CERTIFIED PUBLIC ACCOUNTANTS (CPA)

POST ADVANCED LEVEL

BUSINESS DATA ANALYTICS (PRACTICAL PAPER)

THURSDAY: 4 December 2025. Morning Paper. **Time allowed: 3 hours.**

This paper consists of three sections. SECTION I consists of twenty (20) Multiple Choice Questions each carrying one (1) mark. SECTION II has three (3) practical questions each carrying twenty (20) marks. SECTION III has two (2) practical questions each carrying twenty (20) marks. Answer ALL questions in section I and II and any one (1) question out of the two (2) questions in section III based on your specialisation.

Under SECTION II and SECTION III, you are required to create Ms Excel Worksheets and input your workings and solutions. You may use the Excel template within the question.

SECTION I (20 MARKS)

Answer ALL questions in this section.

1. You have a dataset where the first row (row 1) contains product codes and the second row (row 2) contains product prices as follows:

A	B	C	D	E
Code	P101	P102	P103	P104
Price	50	75	100	125

   You want to retrieve the price of the product with code "P103".
   Which one of the following HLOOKUP formulas will return the correct price?

   A. =HLOOKUP("P103", A1:E2, 2, FALSE)
   B. =HLOOKUP("P103", A1:E3, 1, TRUE)
   C. =HLOOKUP("P103", A1:E2, 1, FALSE)
   D. =HLOOKUP("P103", A1:E2, 2, TRUE) (1 mark)

2. Which one of the following data modeling techniques is primarily used to organise data into entities and relationships in a database?

   A. Time series analysis
   B. Entity-Relationship (ER) modeling
   C. Regression analysis
   D. K-Means clustering (1 mark)

3. Which one of the following statements **BEST** describes the difference between a "Logical Data Model" and a "Conceptual Data Model"?
    A. A logical data model includes specific data types and relationships while a conceptual data model focuses on high-level entities and their relationships
    B. A conceptual data model contains detailed attributes and normalisation whereas a logical data model only defines business rules
    C. A logical data model is independent of technology whereas a conceptual data model is specific to a database system
    D. A conceptual data model includes implementation details while a logical data model is only used for brainstorming ideas   (1 mark)

4. Which one of the following statements is **TRUE** with regard to the data preparation phase in the Cross-Industry Standard Process for Data Mining (CRISP-DM) framework? Data preparation _____.
    A. is optional if raw data is available
    B. is the final phase of CRISP-DM before deployment
    C. does not include data transformation or feature engineering
    D. involves selecting, cleaning, constructing and formatting data for modeling   (1 mark)

5. Which one of the following statements is **FALSE** with regard to the data understanding phase in the Cross-Industry Standard Process for Data Mining (CRISP-DM)?
    A. It involves collecting initial data and assessing its quality
    B. It includes exploring the data through visualisation and statistical summaries
    C. It focuses on transforming and preparing the data for modeling
    D. It aims to understand the data's characteristics and potential patterns   (1 mark)

6. Which one of the following **CORRECTLY** matches one of the Vs of Big Data with its definition?
    A. Volume – The speed at which data is generated and processed
    B. Variety – The different types and sources of data such as structured and unstructured data
    C. Velocity – The trustworthiness and quality of data
    D. Veracity – The large amount of data generated from various sources   (1 mark)

7. The following statements relate to the Vs of Big Data:

    Statement1: Velocity refers to the speed at which data is generated, processed and analysed.

    Statement 2: Veracity refers to the different types and formats of data such as structured, semi-structured and unstructured data.

    Which one of the following choices is **CORRECT**?
    A. Both statements are correct
    B. Both statements are incorrect
    C. Only Statement 1 is correct
    D. Only Statement 2 is correct   (1 mark)

8. Which one of the following data visualisation types is **BEST** suited for comparing values across categories and showing trends over time?
    A. Bar chart and line chart
    B. Scatter plot and pie chart
    C. Histogram and heatmap
    D. Box plot and radar chart   (1 mark)

9. Which one of the following statements **CORRECTLY** describes the principle of "informed consent" in big data ethics?
    A. Users must be told that their data is being collected and give explicit permission
    B. Organisations can use data freely as long as it is anonymised
    C. Companies must collect as much data as possible for better analytics
    D. Data collection must always be hidden to avoid influencing user behaviour   (1 mark)

10. Which one of the following **BEST** describes the data principle of attributable? Data should be _____.
    A. complete, consistent and accurate
    B. recorded in a way that clearly links it to its source or author
    C. accessible and retrievable when needed
    D. protected from unauthorised access and alterations   (1 mark)

11. Which one of the following is **NOT** one of the data protection principles under the Data Protection Act, 2023? Personal data must be _____.
    A. processed lawfully, fairly and transparently
    B. stored indefinitely for future reference
    C. accurate and kept up to date
    D. collected for specified, explicit and legitimate purposes   (1 mark)

12. Which one of the following is considered a key quality of effective data visualisation in data analytics?
    A. Using as many colours and shapes as possible to enhance visual appeal
    B. Displaying data without any labels or legends to keep the focus on the chart itself
    C. Simplifying complex data into clear, easy-to-understand visual formats that convey the message
    D. Including every possible data point even if it overcomplicates the visual representation   (1 mark)

13. Which one of the following statements **BEST** describes the difference between distribution charts and relationship charts?
    A. Distribution charts show how variables correlate while relationship charts display the frequency or spread of data
    B. Distribution charts focus on the frequency or spread of a single variable while relationship charts depict the interaction between two or more variables
    C. Relationship charts show how data is clustered while distribution charts are used for time-series analysis
    D. Distribution charts can only display categorical data while relationship charts are limited to continuous data   (1mark)

14. Which one of the following programming languages is **MOST** commonly used for statistical analysis in data analytics?
    A. C++
    B. R
    C. Java
    D. HTML   (1 mark)

15. Which one of the following statements defines the term "data wrangling" in data analytics?
    A. Archiving old datasets for future use
    B. Combining multiple data sources into a single dataset
    C. Cleaning and organising raw data into a usable format
    D. Creating predictive models using past data   (1 mark)

16. What is the meaning of "ETL" in the context of data analytics?
    A. Extract, Transfer, Load
    B. Evaluate, Transform, Load
    C. Extract, Transform, Load
    D. Export, Transform, Learn   (1 mark)

17. Which one of the following statements defines the term a "logical data model"? It is a model that _____.
    A. shows the physical design of a database
    B. shows the logical design of a database
    C. describes the relationships between data elements
    D. shows the business requirements of a database   (1 mark)

18. What formula should be entered in cell A3 to display the results as shown below?

	A	B	C	D	E	F	G
1	USD $000's		2014A	2015A	2016A	2017A	2018A
2							
3	Income Statement USD $000's						
4	Revenue		150,000	165,000	181,500	199,650	219,615
5	COGS		67,500	74,250	81,000	89,000	98,000
6	Gross Margin		82,500	90,750	100,500	110,650	121,615
7	SG&A		16,500	18,150	20,000	22,000	25,300
8	EBITDA		66,000	72,600	80,500	88,650	96,315
9	Depreciation		6,600	7,260	9,000	9,500	9,800
10	Interest		1,000	1,000	1,000	1,000	1,000
11	EBT		58,400	64,340	70,500	78,150	85,515
12	Taxes		17,500	19,300	20,100	23,600	26,600
13	Net Income		40,900	45,040	50,400	54,550	58,915

    A. ="Income Statement"&A1
    B. ="Income Statement "&A1
    C. ="Income Statement "+A1
    D. ="Income Statement "&"A1"   (1 mark)

19. Which "V" of Big Data analytics refers to the trustworthiness and quality of data?
    A. Volume
    B. Variety
    C. Veracity
    D. Value   (1 mark)

20. Which stage in the CRISP-DM process involves cleaning and transforming data for analysis?
    A. Business understanding
    B. Deployment
    C. Data preparation
    D. Data collection   (1 mark)

CPA POST ADVANCED LEVEL

BUSINESS DATA ANALYTICS (PRACTICAL PAPER)

FRIDAY: 25 April 2025. Morning Paper. **Time allowed: 3 hours.**

Answer ALL questions in SECTION I and SECTION II and any one (1) question in SECTION III. SECTION I has twenty (20) Multiple Choice Questions each carrying one (1) mark. SECTION II has three (3) practical questions each carrying twenty (20) marks. SECTION III has two (2) practical questions each carrying twenty (20) marks. Answer any one (1) question out of the two (2) questions in section III based on your specialisation.

Under SECTION II and SECTION III, you are required to create Ms Excel Worksheets and input your workings and solutions. You may use the Excel template within the question.

SECTION I (20 MARKS)

Answer ALL questions in this section.

QUESTION ONE
Which one of the following Ms Excel features allows you to select or highlight all cells that have formulas?
A. Find it
B. Replace ALL
C. Go To Find
D. Go To Special (1 mark)

ANSWER: D

QUESTION TWO
Assuming cell A1 in Ms Excel is displaying the number "12000.7789". Which one of the following Ms Excel formulas should be used to round this number to the closest integer?
A. =MROUND(A1,100)
B. =MROUND(A1,10)
C. =ROUND(A1,0)
D. =ROUND(A1,1) (1 mark)

ANSWER: C

QUESTION THREE
Which one of the following stages of the CRISP framework involves evaluating the results of the analysis?
A. Interpret
B. Insert
C. Investigate
D. Iterate (1 mark)

ANSWER: A

QUESTION FOUR
Which one of the following types of data visualisation tools is **BEST** suited for showing the composition of a whole?
A.   Scatter plot
B.   Line graph
C.   Bar chart
D.   Pie chart   (1 mark)

ANSWER: D

QUESTION FIVE
Which one of the following is the first stage in the data lifecycle?
A.   Modeling data requirements
B.   Recording data
C.   Identifying data sources
D.   Removing data   (1 mark)

ANSWER: C

QUESTION SIX
In Ms Excel, which one of the following statements **BEST** describes the purpose of the VLOOKUP function?
A.   It multiplies values in a specified range
B.   It searches for a value in the leftmost column of a table and returns a value in the same row from a specified column
C.   It allows you to enter data into multiple cells at once
D.   It automatically generates charts based on selected data   (1 mark)

ANSWER: B

QUESTION SEVEN
In data analytics, which one of the following data models is typically used to organise data in a structure of tables with rows and columns, where each table has a unique key and relationships are established between tables?
A.   Dimensional Model
B.   Entity-Relationship (ER) Model
C.   Document Model
D.   Key-Value Model   (1 mark)

