhabi	24.493598	54.407841
1	Abu Dhabi	24.494022	54.607372
2	Abu Dhabi	24.485931	54.600939

```
[4]: df.info()
```

```

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 73742 entries, 0 to 73741
Data columns (total 17 columns):
#   Column                Non-Null Count  Dtype
---  -
0   Address                73742 non-null  object
1   Rent                   73742 non-null  int64
2   Beds                   73742 non-null  int64
3   Baths                  73742 non-null  int64
4   Type                   73742 non-null  object
5   Area_in_sqft           73742 non-null  int64
6   Rent_per_sqft          73742 non-null  float64
7   Rent_category          73742 non-null  object
8   Frequency              73742 non-null  object
9   Furnishing             73742 non-null  object
10  Purpose                73742 non-null  object
11  Posted_date            73742 non-null  object
12  Age_of_listing_in_days 73742 non-null  int64
13  Location               73742 non-null  object
14  City                   73742 non-null  object
15  Latitude               73023 non-null  float64
16  Longitude              73023 non-null  float64
dtypes: float64(3), int64(5), object(9)
memory usage: 9.6+ MB

```

```
[5]: df["Posted_date"] = pd.to_datetime(df["Posted_date"])
```

```

[6]: numerical_columns = [co for co in df.columns if df[co].dtype != 'object']
categorical_columns = [col for col in df.columns if df[col].dtype == 'object']
print("Numerical Columns : ",numerical_columns)
print("-----")
print("categorical Columns : ",categorical_columns)

```

```

Numerical Columns :  ['Rent', 'Beds', 'Baths', 'Area_in_sqft', 'Rent_per_sqft',
'Posted_date', 'Age_of_listing_in_days', 'Latitude', 'Longitude']

```

```

-----
categorical Columns :  ['Address', 'Type', 'Rent_category', 'Frequency',
'Furnishing', 'Purpose', 'Location', 'City']

```

```
[7]: df.isna().sum()
```

```

[7]: Address                0
Rent                      0
Beds                      0
Baths                    0
Type                     0
Area_in_sqft             0

```

```
Rent_per_sqft      0
Rent_category      0
Frequency          0
Furnishing         0
Purpose            0
Posted_date        0
Age_of_listing_in_days  0
Location           0
City               0
Latitude           719
Longitude          719
dtype: int64
```

```
[8]: df =df.dropna()
```

```
[9]: for i in categorical_columns:
      print(df[i].unique())
      print("-----")
```

```
['The Gate Tower 2, The Gate Tower, Shams Gate District, Shams Abu Dhabi, Al
Reem Island, Abu Dhabi'
```

```
"Water's Edge, Yas Island, Abu Dhabi"
```

```
'Al Raha Lofts, Al Raha Beach, Abu Dhabi' ...
```

```
'Umm Al Quwain Marina, Umm Al Quwain'
```

```
'Al Humrah B, Al Humrah, Umm Al Quwain'
```

```
'Al Huboob 1, Al Salamah, Umm Al Quwain']
```

```
-----
['Apartment' 'Penthouse' 'Villa' 'Townhouse' 'Villa Compound'
```

```
'Residential Building' 'Residential Floor' 'Hotel Apartment'
```

```
'Residential Plot']
```

```
-----
['Medium' 'High' 'Low']
```

```
-----
['Yearly']
```

```
-----
['Unfurnished' 'Furnished']
```

```
-----
['For Rent']
```

```
-----
['Al Reem Island' 'Yas Island' 'Al Raha Beach' 'Al Bateen' 'Al Reef'
```

```
'The Marina' 'Al Khalidiyah' 'Al Raha Gardens' 'Mohammed Bin Zayed City'
```

```
'Shakhbout City' 'Madinat Al Riyadh' 'Corniche Area' 'Al Muroor'
```

```
'Zayed Sports City' 'Khalifa City' 'Hamdan Street' 'Al Bahia'
```

```
'Masdar City' 'Al Najda Street' 'Tourist Club Area (TCA)' 'Al Matar'
```

```
'Al Shamkha' 'Rawdhat Abu Dhabi' 'Al Ghadeer' 'Danet Abu Dhabi' 'Baniyas'
```

```
'Al Mushrif' 'Airport Street' 'Al Jubail Island' 'Al Hosn'
```

