OLLSCOIL NA GAILLIMHE
UNIVERSITY OF GALWAY

Semester 2 Examinations 2023-2024

Course Instance Code(s)	1CSD1, 1MAI1, 1MAO3,1DDV1, 1BMG1, 1GDS1, 1SPD1, 1SPE1, 1ACS1, 3SPS1
Exam(s)	MSc in Computer Science (Data Analytics), MSc in Computer Science (Artificial Intelligence), MSc in Computer Science (Artificial Intelligence) – Online, MSc in Computer Science Adaptive Cybersecurity, PDip. in Data Analytics and Visualisation, MSc (Biomedical Genomics), MSc (Genomics Data Science), Structured PhD(Medicine), Structured PhD(Engineering), Structured Ph.D. (Science)
Module Code(s)	CT5100, CT5136
Module(s)	Data Visualisation, Data Visualisation Online
Paper No.	1
External Examiner(s)	Dr John Woodward
Internal Examiner(s)	Professor Michael G Madden Dr. Conor Hayes*
Instructions:	Answer ALL questions in SECTION 1. Answer TWO questions out of FOUR in SECTION 2. Use a separate answer book for each section.
Duration	2 hours
No. of Pages	7
Discipline(s)	School of Computer Science

Requirements:

Release in Exam Venue	No[☑]	Yes [ ]
MCQ Answer sheet	No [☑]	Yes [ ]
Handout	No [☑]	Yes [ ]
Formulae & Tables*	No [☑]	Yes [ ]
Cambridge Tables 2nd Edition**	No [☑]	Yes [ ]
Graph Paper*** A4 Graph Paper 1mm 0.1cm Squared (Standard)	No [ ]	Yes [☑]
Other Materials	No [☑]	Yes [ ]
Graphic material in colour	No [ ]	Yes [☑]

End of requirements.

PTO

CT5100 Data Visualisation

Exam Duration: 2 Hours

Section 1 60 marks

This section consists of 15 short questions. There are FOUR marks for each short question. Answer ALL questions. *To get marks for each question, you must explain each answer briefly and clearly.*

1.1 Provide a concise definition of data visualisation based on what you have learned in this module. [4]

1.2 List six perceptual features that the human visual system uses to determine if a pattern exists in a scene. [4]

1.3 Briefly explain why Cleveland and McGill's data visualisation paper from 1985 is considered so important. [4]

1.4 Explain the importance of pre-attentive visual perception to data visualisation. [4]

1.5 Explain how the context in which your visualisation will be presented will affect your visualisation design decisions. [4]

1.6 List the key questions you should answer *before* starting a data visualisation solution. [4]

1.7 Compare the primary visual features used to encode quantity in a bar charts vs a Cleveland dot plot? [4]

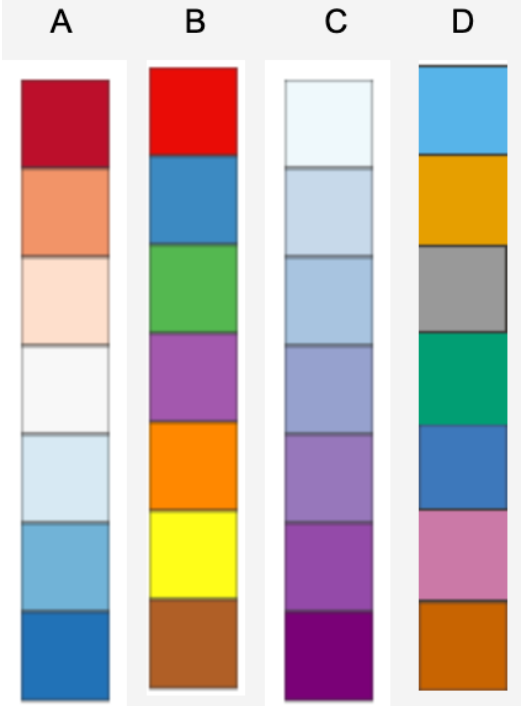
1.8 Explain the difference between a dot plot and a scatter plot. [4]

1.9 You are required to create a scatter plot where each point is one of two colours. The scatterplot will have a white background. The colours are given by their rgb values. Which pair of colours from the options below do you select so that they are accessible to someone with colour vision deficiency (CVD)? [4]

I. rgb(255,255,255) and rgb(255,0,0)
II. rgb(255,255,255) and rgb(0,0,0)
III. rgb(255,0,00 and rgb(0,255,0)
IV. rgb(255,0,0) and rgb(0,0,255)

PTO

1.10 Figure 2. Shows four colour palettes and lists 4 measurement scales. Match each palette (A,B,C,D) to an appropriate measurement scale type [4]

Palettes	**Measurement scale type**
 **Figure 2**	I. nominal (Colourblind friendly) II. nominal III. continuous (diverging) IV. ordinal

