

REPORT

U.S. Apartment Rental Market Analysis

KAYLIE LE

May 31, 2025

Table of Content

<i>Introduction</i>	<i>3</i>
<i>Data Analysis.....</i>	<i>5</i>
<i>Summary.....</i>	<i>12</i>
<i>References.....</i>	<i>13</i>

Introduction

This report explores a dataset of 10,000 U.S. apartment listings sourced from RentLingo via UCI Machine Learning Repository. The purpose of the dataset is to provide comprehensive rental information—such as price, size, location, and amenities—that can support housing market analysis, pricing strategies, and urban planning.

The dataset includes both numerical (e.g., price, square footage, bedrooms, bathrooms) and categorical/textual fields (e.g., city, state, amenities, pets policy). Its broad coverage makes it ideal for identifying rental trends, affordability differences, and correlations between price and apartment features.

The data fields including the title, the data type, the data description, etc.:

Column Name	Data Type	Description
id	int64	Unique listing ID
category	object	Rental category type (e.g., "housing/rent/apartment")
title	object	Listing title (e.g., "Studio Apartment in Boston")
price	int64	Monthly rental price in USD
square_foot	int64	Apartment size in square feet
bedrooms	float64	Number of bedrooms
bathrooms	float64	Number of bathrooms
latitude	float64	Latitude of the listing location
longitude	float64	Longitude of the listing location
amenities	object	Comma-separated list of apartment amenities
cityname	object	City name
state	object	U.S. state abbreviation
source	object	Data provider ("RentLingo")
time	int64	Unix timestamp of the listing

The dataset consists of 10,000 rows and 22 columns, representing individual apartment listings and their associated attributes.

To prepare the dataset for meaningful analysis, several data cleaning steps were performed:

1. Character Encoding Correction

The dataset was imported using an appropriate character encoding (ISO-8859-1) to address potential issues with special or non-English characters. This ensures that all symbols and location names are properly displayed.

2. Removal of Irrelevant Columns

Non-analytically relevant fields, such as indicators for the presence of photos, were removed to reduce noise and focus on meaningful attributes.

3. Cleaning the Rental Price Field

The price_display column often included currency symbols and formatting characters (e.g., \$, ,). These were cleaned and converted into numeric values, making the data usable for quantitative analysis.

4. Filtering Invalid Price Entries

Listings where the rental price could not be extracted or converted (e.g., listings with “Contact for price”) were excluded to avoid analytical distortion.

5. Standardizing Numeric Variables

Key numeric columns such as number of bedrooms, bathrooms, and square footage were converted to proper numeric formats. This ensures statistical functions (like mean, median) can be applied accurately.

6. Handling Missing Values

Rows missing critical data — particularly price, room count, square footage, or location — were removed entirely. This prevents misleading results in summary statistics and visualizations.

7. Standardizing Location Fields

The state and cityname fields were cleaned to ensure consistency. State abbreviations were converted to uppercase, while city names were formatted with proper capitalization. This avoids duplication caused by inconsistent casing or spacing.

8. Excluding Rows with Invalid Categorical Values

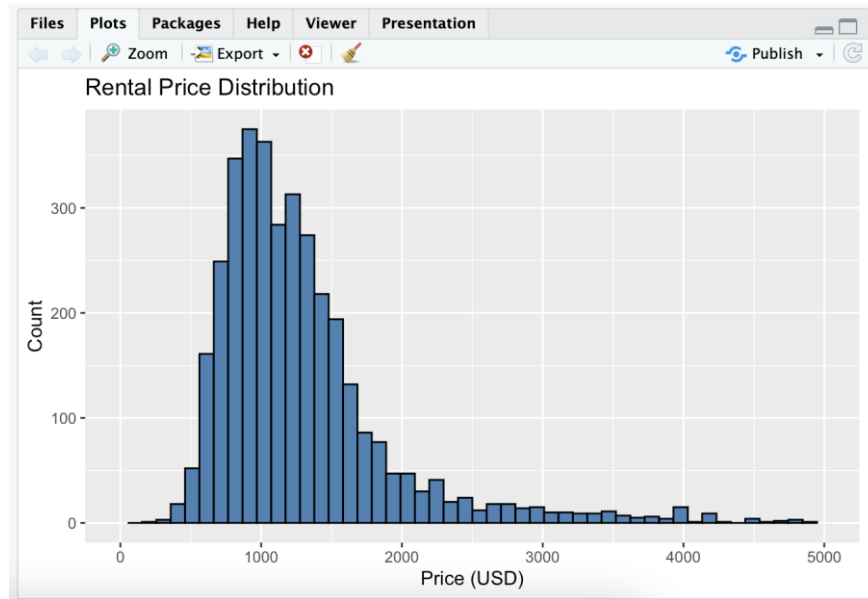
Several categorical columns contained the literal string "null" as a placeholder for missing information. These rows were removed, as "null" is not a valid category and could skew categorical analysis.