ANSWER: B

QUESTION EIGHT
Which one of the following **BEST** describes the difference between a "logical data model" and a "conceptual data model"?
A.   A logical data model defines the structure of data without specifying how it is stored while a conceptual data model defines the specific tables and columns
B.   A logical data model is created to outline the business processes whereas a conceptual data model outlines the technical database structure
C.   A conceptual data model is focused on high-level business concepts, while a logical data model is more detailed and defines specific attributes, relationships and data types
D.   A conceptual data model includes specific database platforms and storage details whereas a logical data model only includes high-level entities   (1 mark)

ANSWER: C

QUESTION NINE
One of the processes in the cross-industry standard process for data mining is data preparation. This includes the following tasks:
1.   Handling missing values
2.   Transforming variables
3.   Selecting features
4.   Combining datasets

Which one of the following choices presents the task in their **CORRECT** order?
A.   1, 2, 3, 4
B.   1, 4, 3, 2
C.   1, 3, 2, 4
D.   1, 2, 4, 3   (1 mark)

ANSWER: A

QUESTION TEN
Which one of the following graphs is the **CORRECT** category for distribution category?
A.   Scatter Plot
B.   Heat map
C.   Nested Pie Charts
D.   Box Plot   (1 mark)

ANSWER: D

QUESTION ELEVEN
The following statement relates to diagnostic analytics and its use in accounting:
Diagnostic analytics helps answer "Why did this happen?" by examining past performance and identifying the underlying reasons for certain outcomes. For example, diagnostic analytics can assist a business decide whether to approve or deny loans based on insights into the likelihood of default.

Which one of the following statements is **CORRECT**?
A.   The definition of diagnostic analytics is the only correct part
B.   The example of diagnostic analytics is the only correct part
C.   The entire statement is correct
D.   The entire statement is false   (1 mark)

ANSWER: B

QUESTION TWELVE
In the field of accounting and finance, prescriptive analytics is primarily used to_____.
A.   identify patterns and trends in historical financial data
B.   recommend actions or strategies to achieve desired financial outcomes
C.   summarise and report past financial transactions for decision-making
D.   predict future financial performance based on historical data   (1 mark)

ANSWER: B

QUESTION THIRTEEN
Which one of the following is considered as a key quality of effective data visualisation in data analytics?
A.   Using as many colours and shapes as possible to enhance visual appeal
B.   Displaying data without any labels or legends to keep the focus on the chart itself
C.   Simplifying complex data into clear, easy-to-understand visual formats that convey the message
D.   Including every possible data point even if it overcomplicates the visual representation   (1 mark)

ANSWER: C

QUESTION FOURTEEN
Which one of the following is **NOT** a core principle of data protection under data protection laws?
A.   Data minimisation
B.   Transparency and accountability
C.   Data retention for unlimited periods
D.   Purpose limitation   (1 mark)

ANSWER: C

QUESTION FIFTEEN

Which one of the following is **NOT** part of the 5Vs of Big Data?
A.   Volume
B.   Variety
C.   Velocity
D.   Validation   (1 mark)

ANSWER: D

QUESTION SIXTEEN

In the context of data analytics, which one of the following statements **CORRECTLY** shows the distinction between "data protection" and "data security" in their approach to managing data?
A.   Data protection focuses on ensuring data privacy and compliance with regulations while data security focuses on preventing unauthorised access and attacks
B.   Data protection and data security are identical since both aim to prevent data loss and ensure privacy
C.   Data protection involves encryption techniques while data security involves access control only
D.   Data protection is concerned with physical security measures while data security involves backup strategies
                                                                                                       (1 mark)

ANSWER: A

QUESTION SEVENTEEN

Which one of the following is the primary purpose of data mapping in data analytics?
A.   To visualise data trends over time
B.   To convert data from one format to another
C.   To organise data into structured tables
D.   To cleanse data by removing errors   (1 mark)

ANSWER: B

QUESTION EIGHTEEN

Which one of the following statements is **TRUE** regarding data portability and data erasure under data protection laws?
A.   Data portability allows individuals to request that their data be erased without restriction while data erasure allows them to request a copy of their data in a structured format
B.   Data portability allows individuals to obtain their data in a commonly used machine-readable format and transfer it to another service provider while data erasure allows individuals to request the deletion of their personal data under certain conditions
C.   Data portability requires companies to retain individuals' data indefinitely while data erasure permits companies to store data without limitations
D.   Data portability and data erasure are not required under data protection laws   (1 mark)

ANSWER: B

QUESTION NINETEEN

Which one of the following is a primary challenge related to the speed of data in data analytics?
A.   Ensuring data security during processing
B.   Handling large volumes of data in real-time
C.   Simplifying data storage architectures
D.   Achieving data consistency across different systems   (1 mark)

ANSWER: B

QUESTION TWENTY

The basic data structure that can be thought of in an SQL table or a spreadsheet data representation is known as
_____.
A.   series
B.   list
C.   dataframe
D.   points   (1 mark)

ANSWER: C

Answer ALL questions in this section.

QUESTION 21
The following financial statements relate to Z Limited for the year ended 31 December 2024:

Z Limited

Statement of profit or loss for the year ended 31 December 2024	Sh."000"
Revenue (Sales of goods)	4,500.00
Cost of Sales:	
Opening inventory	600.00
Purchases	3,260.00
	3,860.00
Closing inventory	(660.00)
Cost of goods sold (Cost of sales)	(3,200.00)
Gross profit	1,300.00
Other expenses:	
Administrative expenses	(210.00)
Distribution expenses	(210.00)
Earnings Before Interest, Tax, Depreciation and Amortisation	880.00
Depreciation and amortisation	(150.00)
Earnings Before Interest and Tax	730.00
Interest expense	(140.00)
Earnings Before Tax	590.00
Income Tax Expense	(160.00)
Earnings After Tax (Net Income)	430.00
Dividends paid	(180.00)
Retained profit for the period	250.00
Retained profit b/f	550.00
Retained profit c/f	800.00

Z Limited

Statement of financial position as at 31 December 2024	Sh."000"
Assets:	
Non-current assets	2,440.00
Current assets:	
Inventory	660.00
Trade receivables	970.00
Cash at bank and in hand	140.00
Total current assets	1,770.00
Total assets	4,210.00
Capital:	
Ordinary share capital	850.00
Share premium	330.00
Retained profits	800.00
Total capital	1,980.00
Liabilities:	
Non-curret liabilities:	
Long-term bank loans	1,300.00
Current liabilities:	
Trade payables	810.00
Tax payable	120.00
	930.00
Total liabilities	2,230.00
Total capital and liabilities	4,210.00

You are required to prepare the forecast financial statements for the next five years commencing 2025, given the following assumptions:

For the statement of profit or loss assume:

1. Revenue and purchases are expected to grow at the rate of 10% per annum.
2. Gross profit is 30% of revenue. (This will make it easier to determine closing inventory)
3. Administration expenses will grow at the rate of 5% per annum.
4. Selling and distribution expenses will likely grow at the rate of 10% per annum in line with revenue.
5. Depreciation and amortisation is on average 10% of previous period non-current assets.
6. Interest expense is 10% of end of last period balance of the long term loans.
7. Income tax expense is 30% of the profit before tax.
8. Dividends is given as 40% of profit after tax.

For the statement of financial position assume:
1. Non-current assets are expected to grow at the rate of 5% annually.
2. Inventory is determined from the profit or loss.
3. Trade receivables will be 2 months of credit sales.
4. Cash at bank are expected to grow at Sh.50,000 per year.
5. Share capital and share premium will remain constant over the period.
6. The longterm bank loan will be the balancing figure in the statement of financial position.
7. Trade payables are 2 months of purchases.
8. Tax payable is 60% of the tax expense.
9. Share price is expected to grow at the rate of 5% per annum.

Required:
(a)   The forecast statement of profit or loss for the five years from 2025 to 2029.   (10 marks)

(b)   The statement of financial position for the five years from 2025 to 2029.   (10 marks)
(Total: 20 marks)

QUESTION 22

Viongozi Enterprise is considering investing in a project with the following details:

	Sh.	Sh."000"
Initial capital outlay		50,000,000
Residual value		15,000,000
Selling price per unit	700	
Variable cost per unit	350	
Fixed costs per year		2,500,000

Additional information:
1. If implemented, the project has a potential sales value of 38,000 units per annum.
2. The project has a useful life of 5 years.
3. Viongozi Enterprise uses a weighted average cost of capital (WACC) of 11%.
4. For better perspective, the management would like to see the variability in the project returns given a 7.5% adverse change in the project details.
5. For each scenario of input variability, the rest of the inputs are restored to the default input values.

Required:

(a)   Compute the base net present value (NPV) of the project.   (5 marks)

(b)   Carry out the NPV sensitivity analysis of all the project inputs except residual value and project life.   (5 marks)

(c)   Prepare an appropriate chart to indicate the different NPVs for each scenario in (b) above.   (5 marks)

(d)   Identify the inputs variability where the project still remain viable.   (5 marks)

(Total: 20 marks)

QUESTION 23

Apollo Limited manufactures fresh juice for both local and export markets.

The company has provided the following data for the year 2024:

	Quantities Produced	Production Costs Sh."000"
1 January 2024	600,000	72,000
1 February 2024	480,000	56,000
1 March 2024	800,000	92,000
1 April 2024	680,000	76,000
1 May 2024	480,000	64,000
1 June 2024	1,000,000	120,000
1 July 2024	880,000	108,000
1 August 2024	360,000	44,000
1 September 2024	720,000	96,000
1 October 2024	1,200,000	128,000
1 November 2024	1,120,000	116,000
1 December 2024	1,400,000	144,000

Required:

(a)   Generate the descriptive statistics for quantities and production costs.   (4 marks)

(b)   Generate the regression output and estimate the cost of producing 900,000 units at the 95% confindence level.   (12 marks)

(c)   If the selling price is Sh.250 per unit, determine by graphical method, the breakeven point.   (4 marks)

(Total: 20 marks)

Answer only ONE (1) question in this section.

