SINGAPORE UNIVERSITY OF SOCIAL SCIENCES

ANL252
Python for Data Analytics

Tutor-Marked Assignment

July 2022 Presentation

TUTOR-MARKED ASSIGNMENT (TMA)

This assignment is worth 18% of the final mark for ANL252 Python for Data Analytics.

The cut-off date for this assignment is **7 August 2022, 2355hrs.**

Up to 25 marks of penalties will be imposed for inappropriate or poor paraphrasing. For serious cases, they will be investigated by the examination department. More information on effective paraphrasing strategies can be found on https://academicguides.waldenu.edu/writingcenter/evidence/paraphrase/effective.

Note to Students:

You are to include the following particulars in your submission: Course Code, Title of the TMA, SUSS PI No., Your Name, and Submission Date.

Question 1

The given dataset comprises 250 staff information records in an organization. The data dictionary of this dataset is depicted in Appendix 1.

Part (a) is to be answered using MS Excel. Use Python and any related libraries for answering Parts (b) to (d).

The Python codes are to be expressed in text format and must be included, with the correct indentation(s), in the answers of Parts (b) to (d), in the report. **Screenshots of the codes are not permitted and will not be marked.**.

The corresponding Jupyter notebook must also be submitted.

(a) Analyse the given data, and provide ***two (2)*** charts and corresponding summarised tables, using MS Excel. Describe the insights obtained from this analysis.

(30 marks)

(b) Read the data and create the same ***two (2)*** charts and their corresponding summarised tables produced in Part (a), using Python.

(40 marks)

(c) Find the length of service of all the staff, using Python. For staff who has not left the organization, use 1st May 2022 as the LeftDate. State the minimum, maximum and average length of service, expressing in years, rounded to 1 decimal place. (1 year is taken as 365 days.)

(15 marks)

(d) Develop an interactive user input which allows the user to query if a particular person was/is a staff of the organization. One can assume the full name is entered by the user. The output should print the answer to the user's question if that person can be found in the organization's records. The interactive user input should also allow the user to continue to make another query, till the user chooses otherwise.

(15 marks)

APPENDIX 1 – DATA DICTIONARY

Variable	Description
ID	Staff Identifier
Staff	Name of Staff
BirthYear	Year of Birth
Gender	Gender (F:Female, M:Male)
Marital	Marital status
Citizenship	Citizenship status
Minority	Racial/Ethnic Minority
JoinDate	Staff's organization join date
LeftDate	Staff's organization left date
Unit	Business Unit
Salary	Salary
PerformanceScore	Performance Score (PIP: Performance Improvement Plan)
Survey	Engagement survey results (Scale of 1 to 5)
Satisfaction	Satisfaction Score (Scale of 1 to 5)
Absence	Number of days of absence in the previous year

---- END OF ASSIGNMENT ----