US Median Rent Data
(2010 - 2017)

By: Angel Lee, Gabriela Delgado, Jovid Rajab, Noah Shore, Jake Moen

Introduction/Analysis Goals

- **Dataset used:** Zillow Rent Index, 2010-Present from Kaggle
- Uncover patterns in the US rental market
- Analyze 4 major states in the 4 regions
- Compare median rent data from the data of years 2010-2017

Research Questions & Hypotheses:

Questions:

- How did the median rent price in the states (NY, CA, IL, TX) change across time (2011-2016) compared to the US as a whole?
- Which of the US states (NY, CA, IL, TX) had the greatest change in median rent price?

Hypotheses:

- **Null hypothesis:** There is no significant statistical difference in rent across US states/regions.
- **Hypothesis:** New York had the most significant difference in rent compared to the rest of the US.

Limitations

- Not enough data in years 2010 and 2017 to include them
- Limited to 2011-2016 data
- Had to remove NYC from data for NY because it had too much missing data
    - In hindsight we should have excluded it in general since 43% of NY's population lives in NYC
- Prices not adjusted for inflation now, only up to 2017

Major Findings

State	2011	2012	2013	2014	2015	2016	Rent Change (%)
TX	1182.50	1166.50	1200.50	1224.5	1310.50	1336.00	12.98
CA	1717.75	1701.50	1759.75	1792.0	1951.50	2098.00	22.14
IL	1399.50	1347.25	1363.75	1390.0	1412.25	1414.75	1.09
NY	1398.00	1413.50	1408.50	1370.5	1433.00	1388.00	-0.72
US	1215.50	1206.00	1222.50	1233.0	1290.50	1306.50	7.49

- The US's median rent prices had an increase of about 7.5% between 2011-2016
- 3 out of the 4 states we selected had an increase in rent between 2011 - 2016.
    - CA had the largest increase and standard deviation in rent between 2011-2016
    - TX also exhibited substantial increase compared to the US but not as high as CA
    - IL experienced tempered rate of escalation
- NY's median rent prices witnessed a decline of .72%

Dataset

	City Code	City	Metro	County	State	Population Rank	November 2010	December 2010	January 2011	February 2011	...	April 2016	May 2016	June 2016	Ju 20
0	6181	New York	New York	Queens	NY	1	NaN	NaN	NaN	NaN	...	2334	2339	2345	23
1	12447	Los Angeles	Los Angeles	Los Angeles	CA	2	2184.0	2184.0	2183.0	2188.0	...	2637	2662	2687	27
2	17426	Chicago	Chicago	Cook	IL	3	1563.0	1555.0	1547.0	1537.0	...	1684	1686	1687	16
3	39051	Houston	Houston	Harris	TX	4	1198.0	1199.0	1199.0	1200.0	...	1444	1446	1446	14
4	13271	Philadelphia	Philadelphia	Philadelphia	PA	5	1092.0	1099.0	1094.0	1087.0	...	1206	1211	1218	12

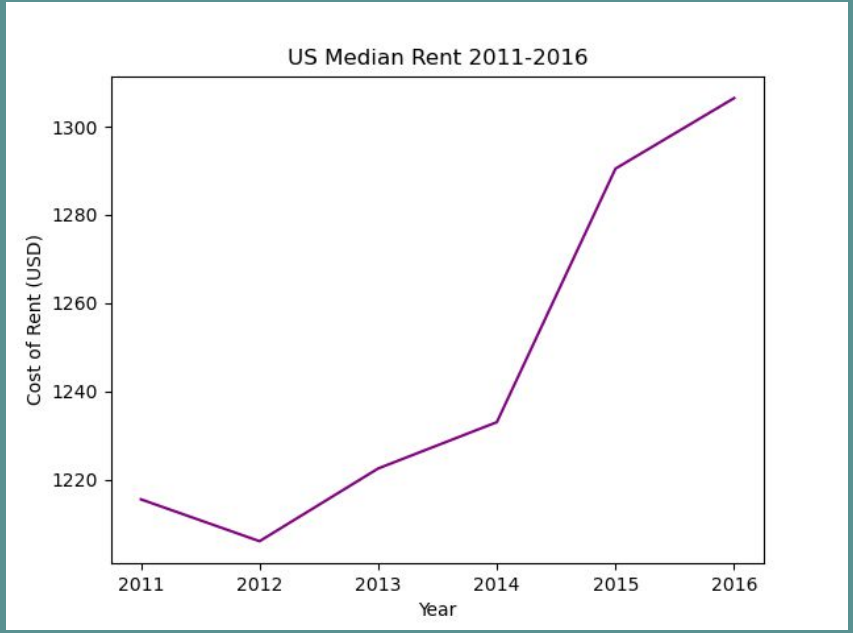
Initial dataset: Median rental prices spanning various U.S. cities, encompassing the period from November 2010 to January 2017.

