UNIVERSITY OF COLOMBO

FACULTY OF GRADUATE STUDIES

Master of Regional Development and Planning 2019/2020

Final Examination (Semester -II)

MRDP 6205: Computer Application in Data Analysis

Answer any FOUR (04) questions

Time: TWO (02) Hours

---

01.   Answer **all** sections.

i.   Calculate the 'Mean' of following data set.
        X= {12, 45, 13, 11, 19, 34, 22, 17, 29, 18, 33, 36}

(01 Mark)

ii.   Explain the difference/s between 'Data' and 'Information' using a suitable example

(03 Marks)

iii.   Name **four (04)** characteristics of valuable information.

(04 Marks)

iv.   Name **one (01)** advantage and **one (01)** disadvantage of Coefficient of Variation (CV).

(04 Marks)

v.   Explain how computer applications are used in the following domains to increase the efficiency of the processes as well as the effectiveness of the procedures.
   a.   Health sector.
   b.   Government sector.

(06 Marks)

vi.   "*Statistics has a lot of potential. Both people and organizations base some of their most important decisions on statistics. Some would argue that numbers do not lie, but statistics can be often misleading.*" Do you agree with this statement? Justify your answer stating relevant examples.

(07 Marks)

02.   Answer **all** sections.

i.   Calculate the 'Range' of the following data set.
Y= {54, 82, 80, 48, 76, 98, 37, 66, 42, 69, 94, 77, 96, 100, 70, 40, 86}

(01 Mark)

ii.   All data that will ever be analyzed could be divided into two categories, Qualitative and Quantitative. Explain the key differences between these two basic data types.

(03 Marks)

iii.   Indicate whether the following items are Qualitative or Quantitative measurements.

Sr No	Description	Quantitative or Qualitative
01	Gender	
02	Distance from home to school	
03	Income of a person	
04	Smell	
05	Number of errors in a document	
06	Religion of a person	
07	Age of a person	
08	Eye Color	

(04 Marks)

iv.   Name **four (04)** methods used to capture Qualitative data.

(04 Marks)

v.   **"Discrete Data are finite or countable and infinite while Continuous Data are uncountable and infinite." E**xplain the terms 'Countable and Infinite' and 'Uncountable and Infinite' given in the statement using suitable examples.

(06   Marks)

vi.   **The four levels of data namely, Nominal, Ordinal, Interval and Ratio are arranged as per the increasing order of information.** Prove that each level consists of more information than its predecessor (when moving from nominal to ratio) citing examples from each data level.

(07 Marks)

03.   Answer **all** sections.

i.   What is the measurement which indicates how tightly the values in a dataset are clustered around the mean value?

(01 Mark)

ii.   Name the **two (02)** basic chart types used in 'Dual-Axis Charts'.

(03 Mark)

iii.   Explain the usage of 'Likert Scales' in surveys and questionnaires.

(04 Marks)

iv.   Explain the importance of 'Data Cleaning' phase in 'Data Analysis' process.

(05 Marks)

v.   Large set of numbers can make us go cross-eyed. It can be difficult to find the significance behind rows of data. Illustrations, charts, descriptive text and engaging designs can be adopted to minimize this difficulty. Explain how 'Data Visualization' helps us to achieve following objectives.

    a.   Identify trends and outliers.
    b.   To highlight important points.

(06 Marks)

vi.   Briefly describe about following chart types. Use illustrations where necessary.

    a.   Scatter plot.
    b.   Pie chart.

(06 Marks)

04.   Answer **all** sections.

i.   What is meant by 'UI'?

(01 Mark)

ii.   Explain the terms 'Entity' and 'Attribute' using a suitable example.

(03   Marks)

iii. Write down the Microsoft Excel keyboard shortcuts used to achieve following objectives.

    a. Insert a new line within a cell.
    b. Go to Format Cells dialog box.
    c. Move to the top of current column.
    d. Create a table.

(04 Marks)

iv. Write short notes on the following two sampling methods.

    a. Systematic sampling.
    b. Snowball sampling.

(05 Marks)

v. Briefly describe the **three (03)** cell referencing styles used in Microsoft Excel. Provide suitable examples where necessary.

(06 Marks)

vi. Compare the differences between Microsoft Excel and Microsoft Access software applications.

(06 Marks)

05. Answer **all** sections.

i. There are two views available in SPSS interface where one is 'Data View'. Name the other view?

(01 Mark)

ii. A 100-watt light bulb has an average brightness of 1650 lumens, with a standard deviation of 60 lumens. What is the probability that a 100-watt light bulb will have a brightness of more than 1740 lumens?
Hint: You can use the 'Standard Normal Distribution' table at the end of the paper.

(03 Marks)

iii. 'Data Collection' is an important step in 'Data Analysis' process. Explain how current technology can be leveraged to increase efficiency of 'Data Collection' phase. Provide suitable examples where necessary.

(04 Marks)

iv.   Write down the steps to be followed in order to obtain a 'Case Processing Summary'in SPSS? What are the details available in above type of summary?