QUESTION 24

You are the audit senior at Quality Professionals Consultancy Limited. You have been furnished with the following sample extracts from the books of XYM Limited to assist in determining the validity and accuracy of the data as well as identify any fraud:

Income statement for the year ended 31 December:

Component	2020 (Sh."000")	2021 (Sh."000")	2022 (Sh."000")	2023 (Sh."000")	2024 (Sh."000")
Revenue	500,000	550,000	400,000	585,000	612,000
Cost of goods sold	300,000	600,000	360,000	200,000	420,000
Gross profit	-	-	-	-	-
Operating expenses	100,000	220,000	120,000	240,000	330,000
Net Profit	-	-	-	-	-

Statement of financial position as at 31 December:

Component	2020 (Sh."000")	2021 (Sh."000")	2022 (Sh."000")	2023 (Sh."000")	2024 (Sh."000")
Current assets	150,000	170,000	190,000	290,000	270,000
Non-current assets	250,000	280,000	270,000	250,000	260,000
Total assets	400,000	450,000	460,000	540,000	530,000
Current liabilities	100,000	110,000	100,000	120,000	96,000
Non-current liabilities	150,000	160,000	170,000	200,000	200,000
Total liabilities	250,000	270,000	270,000	320,000	296,000
Equity	-	-	-	-	-

Purchase Orders Data

Order ID	Order Date	Supplier ID	Item ID	Quantity Ordered	Unit Price (Sh.)	Total Order Value (Sh.)
O01	1/1/2024	S001	I001	500	500	250,000
O02	9/7/2024	S002	I002	1,200	200	240,000
O03	3/15/2024	S003	I003	350	430	150,500
O04	2/9/2024	S004	I004	250	500	125,000
O05	2/19/2024	S005	I005	300	700	210,000

Invoice Data

Invoice ID	PO ID	Invoice Date	Supplier ID	Amount Invoiced (Sh.)
A001	O01	7/10/2024	S001	270,000
A002	O02	12/1/2024	S002	240,000
A003	O03	9/16/2024	S003	150,000
A004	O04	9/2/2024	S004	125,000
A005	O09	3/8/2024	S005	310,000

Payments Data

Payment ID	Invoice ID	Payment Date	Amount Paid (Sh.)
P001	A001	5/5/2024	250,000
P002	A002	6/15/2023	140,000
P003	A003	7/1/2024	150,000
P004	A002	6/21/2024	120,000
P005	A005	5/18/2024	200,000

Required:

(a)   Fill appropriately the details in the Gross Profit, Net Profit and Equity Rows.   (1 mark)

(b)   Undertake the following analysis of trends in key financial statement components: (9 marks)

   (i)   Use a correctly referenced Index and Match nested function to fill in the details in the below data array from the output in (a) above:

Component	2020	2021	2022	2023
Revenue				
COGS				
Operating Expenses				

   (ii)   Create a line graph with title "Revenue and Expenses Analysis" to visually represent the summary per year.

   (iii)   Use a correctly reference Index and Match nested function to fill in the details in the below data array from the output in (a) above Revenue, Cost of Goods Sold and Operating Expenses trends over the years.

Component	2020	2021	2022	2023
Total sssets				
Total liabilities				
Equity				

    (iv)   Create a column chart to visualise the comparison of total assets and total liabilities per year.

(c)   In this sheet, create a Pivot Table of the payments Data to summarise the payments made over the years showing the year and months breakdown

(d)   Calculate the percentage change for each component of the Income Statement year-over-year (YOY).

(e)   Perform a 3-way order match between the Purchase Orders, Invoices and Payments data as below: (10 marks)

    (i)   Using a relevant function, for each purchase order, match the corresponding invoice.

    (ii)   For each invoice, check whether there is a payment made (based on Invoice ID).

    (iii)   Highlight any discrepancies, such as missing invoices, unmatched payments or over/under payments.

    (iv)   Provide a column and give explanations for each discrepancy.   **(Total: 20 marks)**

QUESTION 25

You recently joined the National Treasury as the Financial analyst. You have been furnished with Project data relating to the 47 counties of Kenya to be able to help develop reports that would assist in strategic and data-driven decision making.

The Data available in Sheet name "Question 25 Data".

Required:

(a)   Create a copy of the entire data in a separate sheet with only the values so as to do away with the formattings in the original data. Name the new

Sheet 'Workings'.   (2 marks)

(b)   Data Wrangling - It is important you follow these steps in order.   (12 marks)

In the new sheet:

1.   Delete any top rows to ensure the data is affixed to the grid (Starts at Cell A1).

2.   Delete all blank columns or columns without a header within the data bounds example, Column C.

3.   Using either a formulae or function, create a new row number 3 that will combine the contents of row 1 and 2 per column to create complete and spaced data headers example cell A3 should read "County Name".

4.   Copy the formulae in Cell A1 and paste it in Cell P5 preceeded by a single apostrophe (') to avoid the formula executing.

5.   Delete the original rows 1 and 2 to remain with the combined data headers.

6.   Explore the data and relevantly fill using a formulae the blank cells in Column A.

7.   Create a new Column after "Transaction ID" column. Split into 2 the "Transaction ID" column at the hyphen (-). Name the text column "County ID" and Numbers Column "Loan Serial".

8.   Create 2 blank columns before "Project Start Date". Split the "Sector / Approval_Date / Region" column into 3 maintaining the split names.

9.   Format the "Project Start_Dates" column appropriately as dates.

10.   Calculate the "Pending Disbursment" (Difference between Approved and Disbused Amounts) and (Difference between Disbusment amount and Project xpenses) columns.

11.   In the "Surplus / Deficit" column, create a function that will either return Deficit or Surplus depending on whether the Project expenses have exceeded or fallen short of the approved amounts.

(c)   Data Summarisation and reporting:

(i)   In the "workings" sheet, prepare the following pivot tables: (6 marks)

-     A summary of "disbursed amounts" per "project start dates" (in Year) seperated into Deficit and Surplus. Visualise this summary in a Trend line.

-     A summary of sectoral expenses vs disbursed amounts. Visualise this in a column chart.

-     Grid all the above 2 charts together next to each other.

(ii)   Insert a slicer for "Region" and connect it to be interactive with the two charts above.

(Total: 20 marks)

COUNTY NAME	TRANSACTION ID	LOAN TYPE	SECTOR / APPROVAL_DATE / REGION	PROJECT START_DATE	APPROVED AMOUNT	DISBURSED AMOUNT	PROJECT EXPENSES	PENDING DISBURSMENT	SURPLUS / DEFICIT	SPENDING STATUS
KITUI	DCIN015-1000	NATIONAL	Construction / 17/1/2017 / West	44830	100000	50,000.00	66474			
	DCIN015-2000	NATIONAL	Agricultural / 10/2/2017 / South	43039	100000	100,000.00	98647			
	DCIN015-3000	NATIONAL	Transport / 12/3/2017 / South	44257	100000	100,000.00	97672			
TURKANA	DCIN023-1000	NATIONAL	Mining / 24/4/2017 / South	45085	50000	100,000.00	44285		42917.00	
TRANS-NZOIA	DCIN026-2000	NATIONAL	Construction / 29/7/2017 / South	45657	10000	50,000.00	88720		45657.00	
MIGORI	DCIN044-1000	NATIONAL	Construction / 21/4/2017 / East	44858	30000	10,000.00	84393			
GARISSA	DCIN007-2000	NATIONAL	Construction / 27/7/2017 / West	44711	50000	30,000.00	62923			
EMBU	DCIN014-1000	NATIONAL	Transport / 17/7/2017 / East	45536	70000	50,000.00	31450			
	DCIN014-4000	NATIONAL	Transport / 17/7/2017 / East	43150	50000	70,000.00	27976			
ISIOLO	DCIN011-1000	NATIONAL	Education / 8/7/2017 / West	45381	50000	50,000.00	86577			
MACHAKOS	DCIN016-1000	NATIONAL	Education / 20/7/2017 / East	44475	100000	50,000.00	67868			
	DCIN016-2000	NATIONAL	Education / 19/7/2017 / South	43708	255000	100,000.00	97195			
VIHIGA	DCIN038-1000	PUBLIC PRIVATE	Mining / 8/7/2017 / West	43583	90000	255,000.00	99598			
MERU	DCIN012-1000	NATIONAL	Transport / 11/7/2017 / South	43404	100000	90,000.00	99593			
	DCIN012-2000	NATIONAL	Construction / 2/7/2017 / East	43266	80000	100,000.00	34950			
KIRINYAGA	DCIN020-1000	NATIONAL	Construction / 17/7/2017 / West	45674	100000	80,000.00	31878			
	DCIN020-2000	NATIONAL	Mining / 20/8/2017 / West	43162	100000	100,000.00	88326			
	DCIN020-5000	NATIONAL	Education / 27/11/2017 / South	43146	100000	100,000.00	93951			
NAIROBI	DCIN047-1000	NATIONAL	Education / 21/2/2017 / East	44137	100000	100,000.00	87251			
	DCIN047-2000	NATIONAL	Mining / 22/7/2017 / West	45036	200000	100,000.00	46394			
	DCIN047-3000	NATIONAL	Agricultural / 3/7/2017 / South	44336	100000	200,000.00	26278			
	DCIN047-4000	NATIONAL	Mining / 11/8/2017 / East	43210	200000	100,000.00	32657			
	DCIN047-5000	OWN-SOURCE	Construction / 19/11/2017 / East	44232	500000	200,000.00	75421			
	DCIN047-8000	DONOR PATNERS	Agricultural / 21/7/2017 / West	45194	200000	500,000.00	65169			
BUSIA	DCIN040-3000	PUBLIC PRIVATE	Agricultural / 8/1/2017 / East	43508	40000	200,000.00	74178			
	DCIN040-1000	NATIONAL	Agricultural / 4/7/2017 / West	43611	100000	40,000.00	92639			
UASIN GISHU	DCIN027-1000	NATIONAL	Agricultural / 30/1/2017 / West	45306	100000	100,000.00	99489			
	DCIN027-4000	NATIONAL	Education / 27/3/2017 / South	44026	100000	100,000.00	42231			
	DCIN027-1000	NATIONAL	Transport / 17/4/2017 / East	44253	150000	100,000.00	35246			
BUNGOMA	DCIN039-1000	DONOR PATNERS	Education / 17/7/2017 / West	44936	50000	50,000.00	40061			
	DCIN039-2000	NATIONAL	Construction / 17/9/2017 / East	44672	100000	50,000.00	61575			
KIAMBU	DCIN022-1000	NATIONAL	Transport / 27/7/2017 / South	44262	100000	100,000.00	38543			
	DCIN022-2000	NATIONAL	Agricultural / 31/8/2017 / South	44402	100000	100,000.00	54637			
	DCIN022-3000	NATIONAL	Education / 1/12/2017 / West	45527	100000	100,000.00	65073			
	DCIN022-5000	NATIONAL	Mining / 17/3/2017 / West	43514	100000	100,000.00	78155			
	DCIN022-6000	NATIONAL	Mining / 10/7/2017 / South	45244	40000	100,000.00	82940			
	DCIN022-1000	NATIONAL	Financial / 19/7/2017 / West	44656	100000	40,000.00	34335			
	DCIN022-1000	NATIONAL	Education / 27/7/2017 / East	44507	350000	100,000.00	88535			
	DCIN022-1000	PUBLIC PRIVATE	Agricultural / 31/3/2017 / West	45380	60000	350,000.00	43892			
NANDI	DCIN029-1000	NATIONAL	Financial / 4/4/2017 / West	44762	70000	60,000.00	29867			
	DCIN029-2000	NATIONAL	Transport / 13/7/2017 / South	45001	150000	70,000.00	91918			
	DCIN029-1000	NATIONAL	Transport / 21/8/2017 / West	43589	20000	150,000.00	97468			
NAROK	DCIN033-1000	NATIONAL	Mining / 27/3/2017 / West	45525	10000	20,000.00	91741			
	DCIN033-2000	NATIONAL	Education / 2/2/2017 / South	44193	300000	10,000.00	56455			
	DCIN033-1000	PUBLIC PRIVATE	Financial / 22/10/2017 / West	44930	100000	300,000.00	52620			
	DCIN033-1000	NATIONAL	Agricultural / 7/7/2017 / West	43490	40000	100,000.00	51865			
	DCIN033-1000	NATIONAL	Education / 8/8/2017 / East	44872	50000	40,000.00	68429			
	DCIN033-1000	NATIONAL	Agricultural / 24/4/2017 / East	42940	100000	50,000.00	51011			
	DCIN033-1000	NATIONAL	Agricultural / 27/8/2017 / East	44615	22000	100,000.00	60736			
MOMBASA	DCIN001-1000	NATIONAL	Education / 7/1/2017 / West	43113	15000	22,000.00	35386			
	DCIN001-2000	NATIONAL	Education / 27/4/2017 / East	42941	100000	15,000.00	84383			
	DCIN001-1000	NATIONAL	Education / 19/7/2017 / East	44713	100000	100,000.00	90957			
	DCIN001-2000	NATIONAL	Education / 19/8/2017 / South	44293	400000	100,000.00	83869			
	DCIN001-1000	PUBLIC PRIVATE	Financial / 27/4/2017 / South	44138	50000	400,000.00	49430			
NAKURU	DCIN032-1000	NATIONAL	Agricultural / 10/7/2017 / West	42946	50000	50,000.00	32442			
	DCIN032-5000	NATIONAL	Construction / 14/7/2017 / East	43138	70000	50,000.00	29462			
	DCIN032-6000	NATIONAL	Education / 14/9/2017 / West	44909	200000	70,000.00	59129			
	DCIN032-7000	PUBLIC PRIVATE	Agricultural / 12/1/2017 / South	43612	60000	200,000.00	63764			
	DCIN032-1000	NATIONAL	Transport / 21/7/2017 / East	43867	10000	60,000.00	28206			
	DCIN032-1000	NATIONAL	Agricultural / 27/3/2017 / West	45327	25000	10,000.00	62306			
	DCIN032-2000	NATIONAL	Mining / 13/4/2017 / South	45295	100000	25,000.00	77007			
MANDERA	DCIN009-1000	NATIONAL	Mining / 14/1/2017 / East	44559	100000	100,000.00	27338			
	DCIN009-2000	NATIONAL	Education / 17/7/2017 / South	43293	100000	100,000.00	79208			
	DCIN009-3000	NATIONAL	Financial / 17/7/2017 / South	44970	100000	100,000.00	84837			
	DCIN009-4000	NATIONAL	Education / 20/8/2017 / South	45318	104960	100,000.00	82120			
	DCIN009-7000	NATIONAL	Agricultural / 18/12/2017 / West	43259	50000	104,960.00	58836			
	DCIN009-9000	NATIONAL	Transport / 22/1/2017 / West	43393	40000	50,000.00	45041			
	DCIN009-8000	NATIONAL	Transport / 27/1/2017 / East	45021	100000	40,000.00	44907			
	DCIN009-6000	NATIONAL	Education / 27/7/2017 / East	43765	100000	100,000.00	73415			
	DCIN009-3000	NATIONAL	Transport / 23/7/2017 / South	45222	100000	100,000.00	35886			

