



MERU UNIVERSITY OF SCIENCE AND TECHNOLOGY

P.O. Box 972-60200 – Meru-Kenya

Tel: +254(0) 799 529 958, +254(0) 799 529 959, + 254 (0) 712 524 293,

Website: info@must.ac.ke Email: info@must.ac.ke

University Examinations 2024/2025

FOURTH YEAR FIRST SEMESTER FOR THE DEGREE OF BACHELOR OF SCIENCE IN
COMPUTER SCIENCE AND FORENSIC

CDS 3352/3252: DATA ANALYTICS AND VISUALIZATION

DATE: DECEMBER 2024

TIME: 2 HOURS

INSTRUCTIONS: Answer question *ONE* (Compulsory) and any other *TWO* questions

QUESTION ONE (30 MARKS)

- a) Differentiate between data analysis and data analytics. (4 marks)
- b) Briefly outline process of data analytics as a cycle. (4 marks)
- c) Briefly explain main activities required while developing an effective visualization. (4 marks)
- d) Discuss any three measures that can be applied to improve data visualization tools. (6 marks)
- e) Discuss any three application of document classification as used in text analytics. (6 marks)
- f) Describe the following terms giving example in each case
 - i. Diagnostic analytics
 - ii. Predictive analytics (6 marks)

QUESTION TWO (20 MARKS)

- a) Discuss any three challenges of data visualization (6 marks)
- b) State the difference between
 - i. Text analysis and linguistic analysis (3 marks)
 - ii. Content understanding and content classification (3 marks)

- c) Briefly explain components of data analysis process and illustrate how linear and cyclic analysis process differs. (8 marks)

QUESTION THREE (20 MARKS)

- (a) Briefly compare the following concepts. You may use an example to explain your point(s).
- i. Data mining and data warehousing
 - ii. Data cleaning, data transformation (6 marks)
- (b) Briefly explain process of developing effective visualization. (8 marks)
- (c) Discuss any two big data processing tools; for each tool state its key features. (6 marks)

QUESTION FOUR (20 MARKS)

- a) Discuss any two text analysis methods and outline how they are applied in big data analytics. (4 marks)
- b) Discuss main component of Hadoop ecosystem. (6 marks)
- c) Discuss any two software components of data mining systems (4 marks)
- d) The table below shows summary of grades attained in an examination. write a code in python or R programming language that will enable visualization of the same. (6 mark)

SUMMARY							
	A	B	C	D	E	IC	TOTAL
TALLY	3	2	12	14	3	2	34

QUESTION FIVE (20 MARKS)

- a) Discuss any three statistical aspects of Big Data (6 marks)
- b) Briefly explain why text analysis is considered to be very difficult and outline how that can be addressed. (4 marks)
- c) Briefly discuss any two text analysis tools. (4 marks)
- d) Discuss stages of text analysis outlining results of each. (6 marks)