Assignment 01

September 23,2024

Submission Deadline: 30th September 2024, 11:59 PM

Google form: https://forms.gle/K7fP23NYWtzic5Xz7

Objective:

Perform an explorative analysis of a dataset of your choice from the Kaggle library (https://www.kaggle.com/datasets). The analysis should be a mix of calculating relevant statistical measures, plotting, and drawing applicable conclusions using both numerical and categorical variables.

The analysis should be performed in two steps:

• Univariate analysis

• Multivariate analysis

Data Set:

Find Open Datasets and Machine Learning Projects | Kaggle Download Open Datasets on 1000s of Projects + Share Projects on One Platform. Explore Popular Topics Like Government, Sports, Medicine, Fintech, Food, and more. Flexible Data Ingestion.

Link: https://www.kaggle.com/datasets

Example:

One example of a dataset and its analysis is below:

1. UK Traffic and Accident Analysis: https://www.kaggle.com/code/imankity/uk-trafficand-accident-analysis/data

2. Univariate Plotting with Pandas: https://www.kaggle.com/code/residentmario/univariate-plotting-with-pandas

For multivariate datasets and their analysis on Kaggle:

1. Introduction to Multivariate Analysis: https://www.kaggle.com/code/sanikamal/introduction-to-multivariate-analysis#PlottingMultivariate-Data

2. Univariate Analysis: https://rkabacoff.github.io/datavis/Univariate.html

Submission Instructions:

1. The assignment must be performed in a Jupyter notebook either on Google Colab, directly on Kaggle, or locally on your computer.

2. Upload your notebook file to your GitHub account and share the link using the provided Google form. Ensure that your GitHub repository has public access.