County	Code	Type	Detail				
KISUMU	DCIN042-1000	NATIONAL	Agricultural / 17/3/2017 / East	43473	60000	100,000.00	68244
	DCIN042-3000	NATIONAL	Financial / 23/7/2017 / West	44968	70000	60,000.00	55378
	DCIN042-4000	NATIONAL	Education / 11/9/2017 / South	45493	100000	70,000.00	41211
	DCIN042-6000	NATIONAL	Education / 8/1/2017 / West	43034	240000	100,000.00	57167
	DCIN042-7000	PUBLIC PRIVATE	Transport / 19/4/2017 / West	42973	200000	240,000.00	81790
HOMABAY	DCIN043-1000	NATIONAL	Agricultural / 8/7/2017 / West	44506	50000	200,000.00	91715
	DCIN043-1000	NATIONAL	Education / 27/9/2017 / East	45524	200000	50,000.00	48816
	DCIN043-2000	OWN-SOURCE	Agricultural / 19/11/2017 / West	43969	50000	200,000.00	45196
	DCIN043-1000	NATIONAL	Construction / 21/3/2017 / East	43838	60000	50,000.00	30216
	DCIN043-2000	NATIONAL	Education / 20/7/2017 / West	44245	60000	60,000.00	78805
	DCIN043-3000	NATIONAL	Mining / 24/9/2017 / South	44749	100000	60,000.00	25367
	DCIN043-1000	NATIONAL	Transport / 2/2/2017 / South	44309	100000	100,000.00	62192
	DCIN043-1000	NATIONAL	Education / 22/4/2017 / West	45116	30000	100,000.00	40326
	DCIN043-1000	NATIONAL	Agricultural / 2/7/2017 / West	43854	20000	30,000.00	50797
	DCIN043-1000	NATIONAL	Financial / 19/7/2017 / East	45105	300000	20,000.00	49158
THARAKA-NITHI	DCIN013-2000	PUBLIC PRIVATE	Financial / 17/10/2017 / East	45198	350000	300,000.00	47627
	DCIN013-6000	PUBLIC PRIVATE	Financial / 17/2/2017 / South	45138	30000	350,000.00	36565
	DCIN013-1000	NATIONAL	Mining / 11/7/2017 / South	44557	60000	30,000.00	36743
ELGEYO MARAKWET	DCIN028-1000	NATIONAL	Agricultural / 10/7/2017 / South	43440	80000	60,000.00	79092
	DCIN028-2000	NATIONAL	Mining / 13/7/2017 / East	43554	260000	80,000.00	28029
	DCIN028-2000	PUBLIC PRIVATE	Transport / 10/12/2017 / East	42974	100000	260,000.00	70886
TANA RIVER	DCIN004-1000	NATIONAL	Construction / 4/2/2017 / South	44436	100000	100,000.00	46953
	DCIN004-2000	NATIONAL	Construction / 14/1/2017 / East	43022	10000	100,000.00	28454
	DCIN004-1000	NATIONAL	Construction / 7/7/2017 / East	44074	50000	10,000.00	73450
	DCIN004-1000	NATIONAL	Transport / 29/4/2017 / South	45536	100000	50,000.00	46196
	DCIN004-2000	NATIONAL	Financial / 3/7/2017 / West	43015	100000	100,000.00	93054
	DCIN004-3000	NATIONAL	Mining / 28/7/2017 / East	45375	20000	100,000.00	42263
	DCIN004-1000	PUBLIC PRIVATE	Transport / 28/7/2017 / East	44705	50000	20,000.00	66044
	DCIN004-1000	NATIONAL	Construction / 13/4/2017 / South	44006	200000	50,000.00	63569
LAIKIPIA	DCIN031-1000	NATIONAL	Transport / 17/9/2017 / West	44396	200000	200,000.00	78950
	DCIN031-2000	OWN-SOURCE	Financial / 21/1/2017 / West	43131	70000	200,000.00	50094
	DCIN031-1000	NATIONAL	Transport / 21/7/2017 / West	44013	80000	70,000.00	70511
	DCIN031-3000	NATIONAL	Education / 27/10/2017 / East	45572	50000	80,000.00	30813
	DCIN031-1000	NATIONAL	Mining / 21/7/2017 / West	45393	50000	50,000.00	84375
	DCIN031-2000	NATIONAL	Financial / 1/9/2017 / South	44729	50000	50,000.00	63039
	DCIN031-2000	NATIONAL	Financial / 17/4/2017 / East	44906	100000	50,000.00	83514
	DCIN031-1000	NATIONAL	Construction / 20/4/2017 / East	43002	20000	100,000.00	66062
SAMBURU	DCIN025-1000	NATIONAL	Agricultural / 10/4/2017 / East	42935	50000	20,000.00	63291
	DCIN025-2000	NATIONAL	Education / 13/7/2017 / South	43437	100000	50,000.00	90355
	DCIN025-1000	NATIONAL	Education / 21/3/2017 / East	44768	100000	100,000.00	70704
	DCIN025-1000	NATIONAL	Transport / 7/3/2017 / West	45004	80000	100,000.00	49125
	DCIN025-1000	NATIONAL	Financial / 8/7/2017 / South	43378	100000	80,000.00	55736
	DCIN025-2000	NATIONAL	Financial / 17/7/2017 / East	45077	30000	100,000.00	70514
	DCIN025-1000	NATIONAL	Financial / 27/2/2017 / West	43245	50000	30,000.00	35583
	DCIN025-2000	NATIONAL	Financial / 27/4/2017 / West	44091	100000	50,000.00	54398
WEST POKOT	DCIN024-1000	NATIONAL	Agricultural / 20/7/2017 / South	43545	100000	100,000.00	90011
	DCIN024-7000	NATIONAL	Transport / 7/10/2017 / West	44359	100000	100,000.00	61904
	DCIN024-8000	OWN-SOURCE	Construction / 27/10/2017 / East	44120	20000	100,000.00	44297
	DCIN024-1000	NATIONAL	Construction / 12/7/2017 / East	45473	50000	20,000.00	31103
	DCIN024-2000	NATIONAL	Transport / 13/7/2017 / South	43541	60000	50,000.00	71425
NYANDARUA	DCIN018-1000	NATIONAL	Construction / 3/8/2017 / East	45138	20000	60,000.00	51607
	DCIN018-2000	NATIONAL	Mining / 11/3/2017 / East	45481	50000	20,000.00	68512
	DCIN018-3000	NATIONAL	Agricultural / 7/4/2017 / West	44161	100000	50,000.00	40721
	DCIN018-4000	NATIONAL	Construction / 7/7/2017 / South	44891	5000	100,000.00	48864
	DCIN018-1000	NATIONAL	Mining / 2/7/2017 / West	43517	10000	5,000.00	62222
	DCIN018-2000	NATIONAL	Mining / 21/10/2017 / South	43493	100000	10,000.00	46275
NYAMIRA	DCIN046-1000	NATIONAL	Financial / 4/3/2017 / East	44641	100000	100,000.00	61603
	DCIN046-2000	NATIONAL	Mining / 8/4/2017 / East	45630	100000	100,000.00	88041
	DCIN046-3000	NATIONAL	Education / 11/7/2017 / West	44355	100000	100,000.00	45920
	DCIN046-4000	NATIONAL	Education / 19/7/2017 / South	43837	150000	100,000.00	49319
	DCIN046-5000	NATIONAL	Construction / 17/7/2017 / East	43224	150000	150,000.00	41105
	DCIN046-6000	NATIONAL	Education / 20/8/2017 / West	44182	150000	150,000.00	79036
	DCIN046-7000	NATIONAL	Construction / 22/9/2017 / East	45056	150000	150,000.00	44094
	DCIN046-9000	NATIONAL	Construction / 21/10/2017 / West	42951	10000	150,000.00	25139
SIAYA	DCIN041-1000	NATIONAL	Mining / 30/1/2017 / West	43399	20000	10,000.00	54590
	DCIN041-2000	NATIONAL	Agricultural / 13/3/2017 / East	44511	30000	20,000.00	86059
	DCIN041-3000	NATIONAL	Construction / 18/7/2017 / East	43536	30000	30,000.00	92357
	DCIN041-1000	NATIONAL	Agricultural / 29/7/2017 / South	44238	100000	30,000.00	46481
	DCIN041-1000	NATIONAL	Transport / 20/4/2017 / East	44473	50000	100,000.00	33211
	DCIN041-1000	PUBLIC PRIVATE	Agricultural / 4/2/2017 / West	45270	20000	50,000.00	86591
	DCIN041-2000	NATIONAL	Education / 18/3/2017 / South	43814	170000	20,000.00	66567
	DCIN041-1000	PUBLIC PRIVATE	Agricultural / 19/2/2017 / East	43112	20000	170,000.00	51767
	DCIN041-1000	NATIONAL	Mining / 13/2/2017 / South	43774	50000	20,000.00	34175
	DCIN041-2000	NATIONAL	Education / 17/3/2017 / West	44427	60000	50,000.00	46294