Data/Method

- Clean data
  - Remove unwanted columns & missing values
- Group and organize into new datasets
  - By state (TX, CA, IL, NY) and year
- Calculation
  - US median rent price dataframe (index)
  - Percent change
- Graphs
  - Line & bar, box plot, histogram, and scatter plots
- Statistical analysis
  - Draw conclusions about the data

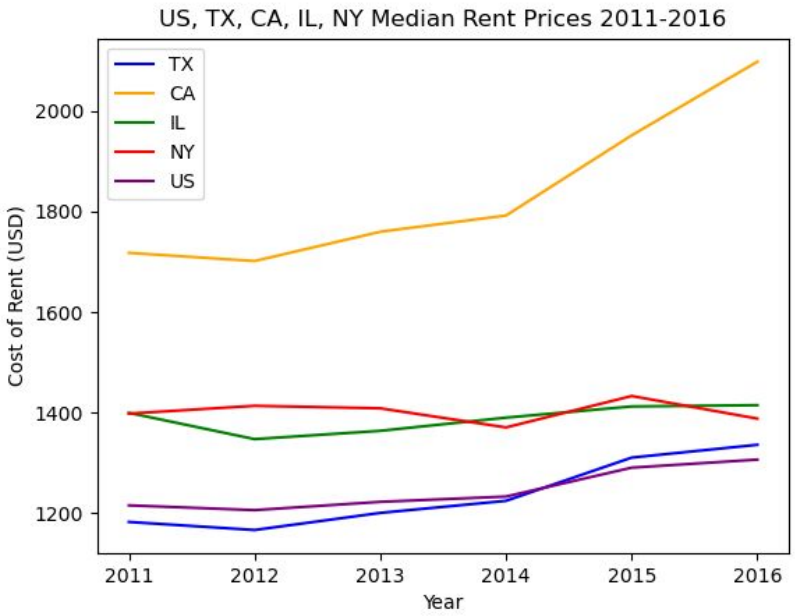
State	2011	2012	2013	2014	2015	2016	Rent Change (%)
TX	1182.50	1166.50	1200.50	1224.5	1310.50	1336.00	12.98
CA	1717.75	1701.50	1759.75	1792.0	1951.50	2098.00	22.14
IL	1399.50	1347.25	1363.75	1390.0	1412.25	1414.75	1.09
NY	1398.00	1413.50	1408.50	1370.5	1433.00	1388.00	-0.72
US	1215.50	1206.00	1222.50	1233.0	1290.50	1306.50	7.49