'Saadiyat Island' 'Corniche Road' 'Sheikh Khalifa Bin Zayed Street'
 'Electra Street' 'Al Falah Street' 'Al Wahdah' 'Madinat Zayed' 'Rabdan'
 'Al Rahba' 'Al Maqtaa' 'Capital Centre' 'Al Karamah' 'Al Markaziya'
 'Defence Street' 'Al Nahyan' 'Between Two Bridges (Bain Al Jessrain)'
 'Al Nasr Street' 'Al Zahraa' 'Al Muntazah' 'Al Zaab' 'Al Samha'
 'Al Shawamekh' 'Hydra Village' 'Sas Al Nakhl Village' 'Al Zahiyah'
 'Al Rawdah' 'Sheikh Rashid Bin Saeed Street' 'Al Maryah Island'
 'Al Manhal' 'Zayed City' 'Al Aman' 'Al Shahama' 'Mussafah' 'Al Khubeirah'
 'Al Mina' 'Al Falah City' 'Al Salam Street' 'Al Qurm' 'Al Ras Al Akhdar'
 'Al Danah' 'Al Dhafrah' 'KIZAD' 'Al Wathba' 'Liwa Street' 'Muwaylih'
 'Al Nahda' 'Al Yasmeen' 'Al Rawda' 'Al Zahya' 'Al Nuaimiya' 'Al Mowaihat'
 'Al Sawan' 'Al Helio' 'Al Alia' 'Corniche Ajman' 'Ajman Downtown'
 'Al Rashidiya' 'Al Nakhil' 'Al Jurf' 'Ajman Industrial' 'Al Hamidiyah'
 'Garden City' 'Masfoot' 'Musherief' 'Al Zorah' 'Al Bustan'
 'Emirates City' 'Al Rumaila' 'Sheikh Maktoum Bin Rashid Street'
 'Ajman Free Zone' 'Al Tallah 2' 'Al Tallah 1' 'Al Ameera Village'
 'Asharij' 'Al Marakhaniya' 'Shiab Al Ashkhar' 'Al Sarouj' 'Al Jimi'
 'Al Tiwayya' 'Zakhir' 'Al Maqam' 'Al Jahili' 'Al Khibeesi' 'Al Hayer'
 'Al Iqabiyyah' 'Central District' 'Hili' 'Al Rawdah Al Sharqiyah'
 'Al Muwaiji' 'Al Yahar' 'Falaj Hazzaa' 'Al Mutarad' 'Al Sidrah' 'Neima'
 'Al Mutaw'ah' 'Al Dhahir' 'Al Qattara' 'Um Ghafah' 'Al Fou'ah'
 'Abu Samrah' 'Al Masoudi' 'Ghnaimah' 'Jumeirah Village Circle (JVC)'
 'Dubai Hills Estate' 'Arabian Ranches 2' 'Dubai Silicon Oasis (DSO)'
 'Dubai Sports City' 'Town Square' 'Meydan City' 'Dubai Creek Harbour'
 'Dubai Harbour' 'Jumeirah Beach Residence (JBR)' 'Palm Jumeirah' 'Mirdif'
 'DAMAC Hills 2 (Akoya by DAMAC)' 'Al Jaddaf' 'Dubailand'
 'Jumeirah Golf Estates' 'Dubai South' 'Dubai Marina' 'Al Furjan'
 'The Valley' 'Downtown Dubai' 'Arjan' 'Ras Al Khor' 'Reem' 'DAMAC Hills'
 'Umm Suqeim' 'Jumeirah Heights' 'Mudon' 'Business Bay'
 'Jumeirah Lake Towers (JLT)' 'Tilal Al Ghaf' 'Za'abeel'
 'Arabian Ranches 3' 'Motor City' 'The Views' 'The Meadows' 'Al Wasl'
 'Jumeirah' 'Sheikh Zayed Road' 'Arabian Ranches' 'Jumeirah Park'
 'Dubai Residence Complex' 'The Springs' 'Mohammed Bin Rashid City'
 'Serena' 'Dubai Production City (IMPZ)' 'Nad Al Sheba' 'The Greens'
 'Sobha Hartland' 'Bur Dubai' 'Dubai Studio City' 'Green Community'
 'Jumeirah Islands' 'Jumeirah Village Triangle (JVT)' 'The Villa'
 'Al Barari' 'Al Barsha' 'Al Nahda (Dubai)' 'DIFC' 'Liwana'
 'Living Legends' 'Discovery Gardens' 'Barsha Heights (Tecom)' 'Al Karama'
 'Remraam' 'Bluewaters Island' 'Dubai Media City' 'International City'
 'The Lakes' 'City of Arabia' 'World Trade Centre' 'Bukadra'
 'Dubai Festival City' 'Culture Village' 'Deira' 'Falcon City of Wonders'
 'Al Safa' 'Majan' 'Liwana 2' 'Al Quoz' 'The Sustainable City'
 'Umm Al Sheif' 'Jebel Ali' 'Nad Al Hamar' 'Al Warqaa' 'Al Qusais'
 'Al Awir' 'Al Satwa' 'The Hills' 'Muhaisnah' 'Al Khawaneej' 'The Gardens'
 'Expo City' 'Dubai Investment Park (DIP)' 'Dubai Waterfront'
 'Dubai Industrial City' 'Dubai Internet City' 'Al Sufouh'
 'Dubai Maritime City' 'Emirates Hills' 'Al Warsan' 'Al Badaa' 'Al Mizhar'
 'Al Jafiliya' 'Al Garhoud' 'Wasl Gate' 'Al Mamzar' 'Wadi Al Shabak'

