**CSC 2304 Intelligent Systems (Section 1)**
**Semester 2 2020/21**
**Group Assignment 4: Data Science and Power BI Essentials [40 MARKS]**
**Due : 13 JUNE 2021 (11.59 pm)**

**INSTRUCTIONS**
This is a group assignment. Each group shall distribute the tasks among group members. Submit your work in pdf format.

## A. Data Science [20 MARKS]

**Read the article** and **watch the video** on Data Science uploaded in italeem. Prepare a summary based on the following questions from the video.

1. A summarized description of the **FOUR (4)** different viewpoints of Data Science. (*Note: Please summarize your thoughts with your own words and not just cut & paste the statements). **[4 marks]**

   a.
   b.
   c.
   d.

2. The **goals** of Data Science are to use data analytic thinking for the purpose of ____ **[3 marks]**

   a.
   b.
   c.

3. Data Science involves _____ **[5 marks]**

   a.
   b.
   c.
   d.
   e.

4. The Data Science process includes the following activities : **[3 marks]**

   a.
   b.
   c.

5. **Draw the diagram** that shows the Data Science process (*Note: Please manually draw using a tool or by hand and not just cut & paste). **[3 marks]**

6. In ONE (1) single paragraph, write a **short summary** of no more than 4 sentences on your own understanding about Data Science. **[2 marks]**

## B. Power Business Intelligence (BI) [20 MARKS]

1. Download and install PowerBI on your laptop/desktop or signup at https://app.powerBI.com

2. Import ('Get Data') the **hurricane-Pacific_2013.xlsx** dataset which has been uploaded in italeem.

3. Play around and explore with Power BI using the lecture slides on "Data Visualization" as a guideline.

4. Answer the following questions:

   Legend for Status:

   - WV - Tropical Wave
   - TD - Tropical Depression
   - TS - Tropical Storm
   - HU - Hurricane
   - EX - Extratropical cyclone
   - SD - Subtropical depression (winds <34 kt)
   - SS - Subtropical storm (winds >34 kt)
   - LO - A low pressure system not fitting any of above descriptions
   - DB - non-tropical Disturbance not having a closed circulation

   a. Find the correlation between the following variables and **capture the screenshots** of your graph plots. Report on the type of relationship you observe and indicate outliers if any: **[4 marks]**
      i. Maximum Wind vs Minimum Pressure
      ii. Low Wind NE vs Moderate Win SE
      iii. Minimum Pressure vs Time
      iv. High Wind SE vs Low Wind NW

   b. Find the correlation between the 2 event conditions, Tropical Storm (TS) and Hurricane (HU) with Maximum Wind and Minimum Pressure. Print the screenshot of the display. When do Maximum Wind and Minimum Pressure reach the peak? **[4 marks]**

   c. Find the **correlation** between 3 types of storms: **BARBARA, GIL** and **HENRIETTE with Maximum Wind** and **Minimum Pressure**. Print the screenshot of the display. **[6 marks]**

   d. **Display the bar graph** showing day by day information of the dataset for all fields. **[4 marks]**

i. On **which day of the month** is the Maximum Wind the lowest and in **which type of storm**?

ii. At **which time of the month** is the Minimum Pressure the highest and in **which type of storm**?

e. You are the **Director of the National Hurricane Center**. As the main stakeholder, suggest **TWO (2)** kinds of insights related to **hurricane forecasts** that will be valuable and critical to your department in the long run based on the available data asset (hurricane-Pacific_2015.csv). **[2 marks]**