How US Median Rent Changed



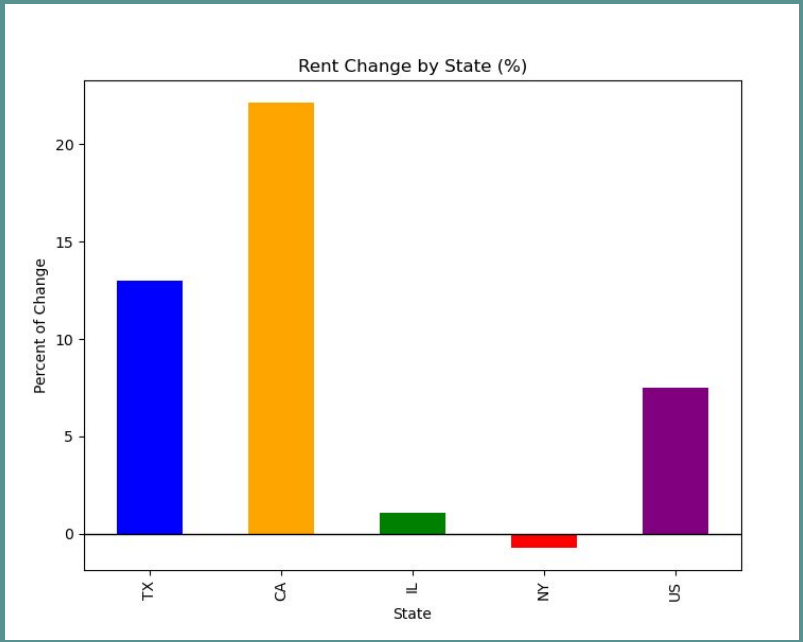
This graphical representation portrays the progressive escalation of rental costs over a span of six years, encompassing the entirety of the United States.

How Each State's Median Rent Changed



This graph elegantly illustrates the dynamic shift in median rental rates among the specifically chosen states, juxtaposed against the national average. Remarkably, every state exhibited an upward trajectory in rental costs during this period, with the notable exception of New York.

How Much Did Each State Increase/Decrease



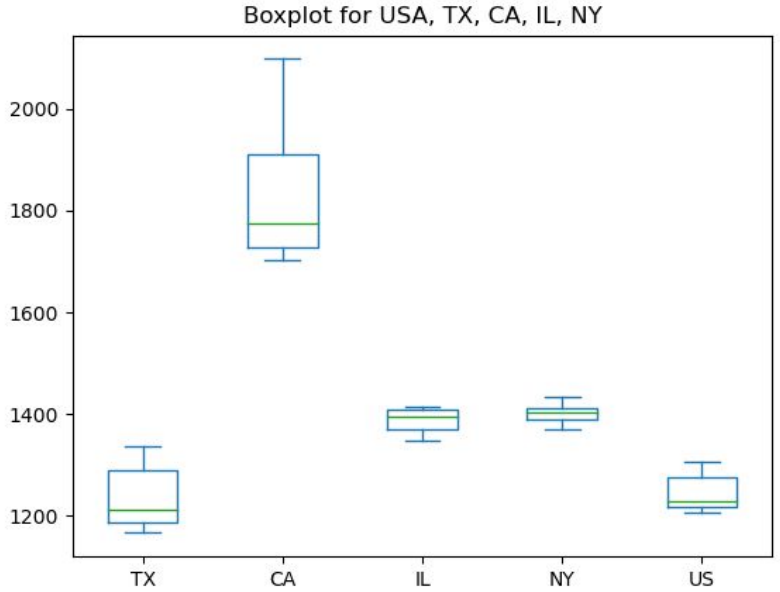
Texas and California exhibited more substantial increase compared to the United States trend. Conversely, Illinois experienced tempered rate of escalation, while New York even witnessed a decline.

Summary Statistics

- CA has the most expensive rent price
  - Rent varies the most
- TX has the cheapest rent price
  - Most similar price to index
- NY has the most uniform prices