County	Code	Source	Sector / Date / Region			Amount	
	DCIN041-3000	NATIONAL	Education / 17/7/2017 / South	42942	100000	60,000.00	44447
KWALE	DCIN002-1000	PUBLIC PRIVATE	Financial / 30/4/2017 / South	44879	100000	100,000.00	88699
	DCIN002-2000	NATIONAL	Financial / 20/7/2017 / South	44710	200000	100,000.00	96280
	DCIN002-2000	NATIONAL	Financial / 22/7/2017 / South	43528	100000	200,000.00	60486
	DCIN002-4000	NATIONAL	Transport / 27/7/2017 / East	43380	400000	100,000.00	41251
	DCIN002-6000	DONOR PATNERS	Financial / 11/8/2017 / South	45460	70000	400,000.00	51596
TAITA-TAVETA	DCIN006-1000	NATIONAL	Construction / 7/7/2017 / South	43343	30000	70,000.00	47126
	DCIN006-2000	NATIONAL	Agricultural / 22/7/2017 / East	43372	100000	30,000.00	45878
	DCIN006-3000	OWN-SOURCE	Agricultural / 7/10/2017 / West	44701	150000	100,000.00	54729
	DCIN006-4000	OWN-SOURCE	Transport / 9/11/2017 / East	43443	64894	150,000.00	53984
	DCIN006-5000	OWN-SOURCE	Education / 18/4/2017 / West	43690	50000	64,894.00	67582
MURANGA	DCIN021-1000	OWN-SOURCE	Mining / 30/10/2017 / South	44453	200000	50,000.00	51657
	DCIN021-1000	NATIONAL	Construction / 10/7/2017 / South	45411	100000	200,000.00	41134
	DCIN021-2000	NATIONAL	Transport / 12/8/2017 / South	44529	200000	100,000.00	92725
	DCIN021-3000	NATIONAL	Mining / 18/9/2017 / East	43063	100000	200,000.00	38034
	DCIN021-4000	OWN-SOURCE	Education / 20/1/2017 / South	43571	100000	100,000.00	56418
MAKUENI	DCIN017-1000	NATIONAL	Financial / 27/7/2017 / West	44120	200000	100,000.00	76059
	DCIN017-2000	NATIONAL	Education / 29/9/2017 / East	44704	400000	200,000.00	55323
	DCIN017-3000	DONOR PATNERS	Construction / 29/3/2017 / East	45250	175792	400,000.00	95768
	DCIN017-4000	NATIONAL	Agricultural / 8/10/2017 / West	44028	300000	175,792.00	98451
	DCIN017-5000	NATIONAL	Mining / 7/4/2017 / East	42951	40000	300,000.00	36849
WAJIR	DCIN008-1000	NATIONAL	Agricultural / 21/1/2017 / West	45415	20000	40,000.00	70187
	DCIN008-2000	NATIONAL	Agricultural / 14/10/2017 / West	43337	150000	20,000.00	61232
	DCIN008-1000	OWN-SOURCE	Financial / 21/12/2017 / West	43302	100000	150,000.00	76419
BOMET	DCIN036-1000	NATIONAL	Construction / 17/1/2017 / East	45517	100000	100,000.00	96692
	DCIN036-2000	NATIONAL	Transport / 17/2/2017 / East	44392	100000	100,000.00	59140
	DCIN036-3000	OWN-SOURCE	Agricultural / 24/3/2017 / East	44532	100000	100,000.00	28636
	DCIN036-4000	NATIONAL	Mining / 27/4/2017 / South	43579	100000	100,000.00	72301
	DCIN036-5000	NATIONAL	Education / 29/7/2017 / East	43859	100000	100,000.00	41172
	DCIN036-6000	NATIONAL	Education / 29/7/2017 / West	43890	100000	100,000.00	26517
	DCIN036-7000	NATIONAL	Financial / 17/7/2017 / West	43456	200000	100,000.00	42016
	DCIN036-8000	OWN-SOURCE	Mining / 24/11/2017 / East	44253	100000	200,000.00	70384
	DCIN036-9000	OWN-SOURCE	Financial / 19/1/2017 / South	44568	50000	100,000.00	73121
KISII	DCIN045-2000	NATIONAL	Transport / 20/7/2017 / West	44715	100000	50,000.00	77237
	DCIN045-2000	NATIONAL	Education / 18/7/2017 / West	43230	200000	100,000.00	91620
	DCIN045-1000	OWN-SOURCE	Education / 12/11/2017 / East	43672	60000	200,000.00	99167
	DCIN045-1000	NATIONAL	Education / 22/7/2017 / East	45186	30000	60,000.00	25167
	DCIN045-1000	NATIONAL	Education / 18/9/2017 / West	43872	100000	30,000.00	55779
NYERI	DCIN019-1000	NATIONAL	Agricultural / 4/2/2017 / South	45550	100000	100,000.00	58262
	DCIN019-2000	DONOR PATNERS	Construction / 14/4/2017 / South	44603	150000	100,000.00	54112
	DCIN019-3000	PUBLIC PRIVATE	Agricultural / 13/7/2017 / South	45618	200000	150,000.00	79203
	DCIN019-1000	NATIONAL	Agricultural / 10/7/2017 / South	43291	200000	200,000.00	89525
	DCIN019-11000	NATIONAL	Agricultural / 30/10/2017 / West	45250	250000	200,000.00	63269
	DCIN019-8000	NATIONAL	Transport / 27/11/2017 / South	44162	200000	250,000.00	32259
	DCIN019-9000	NATIONAL	Mining / 27/9/2017 / West	44727	200000	200,000.00	55582
	DCIN019-1000	DONOR PATNERS	Agricultural / 11/9/2017 / South	44020	100000	200,000.00	70752
KERICHO	DCIN035-10000	NATIONAL	Agricultural / 7/10/2017 / South	43925	100000	100,000.00	31266
	DCIN035-15000	NATIONAL	Mining / 7/11/2017 / South	44971	100000	100,000.00	89935
	DCIN035-16000	NATIONAL	Transport / 8/12/2017 / East	44735	200000	100,000.00	28917
	DCIN035-17000	NATIONAL	Construction / 22/1/2017 / West	45222	200000	200,000.00	25986
	DCIN035-18000	NATIONAL	Financial / 28/7/2017 / West	44459	100000	200,000.00	94542
	DCIN035-2000	OWN-SOURCE	Financial / 7/4/2017 / West	43577	100000	100,000.00	44278
	DCIN035-3000	NATIONAL	Agricultural / 8/7/2017 / East	44308	100000	100,000.00	58770
	DCIN035-5000	NATIONAL	Education / 17/7/2017 / South	43099	100000	100,000.00	69440
	DCIN035-6000	NATIONAL	Education / 3/3/2017 / West	43358	100000	100,000.00	28018
	DCIN035-7000	NATIONAL	Mining / 17/7/2017 / South	44280	100000	100,000.00	81951
	DCIN035-8000	NATIONAL	Construction / 1/9/2017 / West	44182	100000	100,000.00	43885
	DCIN035-9000	NATIONAL	Mining / 30/1/2017 / South	44002	100000	100,000.00	92657
KILIFI	DCIN003-1000	NATIONAL	Financial / 27/7/2017 / South	43902	100000	100,000.00	48207
	DCIN003-5000	NATIONAL	Transport / 14/7/2017 / West	44710	200000	100,000.00	39170
	DCIN003-6000	OWN-SOURCE	Agricultural / 30/9/2017 / East	43912	450000	200,000.00	60832
	DCIN003-7000	DONOR PATNERS	Financial / 7/7/2017 / South	43516	200000	450,000.00	37403
MARSABIT	DCIN010-1000	NATIONAL	Mining / 17/8/2017 / West	43037	50000	200,000.00	56551
	DCIN010-1000	NATIONAL	Agricultural / 13/2/2017 / South	44857	100000	50,000.00	55270
	DCIN010-2000	NATIONAL	Construction / 10/3/2017 / East	45015	30000	100,000.00	84020
	DCIN010-2000	NATIONAL	Agricultural / 30/4/2017 / West	44507	50000	30,000.00	48973
	DCIN010-1000	NATIONAL	Financial / 24/4/2017 / East	43922	200000	50,000.00	60915
	DCIN010-1000	NATIONAL	Financial / 10/7/2017 / South	43200	100000	200,000.00	58335
	DCIN010-2000	NATIONAL	Construction / 17/8/2017 / South	42937	300000	100,000.00	87163
	DCIN010-6000	OWN-SOURCE	Transport / 7/10/2017 / South	45047	100000	300,000.00	78249
KAKAMEGA	DCIN037-1000	NATIONAL	Agricultural / 9/7/2017 / East	45268	20000	100,000.00	97312
	DCIN037-1000	NATIONAL	Financial / 21/3/2017 / West	45485	100000	20,000.00	73171
	DCIN037-1000	NATIONAL	Financial / 3/7/2017 / East	44865	100000	100,000.00	83470
	DCIN037-2000	NATIONAL	Agricultural / 8/8/2017 / South	44873	100000	100,000.00	69880

