

CONFIDENTIAL



UNIVERSITI TUN HUSSEIN ONN MALAYSIA

**FINAL EXAMINATION
SEMESTER I
SESSION 2021/2022**

COURSE NAME : DATA SCIENCE AND APPLICATIONS
COURSE CODE : BFS 41203
PROGRAMME CODE : BFF
EXAMINATION DATE : JANUARY / FEBRUARY 2022
DURATION : 3 HOURS
INSTRUCTION : 1. ANSWER **ALL** QUESTIONS.
2. THIS FINAL EXAMINATION IS AN **ONLINE** ASSESSMENT AND CONDUCTED VIA **CLOSE BOOK**.

THIS QUESTION PAPER CONSISTS OF FIVE (5) PAGES

TERBUKA

CONFIDENTIAL

- Q1** (a) Digital transformation is driven by advance computational power of processor to solve complex problem, with capability of cognitive judgement, such as artificial intelligence technology. Using appropriate examples, justify **SIX (6)** key factors that make data science and big data as part of components in digital transformation. (18 marks)
- (b) List and elaborate main analytics spectrums in Data Science with relevant examples in Civil Engineering. (7 marks)
- Q2** (a) The **Table Q2** shows the tabular data that stored in `data.csv` file.
- (i) Name the appropriate node to import the dataset into KNIME's editor workspace. (2 marks)
- (ii) State the node to calculate the sum of x_3 for each class in variable x_2 . Draw the data table resulted when executing the described node. (5 marks)
- (iii) Name the node to transform all values in x_3 into **ONE (1)** decimal place. Specify the function to perform the action. (3 marks)
- (b) Using data in **Table Q2**, evaluate the descriptive information of the data using procedures required in exploratory data analysis. (15 marks)
- Q3** (a) Discuss **FIVE (5)** major differences between supervised and un-supervised machine learning model. Use relevant example for each point. (10 marks)
- (b) The confusion matrix of classification model is given in **Table Q3**.
- (i) Compute the precision, recall and accuracy of the model. (6 marks)
- (ii) Based on the calculated metrics in **Q3b(i)**, write your conclusion about the model. Propose the strategy to improve the performance of the model. (9 marks)

TERBUKA

- Q4** (a) A huge data from a weather recording device was obtained to study weather patterns in Malaysia. There were two datasets that was gathered from two different locations of weather stations. The first dataset of a csv file comes from a weather station in Pahang and the other comes from an Excel file of Johor.

Both datasets have the same columns. You need to combine the data from both the datasets to create an insightful dashboard in Power BI. Explain in detail the process on which command should you use in order to have both these data presented into a single dashboard.

(7 marks)

- (b) Data visualization using the concept of business intelligence is an essential feature that is used in the whole process of data science. Provide detail explanation on what are the main outcomes that is obtained from the concept of business intelligence.

(6 marks)

- (c) **Figure Q4** shows a dashboard representing the distribution of energy sources for a single year. Based on the dashboard presented, discuss what insights that you can obtain from the trend and patterns of the effects resulting from six type of energy sources recorded.

(12 marks)

– END OF QUESTIONS –

TERBUKA

SEMESTER / SESSION : SEM I 2021/2022 PROGRAMME CODE : BFF
COURSE NAME : DATA SCIENCE AND APPLICATIONS COURSE CODE : BFS 41203

Table Q2

ID	x1	x2	x3
A001		10 A	10.98
A002		20 B	7.09
A003		70 A	7.88
A004		80 B	3.87
A005		30 B	10.1
A006		10 C	8.89

Table Q3

	Predicted Class 1	Predicted Class 2
Actual Class 1	891	21
Actual Class 2	39	589

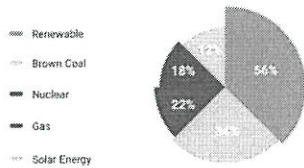
TERBUKA

SEMESTER / SESSION : SEM I 2021/2022
COURSE NAME : DATA SCIENCE AND APPLICATIONS

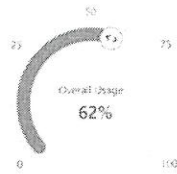
PROGRAMME CODE : BFF
COURSE CODE : BFS 41203

Dashboard

Energy Sources



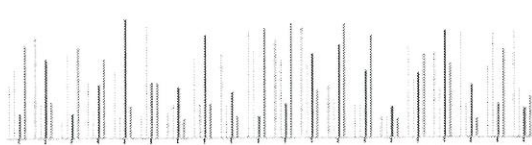
Overall Energy Usage



Trend Analysis

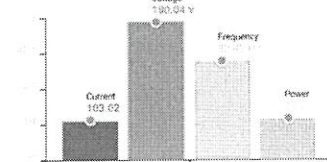


Average Energy Consumption by Sector

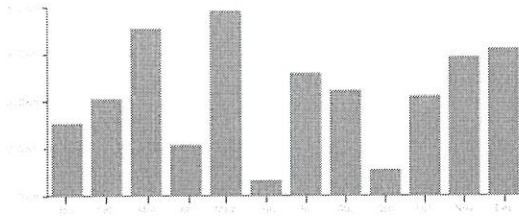


Industry	Value
Industry	\$360
Residential	\$76
Private	\$17.05
Government	\$60

Power Quality Analysis



Monthly Energy Consumption



Monthly Production Costs (per wh)

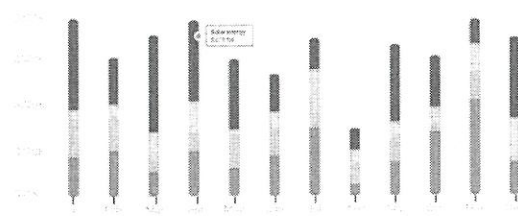


Figure Q4

TERBUKA