Usecase 6 - (Project 3)

By: eng. Esraa Madhi

Utilizing your knowledge of advanced EDA tools, proceed with the analysis of the provided data.

This project must at least satisfy the following minimum requirements:

Usecase 6

For any **Data project** we should go through these steps:

Step 1: Defining the Problem Statement

Define at least 4 questions to answer using the data

Step 2: Collecting Data

- Use the following dataset.
  - https://www.kaggle.com/datasets/abdulmalikm/apartments-in-riyadh
  - https://www.kaggle.com/datasets/myfaisal/riyadh-aqaar-dataset
  - https://www.kaggle.com/datasets/salmanshir/riyadhhousingdata

Step 3: Data Quality Checking and Remediation

Step 4: Exploratory Data Analysis

- For these two steps, make sure to do:
  a. Data Profiling: apply the 7 types of data profiling
  b. Data Cleaning: handle missing values, correcting errors, and dealing with outliers.

c. Univariate Analysis &Bivariate/Multivariate Analysis: to understand their distribution and look at the relationships between variables. For your visualizations make sure to:
   - Drive meaningful insights (at least 10 different charts).
   - Choose a specific style for your charts.
     - Apply one color palette from your choice on all charts.
     - Use the title, x-y labels, font size, figure size, and legends.

Step 5: Building Machine Learning Models

Not applicable

Step 6: Model Evaluation

Not applicable

Step 7: Communicating Results

- Create an interactive data story using Streamlit.
- Report your final conclusion and findings in one page (readme markdown file).
  - Team members.
  - Introduction (Problem, Objectives)
  - Dataset Overview and Source.
  - List of EDA steps that applied on data with description
  - Describe the final ten insights with their charts

Step 8: Model Deployment

Not applicable

Step 9 : Model Performance Maintenance in Production

Not applicable

Note: the red steps means they are Not applicable in the project

- **The Final presentation will be on Sunday.**
- **Due Date: Sun, 19 May, 08:00 AM.**

Final Deliverables:

- Notebook file(.ipynb).
- Streamlit file (.py)
- README.md file (.md).