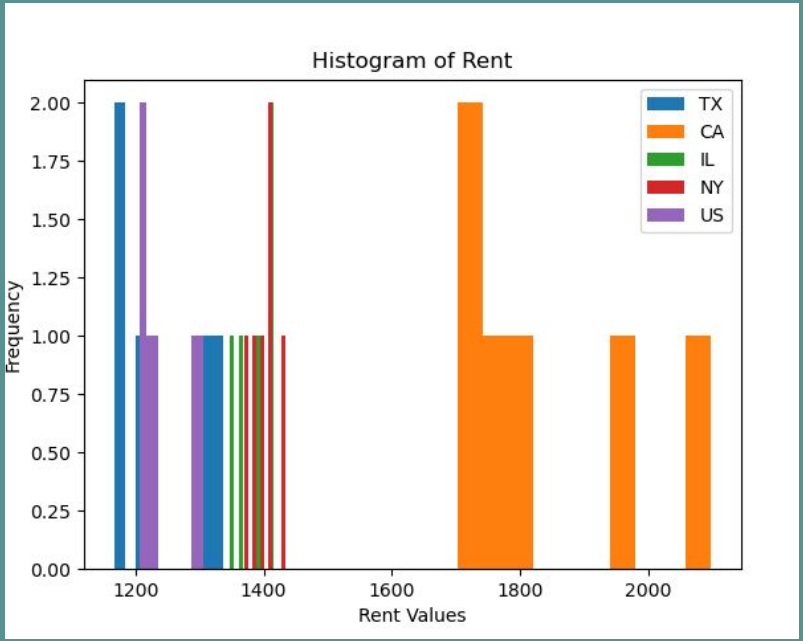
State	TX	CA	IL	NY	US
count	6.00000	6.000000	6.000000	6.000000	6.000000
mean	1236.75000	1836.750000	1387.916667	1401.916667	1245.666667
std	70.18814	156.079627	27.160480	21.631959	42.169499
min	1166.50000	1701.500000	1347.250000	1370.500000	1206.000000
25%	1187.00000	1728.250000	1370.312500	1390.500000	1217.250000
50%	1212.50000	1775.875000	1394.750000	1403.250000	1227.750000
75%	1289.00000	1911.625000	1409.062500	1412.250000	1276.125000
max	1336.00000	2098.000000	1414.750000	1433.000000	1306.500000

How Does the Spread of Rent Look



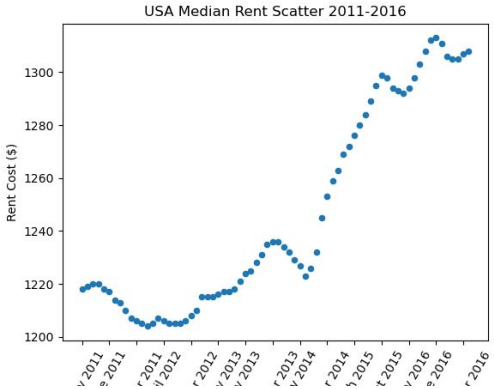
Within this dataset, outliers are conspicuously absent, signifying a notable uniformity in the observations. Additionally, it becomes apparent that California, distinguished by the largest Interquartile Ranges (IQRs) among all states, emerges as a distinctive outlier in terms of data distribution.

