

## **TASK:**

After reading the assigned articles and textbook chapters, provide an example of a dataset you would like to explore using exploratory data analysis techniques. Discuss the potential research questions or hypotheses that could be addressed through this analysis, as well as the types of visualizations and statistical techniques that might be used to explore the data. Finally, provide a brief overview of how you would approach analyzing the data in Python.

Exploring housing price data using exploratory data analysis

Description of the dataset:

I would like to explore a dataset containing information on housing prices. The dataset includes variables such as house size (in square feet), number of bedrooms...

Research questions or hypotheses:

What is the relationship between the size of the house and its selling price?

Do homes with more bedrooms and bathrooms typically sell for higher prices?

Types of visualizations and statistical techniques:

Scatterplots: To visualize the relationship between home size, number of bedrooms/bathrooms, and sales price.

Histograms and Density Plots: To examine the distribution of house prices and other numerical variables.

Box plots: to compare the distribution of house prices in different neighborhoods or zip codes.

Data Analysis Approach in Python:

I would start by loading the dataset into a pandas DataFrame, then generate descriptive statistics to get an overview of the dataset.

Next, I would create visualizations using matplotlib and Plotly to explore relationships between variables and identify patterns or trends.

DataSet: <https://www.kaggle.com/datasets/shree1992/housedata>

Github: <https://github.com/ARMAND-cod-eng/Topic-8-DQ-1/blob/main/Topic%208%20DQ%201.ipynb>