	DCIN037-3000	NATIONAL	Agricultural / 10/12/2017 / East	44948	100000	100,000.00	30849
	DCIN037-4000	NATIONAL	Agricultural / 17/3/2017 / East	44606	100000	100,000.00	95312
	DCIN037-5000	NATIONAL	Agricultural / 22/7/2017 / East	45666	50000	100,000.00	66755
LAMU	DCIN005-1000	NATIONAL	Agricultural / 4/7/2017 / South	44752	100000	50,000.00	25925
	DCIN005-2000	NATIONAL	Education / 9/7/2017 / West	43461	100000	100,000.00	43699
	DCIN005-1000	NATIONAL	Financial / 2/7/2017 / South	43033	100000	100,000.00	51816
	DCIN005-1000	NATIONAL	Education / 10/7/2017 / West	45231	50000	100,000.00	71555
KAJIADO	DCIN034-1000	NATIONAL	Mining / 7/7/2017 / West	44908	80000	50,000.00	94048
	DCIN034-4000	NATIONAL	Education / 21/7/2017 / South	45426	45000	80,000.00	38146

CPA POST ADVANCED LEVEL

BUSINESS DATA ANALYTICS (PRACTICAL PAPER)

FRIDAY: 6 December 2024. Morning Paper.   **Time allowed: 3 hours.**

Answer ALL questions in SECTION I and SECTION II and any one (1) question in SECTION III. SECTION I has twenty (20) Multiple Choice Questions each carrying one (1) mark. SECTION II has three (3) practical questions each carrying twenty (20) marks. SECTION III has two (2) practical questions each carrying twenty (20) marks. Answer any one (1) question out of the two (2) questions in section III based on your specialisation.

Under SECTION II and SECTION III, you are required to create Ms Excel Worksheets and input your workings and solutions. You may use the Excel template within the question.

SECTION I (20 MARKS)

Answer ALL questions in this section.

1.   Which one of the following is **NOT** a phase in the data analytics based on a CRISP lifecycle?
        A.   Data preparation
        B.   Data exploration
        C.   Data manipulation
        D.   Data modelling   (2 marks)

2.   Which one of the following statistical techniques is used to find relationships between variables in data analytics?
        A.   Regression analysis
        B.   Principal component analysis
        C.   Cluster analysis
        D.   Association rule learning   (2 marks)

3.   Which one of the following terms refers to the process of cleaning and preparing data for analysis?
        A.   Data mining
        B.   Data wrangling
        C.   Data extraction
        D.   Data modelling   (2 marks)

4.   Which one of the following tools is commonly used for data summarisation in Ms Excel?
        A.   Pivot tables
        B.   VLOOKUP
        C.   Power query
        D.   Scatter plots   (2 marks)

5. Which one of the following analyses is used to assess a project's vulnerability to different variables?
    A. Break-even analysis
    B. Sensitivity analysis
    C. Scenario analysis
    D. Ratio analysis   (2 marks)

6. Which one of the following tools is commonly used for cloud-based data management in business analytics?
    A. SQL
    B. Oracle
    C. AWS
    D. Excel   (2 marks)

7. A data analyst makes the following comments about functions in Ms Excel:
    1. To determine the variance of a list of 50 numbers we use =VAR.P(A1:A50)
    2. To determine the standard deviation of 50 numbers we use = VAR.P(A1:A50)
    3. To determine the variance of a list of 50 numbers we use = VAR.S(A1:A10)

    Which one of the following statements is **TRUE**?
    A. Only Statement 1 is correct
    B. Only Statement 2 is correct
    C. Only Statement 3 is correct
    D. All the statements are correct   (2 marks)

8. Which one of the following choices provides the **CORRECT** order of data models?
    A. Physical, logical and conceptual
    B. Physical, conceptual and logical
    C. Conceptual, logical and physical
    D. Conceptual, physical and logical   (2 marks)

9. One of the processes in the cross-industry standard process for data mining is data understanding. This involves the following tasks:
    1. Collect data
    2. Verify data quality
    3. Explore data
    4. Describe data

    Which one of the following choices present the task in the **CORRECT** order?
    A. 1, 2, 3, 4
    B. 1, 4, 3, 2
    C. 1, 3, 2, 4
    D. 1, 2, 4, 3   (2 marks)

10. Susan Keni makes the following statement about data modelling and data evaluation in the cross-industry standard process for data mining:

    "The modelling phase focuses on technical model assessment, the evaluation phase looks more broadly at which model best meets the business and what to do next."

Which one of the following choices is **CORRECT**?

A.   Only the statement on modelling phase is correct

B.   Only the statement on evaluation phase is correct

C.   The statement is correct for both modelling and evaluation phases

D.   The statement is not correct for both modelling and evaluation phases   (2 marks)

11.   Which one of the following 'V's of big data means the ability to forecast data?

A.   Veracity

B.   Vaticination

C.   Viscosity

D.   Volatility   (2 marks)

12.   What does "visualisation of data relationships" refer to in business data analytics?

A.   Comparing different variables

B.   Presenting data in tabular form

C.   Creating a dashboard

D.   Highlighting associations between data points   (2 marks)

13.   One of the challenges of data analytics is obtaining real-time data from multiple sources. Which one of the following data systems would be **MOST** effective in assisting to resolve this challenge?

A.   A data system that enables report building

B.   A comprehensive and centralised data system

C.   A data system that enables data integration

D.   A data system that collects, organises and automatically alerts on trends   (2 marks)

14.   Victoria Morosa has been asked to pair the following accounting and finance tasks with the type of data analytics technique:

Task	Technique
Analysing ratios	Descriptive
Computing ratios	Prescriptive
Preparing forecast financial statements	Diagnostic
Preparing a budget	Predictive

Which one of the following tasks is **CORRECTLY** paired with the data analytics technique?

A.   Analysing ratios

B.   Computing ratios

C.   Preparing forecast financial statements

D.   Preparing a budget   (2 marks)

15.   Which one of the following is **NOT** a type of diagnostic activity in accounting and finance?

A.   Estimating customer defaults in a certain period

B.   Estimating the probability of stockouts

C.   Computing the tax payable for a period

D.   Identifying the appropriate stock levels   (2 marks)

16. The following statements are qualities of good data visualisation:
    (i)   The visualisation should accurately represent the data and its trends
    (ii)   The reader should know what action to take after viewing your visualisation
    (iii)   Your visualisation should be easy to understand
    (iv)   Your message shouldn't take long to resonate

    Which of the statements above describes the quality of empowering?
    A.   Statement (i)
    B.   Statement (ii)
    C.   Statement (iii)
    D.   Statement (iv)   (2 marks)

17. The following statement applies to one of the data protection principles:
    "Any processing of personal data should be explained to the individual concerned on how the data is being collected, used, consulted, and the extent to which the data will be processed". This is the principle of_____.
    A.   lawfulness
    B.   fairness
    C.   transparency
    D.   integrity   (2 marks)

18. The following statement highlights a principle of data protection:
    "Personal data should be processed in a manner that ensures appropriate security and confidentiality of the personal data, including protection against unauthorised or unlawful access to or use of personal data and the equipment used for the processing and against accidental loss, destruction or damage, using appropriate technical or organisational measures.

    Which one of the following statements is **CORRECT**?
    A.   The statement only applies to data integrity
    B.   The statement only applies to data confidentiality
    C.   The statement applies to both integrity and confidentiality
    D.   The statement applies to neither integrity nor confidentiality   (2 marks)

19. Under the data protection laws, a data subject has the right to receive their data in a structured, commonly used, machine-readable format, to transmit this ported data to another data controller or processor or to request the transfer to another data controller or processor where possible.
    This is referred to as the right to data _____.
    A.   access
    B.   portability
    C.   data erasure
    D.   data rectification   (2 marks)

20. Data mapping is an essential aspect of managing the risks associated with data and it is recommended for data protection. Which one of the following is **NOT** a component of data mapping?
    A.   Data modelling
    B.   Data migration
    C.   Data integration
    D.   Data transformation   (2 marks)

CPA POST ADVANCED LEVEL

BUSINESS DATA ANALYTICS (PRACTICAL PAPER)

FRIDAY: 23 August 2024. Morning Paper.   **Time allowed: 3 hours.**

Answer ALL questions in SECTION I and SECTION II and any one (1) question in SECTION III. SECTION I has twenty (20) Multiple Choice Questions each carrying one (1) mark. SECTION II has three (3) practical questions each carrying twenty (20) marks. SECTION III has two (2) practical questions each carrying twenty (20) marks. Answer any one (1) question out of the two (2) questions in section III based on your specialisation.

Under SECTION II and SECTION III, you are required to create Ms Excel Worksheets and input your workings and solutions. You may use the Excel template within the question.

SECTION I (20 MARKS)

Answer ALL questions in this section.

