



**Riara School of Computing
Nurturing business innovators**

**JAN- APRIL 2026 TRIMESTER
EXAMINATION FOR MASTER OF SCIENCE IN DATA SCIENCE AND.
COMPUTATIONAL
INTELLIGENCE
MRDC 9215: VISUAL ANALYTICS**

DATE: APRIL 2026

TIME: 2 HOURS

GENERAL INSTRUCTIONS:

Students are NOT permitted to write on the examination paper during examination time.

This is a closed book examination. Text book/Reference books/notes are not permitted.

SPECIAL INSTRUCTIONS:

This examination paper consists Questions in Section A followed by section B.

Answer **Question 1 and any Other Two** questions.

QUESTIONS in ALL Sections should be answered in answer booklet(s).

- 1. PLEASE start the answer to EACH question on a NEW PAGE.**
- 2. Keep your phone(s) switched off at the front of the examination room.**
- 3. Keep ALL bags and caps at the front of the examination room and DO NOT refer to ANY unauthorized material before or during the course of the examination.**
- 4. ALWAYS show your working.**
- 5. Marks indicated in parenthesis i.e. () will be awarded for clear and logical answers.**
- 6. Write your REGISTRATION No. clearly on the answer booklet(s).**
- 7. For the Questions, write the number of the question on the answer booklet(s) in the order you answered them.**
- 8. DO NOT use your PHONE as a CALCULATOR.**
- 9. YOU are ONLY ALLOWED to leave the exam room 30minutes to the end of the Exam.**
- 10. DO NOT write on the QUESTION PAPER. Use the back of your BOOKLET for any calculations or rough work.**

SECTION A – COMPULSORY Answer ALL parts of Question 1
--

QUESTION 1 (30 MARKS)

- (a) Florence Nightingale's 1858 polar area diagram is one of the most celebrated examples of data visualization in history. Explain how her visualization influenced decision-makers who had no statistical background. In your answer, identify TWO specific design choices she made and explain why each one made her argument more persuasive to a general audience. [4 mks]
- (b) A data analyst at Nairobi City County redesigns a bar chart by adding a 3D effect, a patterned background, and gridlines on every axis. Using the concept of the data-ink ratio, explain what is wrong with these additions and describe what the analyst should do to improve the chart. [3 mks]
- (c) During the course, misleading COVID-19 visualizations were examined as a case study. Describe TWO ways in which a chart can mislead its audience. For each, give a specific example related to health or public data in Kenya, and explain what a viewer would wrongly conclude. [3 mks]

SCENARIO

A data analyst at Naivas Supermarket has collected the following information:

- Weekly sales revenue (KSh) recorded every week for the past 12 months
- The percentage contribution of five departments (Groceries, Bakery, Dairy, Electronics, and Clothing) to total monthly revenue
- The ages of customers enrolled in the Naivas loyalty card programme
- The relationship between weekly promotional spend (KSh) and weekly sales

- (d) For each of the four datasets described above, recommend the most appropriate chart type and give ONE reason why it suits that particular data. Write your answer as four clearly labelled paragraphs, one per dataset. [4 mks]
- (e) The analyst decides to use a pie chart to show revenue contribution across twelve product sub-categories. The store manager finds it very difficult to compare the slices. Explain why a pie chart struggles when there are many categories and recommend a better chart type, giving ONE reason for your choice. [3 mks]
- (f) Hans Rosling's Gapminder bubble chart was used in the course to show global development trends. Explain how a bubble chart can display more than two variables at the same time. Name

at least THREE of the variables Gapminder encoded and state which visual channel was used for each. [3 mks]

(g) An e-commerce customer journey Sankey diagram was covered in the course. Explain what a Sankey diagram shows that a simple bar chart cannot. Describe what the width of each link represents and give ONE specific business insight a Kenyan online retailer could draw from such a diagram. [3 mks]

(h) Compare animation and small multiples as two different ways of showing how data changes over time. Give ONE advantage and ONE disadvantage of each approach, using examples relevant to Kenyan data. [4 mks]

SCENARIO

A Kenyan SME retailer with branches in Nairobi, Mombasa, Kisumu, and Nakuru has commissioned a Power BI executive dashboard. The brief requires:

- Total monthly revenue across all branches
- Top ten best-selling products
- A map showing sales performance by county
- A year-over-year revenue comparison

(j) Describe the layout you would design for this Power BI dashboard. For each of the four required elements, state which visual type you would use and where on the page you would place it, giving a reason for each placement decision. [3 mks]

SECTION B – ANSWER ANY TWO QUESTIONS

Each question carries 15 marks

QUESTION 2 (15 MARKS)