'Al Hudaiba' 'Pearl Jumeirah' 'Al Manara' 'Wadi Al Safa 2' 'Al Twar'
 'Hadaeq Sheikh Mohammed Bin Rashid' 'Knowledge Village' 'Umm Ramool'
 'Al Hebiah 2' 'Al Lisaili' 'Oud Al Muteena' 'Fujairah Tower'
 'Fujairah Freezone' 'Dibba' 'Tawyeen' 'Mirbah'
 'Address Fujairah Beach Resort' 'Sakamkam' 'Al Marjan Island'
 'Al Hamra Village' 'Al Nakheel' 'Mina Al Arab' 'Al Qusaidat'
 'The Cove Rotana Resort' 'Dafan Al Nakheel' 'Rak City' 'Al Seer'
 'Yasmin Village' 'Al Mairid' 'Al Dhait' 'Khuzam' 'Dafan Al Khor'
 'Seih Al Uraibi' 'Sidroh' 'Al Uraibi' 'Wadi Ammar' 'Dahan' 'Al Ghubb'
 'Al Nudood' 'Al Kharran' 'Julfar' 'Al Sharisha' 'Al Rams' 'Aljada'
 'Al Khan' 'Muwaileh' 'Al Nahda (Sharjah)' 'Al Tai' 'Al Taawun'
 'Muwailih Commercial' 'Al Majaz' 'Al Wahda Street' 'Industrial Area'
 'Al Qasimia' 'Tilal City' 'Al Rahmaniya' 'Sharqan' 'Al Qasba' 'Al Fisht'
 'Abu Shagara' 'Al Mujarrah' 'Barashi' 'Al Mareija' 'Al Dhaid'
 'Al Ramaqiya' 'Al Nabba' 'Al Sharq' 'Al Nasserya' 'Rolla Area'
 'Al Ghuwair' 'Al Ramtha' 'Um Tarafa' 'Bu Tina' 'Al Mahatah' 'Al Musalla'
 'Al Soor' 'Al Ramla' 'Al Falaj' 'Al Jazzat' 'Maysaloon' 'Al Nekhailat'
 'Al Mansoura' 'Hoshi' 'Al Abar' 'Al Fayha' 'Samnan' 'Al Yarmook'
 'Al Jubail' 'Al Noaf' 'Al Sajaa' 'Al Shahba' 'Al Manakh' 'Khor Fakkan'
 'Al Ghafia' 'Al Sabkha' 'Kalba' 'Dasman' 'Al Darari' 'Al Juraina'
 'Sharjah University City' 'Al Gharb' 'Al Sajaa Industrial' 'Al Tay East'
 'Al Mirgab' 'Al Riqaiabah' 'Al Rifa' 'Al Ghubaiba' 'Al Riqqa Suburb'
 'Al Yash' 'Al Bataeh' 'Al Ramlah' 'Al Butain' 'Al Salamah' 'Al Abra q 1'
 'Al Qarayen' 'Old Town Area' 'Umm Al Quwain Marina' 'Al Hawiyah'
 'Al Humrah']

['Abu Dhabi' 'Ajman' 'Al Ain' 'Dubai' 'Fujairah' 'Ras Al Khaimah'
 'Sharjah' 'Umm Al Quwain']

[10]: df.describe()

```
[10]:
```

	Rent	Beds	Baths	Area_in_sqft	Rent_per_sqft \
count	7.302300e+04	73023.000000	73023.000000	73023.000000	73023.000000
mean	1.483723e+05	2.154458	2.638771	2035.634471	88.537296
min	0.000000e+00	0.000000	1.000000	74.000000	0.000000
25%	5.499900e+04	1.000000	2.000000	850.000000	40.000000
50%	9.800000e+04	2.000000	2.000000	1329.000000	71.813285
75%	1.700000e+05	3.000000	3.000000	2101.000000	119.047619
max	5.500000e+07	12.000000	11.000000	210254.000000	2182.044888
std	3.082652e+05	1.571260	1.620881	2976.159891	66.627532

	Posted_date	Age_of_listing_in_days	Latitude \
count	73023	73023.000000	73023.000000
mean	2024-02-07 02:45:58.900620544	73.884735	24.918929
min	2018-01-27 00:00:00	11.000000	15.175847
25%	2024-01-17 00:00:00	30.000000	24.493598

50%	2024-03-01 00:00:00	51.000000	25.078641
75%	2024-03-22 00:00:00	95.000000	25.197978
max	2024-04-10 00:00:00	2276.000000	25.920310
std	NaN	71.837749	0.569356

	Longitude
count	73023.000000
mean	55.053133
min	43.351928
25%	54.607372
50%	55.238209
75%	55.367138
max	56.361294
std	0.653722

```
[11]: vacant_buildings = df["Type"][df["Rent"]<=1].count()
print("Count of Building =",df["Rent"].count())
print("Count of Occupied Building =",len(df)-vacant_buildings,"percentage %\n
↳",round((len(df)-vacant_buildings)*100/len(df),3))
print("Count of Vacant Building =",vacant_buildings,"percentage %\n
↳",round(vacant_buildings*100/len(df),3))
```