QUESTION ONE
Which one of the following is the **BEST** reason why an organisation may avoid cloud computing as a means of data management?
A.   Data costs
B.   Data scalability
C.   Data integrity
D.   Data safety

QUESTION TWO
The following are the main examples of data visualisation:
(i)   Comparison
(ii)   Composition
(iii)   Relationship
(iv)   Distribution

Which one of the following summarises the order of the examples from simple to complex?
A.   Distribution, comparison, composition, relationship
B.   Distribution, composition, comparison, relationship
C.   Comparison, composition, distribution, relationship
D.   Relationship, composition, distribution, comparison

QUESTION THREE
The following activities apply to data visualisation:
(i)   The format
(ii)   The content
(iii)   The structure
(iv)   The purpose

Which one of the following is the correct order of activities?
A.   (iii), (iv), (ii) and (i)
B.   (iv), (iii), (i) and (ii)
C.   (iii), (ii), (iv) and (i)
D.   (iv), (ii), (iii) and (i)

QUESTION FOUR
A method of storing data within a system that facilitates collocation of data in various schemata and structural forms is known as _____.
A.   data visualisation
B.   data lake
C.   big data management
D.   deep analytics

QUESTION FIVE
The Kenya Data Protection Act requires that a Data Protection Impact Assessment (DPIA) be carried out where a data processing activity creates high risk to the rights and freedoms of a data subject. Which one of the following is an example of high-risk data activity?
A.   Customer profiling
B.   Customer scoring
C.   Customer inquiry
D.   Customer feedback

QUESTION SIX
The following statements are on data security and data privacy:

(i)   Data security is the process of protecting corporate data and preventing data loss through unauthorised access.
(ii)   Data privacy aims to ensure that data is not shared within and outside the organisation.

Which one of the following statements is **TRUE**?
A.   Only statement (i) is correct
B.   Only statement (ii) is correct
C.   Both statements are correct
D.   None of the statements is correct

QUESTION SEVEN
Even when intentions are good, the outcome of data analysis can cause inadvertent harm to individuals or groups of people. In data ethics, this concept is referred to as _____.
A.   diverse impact
B.   disparate impact
C.   dissimilar impact
D.   contrasting impact

QUESTION EIGHT
According to cross-industry standard process for data mining, data modelling involves:
A.   Obtaining data and information from different sources, processing and storing for future reference
B.   Fixing or removing incorrect, corrupted, incorrectly formatted data and information
C.   Collecting data and information about business requirements from stakeholders and end users
D.   Creating a visual representation of either a whole information system or parts of it to communicate connections between data points and structures

QUESTION NINE
Which one of the following is **NOT** an array function in Ms Excel?
A.   MMULT
B.   TRANSPOSE
C.   VLOOKUP
D.   FREQUENCY

QUESTION TEN

Miriam Wambui is the Chief Finance Officer of Zelta Company Limited. She is using data analytics in estimating future risks that the company is facing and also cash budgeting with scenario analysis.

Which one of the following choices is **CORRECT**?

By carrying out risk management and cash budgeting, the category is classified as _____.
A.   predictive analytics for risk management and cash budgeting
B.   predictive analytics for risk management and prescriptive analytics for cash budgeting
C.   predictive analytics for cash budgeting and prescriptive analytics for risk management
D.   prescriptive analytics for risk management and cash budgeting

QUESTION ELEVEN

Which one of the following is **NOT** a strategy the data analyst will use in handling missing data to enable data analysis?
A.   Drop observations that have missing values
B.   Estimate missing value using other observations
C.   Assign an integer that will enable the values be processed as null values
D.   Highlighting the observation with a tag such as not available

QUESTION TWELVE

Which one of the following keyboard shortcuts is used to display the "function Arguments" dialog box in Ms Excel?
A.   Ctrl + F2
B.   Ctrl + Shift + F3
C.   Ctrl + Shift + A
D.   Ctrl + Shift + #

QUESTION THIRTEEN

Which one of the following software is **NOT** used in data visualisation?
A.   Python
B.   Power BI
C.   Tableau
D.   Datawrapper

QUESTION FOURTEEN

Which one of the following is the **BEST** way to address the issue of transparency in data analytics?
A.   Use proprietary algorithms for analysis
B.   Only use numerical data for analysis
C.   Document the analysis process and results
D.   Use a single data source for analysis

QUESTION FIFTEEN

Which one of the following Ms Excel functions is used to count the number of cells in a range that meet a certain criteria?
A.   COUNTA
B.   COUNTIF
C.   COUNT
D.   COUNTBLANK

QUESTION SIXTEEN

Which one of the following software will a data analyst use to clean data?
A.   Data stage
B.   SQL
C.   Tableu
D.   Microstrategy

QUESTION SEVENTEEN

Which one of the following Ms Excel features would allow the creation of a drop-down list of values that users can select from?
A.   Trace precedents
B.   Data validation
C.   Data table
D.   Text to column

QUESTION EIGHTEEN

Which one of the following cloud computing platforms is known for its data management and analytics capabilities?

A.   Azure
B.   AWS
C.   IBM Cloud
D.   Google Cloud

QUESTION NINETEEN

In accounting, which type of analytics is used to detect anomalies or irregularities in financial transactions?

A.   Predictive analytics
B.   Descriptive analytics
C.   Diagnostic analytics
D.   Prescriptive analytics

QUESTION TWENTY

Which Ms Excel tool can Alex Mutua use to figure out the interest rate at which he would make monthly payments of Sh.50,000 for the Sh.2 million bank loan he took for 5 years?

A.   NPV
B.   Vlookup
C.   Scenario manager
D.   Goal seek

CPA POST ADVANCED LEVEL

BUSINESS DATA ANALYTICS (PRACTICAL PAPER)

FRIDAY: 26 April 2024. Morning Paper.   **Time allowed: 3 hours.**

Answer ALL questions in SECTION I and SECTION II. SECTION I has twenty (20) Multiple Choice Questions each carrying one (1) mark. SECTION II has three (3) practical questions each carrying twenty (20) marks. SECTION III has two (2) practical questions each carrying twenty (20) marks. Answer any one (1) question out of the two (2) questions in section III.

Under SECTION II and SECTION III, you are required to create Ms Excel Worksheets with the name of the entity in each question and input your workings and solutions. You may use the Excel template within the question.

SECTION I (20 MARKS)

Answer ALL questions in this section.

QUESTION ONE
Emma Atieno would like to determine the number of times revenues have exceeded Sh.12 million over the past 15 years. If the revenues are listed vertically in column A (From Cell A2 to Cell A16) of Excel, which of the following formulas will provide the correct output?
A.   COUNTIF(A2:A11,"=12")
B.   COUNTIF(A2:A11,">12")
C.   COUNTIF(A2:J11, ">12")
D.   COUNTIF(">12", A2:A11)

ANSWER: B

QUESTION TWO
A data analyst would like to determine the total revenue generated by a business in the first quarter of the year provided in Excel. Which of the following syntax in Excel will provide the correct output?
A.   =SUM([sum_range] range, criteria)
B.   = SUMIF(range, criteria, [sum_range])
C.   =SUMIFS([sum_range] range, criteria)
D.   = SUMIF(Criteria, range, [sum_range)

ANSWER: B

QUESTION THREE
Which one of the following data analytical tools will require the **LEAST** programming activity by a data analyst?
A.   Tableau
B.   Python
C.   R
D.   Power BI

ANSWER: A

QUESTION FOUR

Which one of the following "Vs" of big data means the ability to generate more copies of data?
A.   Value
B.   Validity
C.   Veracity
D.   Vagueness

ANSWER: C

QUESTION FIVE

The following statements are made by two data analysts regarding the "Vs" of big data:
Jane: "Visibility: Often the only way customers interact with models"
John: "Visualisation: Data science provides images into complex big data problems"

Which one of the following statements is **TRUE**?
A.   Only Jane is correct
B.   Only John is correct
C.   Both Jane and John are correct
D.   Both Jane and John are not correct

ANSWER: D

QUESTION SIX

Which one of the following approaches to data collection is more **EFFICIENT** in data cleaning?
A.   Online administered questionnaire
B.   Email administered questionnaire
C.   Physical administered questionnaire
D.   Interview

ANSWER: A

QUESTION SEVEN

Which one of the following ethical issues in data analytics pertains to the potential misuse of personal data for profiling or targeting?
A.   Data security
B.   Data cleaning
C.   Data privacy
D.   Data aggregation

ANSWER: C

QUESTION EIGHT

Which one of the following data visualisation charts presents comparison over many periods of time?
A.   Scatter graph
B.   Bubble chart
C.   Circular area chart
D.   Line chart

ANSWER: C

QUESTION NINE

The following activities are involved in the data understanding stage in the cross-industry standard process for data mining, **EXCEPT**_____.
A.   Checking for data completeness
B.   Checking for errors in data
C.   Checking for missing values in data
D.   Identifying appropriate data modelling

ANSWER: D

QUESTION TEN

You are given the following statements about data mining:

1.   Descriptive data mining is a type of analysis that extracts data that may help determine an outcome.
2.   Prescriptive data mining is a type of analysis that informs users of data of a given outcome.

Which one of the following statements is **CORRECT**?

A.   Only Statement 1 is correct
B.   Only statement 2 is correct
C.   Both statements are correct
D.   Both statements are not correct

ANSWER: D

QUESTION ELEVEN

The following statements are made about the qualities of good data visualisation:

1.   The visualisation should accurately represent the data and its trends
2.   The reader should know what action to take after viewing your visualisation
3.   Your visualisation should be easy to understand
4.   Your message should not take long to resonate

Which one of the following statements describes the quality of empowering?

A.   Statement 1
B.   Statement 2
C.   Statement 3
D.   Statement 4

ANSWER: B

QUESTION TWELVE

A scatter graph will likely be classified as a type of _____ data visualisation.

A.   Distribution
B.   Relationship
C.   Comparison
D.   Composition

ANSWER: B

QUESTION THIRTEEN

Which one of the following statements is the **CORRECT** definition of "data transformation"?

A.   The process that removes data that does not belong in a dataset
B.   The process of checking the integrity, accuracy and structure of data
C.   The process of converting data from one format or structure to another
D.   The process of creating a visual representation of data elements

ANSWER: C

QUESTION FOURTEEN

The following are various approaches to data risk management:

1.   Data risk sharing
2.   Data risk avoidance
3.   Data risk transfer
4.   Data risk acceptance
5.   Data risk reduction

Which one of the following answers provides the order of the **MOST** to the **LEAST** effective data risk management strategy?