1.11 You have created a diverging bar chart to present time series data. The data values in your visualisation represent the difference from the mean value. In your plot, you have coloured the bars representing positive values with a green fill and the bars representing negative values with a red colour. Explain whether this plot readable to a person with CVD. [4]

PTO

1.12 Identify the different visual scales used to represent data values in Figure 3. [4]

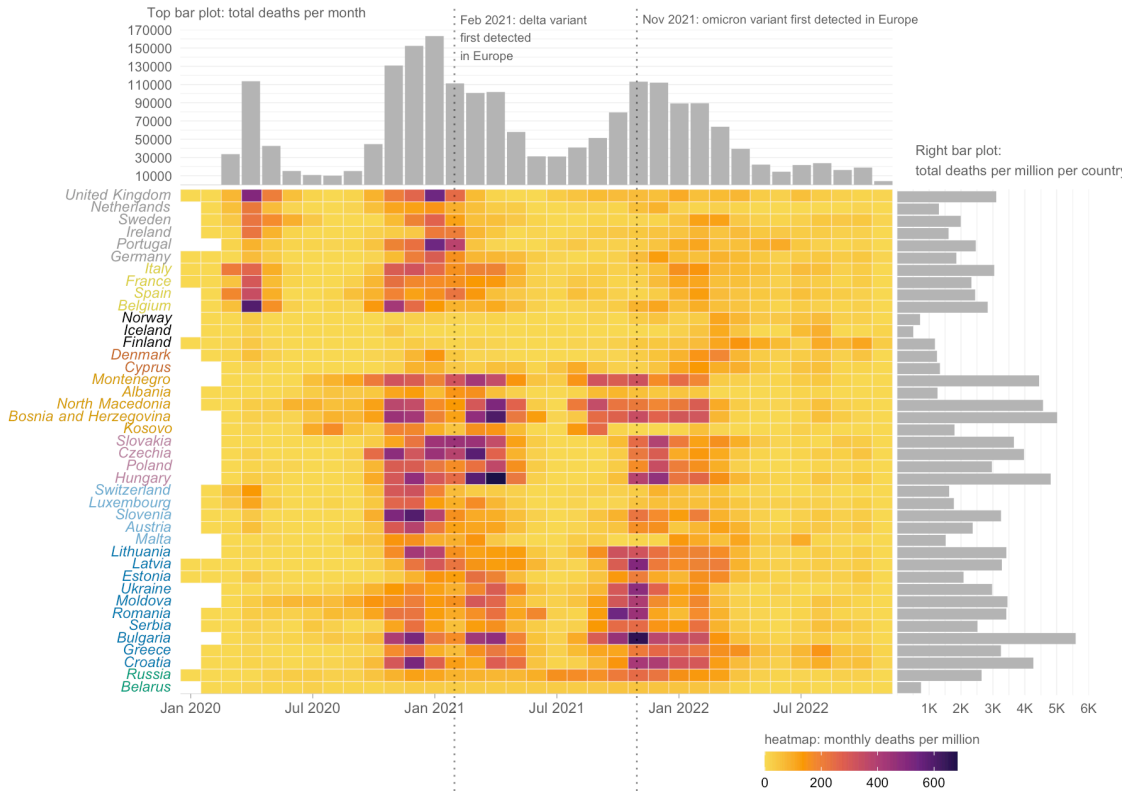


Figure 3

1.13 Your data consists of a table with 50 rows and three columns containing continuous quantitative values such as that show by Table 1. You want to show the *relationship* between variable A and variable B and the size of variable R. What type of plot would you use. Explain why. [4]

A	B	R
7.2	9.3	20.1
3.4	7.0	18.4
2.2	5.1	23.5
10.5	11.4	11.3
…	…	…

1.14 What is overplotting? Identify two techniques for handling overplotting. [4]

1.15 Explain the difference between exploratory and explanatory data visualisation [4]

PTO FOR SECTION 2

Section 2 40 marks

Answer any TWO questions out of FOUR. 20 marks for each question

2.1 **[20]**

I. Explain the type of information a visualisation of a data distribution might communicate. [4]

II. List *three* visualisation types suitable for visualising a data distribution. At least one visualisation type should be suitable for a general audience and another more suited to a scientific audience. Briefly explain your answer. [6]

III. An air quality sensor captures ozone values every hour. You are required to produce a visualisation that allows for the comparison of the *distribution* of ozone value for each of three months: May, June and July. Note that this is not a time-series question. Produce a drawing of a suitable visualisation to represent this data. You should explain the visualisation design decisions you have made in your drawing. [10]

2.2 Figure 4 shows a visualisation of the number of people in four meeting rooms labelled A,B,C and D in a single building on one particular day. **[20]**

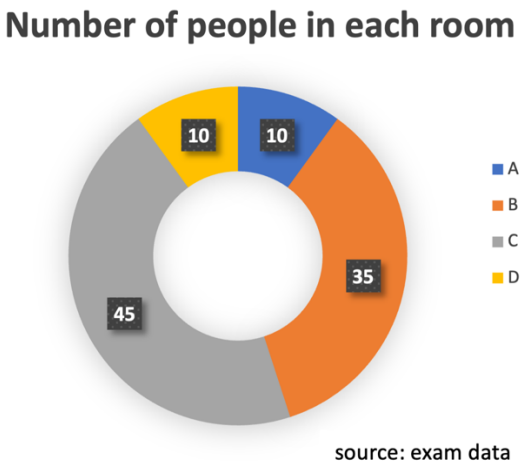


Figure 4.

