

EDA Presentation

An EDA Project on the housing data from Kings County,
Seattle USA

Who Am I

- Rahul George, 25 Years Old
- Aspiring Data Scientist
- My 4 F's
 - Fitness
 - Food
 - Family
 - Friends.



Rahul George, 20.12.2024

Berlin, Germany

Let Me Introduce My Client

- **Mr. Jacob Phillips**
- A Buyer who has an unlimited budget and is looking for:
 - At least 4 bathrooms or a smaller house nearby
 - Large Lot for a swimming pool and tennis court
 - Close to golf courses
 - Interested in historical properties
 - Not interested in waterfront properties



What Data Am I Working On

- **Dataset Overview**
 - **Time Span:** 02.05.2014 - 27.05.2015
 - **Location:** King County, Seattle
 - **Size:** 21597 rows x 21 columns

Data Overview

- **Key Statistics**
 - **Total Properties:** 21420
 - **Price Range:** 78,000 - 7,700,000
 - **Average Lot Size:** 15,100 Sqft (Approx)
 - **Average Living Space:** 2080.32 Sqft

My Hypothesis

- **Hypothesis 1:** Houses with more floors have larger living spaces
- **Hypothesis 2:** Historical houses (Built on or before 1940) that have been renovated have larger basement areas compared to historical houses that have not been renovated.
- **Hypothesis 3:** Houses located closer to the city center have higher price per square foot compared to houses on the outskirts

Hypothesis Results

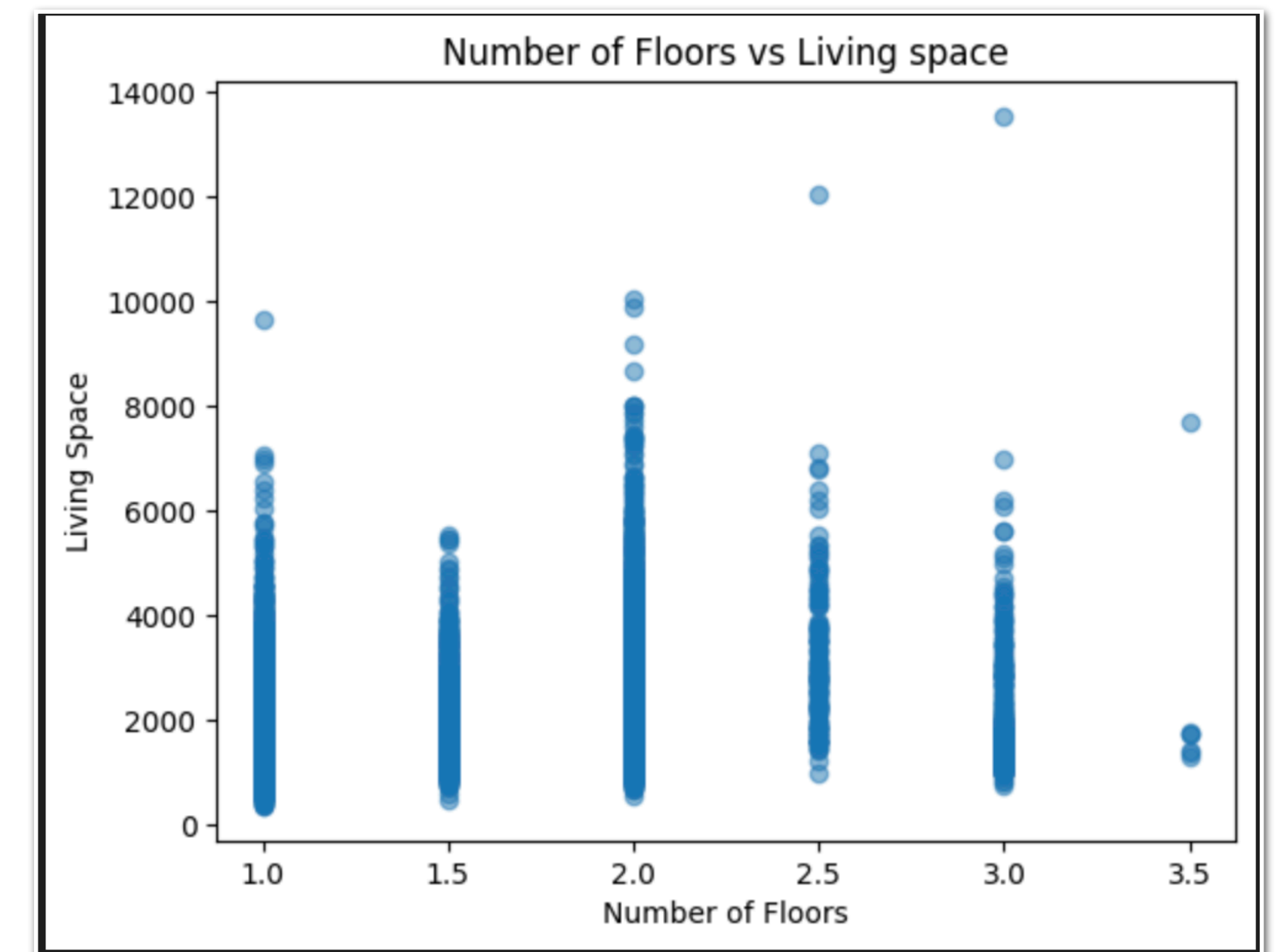
Hypothesis 1: Houses with more floors have larger living spaces

