

NAME: Vijesh Hingu  
UID: 2021300042  
CLASS: BE COMPS A  
BATCH: ADV BATCH F

## ADV EXPERIMENT 5

### DATASET:

Covid-19 dataset - <https://www.kaggle.com/datasets/arnavkulkarni/housing-prices-in-london>

### Objectives:

1. To visualize the distribution and relationship between various features in the housing dataset.
2. To identify potential outliers and understand the spread of the data.
3. To explore the relationship between independent variables and the target variable (e.g., house prices).
4. To create informative visualizations that can guide decision-making in the housing market.

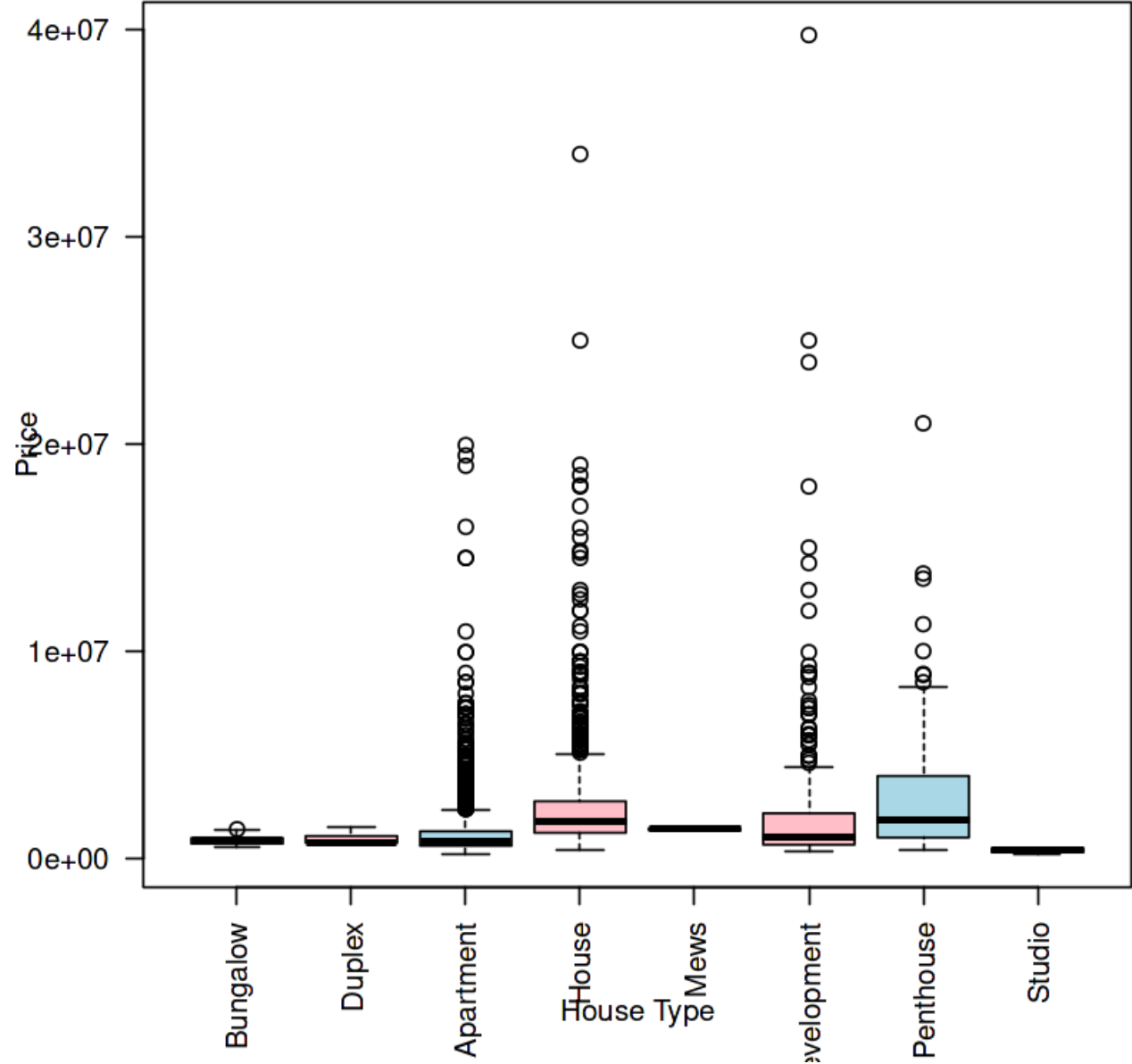
### DATASET DESCRIPTION:

This dataset comprises of various house listings in London and neighbouring region. It also encompasses the parameters listed below, the definitions of which are quite self-explanatory.

- Property Name
- Price
- House Type - Contains one of the following types of houses (House, Flat/Apartment, New Development, Duplex, Penthouse, Studio, Bungalow, Mews)
- Area in sq ft
- No. of Bedrooms
- No. of Bathrooms
- No. of Receptions
- Location
- City/County - Includes London, Essex, Middlesex, Hertfordshire, Kent, and Surrey.
- Postal Code

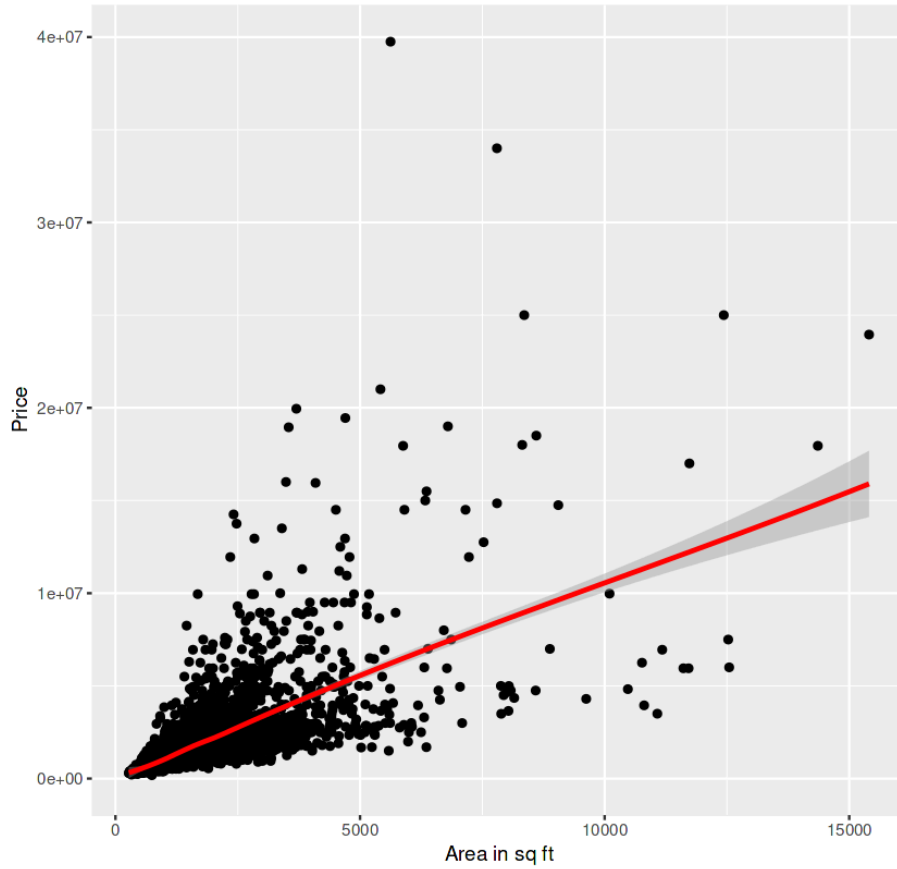
REPORT:

Boxplot of Property Price by House Type

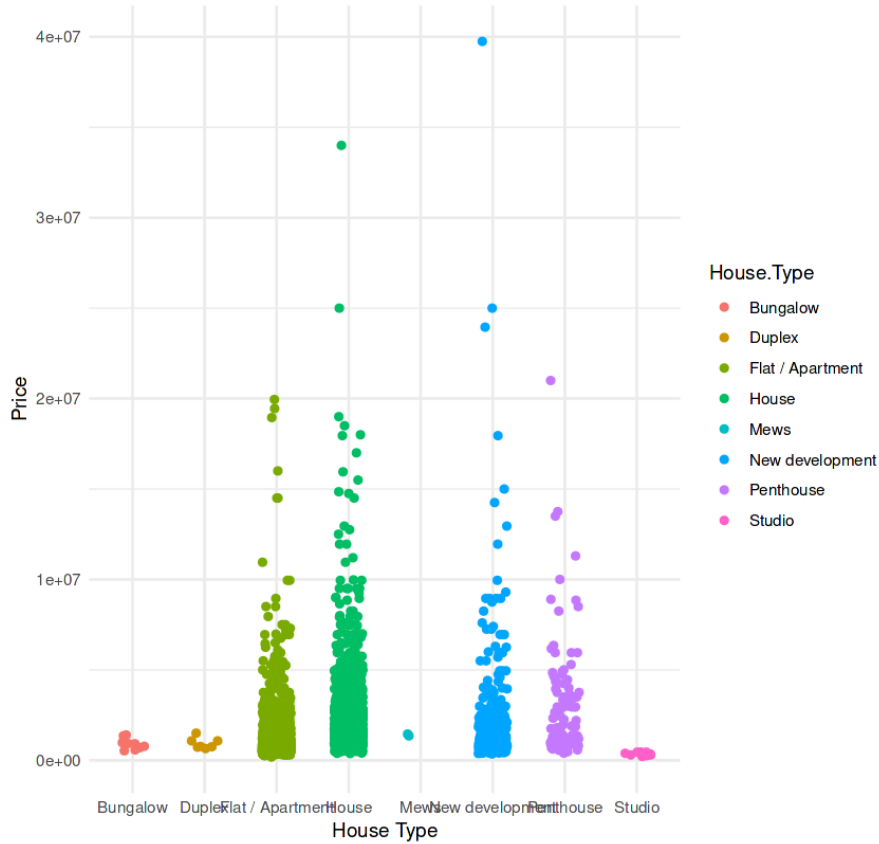




Non-Linear Regression of Price vs Area in sq ft



Jitter Plot of Price by House Type



Question: Which locations are mentioned most frequently in the dataset?

Answer: The word cloud reveals the most frequent locations. Larger words represent locations that appear more often in the dataset.

Box and Whisker Plot

Question: How do property prices vary between different house types?

Answer: The box plot shows the distribution of property prices for different house types, with the range of prices being wider for houses than for flats/apartments.

Violin Plot

Question: What is the density of property prices across different house types?

Answer: The violin plot indicates the distribution and density of property prices. For example, houses tend to have a higher density of prices in the higher range compared to flats/apartments.

Linear and Non-Linear Regression Plots

Question: What is the relationship between property price and area in square feet?

Answer: The linear regression shows a positive correlation between area and price. The non-linear regression better captures any non-linearities in the data.

3D Chart

Question: How do price, area, and number of bedrooms relate to each other?

Answer: The 3D scatter plot allows you to explore the relationship between these three variables, showing that larger homes with more bedrooms tend to be priced higher.

Jitter Plot

Question: Is there any overlap in property prices between different house types?

Answer: The jitter plot shows that there is some overlap in prices between house types, but houses generally have higher prices compared to flats/apartments.

**CONCLUSION:** I have successfully plotted advanced graphs using R language and answered all questions regarding the dataset.