

**EXAMINATION PAPER**

**FACULTY: COMPUTER SCIENCE & MULTIMEDIA**

**COURSE: MASTER OF COMPUTER SCIENCE**

**YEAR/SEMESTER: FIRST / SECOND**

**MODULE TITLE & CODE: Data Analytics Visualization: DAV234**

**DATE:**

**TIME ALLOWED: 3 HOURS  START: 1:00 PM FINISH: 4:00 PM**

**Instruction to candidates**

1. This question paper has THREE (3) Sections.
2. Answer all **7** questions in Section A, VSAQ
3. Answer **7** questions out of 9 in Section B, SAQ.
4. Answer **2** questions out of 3 in Section C, LAQ.
5. No scripts or answer sheets are to be taken out of the Examination Hall.
6. Lengthy answers do not win more marks. Students are advised to write clear and concise answers with valid relevant examples

***Do not open this question paper until instructed***

*(Candidates are required to give their answers in their own words as far as practicable)*

**SECTION A**

**Very Short Answer Questions**

**Attempt all seven (7) questions [2 × 7 = 14]**

1. *[Ch5: Introduction to Data Visualization Using R]* What are vectors in R? What is the primary purpose of the **ggplot** library? State one advancement in **ggplot2** from **ggplot1**.
2. *[Ch4: Introduction to Data Visualization Using Tableau]* What are context filters in Tableau? List one use case.
3. *[Ch5: Introduction to Data Visualization Using R]* Why do we use “**install.packages(file.choose(), repos=NULL)**” in R?
4. *[Ch6: Introduction to Programming with Python]* State two differences between a **list** and a **tuple** in Python.
5. *[Ch6: Introduction to Programming with Python]* What do you mean by **List Comprehension** in Python? Show an example.
6. *[Ch3: Analysing Data Using Excel]* List two shortcomings of the **VLOOKUP** function in Excel.
7. *[Ch3: Analysing Data Using Excel]* What are **Macros** in Excel? List two of its uses.

**SECTION B**

**Short Answer Questions**

**Attempt only Seven (7) questions out of Nine (9) questions [7 × 8 = 56]**

Q. 1   *[Ch4: Introduction to Data Visualization Using Tableau]* What is **Data Blending** in Tableau? How does **Blending** differ from traditional **Joins**? What do you think: Which one is faster; **blended data** or **Joins**? Why?

Q. 2   *[Ch4: Introduction to Data Visualization Using Tableau]* What is the significance of the **LOD Expressions** in Tableau? List a few of them with the syntax and elaborate on each of them.

Q. 3   *[Ch6: Introduction to Programming with Python]* What is the striking difference between an **array** and a **list** in Python? Which library is more complex(in the context of the amount of code required) to use; **Matplotlib** or **Seaborn**? What must be the output of the following code of Python:

|  |
| --- |
| **import array as arr**  **My\_Array=arr.array('i',[1,2,3,4,5])**  **My\_Array[::-1]** |

Q. 4   *[Ch6: Introduction to Programming with Python]* What does the following code do in Python? Explain the role of each argument.

**plt.subplot(2, 1, 1)**

What does the function **hist()** do in Matplotlib? Write a Python program to create a scatter plot for the following points using Matplotlib:

**x = [1,2,3,4,5]** and **y = [5,8,6,4,1]**

Q. 5   *[Ch3: Analysing Data Using Excel]* What are different **wildcards** available in Excel? Explain in Detail. What is the exact match and approx match in VLOOKUP? Is it possible to append comments to a cell? If yes, How? If not, why not?

Q. 6   *[Ch3: Analysing Data Using Excel]* Differentiate between **SUM()** and **SUMIF()** functions in **Excel**. Write the role of each of the following functions:

1. **COUNT()**
2. **COUNTA()**
3. **COUNTBLANK()**
4. **COUNTIF()**
5. **COUNTIFS()**
6. **INDEX()**

What function will you use to extract the date from the “20-12-2024”?

Q.7 *[Ch5: Introduction to Data Visualization Using R]* Is it true that in **R**, you must declare the data type of a variable? How do you create a variable named **x** with the numeric value 5? How can you assign the same value to multiple variables in one line? Show an example in **R**. Which statement is used to stop a loop, and which operator is used to add together two values in **R**?

Q. 8   *[Ch5: Introduction to Data Visualization Using R]* Write statements in **R** to install and load the **dplyr** library. What is the **geom** function in **R**? What do you understand by **Aesthetic Mappings** in **R**? Write its role in the visualization.

Q. 9 *[Ch3: Analysing Data Using Excel]* What are the different types of visualization charts we can create in Excel? List and elaborate five(5) of them illustrating the use cases. How can you integrate such charts into a dashboard?

**Group C**

**Long Analytical or Case Question**

**Attempt any two (2) questions out of three (3) questions [2 × 15 = 30]**

Q. 1 *[Ch1: Introducing Data Visualization]* As per your understanding, what is **Data Visualization**? What do you think should be the **traits** of a good visualization? What are some of the **impacts** that a good visualization can make? Also, discuss what role do **ethics** play while performing the task of Data Visualization, with **examples**. (2+5+5+3)

Q. 2 *[Ch2: Exploring Common Types of Data Visualization]* Point out the differences between **Data Visualization** and **Infographics**. Being a student of **Computer Science**, explain the role **Data Science and Visualization** play in comprehending technology, and in its advancement. (4+5+12)

Q. 3*[Ch2: Exploring Common Types of Data Visualization]* While employing the appropriate **Data Visualization Chart** is crucial for representing data effectively, ensuring the data selected for visualization is **ideal** raises an essential question. In your perspective, what criteria(s) determine the **quality of data**? How can the **best data for visualization** be selected? Discuss in perspective of **Domain Knowledge**, **Data Formats**, **Data Storage Techniques**, **Data Extraction**, and **Knowledge Discovery**. (1.5+1+2.5\*5)

**Good Luck**