Correlation between Floors and Living Space: 0.3539

P-Value: 0

Conclusion

Positive Correlation, but not very strong. The P-Value indicates that the result is statistically significant



Hypothesis Results

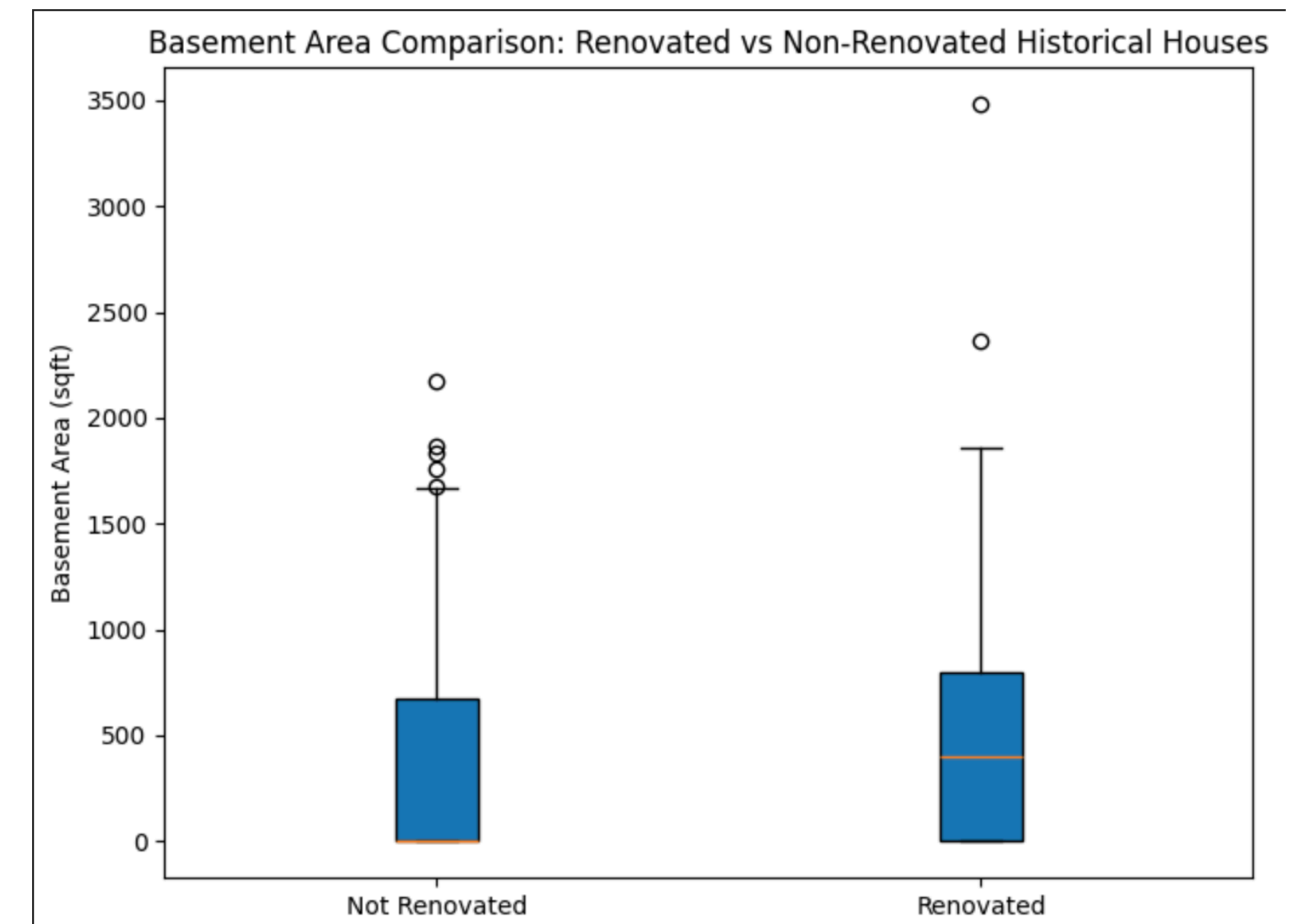
Hypothesis 2: Historical houses (Built on or before 1940) that have been renovated have larger basement areas compared to historical houses that have not been renovated

