Assignment (40%)

Instructions:

1. This assignment will be done individually. Due date 27 Aug 2359hr.
2. Assignment Tasks:
    - Data Preparation, Exploration & Modeling
    - Submission of assignment report with screenshots to LMS
3. Module Assessment Weightage: 40%.

The **Hotel** data set contains 63 variables and over 52,000 observations. The variables in the data set are shown below with the appropriate roles and levels. Text mining has been performed on the text reviews and the keywords are extracted. Each keyword is a variable starting with the word "Concept_" and if the review contains the word, that matrix cell will indicate a 1, if the review does not contain the word, the matrix cell will indicate a 0.

Column Name	Measurement Level	Description
reviewID	Nominal	ID of each hotel review
custID	Nominal	ID of customer
custName	Nominal	Name of customer who wrote the review
Cust_travl_type	Nominal	Type of customers
Travel_custer_no	Nominal	Travel cluster of customers
custCountry	Nominal	Country where the reviewers are from
countryRegion	Nominal	Region where the reviewers are from
datePosted	Date	Date of review posted
Cust_review_title	Text	Title of the review
Star_rating	Numerical	Rating of the review
Comment	Text	Text comment of the hotel review
hotelName	Nominal	Name of hotel that the review was for
Source	-	Not relevant
Concept_bed	Binary	Reviews with the word "bed" will be indicated with 1. Reviews without the word "bed" will be indicated with 0.
Concept_bath room	Binary	Reviews with the word "bath room" will be indicated with 1. Reviews without the word "bath room" will be indicated with 0.
Concept_.... . . .		
Concept_xx	Binary	Reviews with the word "xx" will be indicated with 1. Reviews without the word "xx" will be indicated with 0.

Task 1 Association Rule Mining (20%)

In your first task, use Rapidminer to perform association of keywords starting with the words "Concept".

In your report, provide a screenshot of your workflow. State the parameters you have set.
Then identify Top 10 association rules with highest confidence and with minimum 3 itemset in each rule. Give your recommendations on what customers usually look for in a hotel stay.

Task 2 Clustering (20%)

In your second task, use SAS Viya to perform k-means clustering on the hotel data. Your objective is to identify the profile of customers that gives the **highest average ratings** for the hotel reviews. Use the following inputs for your clustering analysis:

- Cust_travel_type
- Star_rating
- Travel_custer_no
- countryRegion

Report Contents (Max 10 pages):

- Cover Page: Indicate your name and admin number in your report
- Print screen of each step:
  - including data preparation
  - modeling process, modeling techniques
  - model evaluation
- Give your recommendations based on the analyses.

Assessment Rubric (40%):

	Not Competent	Developing	Functional	Proficient	Advanced
Data Prep 5%	No evidence of data preparation	Unable to conduct proper data preparation	Able to conduct simple data preparation	Able to conduct proper data preparation	Able to conduct comprehensive data preparation.
Modelling 20%	No evidence of modelling techniques used	Unable to select correct modelling techniques and build models with basic test design and considerations	Able to select correct modelling techniques and build models with basic test design and considerations	Able to select correct modelling techniques and build models with appropriate test design and considerations.	Able to select correct modelling techniques and build models with comprehensive test design and considerations
Evaluation 10%	No evaluations or recommendations made	Provide inappropriate evaluation and recommendation	Provide simple evaluation and recommendation	Provide justified evaluation and recommendation	Provide comprehensive evaluation and recommendation
Report 5%	Report was not submitted	Report content is not well organized and hard to follow	Report content is organized and easy to follow	Report content is well-organized with ideas clearly	Report content is well-organized with ideas clearly conveyed