These steps helped preserve data integrity while ensuring that the analysis focused on complete and relevant entries.

Data Analysis

1. Continuous Data

Rental Price Distribution



What the chart shows:

This histogram displays the distribution of monthly rental prices across all apartment listings in the dataset.

Interpretation:

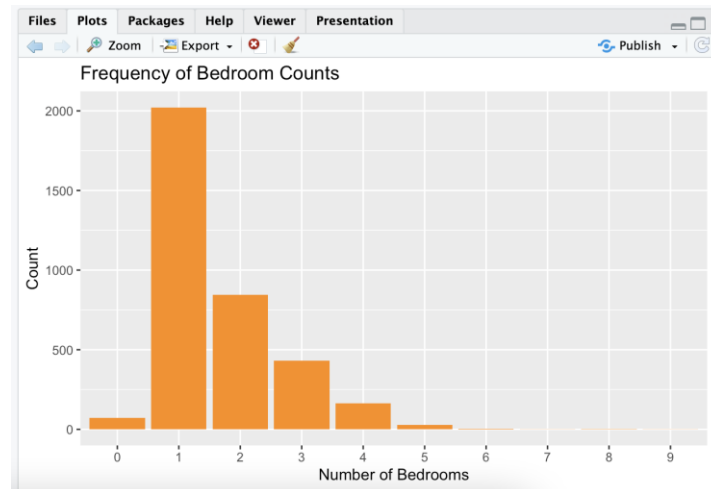
- The majority of rental prices fall between \$800 and \$1,500, indicating a typical price range for most listings.
- The distribution is right-skewed (positively skewed), with a long tail stretching to the right.
- A small number of listings exceed \$3,000/month, possibly representing luxury or high-demand urban apartments.
- There is a visible peak around \$1,000, suggesting this is the most common rental price point.

Insight:

- The skewed shape highlights income inequality or market segmentation in the rental market.
- Analysts should consider this skewness when applying statistical models (e.g., mean vs. median).
- Data cleaning helped remove invalid or extreme entries, which ensures that the distribution reflects realistic prices.

2. Discrete Data

Bar chart: Bedroom Frequency



What the chart shows:

This bar chart illustrates how many listings in the dataset fall under each bedroom count category (from studios to multi-bedroom apartments).

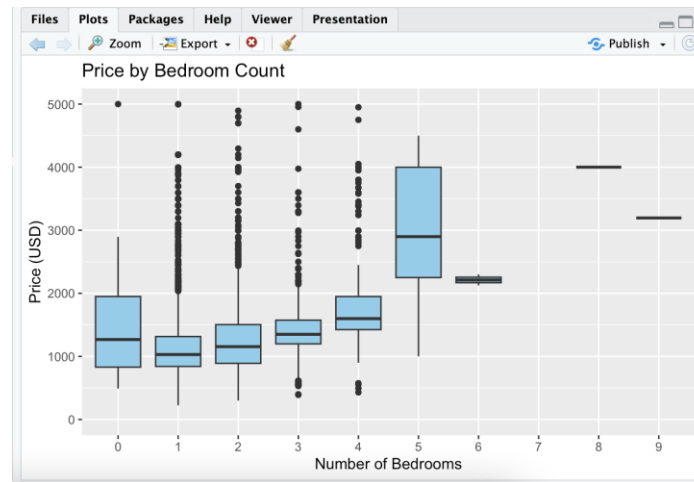
Interpretation:

- 1-bedroom apartments are the most common type of listing, with over 2,000 entries, indicating strong market availability or demand.
- 2-bedroom apartments are the second most frequent, followed by 3-bedrooms.
- Studios (0-bedroom listings) do exist but are relatively rare.
- Apartments with 5 or more bedrooms are extremely uncommon, suggesting they serve a niche or luxury market.

Insight:

- This distribution can inform rental companies or developers where demand is concentrated.
- The dataset skews heavily toward small-to-medium sized living spaces, which aligns with urban housing trends.
- Knowing the most prevalent apartment types is helpful for segmentation, pricing strategy, and marketing efforts.

Boxplot: Price by Bedrooms



What the chart shows:

This boxplot illustrates the distribution of rental prices across different bedroom counts, providing insights into price variability and outliers.

Interpretation:

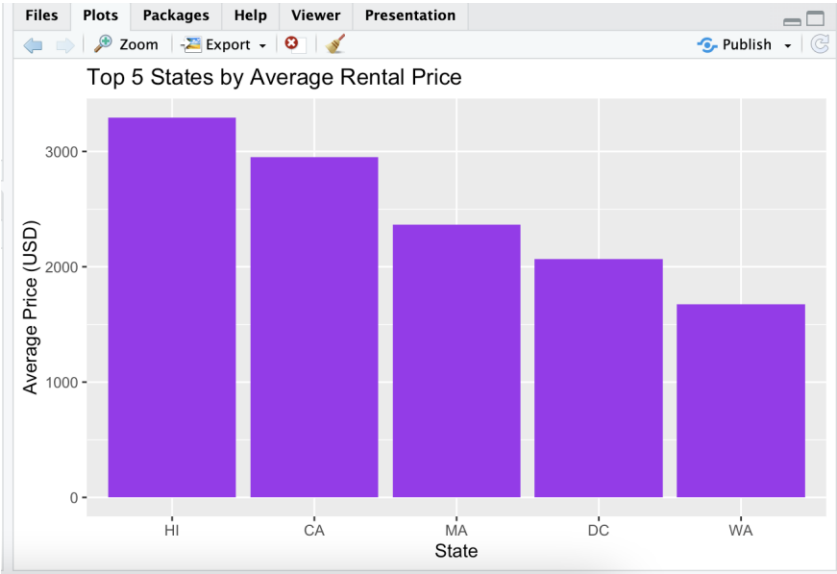
- In general, the median rental price increases with the number of bedrooms — indicating that more bedrooms typically command higher rent.
- Studios (0 bedrooms) and 1-bedroom apartments have lower median prices and narrower interquartile ranges (IQR), suggesting relatively consistent pricing.
- 3- and 4-bedroom units show greater price variability, while 5-bedroom apartments and above exhibit significantly higher medians and broader spreads — possibly reflecting luxury units or houses.
- The chart includes many outliers, especially in lower-bedroom categories, indicating some high-priced listings even among small apartments.

Insight:

- There is a positive correlation between bedroom count and rental price, but the relationship is not linear — prices spike notably from 4 to 5 bedrooms.
- The presence of outliers may reflect high-demand urban locations, premium amenities, or mispriced listings.
- This visualization supports pricing analysis and can be useful for property owners or renters comparing value per room.

3. Categorical Data

Top 5 States by Average Price



What the chart shows:

This bar chart presents the five U.S. states with the highest average monthly rental prices in the dataset.

Interpretation:

- Hawaii (HI) has the highest average rent, exceeding \$3,200/month. This likely reflects the state's high cost of living and limited housing supply.
- California (CA) ranks second, consistent with its expensive metropolitan areas like San Francisco and Los Angeles.
- Massachusetts (MA) and Washington, D.C. (DC) follow, both known for dense urban centers and competitive housing markets.
- Washington (WA) rounds out the top five, indicating relatively high prices possibly driven by tech industry hubs like Seattle.

