DATA SCIENCE-(BSCS-2) ASSIGNMENT TWO

INSTRUCTIONS [TOTAL: 25 MARKS]

DATA DESCRIPTION

(The data “Assignment 2.xlsx” shows the life-cycle savings of people from the years 1960-1970. The description of the variables are shown in the sheet “Descriptors” within the excel workbook)

Questions

1. Does the dataset have any anomalies such as missing data or outliers? [5 MARKS]
2. Calculate the descriptive statistics of all the variables and explain their implication? [5 MARKS]
3. Generate a graph showing the distribution of the population under 15 (“pop15”). [5 MARKS]
4. What does the relationship between the population over 75 (“pop75”) and the growth rate of dpi (“ddpi”) look like? Describe the relationship using graphs and statistics. [5 MARKS]
5. What is the relationship between aggregate personal saving (“sr”) and real rear disposal income (“dpi)”? [5 MARKS]

Submission Instructions

Please submit a Python notebook with the file extension ‘.ipynb’. Please use the naming convention for your work in this format: ACCESS NO_PROGRAM_COURSE & YEAR_COURSE UNIT NAME (e.g., A12345_BBA3_Data Science).