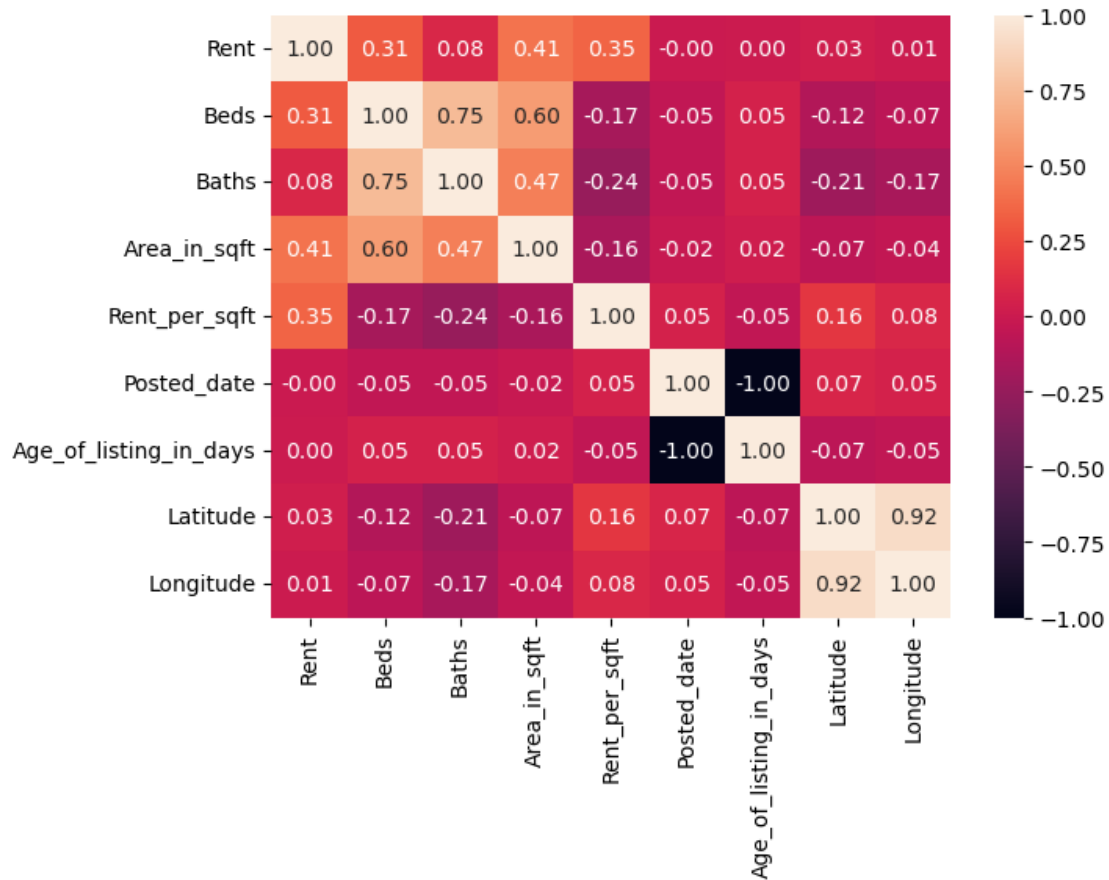
Count of Building = 73023

Count of Occupied Building = 73006 percentage % = 99.977

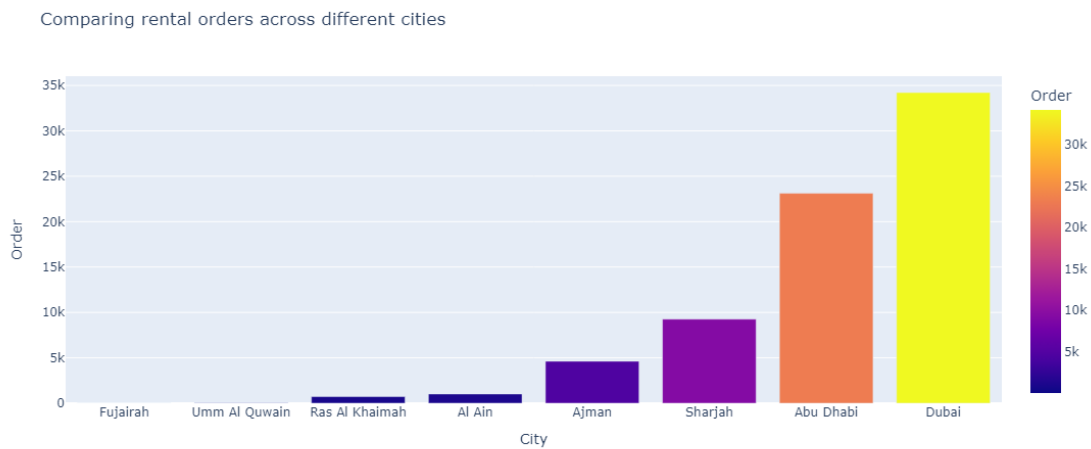
Count of Vacant Building = 17 percentage % = 0.023

```
[12]: df_dummies =df.drop(categorical_columns,axis=1)
plt.figure(figsize=(7,5))
sns.heatmap(df_dummies.corr(),annot=True,fmt="0.2f")
```

```
[12]: <Axes: >
```

