

**ANL252**

# End-of-Course Assessment – July Semester 2022

**Python for Data Analytics**

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Name: Lee Zhu

PI No. Q2110768

Date of Submission: 5 September 2022

# Question 1

Categorical variables are columns that have a limited number of possibilities of values. A numeric variable covers a range of values.

From the definitions provided in the ECA document, the following are categorical variables:

* Rating can only take 2 values (0 and 1), which corresponds to 'Good' and 'Bad' respectively. Hence, it's a categorical variable.
* Gender can only take 2 values (0 and 2), which corresponds to 'Male' and 'Female' respectively. Hence, it's a categorical variable.
* Education can only take 4 values (0, 1, 2, 3), which corresponds to 'Others', 'Postgraduate', 'Tertiary' and 'High School' respectively. Hence, it's a \* categorical variable.
* Marital can only take 3 values (0, 1 and 2), which corresponds to 'Others', 'Single' and 'Married' respectively. Hence, it's a categorical variable.
* Age can only take a finite ordered numeric values as categorical variables.
* S(n) is a mixed of fixed values (-1 and 0) corresponding to 'Prompt Payment' and 'Minimum Sum Payment', and finite ordered numeric variable. Hence, it's a categorical variable

The following are numeric variables:  
Limit, Balance, Income, R(n), B(n)

**Question 2 – Question 6**

Answered in Juypter Notebook