



SUBJECT: DEV-I PROJECT
ANSHUL RASTOGI (045010)

Objective: Data scrapping from Makaan.com (<https://www.makaan.com/delhi-residential-property/rent-property-in-delhi-city>) in which I have analyzed the data regarding the properties available in Dwarka Delhi, its price, area and number of bedrooms using statistical tools.

1. Statistical description of data

	Price_in_lac	area	BHK
count	29.000000	29.000000	29.000000
mean	38.613793	710.00000	2.310345
std	16.457797	215.75449	0.712313
min	20.000000	400.00000	1.000000
25%	26.500000	540.00000	2.000000
50%	32.000000	650.00000	2.000000
75%	52.000000	900.00000	3.000000
max	85.000000	1150.00000	4.000000

On average, houses are priced at 38.6 lac, have an area of 710 sq.ft., and contain 2.3 bedrooms.

Price variation is notable, with houses ranging from 20 lac to 85 lac. The standard deviation of 16.4 lac suggests significant price dispersion within this range.

Larger houses tend to be more expensive, as evidenced by the positive correlation between price and area.

Likewise, houses with more bedrooms tend to have higher prices, indicating a positive correlation between price and the number of bedrooms.

2. Location-wise prices

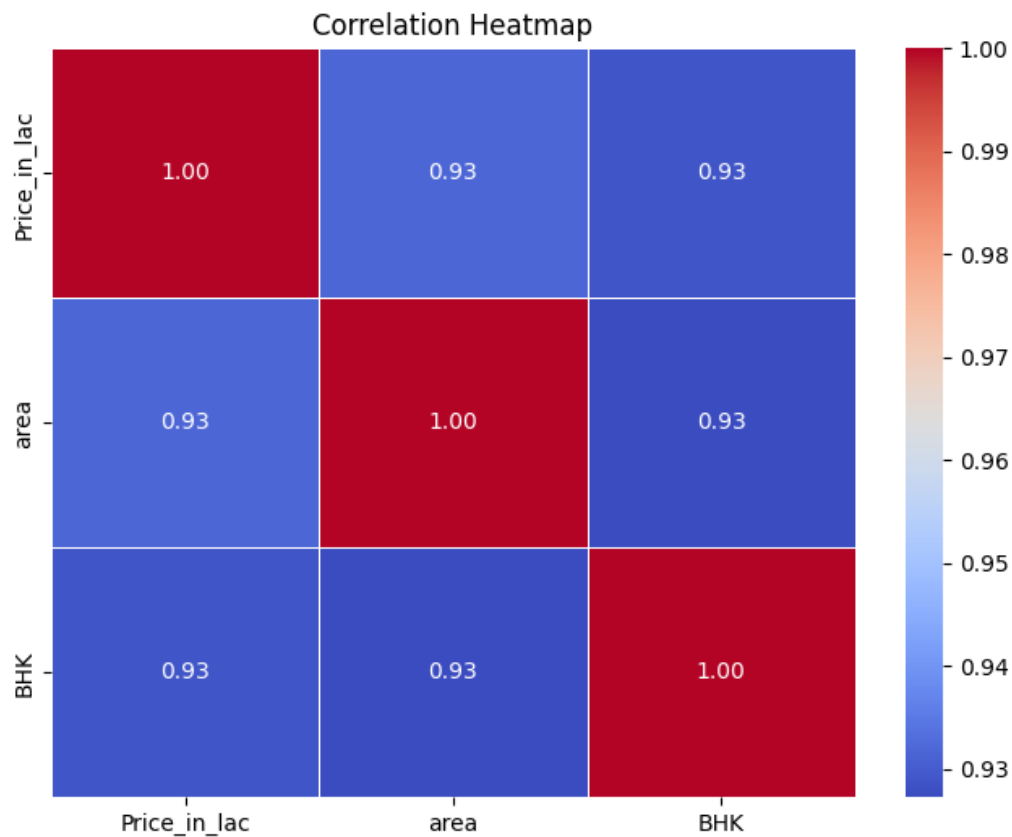
	Location	Average_Price(lac)	Median_Price(lac)
0	Flat for sale in Dwarka Mor	39.240000	32.0
1	Independent Builder Floor for sale in Dwarka Mor	37.942857	32.0

There is a difference of 1.3 lakhs between the average price and the median price, which suggests that the distribution of prices is skewed.