Insight:

- The chart reveals a geographic concentration of high rent in coastal and urbanized states, where housing demand tends to exceed supply.
- This kind of geographic breakdown is useful for market comparison, regional policy design, or relocation planning.
- It can also inform investment decisions for developers and landlords looking to enter high-yield markets.

4. Descriptive statistical tables

Table 1 — Apartment Features (Continuous & Discrete Variables)

This table summarizes core physical attributes of the apartments: square footage, number of bedrooms, and number of bathrooms.

- Purpose: Understand the typical physical characteristics of rental units.
- Metric: Descriptive statistics for square footage, bedrooms, and bathrooms.

- Use case: Identify standard apartment layouts and detect outliers.

	Mean	SD	Count	Min	Median	Max	Range
square_feet	845.338940	450.9594648	3567	138	719	5700	5562
bedrooms	1.638912	0.9713967	3567	0	1	9	9
bathrooms	1.228063	0.4973911	3567	1	1	5	4

The table provides a descriptive summary of three key apartment characteristics: square footage, number of bedrooms, and number of bathrooms. Here's what these statistics tell us:

Square Footage

- Mean = 845.34 sq. ft., Median = 719 sq. ft.
This shows the average apartment size is around 845 sq. ft., but the median is lower, suggesting the distribution is right-skewed — some very large apartments increase the mean.
- Range = 5,562 sq. ft., Max = 5,700 sq. ft.
There is a huge spread in apartment sizes, indicating a wide diversity in property types — from small studios to luxury penthouses.

Bedrooms

- Mean = 1.64, Median = 1, Range = 9
Most apartments have 1–2 bedrooms, but some listings go as high as 9 bedrooms, which is likely rare and pulls the average up.
- Minimum = 0 bedrooms
Indicates the presence of studio apartments, which can be a major market segment in urban areas.

Bathrooms

- Mean = 1.23, Median = 1
Most apartments offer just one bathroom, consistent with the trend toward compact and affordable housing.
- Max = 5 bathrooms
Some listings have multiple bathrooms — possibly targeting families or luxury rentals.

Overall Patterns:

- All variables show positive skew, with mean > median.
- High ranges across all features indicate that the dataset contains a diverse mix of properties — from budget-friendly to luxury units.
- The consistency of medians (1 bedroom, 1 bathroom) reflects the standard structure of rental apartments.

Table 2: Rental Price by State or Region

- Purpose: Identify geographic trends in rental pricing.
- Metric: price_display_clean grouped by state or region.

- Use case: Understand where rent is most expensive or affordable.

	state	Mean	Median	Count
1	HI	3295.0000	3295.0	1
2	CA	2951.7477	2612.0	222
3	MA	2363.7213	2150.0	61
4	DC	2067.0000	1818.0	29
5	WA	1674.0847	1484.5	118
6	NJ	1634.8000	1487.5	10
7	CT	1564.9412	1500.0	17
8	RI	1501.0000	1315.0	9
9	NULL	1462.9583	1325.0	24
10	VA	1460.8519	1450.0	54
11	NH	1447.8000	1475.0	5
12	IL	1426.1914	1249.5	162
13	MT	1380.0000	1380.0	2
14	NY	1350.9130	1040.0	23

	state	Mean	Median	Count
15	MI	1343.2955	1001.5	88
16	CO	1336.0492	1250.0	61
17	FL	1325.9592	1180.0	147
18	MD	1323.6667	1295.0	39
19	SC	1321.3429	1249.0	35
20	OR	1313.3846	1284.5	52
21	GA	1297.8556	1315.5	90
22	WI	1292.3397	1185.0	262
23	NC	1262.0230	1299.0	87
24	TN	1239.8605	1211.0	43
25	MN	1219.2653	1203.0	98
26	UT	1200.4348	1185.0	23
27	NV	1200.3478	986.0	23
28	OH	1157.1026	1034.5	78