(05 Marks)

v.   Briefly describe about positive correlation and negative correlation using suitable examples. You may use diagrams if necessary.

(05 Marks)

vi.   Regression analysis is a powerful statistical method that allows to examine the relationship between two or more variables of interest. Explain how we could utilize regression analysis for business cases using a suitable example.

(07   Marks)

06.   Answer **all** sections.

i.   The 'Mean' of 6, 8, x + 2, 10, 2x - 1, and 2 is 9. Find the value of x.

(01 Mark)

ii.   Name **three (03)** chart types which can be used to show the composition of a data set.

(03 Marks)

iii.   Name **three (03)** facts that needs to be considered when preparing a questionnaire.

(06 Marks)

iv.   Statistics is not only for business; we can take advantage of statistics when performing our day-to-day tasks. Briefly explain about **two (02)** real life scenarios to demonstrate the practical usage of statistics.

(06 Marks)

v.   Anne scored 55 in an English test for which the mean was 50 and the standard deviation was 6. She scored 64 in a Mathematics test for which the mean was 59 and the standard deviation 9.

   a.   Calculate her standardized score (Z-Score) for each subject.
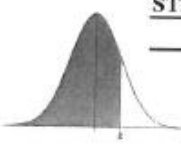   b.   In which subject did Anne perform better, given that the classes are of equal   ability? Justify your answer.
   c.   What mark would Anne has had to score in Mathematics for her performance   to   be equivalent to that for English?

(09 Marks)

****

Standard Normal Distribution Table

STANDARD NORMAL DISTRIBUTION: Table Values Represent AREA to the LEFT of the Z score.

Z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	.50000	.50399	.50798	.51197	.51595	.51994	.52392	.52790	.53188	.53586
0.1	.53983	.54380	.54776	.55172	.55567	.55962	.56356	.56749	.57142	.57535
0.2	.57926	.58317	.58706	.59095	.59483	.59871	.60257	.60642	.61026	.61409
0.3	.61791	.62172	.62552	.62930	.63307	.63683	.64058	.64431	.64803	.65173
0.4	.65542	.65910	.66276	.66640	.67003	.67364	.67724	.68082	.68439	.68793
0.5	.69146	.69497	.69847	.70194	.70540	.70884	.71226	.71566	.71904	.72240
0.6	.72575	.72907	.73237	.73565	.73891	.74215	.74537	.74857	.75175	.75490
0.7	.75804	.76115	.76424	.76730	.77035	.77337	.77637	.77935	.78230	.78524
0.8	.78814	.79103	.79389	.79673	.79955	.80234	.80511	.80785	.81057	.81327
0.9	.81594	.81859	.82121	.82381	.82639	.82894	.83147	.83398	.83646	.83891
1.0	.84134	.84375	.84614	.84849	.85083	.85314	.85543	.85769	.85993	.86214
1.1	.86433	.86650	.86864	.87076	.87286	.87493	.87698	.87900	.88100	.88298
1.2	.88493	.88686	.88877	.89065	.89251	.89435	.89617	.89796	.89973	.90147
1.3	.90320	.90490	.90658	.90824	.90988	.91149	.91309	.91466	.91621	.91774
1.4	.91924	.92073	.92220	.92364	.92507	.92647	.92785	.92922	.93056	.93189
1.5	.93319	.93448	.93574	.93699	.93822	.93943	.94062	.94179	.94295	.94408
1.6	.94520	.94630	.94738	.94845	.94950	.95053	.95154	.95254	.95352	.95449
1.7	.95543	.95637	.95728	.95818	.95907	.95994	.96080	.96164	.96246	.96327
1.8	.96407	.96485	.96562	.96638	.96712	.96784	.96856	.96926	.96995	.97062
1.9	.97128	.97193	.97257	.97320	.97381	.97441	.97500	.97558	.97615	.97670
2.0	.97725	.97778	.97831	.97882	.97932	.97982	.98030	.98077	.98124	.98169
2.1	.98214	.98257	.98300	.98341	.98382	.98422	.98461	.98500	.98537	.98574
2.2	.98610	.98645	.98679	.98713	.98745	.98778	.98809	.98840	.98870	.98899
2.3	.98928	.98956	.98983	.99010	.99036	.99061	.99086	.99111	.99134	.99158
2.4	.99180	.99202	.99224	.99245	.99266	.99286	.99305	.99324	.99343	.99361
2.5	.99379	.99396	.99413	.99430	.99446	.99461	.99477	.99492	.99506	.99520
2.6	.99534	.99547	.99560	.99573	.99585	.99598	.99609	.99621	.99632	.99643
2.7	.99653	.99664	.99674	.99683	.99693	.99702	.99711	.99720	.99728	.99736

The formula for calculating a z-score is $z = (x-\mu)/\sigma$, where **x** is the raw score, **μ** is the population mean, and **σ** is the population standard deviation.

Table 1