T-Static: 4.818

P-Value: 0.00000203

Conclusion

Renovated historical houses tend to have larger basement areas.



Hypothesis Results

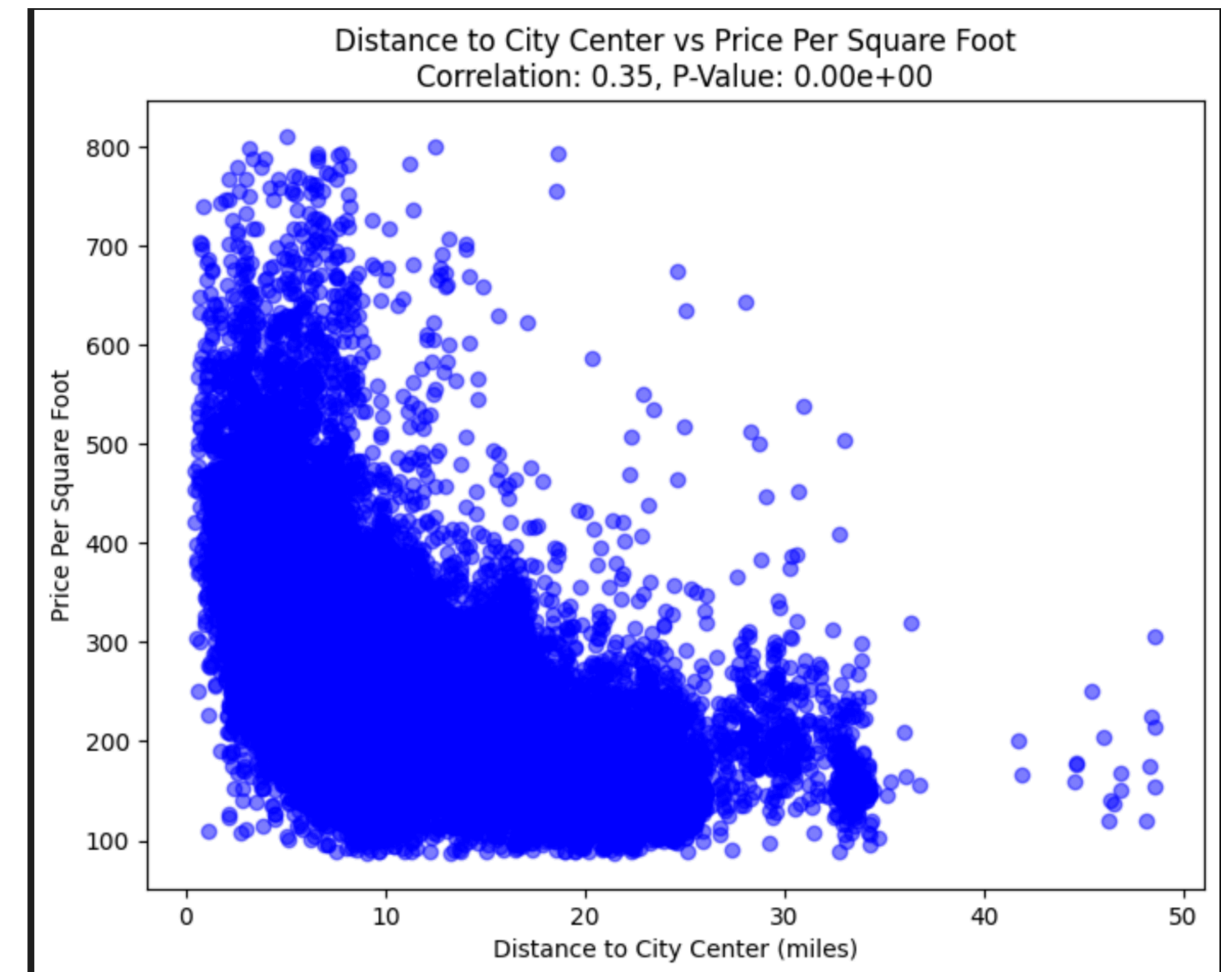
Hypothesis 3: Houses located closer to the city center have higher price per square foot compared to houses on the outskirts