CASE STUDY

KCB Bank Kenya has a team of data analysts who have built a customer analytics platform. The platform holds five years of data on 2 million customers covering:

- *Loan repayment behaviour (on time, late, defaulted)*
- *Monthly account balance trends*
- *Branch location and county of residence*
- *Product usage (mobile banking, savings, insurance, loans)*

The Head of Retail Banking needs visualizations to support THREE decisions:

- *Decision 1: Which counties have the highest loan default rates?*

- *Decision 2: How have average account balances changed over the past five years?*
- *Decision 3: Which combination of products is most commonly used together?*

- (a) For each of the three decisions above, recommend the most appropriate chart type, identify the specific variables from the dataset that should be encoded, and explain how the chart directly supports that decision. Write your answer as three clearly labelled paragraphs, one per decision. **[6 mks]**
- (b) The Head of Retail Banking wants to be able to click on any county on the map and immediately see that county's loan default trend over the past five years on a separate panel. Explain what interactive visualization feature makes this possible, describe how it works, and explain what analytical insight becomes available to the Head of Banking that would not be possible from two separate static charts. **[4 mks]**
- (c) A junior analyst presents the Head of Retail Banking with a chart showing that KCB branches in Nairobi county have the highest number of loan defaults in Kenya. The Head of Banking concludes that Nairobi is KCB's riskiest region and proposes cutting lending there. Identify what is wrong with both the visualization and the conclusion, explain what additional variable the analyst should have included in the chart, and describe what the corrected visualization would look like. **[5 mks]**

QUESTION 3 (15 MARKS)

SCENARIO

A research team at the Kenya Agricultural and Livestock Research Organisation (KALRO) is building a visualization of crop yields across all 47 counties. Their data comes from three different sources:

- A CSV file of yield records downloaded from the Kenya Open Data Portal
- Seasonal rainfall figures stored in a PostgreSQL database at the Kenya Meteorological Department
- Soil quality readings received from an external satellite imagery API

- (a) Describe the ETL process the KALRO team should follow to combine these three data sources into a single dataset ready for visualization. In your answer, explain what each of the three stages involves, and give ONE specific task for each stage based on this scenario. **[4 mks]**
- (b) During preprocessing, the team discovers two problems. First, the county name 'Murang'a' is spelled as 'Muranga' in the meteorological database. Second, rainfall readings are missing for 15% of the county-month combinations. For each problem, explain how it should be handled before visualization, and describe what would go wrong in the final chart if it were left unresolved. **[4 mks]**
- (c) The team wants to show the distribution of maize yields across counties using either a box plot or a violin plot. Explain what each chart shows, identify ONE piece of information the

violin plot provides that the box plot does not, and state which you would recommend for this KALRO report and why. **[3 mks]**

(e) A data scientist at the Ministry of Health wants to communicate findings on maternal mortality rates across Kenyan counties to Cabinet Secretaries. Describe the THREE key elements of an effective data story and explain how each element helps the Cabinet Secretaries understand and act on the findings. **[4 mks]**

QUESTION 4 (15 MARKS)

SCENARIO

Kenya Power wants to use data visualization to improve how it monitors electricity distribution and communicates with its customers and regulators. The company has five years of data covering:

- *Hourly electricity consumption by county and customer category* (residential, commercial, industrial)
- *Transformer fault incidents — location, date, duration of outage, and number of customers affected*
- *Monthly revenue collected versus revenue billed by county*
- *Customer complaints logged by type (billing, outage, connection request)*

The CEO wants a dashboard for the executive team, a separate operational dashboard for field engineers, and a public-facing visualization for customers.

(a) (The CEO's executive dashboard and the field engineers' operational dashboard serve very different audiences and purposes. For each dashboard, describe the audience, state TWO specific decisions each audience needs to make, recommend the most appropriate chart types for those decisions, and explain why the design and level of detail would differ between the two dashboards. **[6 mks]**

(b) The data team wants to show how transformer faults are distributed geographically across Kenya and whether counties with more faults also have more customer complaints. Recommend TWO chart types that together would answer this question — one for geographic distribution and one for the relationship between faults and complaints. For each chart, identify the variables encoded and explain what pattern or insight it would reveal to a field engineering manager. **[5 mks]**

(c) Kenya Power wants to create a public-facing visualization showing customers in each county how their electricity consumption compares to the county average, and how it has changed over the past 12 months. Explain what makes a visualization appropriate for a general public audience rather than a technical one, recommend a suitable chart type for this purpose, and describe how Kenya Power could make this visualization accessible to customers accessing it on a basic smartphone. **[4 mks]**

END OF EXAMINATION PAPER: ALL BEST

MRDC 9215 – Visual Analytics / Data Visualization | Riara University | Jan–April 2026