

UNIVERSITI MALAYA
UNIVERSITY OF MALAYA

PEPERIKSAAN IJAZAH SARJANA SAINS DATA
EXAMINATION FOR THE DEGREE OF MASTER OF DATA SCIENCE
SESI AKADEMIK 2022/2023 : SEMESTER I
ACADEMIC SESSION 2022/2023 : SEMESTER I

WQD7001 : Prinsip Sains Data
Principles of Data Science

Januari 2023
January 2023

Tempoh Masa: 48 jam
Duration: 48 hours

ARAHAN KEPADA CALON:
INSTRUCTIONS TO CANDIDATES:

Calon dikehendaki menyediakan dan menghantar SEMUA penilaian alternatif berikut:
Candidate needs to prepare and submit ALL Alternative Assessments as follows:

1. Alternative Assessment 1: (submit via SPECTRUM and use your Name and Matric Number as the file name, for example: Anne171100.pdf)

(Kertas soalan ini mengandungi 2 Soalan dalam 4 mukasurat yang bercetak)
(This question paper consists of 2 Questions in 4 printed pages)

ALTERNATIVE ASSESSMENT ONE (AA1)

Question 1. Examine the Table 1.0 below. Answer the following questions.

Table 1.0 : Sales of Electronic Products for the year of 2021

Products	Months	Item Sold
Ipad	Aug	45
Ipad	Sep	56
Ipad	Oct	39
Ipad	Nov	48
Ipad	Dec	44
Smart Phone	Jan	100
Smart Phone	Feb	60
Smart Phone	Mar	95
Smart Phone	April	75
Smart Phone	May	100
Smart Phone	June	60
Smart Phone	July	95
Smart Phone	Aug	75
Ipad	Aug	45
Playstation 5	June	65
Playstation 5	July	85
Playstation 5	Aug	90
Playstation 5	Sep	80
Playstation 5	Oct	75

Products	Months	Item Sold
Playstation 5	Nov	65
Playstation 5	Dec	80
Smart Phone	Sep	109
Ipad	Jan	50
Ipad	Feb	55
Ipad	Mar	51
Ipad	April	40
Ipad	May	450
Ipad	June	30
Ipad	July	39
Smart Phone	Oct	80
Smart Phone	Nov	109
Smart Phone	Dec	75
Playstation 5	Jan	180
Playstation 5	Feb	199
Playstation 5	Mar	93
Playstation 5	April	80
Playstation 5	May	70

- State THREE sharp questions so that an early discovery can be made based on the given dataset in Table 1.0.
(3 marks)
- Based on the questions answered above, what are the observations made? Support your observation with an appropriate graph.
(5 marks)
- Provide TWO plausible assumptions in terms of the most profitable and less profitable products based on Table 1.0
(2 marks)

Question 2.

You have these important questions to answer based on the data of Covid-19 issues.

Q1	Which FIVE countries of the world are the earliest hit by the Covid-19 pandemic?
Q2	How many categories of patients do we have from the above Q1 countries?
Q3	What are the most expensive prices of the vaccines?

- a. As a Data Scientist, create a relevant dataset by using any imaginary data and variables. How are you going to deal with the noisy data? Justify your answer and provide the relevant references.
- (5 marks)

- b. Explain FOUR observations that can be made based on the following example findings. How will the observed information used by the authority for decision making purpose?

Country	Crisis	Covid Impact (losses - USD)	Intervention
Malaysia	Insufficient funding for Food	USD20M	WHO
Indonesia	Lost of jobs	USD30M	Local Gov
Thailand	Flash flood – increase covid cases	USD50M	WHO and International NGO

(5 marks)