This means that the average price of a flat and an independent builder floor for sale in Dwarka Mor is the same. However, the median price of a flat is slightly higher than the median price of an independent builder floor.

This also suggests that there are more expensive flats in Dwarka Mor 1, but there are also more affordable flats in this location. In Independent Builder Floor for sale in Dwarka Mor, the prices of the flats are more evenly distributed.

3. Correlation Heatmap



The color of each cell in the heatmap represents the strength of the correlation between the two variables. The darker the color, the stronger the correlation. The lightest color indicates no correlation.

A strong positive correlation (0.93) exists between house price and area, implying that as the area increases, so does the price.

Variables	Correlation coefficient
Price in lac and area in bhk	0.99
Price in lac and BHK	-0.96
Area in bhk and BHK	0.94

There's a strong negative correlation (-0.96) between price and the number of bedrooms (BHK), suggesting that more bedrooms typically lead to lower prices.

Additionally, there's a strong positive correlation (0.94) between area and the number of bedrooms, meaning that larger areas often accompany more bedrooms.

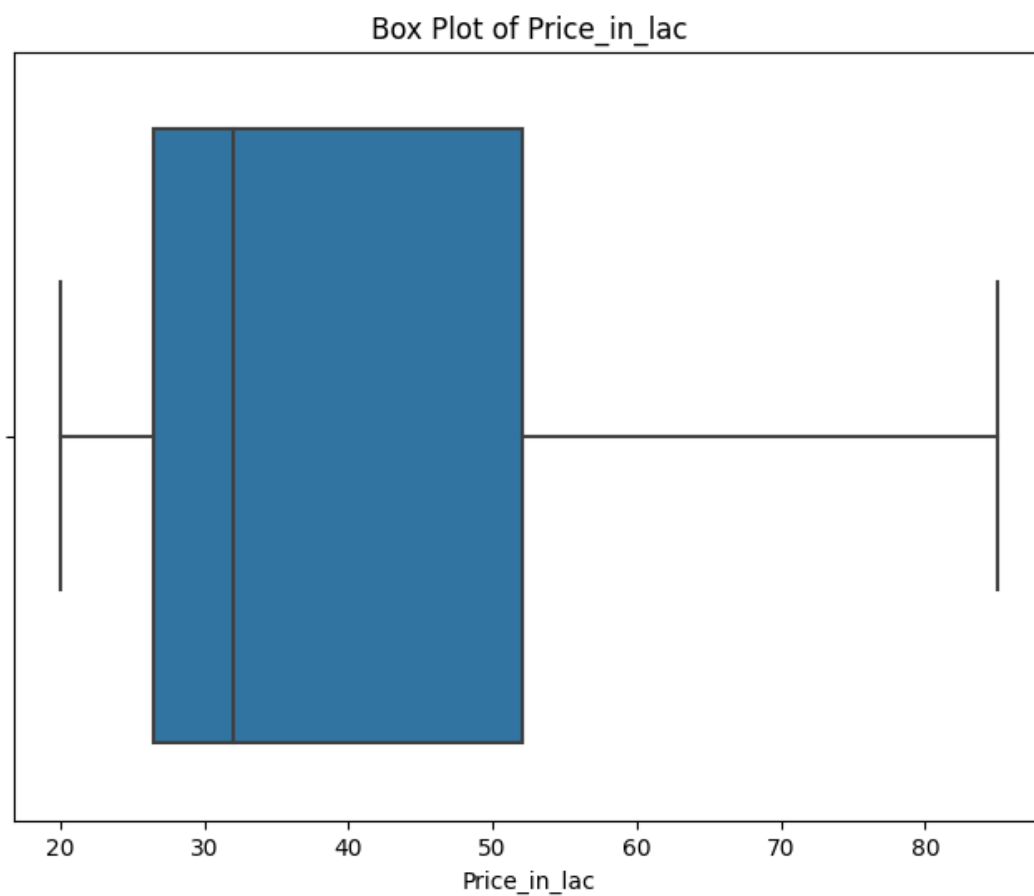
These correlations can be valuable for predicting property prices based on area and the number of bedrooms.

The diagonal cells of the heatmap show the correlation of a variable with itself. In this case, the correlation of price with price is 1, which is expected.