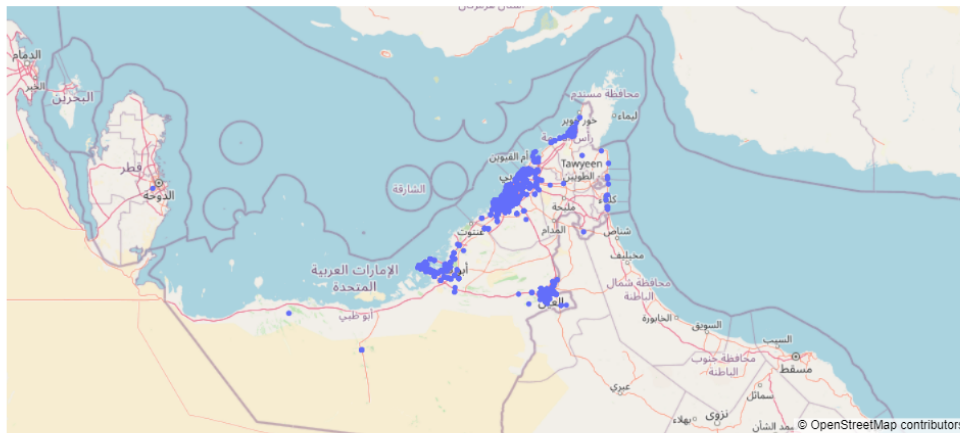


```
[13]: category = df.groupby(df["City"])[ "Type" ].count().sort_values()
px.bar(category,y="Type",color="Type",title = "Comparing rental orders across
different cities",
labels={'Type': 'Order'},width=1000,height=500)
```



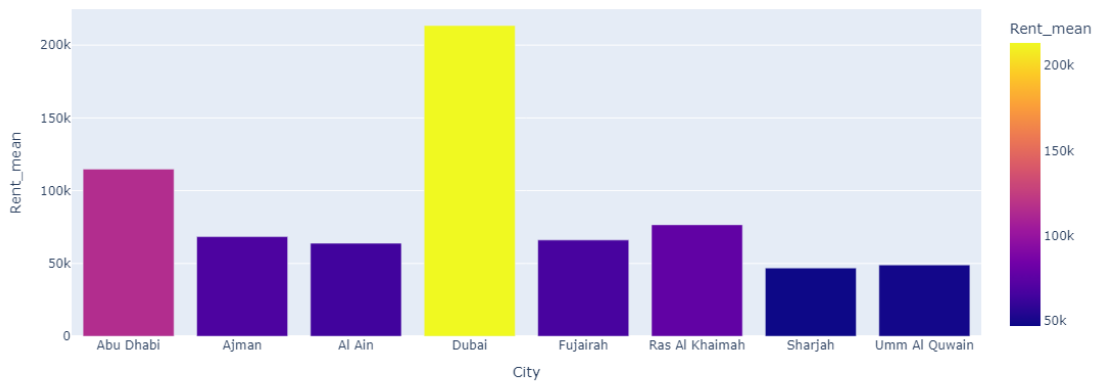
```
[14]: px.scatter_mapbox(df,lat="Latitude", lon="Longitude",title='Rental Properties Locations in UAE',
    ↪Locations in UAE',
    mapbox_style="open-street-map",zoom=6, height=600)
```

Rental Properties Locations in UAE

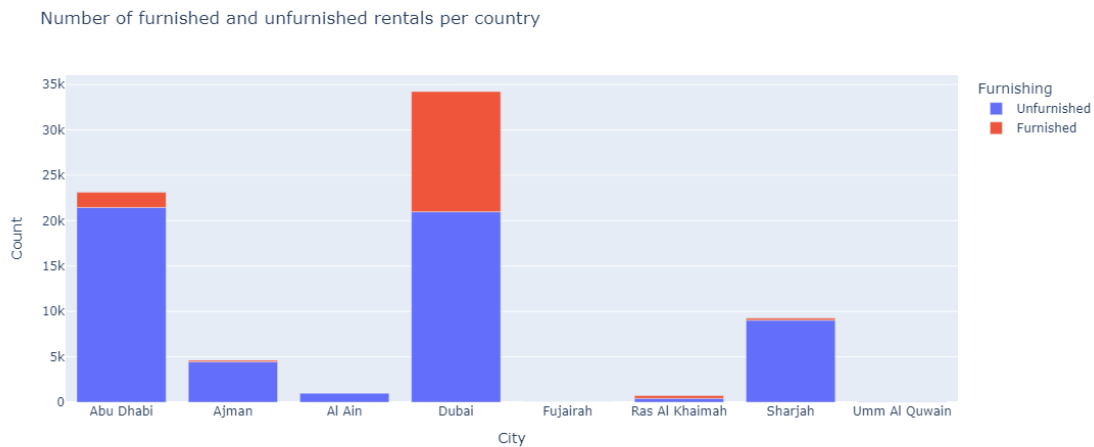


```
[15]: category = df.groupby("City")["Rent"].mean().reset_index()
category.columns = ['City', 'Rent_mean']
px.bar(category,x="City",y="Rent_mean",color="Rent_mean",title = "Comparing rental prices across different cities",width=1000,height=500)
```