	state	Mean	Median	Count
28	OH	1157.1026	1034.5	78
29	NM	1155.0000	1155.0	1
30	AZ	1144.0641	1162.5	78
31	PA	1120.0526	950.0	38
32	TX	1119.5060	1040.0	911
33	KY	1061.6875	795.5	16
34	MS	983.3333	830.0	3
35	AL	965.0000	922.5	14
36	AK	955.8788	895.0	33
37	ID	955.7500	973.5	4
38	IN	945.0000	850.0	189
39	SD	944.0000	835.0	31
40	KS	931.1750	782.5	40
41	AR	909.1818	794.5	22
41	AR	909.1818	794.5	22
42	NE	892.8571	840.0	35
43	MO	890.5179	810.0	168
44	IA	866.9861	815.0	72
45	OK	866.0000	792.0	22
46	DE	865.0000	865.0	2
47	LA	841.5333	825.0	15
48	ND	770.0000	805.0	9
49	WY	713.0000	713.0	1

The analysis shows significant variation in average rental prices across U.S. states. Hawaii and California have the highest average rents, exceeding \$2,900/month, reflecting their high living costs and limited housing supply. In contrast, states like Wyoming, North Dakota, and Louisiana show the lowest averages, around \$700–\$900/month.

Texas, despite having the largest number of listings, maintains relatively low average prices (~\$1,100), suggesting greater housing availability. In most states, the mean price is higher than the median, indicating the presence of luxury listings that pull the average up.

Overall, geographic location plays a critical role in determining rental affordability, with urban and coastal states being the most expensive

Summary

This exploratory analysis of 10,000 U.S. apartment listings uncovers several important patterns in the rental housing market. The majority of listings are 1-bedroom, 1-bathroom units priced between \$800 and \$1,500/month, reflecting a strong demand for compact, affordable urban housing. The process of cleaning, visualizing, and interpreting data reinforced the importance of rigorous preprocessing and thoughtful statistical analysis when working with real-world datasets.

Key Takeaways:

- **Rental Pricing:** Most listings fall between \$800–\$1,500/month. The distribution is positively skewed, with luxury listings driving up the average.
- **Apartment Characteristics:** One-bedroom, one-bathroom apartments dominate the dataset. Median square footage is around 719 sq. ft., reflecting a preference for compact urban housing.
- **Price vs. Bedrooms:** There is a positive correlation between the number of bedrooms and price, but the price increases disproportionately for 5+ bedrooms, likely reflecting high-end properties.
- **Geographic Trends:** Hawaii and California lead in average rent, while inland states like North Dakota and Wyoming are more affordable. Urban and coastal areas tend to be more expensive.

Future Exploration:

To build upon the initial findings, several directions for deeper analysis can be pursued:

- Evaluate price per square foot to assess value-for-money across different states and cities. This metric can offer a standardized way to compare listings of varying sizes and uncover regional efficiency in pricing.
- Analyze how specific amenities influence rental prices and whether some features (e.g., pools, gyms, parking) are more strongly associated with price increases than others. This can help uncover what renters truly pay extra for.
- Explore temporal trends in rental pricing if timestamps can be transformed into human-readable dates. Time-based analysis would reveal how the market evolves seasonally or year over year..

Questions to Answer Next:

These exploratory directions naturally lead to several questions that could guide future analysis:

- How much do amenities like a gym, parking, or in-unit laundry affect rental price? Understanding this could help renters prioritize and landlords price more strategically.
- Which cities offer the best value in terms of space and pricing? Identifying high space-to-price ratios could benefit budget-conscious renters or investors seeking undervalued markets.
- Are there differences in pricing trends between corporate listings and private landlords? This could uncover patterns in pricing strategies, negotiation flexibility, or listing transparency.

References

RentLingo. (2018). *Apartments for Rent Classifieds Dataset* [Data set]. UCI Machine Learning Repository.
<https://archive.ics.uci.edu/dataset/555/apartment+for+rent+classified>