I. Write a brief critique of this visualisation based on your understanding of Cleveland and McGill's insights on visual perception. [4]

II. List three *better* visualisation alternatives to the plot in Figure 4. Explain why each is better than the plot in Figure 4 [6]

III. Let's assume that you are required to show the number of people occupying these rooms for every day in week. Produce a drawing of a suitable visualisation to represent this data. You should explain the visualisation design decisions you have made in your drawing. [10]

PTO

2.3 Table 2 shows a dataset that gives race times and rankings of 10 of the fastest female 100 metre sprinters in Ireland for the years 2017, 2018 and 2019. The times are given in seconds. **[20]**

Name	2017_time	2017_rank	2018_time	2018_rank	2019_time	2019_rank
Ciara Neville	11.52	2	11.54	4	11.33	1
Gina Akpe-Moses	11.56	3	11.46	2	11.45	2
Phil Healy	11.66	6	11.28	1	11.51	3
Molly Scott	11.89	12	11.76	9	11.58	4
Joan Healy	11.63	4	11.57	6	11.59	5
Patience Jumbo-Gula	11.74	9	11.51	3	11.61	6
Amy Foster	11.42	1	11.54	4	11.64	7
Rhasidat Adeleke	11.98	15	11.6	7	11.69	8
Niamh Whelan	11.63	4	11.76	9	11.74	9
Sarah Lavin	12.09	18			11.74	9

Table 2.

I. What are the scale types of the ‘Name’, ‘2017_rank’ and ‘2017_time variables shown in Table 2. Explain your answer. [4]

II. You are asked to produce a visualisation that emphasises the change in rankings for these 10 athletes over the three years in the data set. List two valid approaches and explain why each one may be an appropriate solution to the visualisation requirements. [6]

III. Produce a sketch of a suitable visualisation to represent this data. To get full marks, your sketch should include the visual elements expected in this type of visualisation. You should explain the visualisation design decisions you have made in your drawing. [10]

PTO

2.4 You've just been hired as a data visualisation analyst for a large firm. Your boss is due to present Figure 5 (designed by your predecessor) to an important meeting tomorrow. **[20]**

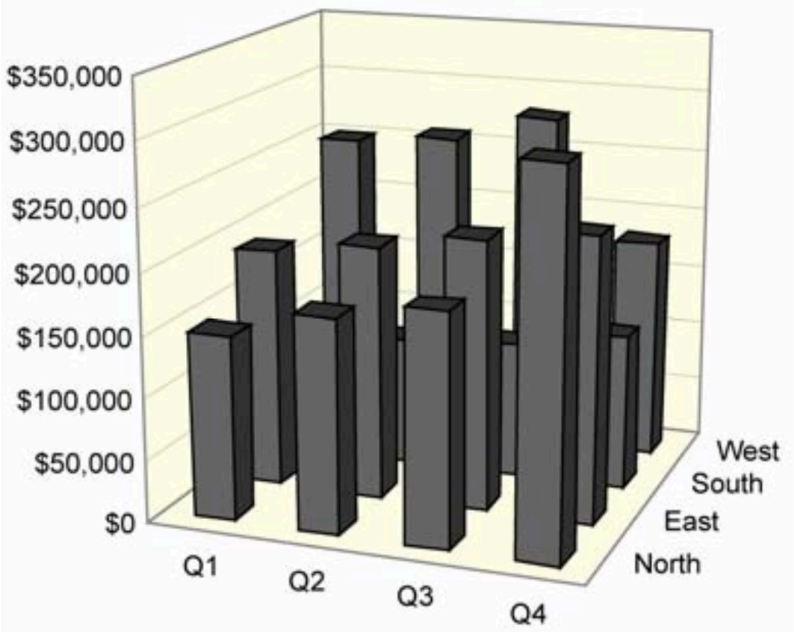


Figure 5: sales per quarter for each of the company's regions.

I. Write a brief critique of this visualisation to explain why you need to change it in advance of the presentation. [4]

II. List and briefly explain three alternatives to presenting this data that would be preferable to the approach shown in Figure 5. [6]

III. Produce a drawing of a suitable visualisation to represent this data. Your boss requests that you emphasise the sales from the West region in your redesign. You should explain the visualisation design decisions you have made in your drawing. [10]

END