Comparing rental prices across different cities



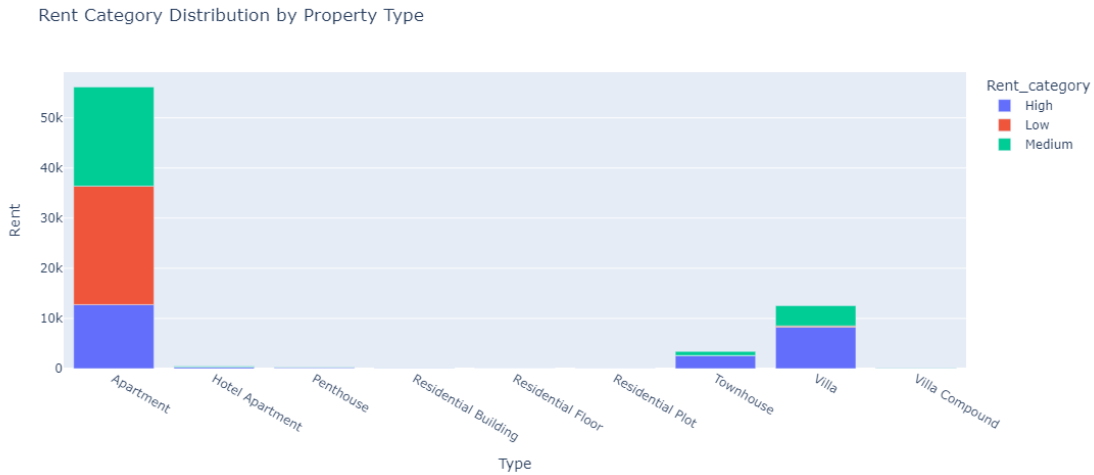

```
[16]: category = df[["City","Furnishing"]].groupby("City").value_counts().
      ↪reset_index()
category.columns=["City","Furnishing","Count"]
px.bar(category,x="City",y="Count",color="Furnishing",title = "Number of
      ↪furnished and unfurnished rentals per country",width=1000,height=500)
```



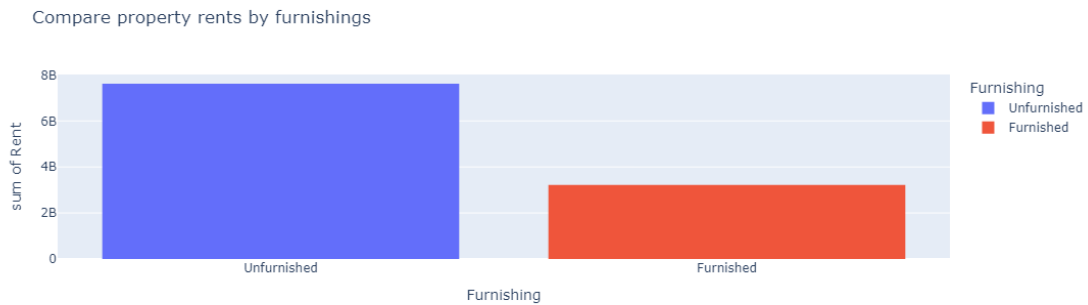
```
[17]: category = df.groupby(df["Type"])[["Rent"]].count().sort_values()
px.bar(category,y="Rent",color="Rent",title = "Comparing rental orders across
      ↪different property types",
      labels={'Rent': 'Order'},width=1000,height=500)
```



```
[18]: category = df.groupby(["Type", "Rent_category"])["Rent"].count().reset_index()
px.bar(category, x="Type", y="Rent", color="Rent_category", title='Rent Category_
↳ Distribution by Property Type', width=1000, height=500)
```

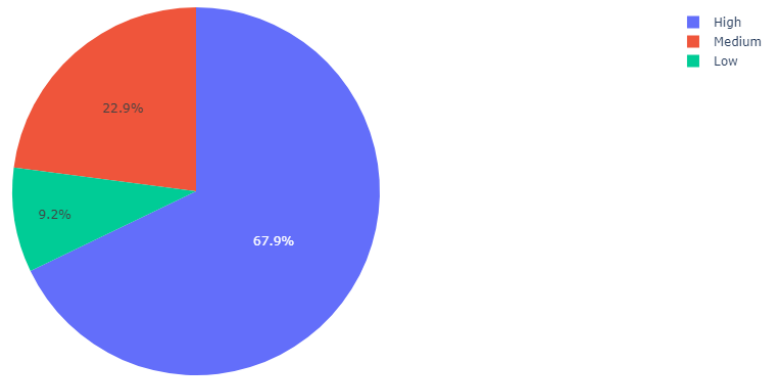


```
[19]: plt.figure(figsize=(20,20))
px.
↳ histogram(data_frame=df, x="Furnishing", y="Rent", color="Furnishing", title="Compare_
↳ property rents by furnishings")
```