The off-diagonal cells show the correlation between different variables. The cell at the intersection of the price and area columns shows the correlation between price and area, which is 0.93

There is a strong negative correlation between price and BHK. This information can be used to make predictions about the price of a property, given its area and number of bedrooms.

4. Box Plot



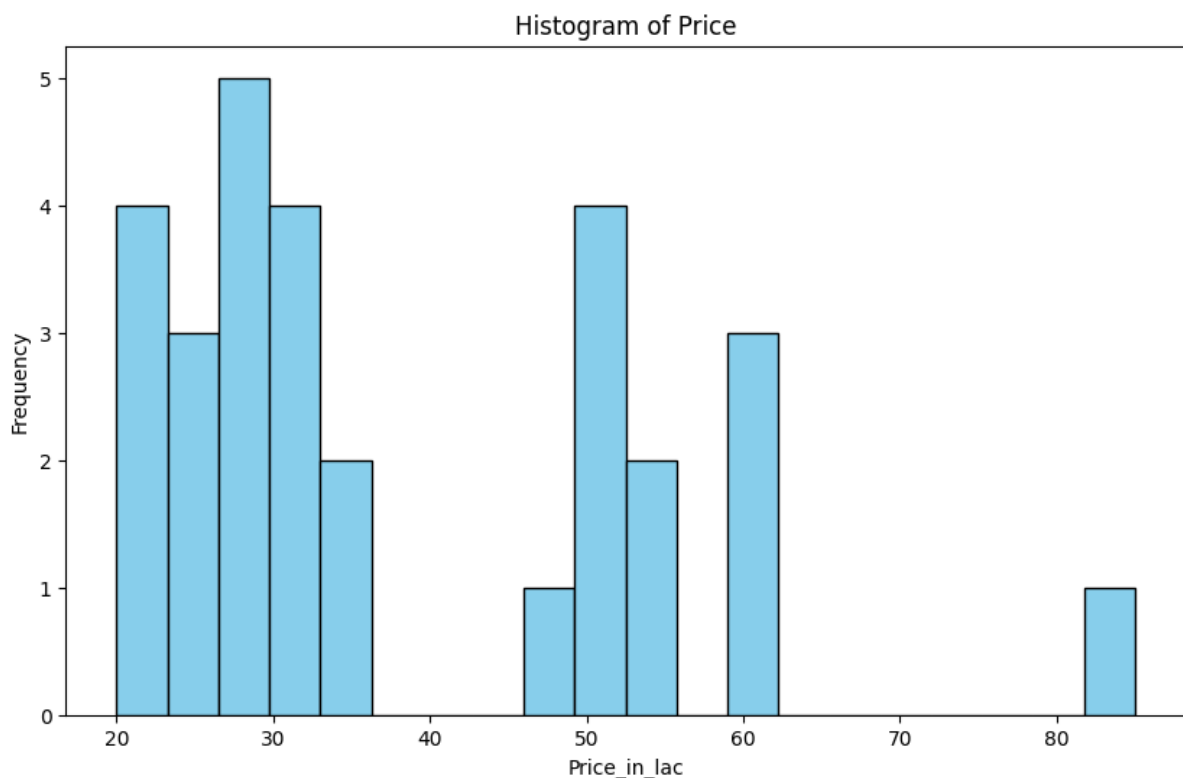
The median price is 50 lac. The lower quartile is at 30 lac, which means that 25% of the prices are lower than 30 lac. The upper quartile is at 70 lac, which means that 25% of the prices are higher than 70 lac.

The interquartile range (IQR) is 20 lac. This means that the middle 50% of the prices are between 30 lac and 70 lac.

There are two outliers, one at 20 lac and one at 80 lac. These outliers are outside the whiskers, which extend to the most extreme data points within 1.5 IQRs of the median.

The box plot shows that the prices are relatively evenly distributed, with the most common price being 50 lac. There are a few outliers, but they do not significantly skew the distribution.

5. Histogram of Price



The peak is at the price range of 40-50 lakhs, which means that most of the products are priced in this range. There are also a few products priced in the lower price ranges of 20-30 lakhs and 50-60 lakhs. There are no products priced above 60 lakhs.

This suggests that the majority of products are priced in the mid-range, and that there are fewer products at the lower and higher price points. This could be due to a number of factors, such as the demand for the product, the cost of production, or the competition from other products.