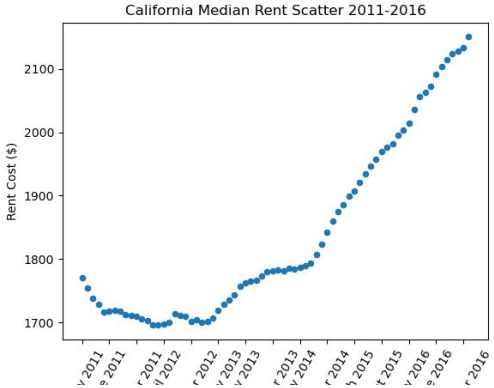
How Frequent Rent Prices Are

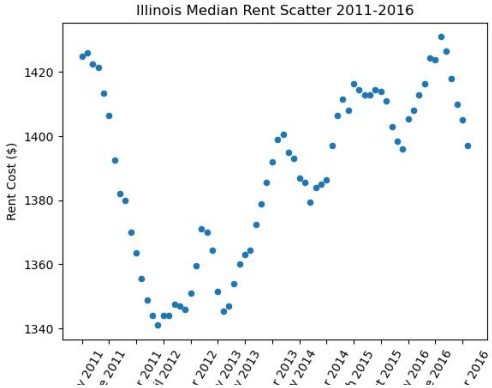


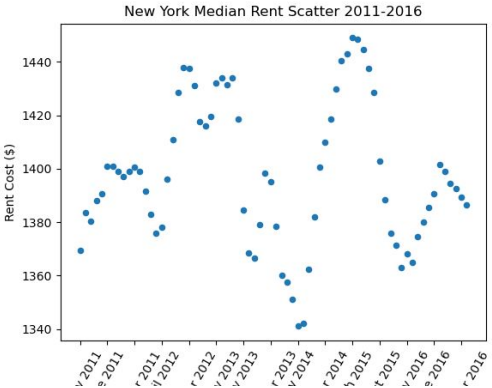
This histogram represents the data distribution of the data by state and year. It shows how CA has a skewed distribution to the right compared to the rest of the states compared to.

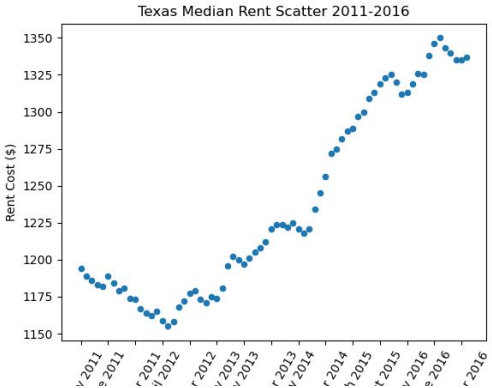
How Rent and Time Correlate











Chi-Square Test

Critical Value = 11.07

Chi Sq. Stat = CA: 0.13 , NY: -5.82, IL: -2.92, TX: -0.07

P-Values ~ 1.0

	2011	2012	2013	2014	2015	2016
CA Observed	1717.75	1701.5	1759.75	1792.0	1951.5	2098.0
Expected	1215.50	1206.0	1222.50	1233.0	1290.5	1306.5

	2011	2012	2013	2014	2015	2016
NY Observed	1398.0	1413.5	1408.5	1370.5	1433.0	1388.0
Expected	1215.5	1206.0	1222.5	1233.0	1290.5	1306.5

	2011	2012	2013	2014	2015	2016
IL Observed	1399.5	1347.25	1363.75	1390.0	1412.25	1414.75
Expected	1215.5	1206.00	1222.50	1233.0	1290.50	1306.50

	2011	2012	2013	2014	2015	2016
TX Observed	1182.5	1166.5	1200.5	1224.5	1310.5	1336.0
Expected	1215.5	1206.0	1222.5	1233.0	1290.5	1306.5

Conclusions

Questions:

How did the median rent price in the states (NY, CA, IL, TX) change across time (2011-2016) compared to the US as a whole?

- CA and TX had higher increases
- IL had a small increase
- NY surprisingly had a decrease

Which of the US states (NY, CA, IL, TX) had the greatest change in median rent price?

- CA

Hypotheses:

Null hypothesis: There is no significant statistical difference in rent across US states/regions.

- Failed to reject per p-value.

Hypothesis: New York had the most significant difference in rent compared to the rest of the US.

- Disproven with the data used but still unclear due to missing data from NYC.

Reference Citations

Landup, D. (n.d.) Change Tick Frequency in Matplotlib. StackAbuse. https://stackabuse.com/change-tick-frequency-in-matplotlib/

StackOverflow. (2021, May 10). "Unicode Error "unicodeescape" codec can't decode bytes... Cannot open text files in Python 3 [duplicate]. StackOverflow. https://stackoverflow.com/questions/1347791/unicode-error-unicodeescape-codec-cant-decode-bytes-cannot-open-text-file

Zillow. (2017, March 3). Zillow Rent Index, 2010-Present. Kaggle. https://www.kaggle.com/datasets/zillow/rent-index

Acknowledgements

- Instructor and TAs
- Tutor: Jesse Wright

Thank you!