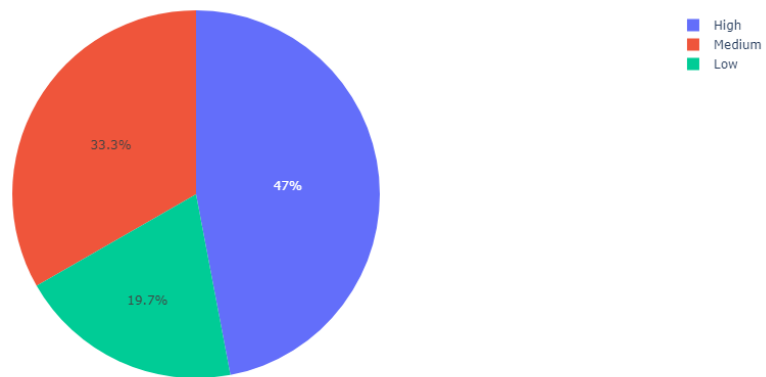


<Figure size 2000x2000 with 0 Axes>

```
[20]: px.pie(df, values="Rent", names="Rent_category", width=1000, height=500)
```



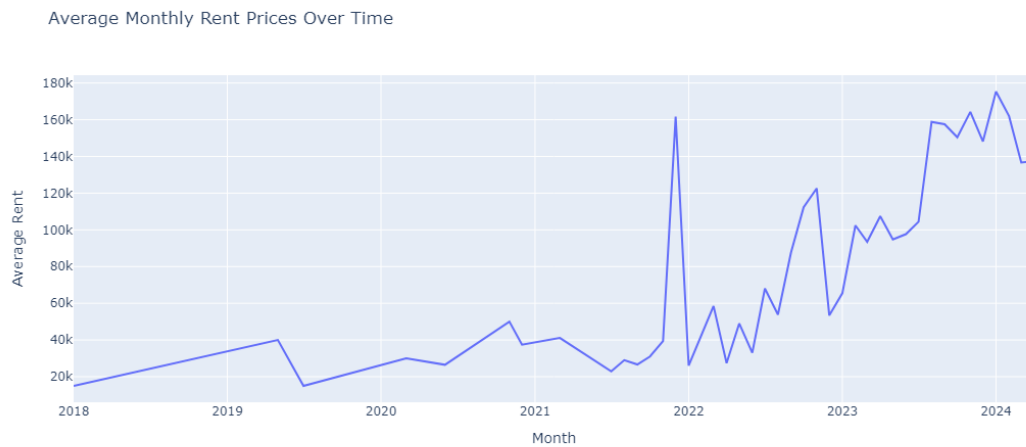
```
[21]: px.pie(df, values="Rent_per_sqft", names="Rent_category", width=1000, height=500)
```



```
[22]: px.violin(data_frame=df, x="Rent_category", y="Area_in_sqft", color="Rent_category", title="Relationship between rent category and area")
```

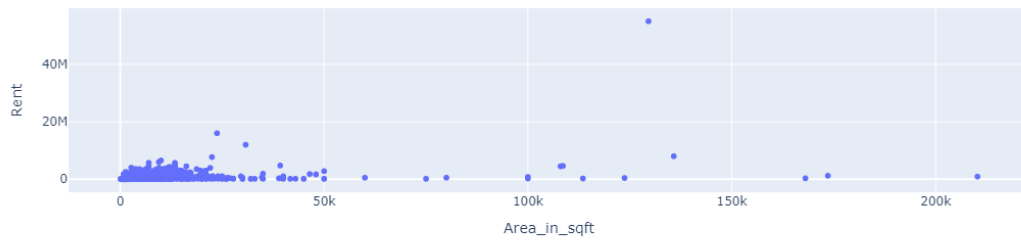


```
[23]: avg_df = df.groupby(df["Posted_date"].dt.to_period("M"))["Rent"].mean().
      ↪reset_index()
avg_df["Posted_date"] = avg_df["Posted_date"].astype(str)
px.line(avg_df, x='Posted_date', y='Rent', title='Average Monthly Rent Prices_
      ↪Over Time',
        labels={'Posted_date': 'Month', 'Rent': 'Average_
      ↪Rent'}, width=1100, height=500)
```