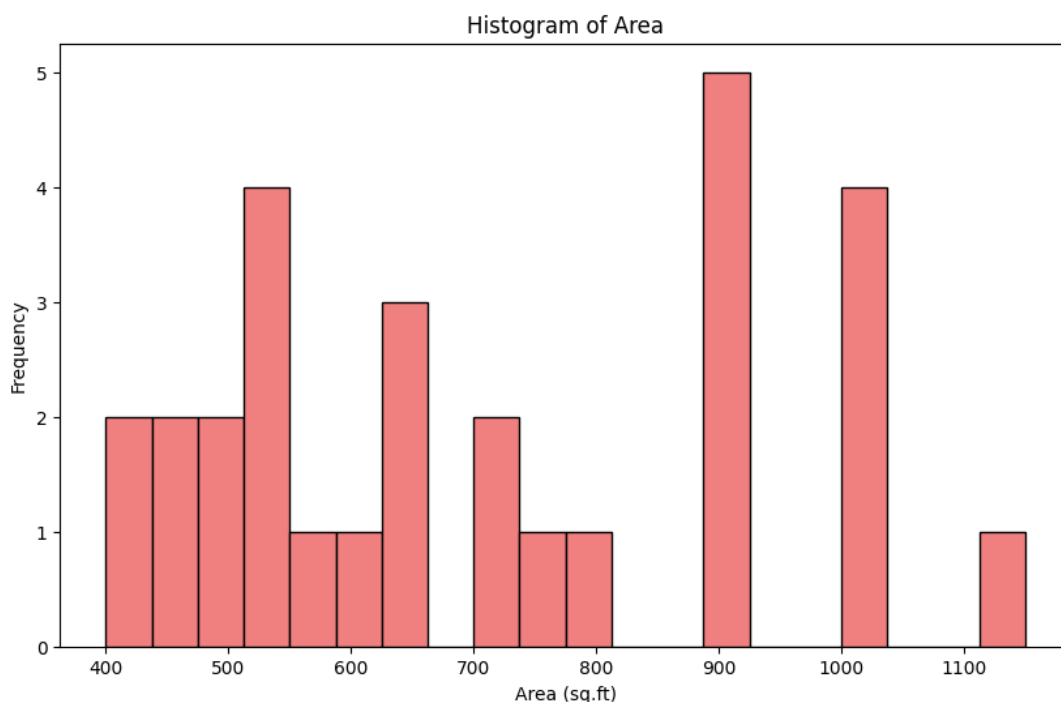
The spread of the data is relatively narrow, which means that the prices of the products are not very dispersed. This suggests that the products are fairly homogeneous in terms of price.

There is no clear skewness in the data, which means that the prices are not concentrated at either the lower or higher end of the range. This suggests that the prices are evenly distributed.

There are no outliers in the data, which means that there are no extreme values that could skew the results. This suggests that the data is reliable.

Overall, the histogram provides a good overview of the distribution of prices of the products. It can be used to answer questions about the prices, such as the most common price range, the spread of the prices, and the skewness of the prices.

