

**FINAL
ALTERNATIVE ASSESSMENT**

(COVER PAGE)

Session : APRIL 2022

Programme : DIPLOMA IN COMPUTER SCIENCE (DCS)

Course : DCS2104: DATA VISUALIZATION

Date of Examination : 05 August 2022

Time : 04.00pm- 06.30pm Reading Time : Nil

Duration : 02 hours 30 minutes

Special Instructions :

This paper consists of **FIVE (5)** questions. Answer **ALL** questions in the answer booklet provided.

Material permitted : Non-Programmable Scientific Calculator

Materials provided : Nil

Examiner(s) : Nadhrah Abdul Hadi
Yogeswari Suppiah

Chief Moderator : Siti Hajar Khairuddin

This paper consists of 5 printed pages, including the cover page

DIPLOMA IN COMPUTER SCIENCE PROGRAMME (DCS)
DCS2104: DATA VISUALIZATION
FINAL ALTERNATIVE ASSESSMENT: APRIL 2022 SESSION

Instructions: This paper consists of **FOUR (4)** questions. Answer **ALL FOUR (4)** questions in the answer booklet provided. All questions carry equal marks.

Question 1

- (a) Write a program for the following statements:
- i. Divide the observations of **sashelp.heart** into three data sets, **work.highchol**, **work.lowchol**, and **work.misschol**.
 - ii. Only observations with **cholesterol** below 200 should be in the **work.lowchol** data set.
 - iii. Only Observations with **cholesterol** that is 200 and above should be in the **work.highchol** data set.
 - iv. Observations with missing cholesterol values should only be in the **work.misschol** data set.

(15 marks)

- (b) Below is **class_birthdate** table from library **pg1**. Write SQL statement that extract students who are more than 14 years old.

(5 marks)

Name	Age	HeightCM	Birthdate
Alfred	14	175.3	26OCT2004
Alice	13	143.5	16NOV2005
Barbara	13	165.9	15JAN2005
Carol	14	159.5	04JUL2004
Henry	14	161.3	01DEC2004
James	12	145.5	15JUN2006
Jane	12	151.9	13MAR2006
Janet	15	158.8	02APR2003
Jeffrey	13	158.8	26APR2005
John	12	149.9	23AUG2006
Joyce	11	130.3	03JAN2007
Judy	14	163.3	05DEC2004
Louise	12	143.0	08AUG2006
Mary	15	168.9	26MAR2003
Philip	16	182.9	21NOV2002
Robert	12	164.6	06JUN2006
Ronald	15	170.2	14OCT2003
Thomas	11	146.1	18MAR2007
William	15	168.9	28DEC2003

- (c) Write a program for the following statements:
- i. Write a PROC SORT step to sort **pg1.eu_occ** and create an output table named **countrylist**.
 - ii. Remove duplicate key values.
 - iii. Sort by **Geo** and then **Country**.

(5 marks)

Question 2

- (a) Modify the following program to use **SELECT** and **WHEN** statements instead of **IF-THEN** statements.

When **Type** is *NP*, create a new column named **ParkType** that is equal to **Park**, and write the row to the **parks** table. When **Type** is *NM*, assign **ParkType** as Monument and write the row to the **monuments** table.

```
data parks monuments;
  set pg1.np_summary;
  where type in ('NM', 'NP');
  Campers=sum(OtherCamping, TentCampers,
              RVCampers, BackcountryCampers);
  format Campers commal7.;
  length ParkType $ 8;
  if type='NP' then do;
    ParkType='Park';
    output parks;
  end;
  else do;
    ParkType='Monument';
    output monuments;
  end;
```

```
keep Reg ParkName DayVisits OtherLodging  
Campers ParkType;  
run;
```

(15 marks)

- (b) Write a program that creates one Excel workbook named **annual** that contains two worksheets, and it uses two data sets, **qtr1_2011** and **qtr2_2011**, as input data.

(10 marks)

Question 3

The **pg1.np_summary** table contains public use statistics from the National Park Service. Use the **PRINT, MEANS, UNIVARIATE, and FREQ** procedures to explore the data.

- (a) Write PROC PRINT statement to list the first 20 observations in **pg1.np_summary**.

(5 marks)

- (b) Add a VAR statement to include only the following variables: **Reg, Type, ParkName, DayVisits, TentCampers,** and **RVCampers**.

(5 marks)

- (c) Write statement to calculate **Summary Statistics** for **DayVisits** **TentCampers** and **RVCampers**.

(5 marks)

(d) Write statement to calculate **Univariate Statistics** for **DayVisits** .
(5 marks)

(e) Write statement to list **unique** values and **frequency** counts for **Reg** table.
(5 marks)

Question 4

(a) Write a statement to generate a report of the **class_update** from a library called "PG1".
(5 marks)

(b) Write a statement to sort the data by **Age**.
(5 marks)

(c) Write a statement to create two-way frequency report of attribute "**height**" and "**weight**" from table "class_update".
(5 marks)

(d) Create a comma-delimited file named **class_update.csv**. Use &outpath to substitute the path of the output folder.
(5 marks)

(e) In SAS **Output Delivery System** terminology, each of the output formats are called destination. Write a steps in exporting output to PowerPoint format and Microsoft Word.
(5 marks)

~ **THE END** ~