Correlation between Distance to city centre and Price per Sqft: -0.566

P-Value: 0

Conclusion

Houses closer to the city center have higher price per square foot



Insights

- **Factors Influencing Price:**

- Location (proximity to city center) has the strongest impact.
- Living space and basement size have moderate effects.
- Historical status adds value.

- **Trends**

- Higher property grades significantly correlate with price.
- Larger lot sizes alone do not drive price unless paired with desirable features.

Revisiting My Client

- **Mr. Jacob Phillips**
- A Buyer who has an unlimited budget and is looking for:
 - At least 4 bathrooms or a smaller house nearby
 - Large Lot for a swimming pool and tennis court
 - Close to golf courses
 - Interested in historical properties
 - Not interested in waterfront properties



Recommendations For Mr. Phillip

	id	pricePrice	bathroomsNumber	sqft_lotsquare	yr_built	lat	long	distance_to_golf	grade	condition
18314	5317100750	2920000.0	4.75	24085.0	1926	47.6263	-122.284	0.941743	10	5
14172	1333300145	2230000.0	4.00	30120.0	1933	47.6379	-122.311	1.091845	11	4
7245	6762700020	7700000.0	8.00	27600.0	1910	47.6298	-122.323	1.780191	13	4
5961	5249800010	2730000.0	4.25	43838.0	1906	47.5703	-122.280	4.572270	12	4

	id	pricePrice	bathroomsNumber	sqft_lotsquare	sqft_livingsquare	yr_built	lat	long	grade	condition
16506	9185700414	1180000.0	1.75	7200.0	1610.0	1973	47.6279	-122.287	8	3
6019	2158900290	920000.0	1.50	3600.0	1850.0	1929	47.6376	-122.307	8	3
8993	2125049131	680000.0	1.75	5500.0	1620.0	1950	47.6393	-122.308	7	3
18286	2125049133	715000.0	1.75	6500.0	1920.0	1951	47.6394	-122.308	7	3
5183	2663000050	525000.0	1.00	4000.0	1570.0	1920	47.6275	-122.321	7	3
20755	6762700376	650000.0	2.75	1251.0	1540.0	2002	47.6298	-122.321	8	3
20901	6762700452	575000.0	3.00	1287.0	1384.0	2006	47.6295	-122.320	8	3
21055	2163900081	1080000.0	2.50	1891.0	1990.0	2012	47.6271	-122.324	9	3
5531	7950300005	681000.0	1.00	6356.0	1700.0	1907	47.5677	-122.281	7	3

Rahul George, 20.12.2024

Berlin, Germany

Questions?

Feedback is also highly appreciated