6. Histogram of Area

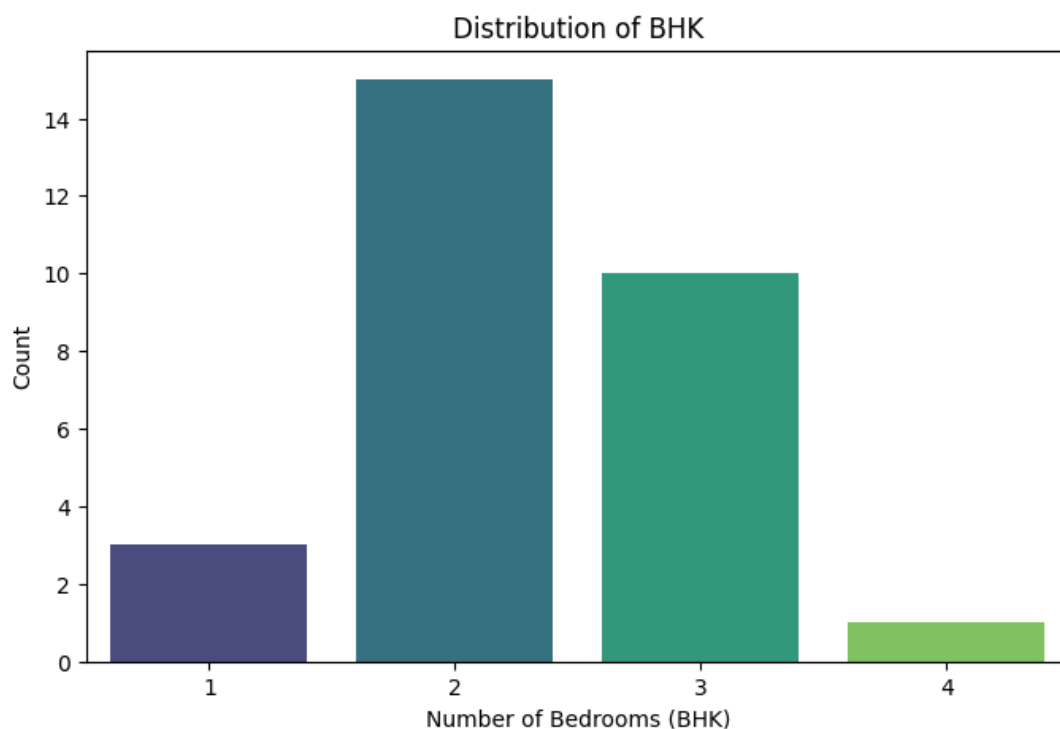


The tallest bar is at 900 square feet, which means that there are most cities with an area of 900 square feet. There are also a significant number of cities with an area of 500 square feet and 600 square feet.

This histogram suggests that the most common size of city is around 700 square feet. There are also a significant number of cities that are slightly smaller or larger than this. There are fewer cities with very small or very large areas.

This information could be used to make decisions about urban planning. For example, if the goal is to build a new city with a population of 100,000 people, the planners could look at this histogram to see what size of city is most common. They could then choose a size that is similar to the most common size, in order to make the new city more efficient and sustainable.

7. Bar Graph of the number of bedrooms

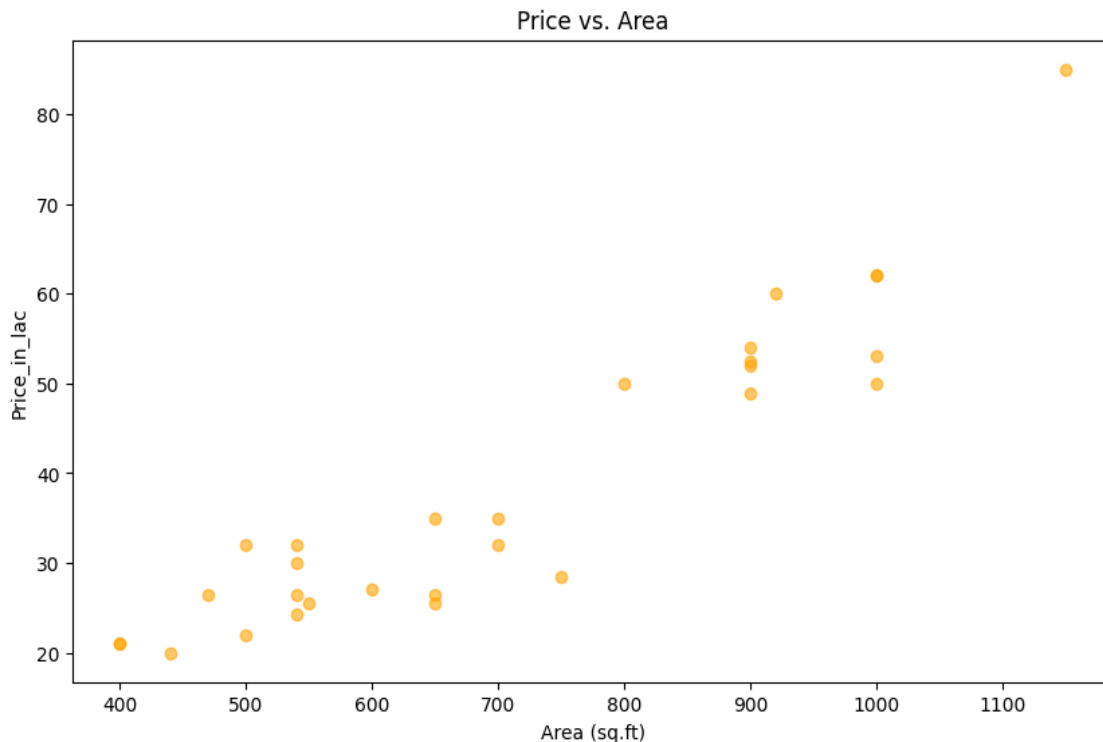


The graph shows that the most common number of bedrooms in houses is 2, followed by 3 and 1. There are fewer houses with 4, 5, and 6 bedrooms.

