Python Project - Data Analytics and Machine Learning

Project Report Template

1. Abstract
2. Introduction
   - Briefly introduce the project topic and its significance.
   - Provide an overview of the dataset selected and its data sources.
3. Data
   - Describe the selected real-world dataset and its relevance to the project.
   - Provide a reference to the source of the dataset. You can find the dataset [here](INSERT LINK TO DATA SOURCE).
4. Importing
   - Explain the method used to import the data into Python.
     - Import data from a flat file (e.g., CSV, XLS, XLSX, TXT).
     - Retrieve data using online SQL, APIs, or web scraping.
5. Data Preparation
   - Discuss the steps taken to prepare the data for analysis.
   - Explain the creation of pandas DataFrames.
   - Describe the sorting, indexing, filtering, and grouping operations performed on the data.
   - Explain how duplicate entries and missing values were handled.
   - Discuss the definition of custom functions for reusable code.
   - Provide details on how multiple DataFrames were merged, if applicable.
6. Data Visualization
   - Generate at least four charts using the Matplotlib library.
   - Generate at least four charts using the Seaborn library.
   - Conduct univariate and bivariate analysis using appropriate charts and techniques.
7. Machine Learning
   - Predict a target variable with Supervised or Unsupervised algorithm.
   - Implementation of Supervised or Unsupervised Model/s.
   - Perform Model Evaluation using metrics suitable for the choice of ML model/s.
   - Perform hyper parameter tuning or boosting, whichever is relevant to the model. If it is not relevant, justify that in your report and Python comments
8. Insights
   - Derive at least eight valuable insights from your data analysis.
   - Justify each insight with reference to the charts or analysis performed.
9. Results and Conclusion
   - Summarize the key findings and insights obtained from the project.