HACKATHON 2019 - QUALIFICATION ROUND

Problem: Vietnam Macroeconomics Analysis

I.   Context

Macroeconomics is the study of the behaviour of the economy as a whole. It is very complicated, with many factors that influence it. These factors are analyzed with various economic indicators that tell us about the overall health of the economy.

In this round, we provide Vietnam macroeconomic data. Could you do exploratory data analysis on the dataset and give us some useful insights as described in part III?

II.   Dataset

The dataset includes only 1 file:

- `Vietnam-Macroeconomic-Data.xls`: with 3 rows and 40 columns,

The dataset has the following schema (see Table 1)

Field	Type	Description
Year	integer	current year of a data-point
GDP, current prices (Billions of U.S. dollars)	float	a Gross Domestic Product which refers to the total amount of goods and services a country produces
Unemployment rate (Percent)	float	a percentage of how many people from the available pool of labor (the labor force) are unable to find work

Table 1: List of features

III.   Requirements

The whole process includes multiple steps from preprocessing to evaluation. You are requested to do and report your work step by step:

1. **Data Imputation (20 points)**: There are several missing values in the given dataset, can you propose a way to fill in those missing values?

2. **Data Exploration (40 points)**: You need to explore the data to show some statistical and analysis information of given dataset, for example, is there any relationship between GDP and Unemployment rate. The insight information will be useful for further steps.

3. **Features Extraction (20 points)**: List of generated features from the given dataset. Provide your reasons and evidence.

4. **Discussion (20 points)**: As a task for data science when tackling a challenge, you are asked to provide your opinion on the quality of your current data for this challenge. Please provide your discussions and ideas for new, feasible features which should be collected to increase model performance.

IV.   Report

- The report should be presented in good format and English usage.

- The report must have a clear structure: introduction, approaches, evaluation, discussion, etc which are suitable for delivering your work efficiently.

V.   Submission

Please submit your report (as pdf format) and your code (in Python) that can reproduce your results.

<div align="center">

END OF PROBLEM

</div>