Data Science Case Study

Description of the Data
The dataset shared with you is an anonymized dataset of different properties in Dubai. The dataset contains the size, number of bedrooms, number of bathrooms, neighborhood name, and building name and the listing price for different properties in Dubai. You can find the data under **data_science_challenge_data.csv**

Assignment

1. Build a **Property Price Prediction Model**. Please use the Python programming language.
    a. Explore the dataset, and share with us any insights that you may find which can help you create the price evaluation tool. Please summarize your findings in terms of the relationship between the different features, the price, and feature importance.
    b. Build a model which predicts the listing price of the property based on the property's features
        i. Model Input: Features of property
        ii. Model Output: Predicted Price
    c. How do you evaluate the quality of your results?
    d. What are the possible shortcoming & extensions of your approach?

2. Build a **Property Price Valuation Tool**, which would take as input the features of a property and its listing price and determine whether the property is under-priced, fairly priced or overpriced.
    a. Implement a program to determine whether a property is underpriced, fairly priced or overpriced. This is your chance to show us the process that you would follow to solve this problem, and how you would model the data.
        i. Program Input: Property features and its price
        ii. Program Output: Whether the prediction is Underpriced, Fairly Priced, Overpriced
    b. How do you evaluate the quality of your results?
    c. What are the possible shortcoming & extensions of your approach?

3. Explain the Price predicted by prediction model.
   This is an open ended question. Draft a plan on how we can validate the price suggested by the model. You can assume any other data point that is not present currently or can use any approach to define metrics which helps Users understand how valid the predictions are.
4. Build an architecture to serve the developed property prediction + valuation tool to users
    a. Draw an architectural diagram and explain your approach on how we can serve this tool to the users.

b. How can you scale this tool?
c. What could be some of the KPI's for this tool.
d. How to monitor the tool performance to make sure it's working well

Deliverables

Please share with us

- your program code
- a short presentation (5-10 slides) or 1 page pdf with the results of your work.