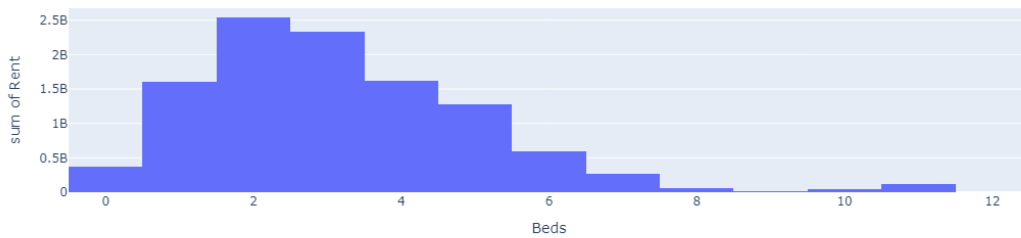
```
[24]: px.scatter(data_frame=df, x="Area_in_sqft", y="Rent", title="Relationship between_
      ↪rent and area")
```

Relationship between rent and area



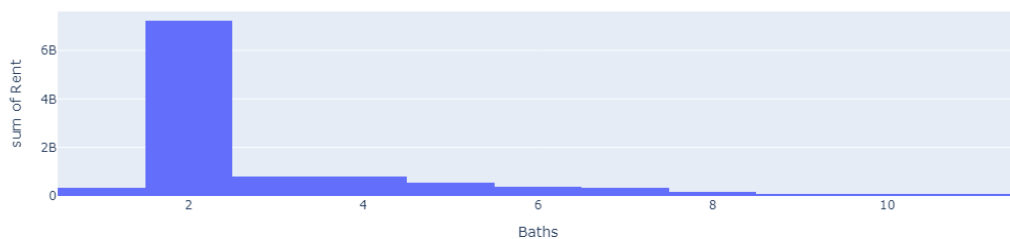
```
[25]: px.histogram(data_frame=df,x="Beds",y="Rent",title="Relationship between rent_
↳and beds")
```

Relationship between rent and beds

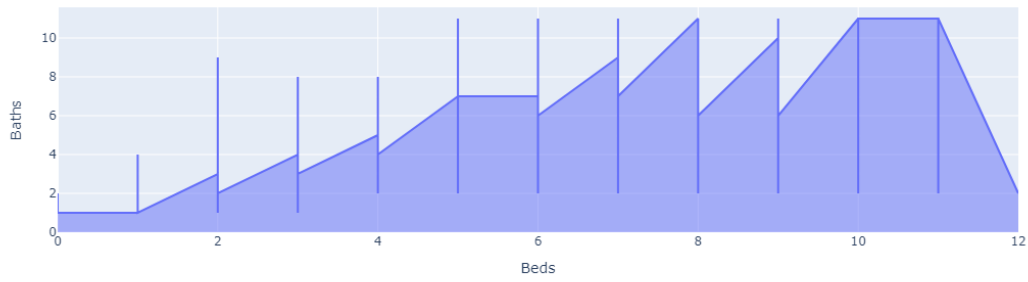


```
[26]: px.histogram(data_frame=df,x="Baths",y="Rent",title="Relationship between rent_
↳and baths")
```

Relationship between rent and baths

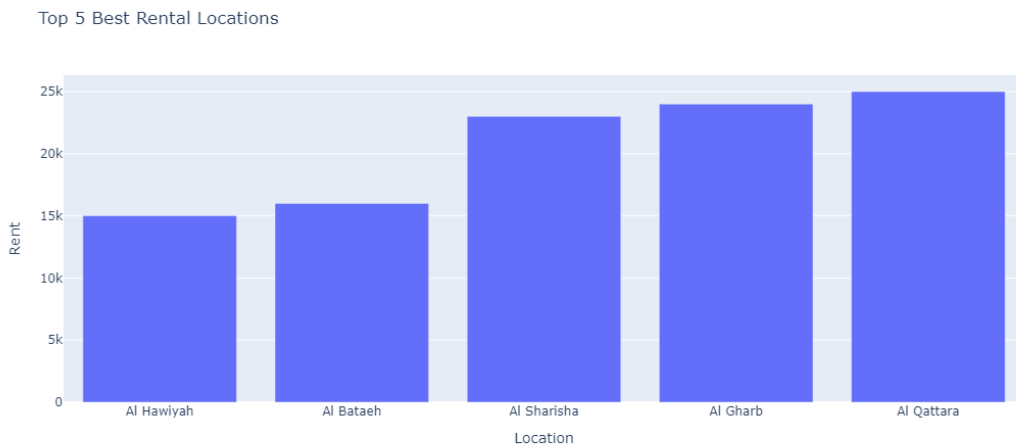


```
[28]: px.area(df,x="Beds",y="Baths")
```



- Top 5 Best and Worst Rental Locations

```
[40]: top_5_best = df.groupby("Location")["Rent"].sum().sort_values().reset_index().
      ↪head(5)
px.bar(top_5_best,x="Location",y="Rent",title ="Top 5 Best Rental_
      ↪Locations",width=1000,height=500)
```



```
[42]: top_5_worst = df.groupby("Location")["Rent"].sum().sort_values().reset_index().
      ↪tail(5)
px.bar(top_5_worst,x="Location",y="Rent",title ="Top 5 Worst Rental_
      ↪Locations",width=1000,height=500)
```

Top 5 Worst Rental Locations