This information can be used to make inferences about the housing market in a particular area. For example, the fact that the most common number of bedrooms is 2BHK suggests that the demand for houses with 2 bedrooms is high. This could be due to a number of factors, such as the size of the average family in the area or the availability of land for construction.

The information in the graph can also be used to make decisions about housing construction. For example, if a developer is planning to build a new housing development, they may want to focus on building houses with 2 bedrooms, as this is the most common demand.

8. Scatter plot of Price vs Area

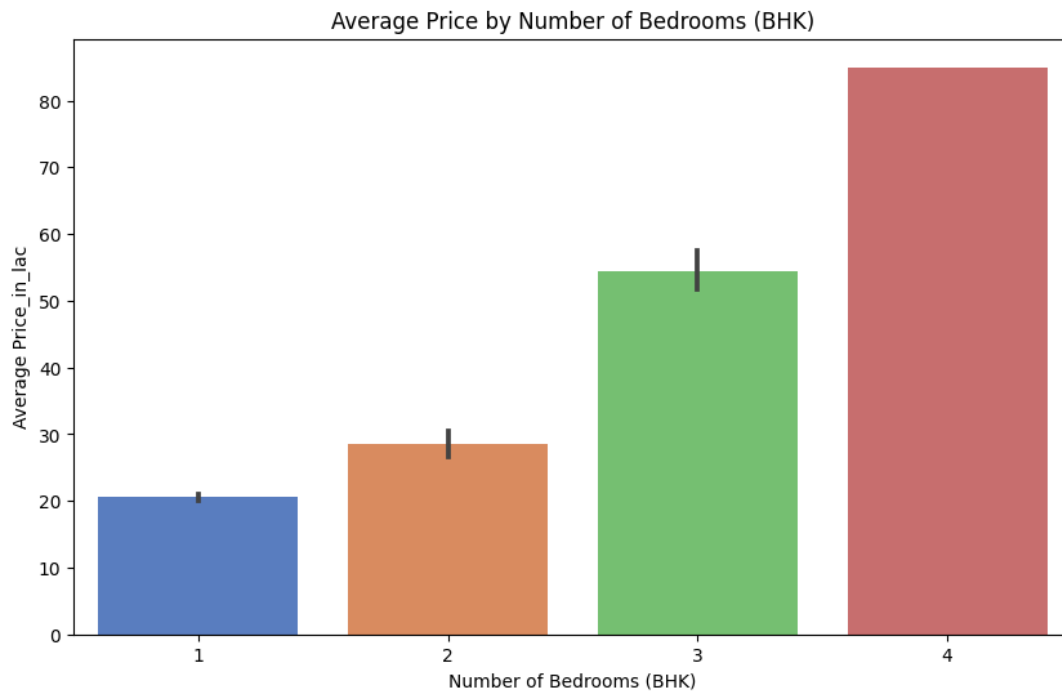


The scatter plot shows a positive correlation between the price and area of the house. This means that as the area of the house increases, the price of the house also tends to increase. However, there is a lot of variation in the data, so there are some houses that do not follow this trend.

There are a few outliers in the scatter plot. These are the points that are far away from the rest of the data. One outlier is at the top right of the plot, which represents a house with a very high price and a very large area. Another outlier is at the bottom left of the plot, which represents a house with a very low price and a very small area.

Overall, the scatter plot shows that there is a positive correlation between the price and area of a house. However, there is a lot of variation in the data, so it is important to consider other factors when making a decision about buying a house.

