UGANDA CHRISTIAN UNIVERSITY
A Centre of Excellence in the Heart of Africa

FACULTY OF ENGINEERING, DESIGN AND TECHNOLOGY
DEPARTMENT OF COMPUTING AND TECHNOLOGY
EASTER 2024 SEMESTER **EXAMINATION**

PROGRAM: *[MASTER'S OF SCIENCE IN DATA SCIENCE]*

YEAR: 1 SEMESTER: 1

COURSE CODE: *[DSC8201]*

COURSE NAME: *[DATA SCIENCE LIFECYCLE]*

EXAMINATION TYPE: *PROJECT-BASED EXAM*

PROJECT DURATION: APRIL 2024

TIME ALLOWED: *[14 DAYS]*

Examination Instructions

1. The general Uganda Christian University examination guidelines and academic & financial policies apply to this examination. Violating any of the policies by the student automatically makes this examination attempt void, even if you have completed and submitted the answer booklet.
2. This exam consists of a project to be executed in [14] days.
    i. There are three parts to this project-based exam; Part A carries 40 Marks, Part B carries 40 Marks and Part C carries 20 Marks.
    ii. Assessment of the project shall be based on five milestones, evaluated during the duration of the project. Each milestone has assigned 20 marks as stated below.
    iii. The last milestone requires a presentation to be done in the presence of a diverse academic cohort a day after the submission deadline.
3. Every student has a responsibility to prove their contribution towards every milestone, and marks will be awarded to every student individually.

Project-based assessment guidelines (This project counts for 30% of your final grade)

S/N	Milestone Description	Maximum Marks
1	**MILESTONE ONE** [Ability to distinctly identify the different data/variable types and show data wrangling using python notebooks]	**20 %**
2	**MILESTONE TWO** [Ability to explore the datasets provided using python notebooks]	**20 %**
3	**MILESTONE THREE** [Ability to respond and make comprehensive interpretations that can analytically respond to the pertinent research questions using python notebooks]	**20 %**
4	**MILESTONE FOUR** [Ability to generate machine learning models that can make accurate predictions]	**20 %**
5	**MILESTONE FIVE** [Ability to compile the results stated from milestone 1 to 4 in a dashboard and powerpoint presentation to be made for a diverse audience]	**20 %**
	TOTAL MARKS	**100 %**

PART A: [TOTAL 40 MARKS]

DATASET DESCRIPTION (*Landslide_Factors_IRAN.csv*)

(The above dataset was sourced from Kaggle at Landslide Kaggle)

This dataset includes information and factors affecting landslides in Iran. The dataset contains over 4,000 landslide hazards in Iran, each of which has been characterized by a variety of factors that may contribute to landslide occurrence. These factors include natural factors such as slope, climate, tectonic activity, etc., and human factors such as land use, etc. These factors can have a significant impact on the likelihood of a landslide occurring. For example, slope is one of the most important natural factors affecting landslides. Note that landslides are more likely to occur on steep slopes.

1. Use the dataset to generate various insights on landslide occurrences in Iran. Generate and publish your results in a python notebook on your Kaggle or Github account.

PART B: [TOTAL 40 MARKS]

DATASET DESCRIPTION (*Traffic_Data.csv*)

(The above dataset provides a comprehensive look at traffic data in a hypothetical urban setting. It includes over 1.2 million records, each representing a unique snapshot of various factors influencing traffic conditions in six fictional cities.)

List of Variables:

City: The name of the city

Vehicle type: type of vehicle in use

Weather: Weather conditions at the time of data capture

Economic Condition: Economic state of the city at time of data capture

Day of Week: The day of the week at time of data capture

Hour of Day: The hour of the day at time of data capture

Speed: Recorded speed of the vehicle

Is Peak Hour: Indicator of whether data was captured during peak traffic hours (Not peak hour: 0, Peak hour: 1)

Random Event Occurred: Indicator of whether a random event (like accidents or road closures) occurred (No random event: 0, random event occurred: 1)

Energy consumption: An estimate of energy consumption based on vehicle type and speed

Traffic Density: The density of traffic at the time of data capture

Use the dataset to investigate a range of research objectives;

1. Understanding the traffic density on various days of the week in the urban environments. **[20 MARKS]**

2. Analyzing the impact of various factors like weather, economic conditions, and vehicle types on traffic density and energy consumption. **[10 MARKS]**

3. Developing and testing traffic density prediction algorithms for smart city solutions. **[10 MARKS]**

PART C: Presentation skills. **[TOTAL 20 MARKS]**

1. Present the results generated in part A in a dashboard to exhibit your interpretation of the data analytics. **[10 MARKS]**

2. Present the results generated in part B using a powerpoint presentation targeted at informing traffic officers, policymakers, traffic users and other relevant stakeholders in Uganda. **[10 MARKS]**

~END OF PROJECT EXAM~