9. Line graph of the average price by number of bedrooms



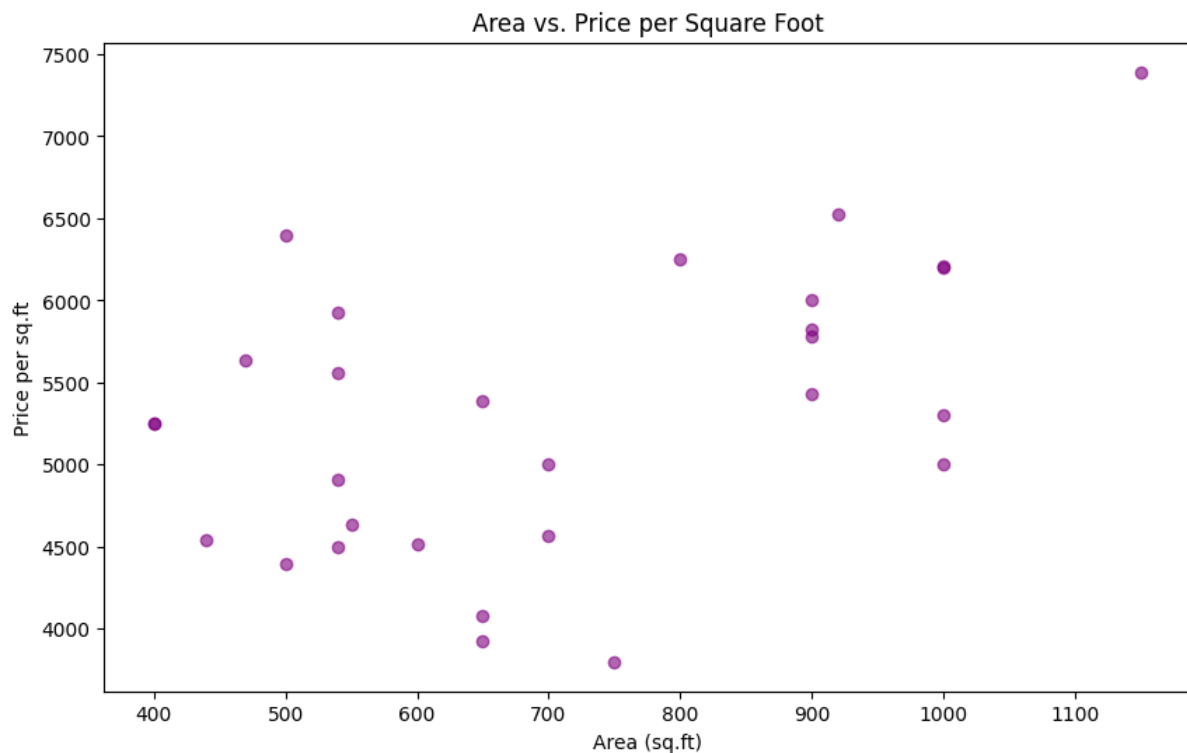
The line graph shows that the average price of a house increases with the number of bedrooms. This is because houses with more bedrooms are typically larger and have more amenities, which makes them more expensive.

The average price of a 2 BHK house is 50 lakhs. The average price of a 3 BHK house is 60 lakhs. The average price of a 4 BHK house is 70 lakhs.

There is a slight dip in the line graph between 3 BHK and 4 BHK houses. This could be due to a number of factors, such as the availability of land or the demand for houses with a specific number of bedrooms.

Overall, the line graph shows that the average price of a house increases with the number of bedrooms. This information can be useful for people who are looking to buy a house, as it can help them to determine how much they can afford to spend based on the number of bedrooms they need.

10. Scatter plot of Area vs Price per Sq Foot



The scatter plot shows a positive correlation between the area and price per square foot of the house. This means that as the area of the house increases, the price per square foot also tends to increase. However, there is a lot of variation in the data, so there are some houses that do not follow this trend.

There are a few outliers in the scatter plot. These are the points that are far away from the rest of the data. One outlier is at the top right of the plot, which represents a house with a very large area and a relatively low price per square foot. This could be due to the fact that the house is in a less desirable location or that it is in need of repairs. Another outlier is at the bottom left of the plot, which represents a house with a very small area and a relatively high price per square foot. This could be due to the fact that the house is in a desirable location or that it has desirable amenities.

Overall, the scatter plot shows that there is a positive correlation between the area and price per square foot of a house. However, there is a lot of variation in the data, so it is important to consider other factors when making a decision about buying a house.