A.   1, 2, 3, 4, 5
B.   4, 3, 1, 2, 5
C.   2, 1, 3, 5 ,4
D.   2, 3, 1, 5, 4

ANSWER: D

QUESTION FIFTEEN

The following are examples of data protection principles:
1.   Only collect what is sufficient for your specified purpose
2.   Process data bearing in mind the safety of individuals
3.   Explain why you are collecting data

Which statement among the three is the principle of retention?
A.   Statement 1
B.   Statement 2
C.   Statement 3
D.   None of the statements

ANSWER: D

QUESTION SIXTEEN

Limo Limited obtains data from clients to provide business consulting services. Recently, a client provided data for analysis, but a dispute arose and the process to provide consultancy to the client did not proceed. The client has requested Limo Limited to delete all the data and confirm that this is done. Limo Limited should_____.
A.   Delete the data
B.   Keep the data as per the requirements of the data protection laws
C.   Provide the data to a legal custodian for safe custody
D.   Seek legal advice before deleting the data

ANSWER: A

QUESTION SEVENTEEN

When analysing corporate financial ratios in Excel, which function will return the highest or lowest value in a dataset?
A.   MAX() and MIN()
B.   LARGE() and SMALL()
C.   TOP() and BOTTOM()
D.   HIGH() and LOW()

ANSWER: B

QUESTION EIGHTEEN

Based on the principles in the Unified Ethical Frame for Big Data Analytics (Abrahams, 2015), the attribute of data analysis being transparent and inclusive is provided by which of the following?
A.   Beneficial
B.   Fairness
C.   Progressive
D.   Respectful

ANSWER: D

QUESTION NINETEEN

What is the Excel function used to calculate the covariance between two data sets stored in cells B1:B10 and C1:C10?
A.   =SD(B1:B10, C1:C10)
B.   =VAR(B1:B10, C1:C10)
C.   =COVAR(B1:B10, C1:C10)
D.   =CORREL(B1:B10, C1:C10)

ANSWER: C

QUESTION TWENTY

In the context of data analytics, what does skepticism primarily refer to?
A.   A bias in favour of accepting data findings at face value
B.   The need to question and critically evaluate data sources and results
C.   The over-reliance on machine learning algorithms
D.   A positive attitude toward data-driven decision-making

ANSWER: B

RATES OF TAX (for employment income including wife's employment, self-employment and professional income).

Year of income 2023.

Assume the following rates of tax applied throughout the year of income 2023:

Monthly taxable pay (Sh.)			Annual taxable pay (Sh.)			Rate of tax % in each Sh.
1	-	24,000	1	-	288,000	10%
24,001	-	32,333	288,001	-	388,000	25%
32,334	-	500,000	388,001	-	6,000,000	30%
500,001	-	800,000	6,000,001	-	9,600,000	32.5%
Excess over		800,000	Excess over		9,600,000	35%

Personal relief Sh.2,400 per month (Sh.28,800 per annum).

Investment allowance: Capital expenditure incurred on:	Rate of investment allowance	Residual value (25% per year on equal instalments)	Prescribed benefit rates of motor vehicles provided by employer (i) Saloons, Hatch Backs and Estates		
				Monthly rates (Sh.)	Annual rates (Sh.)
(a) Buildings:			Up to — 1200 cc	3,600	43,200
• Hotel building	50% in the first year of use	25%	1201 - 1500 cc	4,200	50,400
• Building used for manufacture	50% in the first year of use	25%	1501 - 1750 cc	5,800	69,600
• Hospital buildings	50% in the first year of use	25%	1751 - 2000 cc	7,200	86,400
• Petroleum or gas storage facilities	50% in the first year of use	25%	2001 - 3000 cc	8,600	103,200
• Educational/hostels building	10% per year on straight line basis		Over - 3000 cc	14,400	172,800
• Commercial building	10% per year on straight line basis				
(b) Machinery:			(ii) Pick-ups, Panel Vans (unconverted)		
• Machinery used for manufacture	50% in the first year of use	25%	Up to - 1750 cc	3,600	43,200
• Hospital equipment	50% in the first year of use	25%	Over - 1750 cc	4,200	50,400
• Ships or aircraft	50% in the first year of use	25%			
• Motor vehicles and heavy earthmoving equipment	25% per year on straight line basis		(iii) Land Rovers/Cruisers	7,200	86,400
• Computer software, calculators, copiers and duplicating machines	25% per year on straight line basis				
• Furniture and fittings	10% per year on straight line basis				
• Telecommunication equipment	10% per year on straight line basis				
• Film equipment by a local producer	25% per year on straight line basis				
• Machinery used to undertake operations under prospecting rights and exploration under mining rights	50% in the first year of use	25%			
• Other machinery	10% per year on straight line basis				
(c) Purchase/acquisition of right to use fibre optic cable by telecommunication operation	10% per year on straight line basis				
(d) Farm works	50% in the first year of use	25%			

Commissioner's prescribed benefit rates: Services	Monthly rates (Sh.)	Annual rates (Sh.)
(i) Electricity (Communal or from a generator)	1,500	18,000
(ii) Water (Communal or from a borehole)	500	6,000
Agriculture employees: Reduced rates of benefits		
(i) Water	200	2,400
(ii) Electricity	900	10,800

…………..………………………………………………………………

CPA POST ADVANCED LEVEL

BUSINESS DATA ANALYTICS (PRACTICAL PAPER)

FRIDAY: 8 December 2023. Morning Paper.   **Time allowed: 3 hours.**

SECTION I has twenty (20) Multiple Choice Questions each carrying one (1) mark. SECTION II has two (2) practical questions each carrying thirty (30) marks. SECTION III has three (3) practical questions each carrying twenty (20) marks. Answer ALL questions in SECTION I and SECTION II. Answer any one (1) question out of the three (3) questions in section III.

Under SECTION II and SECTION III, you are required to create Ms Excel Worksheets with the name of the entity in each question and input your workings and solutions. You may use the Excel template within the question.

RATES OF TAX (Including wife's employment, self-employment and professional income rates of tax).

Year of income 2022.

Monthly taxable pay (Sh.)			Annual taxable pay (Sh.)			Rate of tax % in each Sh.
1	-	24,000	1	-	288,000	10%
24,001	-	32,333	288,001	-	388,000	25%
Excess over	-	32,333	Excess over	-	388,000	30%

Personal relief Sh.2,400 per month (Sh.28,800 per annum).

Investment allowance: Capital expenditure incurred on:	Rate of investment allowance	Residual value (25% per year on equal instalments)	Prescribed benefit rates of motor vehicles provided by employer (i) Saloons, Hatch Backs and Estates		
				Monthly rates (Sh.)	Annual rates (Sh.)
(a) Buildings:			Up to 1200 cc	3,600	43,200
• Hotel building	50% in the first year of use	25%	1201 - 1500 cc	4,200	50,400
• Building used for manufacture	50% in the first year of use	25%	1501 - 1750 cc	5,800	69,600
• Hospital buildings	50% in the first year of use	25%	1751 - 2000 cc	7,200	86,400
• Petroleum or gas storage facilities	50% in the first year of use	25%	2001 - 3000 cc	8,600	103,200
• Educational/hostels building	10% per year on straight line basis		Over - 3000 cc	14,400	172,800
• Commercial building	10% per year on straight line basis				
(b) Machinery:			(ii) Pick-ups, Panel Vans (unconverted)		
• Machinery used for manufacture	50% in the first year of use	25%	Up to - 1750 cc	3,600	43,200
• Hospital equipment	50% in the first year of use	25%	Over - 1750 cc	4,200	50,400
• Ships or aircraft	50% in the first year of use	25%			
• Motor vehicles and heavy earth moving equipment	25% per year on straight line basis				
• Computer software, calculators, copiers and duplicating machines	25% per year on straight line basis				
• Furniture and fittings	10% per year on straight line basis				
• Telecommunication equipment	10% per year on straight line basis				
• Film equipment by a local producer	25% per year on straight line basis				
• Machinery used to undertake operations under prospecting rights and exploration under mining rights	50% in the first year of use	25%			
• Other machinery	10% per year on straight line basis				
(c) Purchase/acquisition of right to use fibre optic cable by telecommunication operation	10% per year on straight line basis		(iii) Land Rovers/Cruisers	7,200	86,400
(d) Farm works	50% in the first year of use	25%			

Commissioner's prescribed benefit rates: Services	Monthly rates (Sh.)	Annual rates (Sh.)
(i) Electricity (Communal or from a generator)	1,500	18,000
(ii) Water (Communal or from a borehole)	500	6,000
Agriculture employees: Reduced rates of benefits		
(i) Water	200	2,400
(ii) Electricity	900	10,800

Answer ALL questions in this section.

QUESTION ONE
Which of the following formulas in Excel will provide a subtotal of variables provided in a list of vertically listed cells?
A.   =sum (A1:H20)
B.   =sum (A1:A20)
C.   =sum (A1;H20)
D.   =sum (A;A20)

ANSWER: B

QUESTION TWO
Which of the following shortcut keys will enable an analyst to view the formula in a cell?
A.   Ctrl + ;
B.   Ctrl +:
C.   Ctrl + '
D.   Ctrl +]

ANSWER: C

QUESTION THREE
Which of the following is the **CORRECT** one for the HLOOKUP function in excel?
A.   = HLOOKUP(lookup_value, [range_lookup], table_array, column_index_num)
B.   = HLOOKUP(lookup_value, table_array, row_index_num, [range_lookup])
C.   = HLOOKUP(lookup_value, column_index_num, [range_lookup], table_array)
D.   = HLOOKUP([range_lookup], lookup_value, table_array, column_index_num)

ANSWER: B

QUESTION FOUR
The following statements relate to the 'Vs' of big data:
1.   **Volume:** It refers to the size of big data.
2.   **Veracity:** It refers to the assurance of integrity or credibility of the data.
3.   **Value:** How useful the data is in decision making.

Which of the following statements is **CORRECT**?
A.   Only statement 1 is correct
B.   Only statement 2 is correct
C.   Only Statement 3 is correct
D.   All the statements are correct

ANSWER: A

QUESTION FIVE
Which of the following 'Vs' of data provides for data as being authentic, trustworthy and original?
A.   Veracity
B.   Value
C.   Variability
D.   Variety

ANSWER: A

QUESTION SIX
The following are activities that are included in data modelling:
1.   Physical
2.   Conceptual
3.   Logical
4.   Implementation

Which of the following represents the **CORRECT** steps in data modelling?
A.   1,2,3,4
B.   2,3,4,1
C.   2,3,1,4
D.   3,1,2,4

ANSWER: C

QUESTION SEVEN
Which of the following is **NOT** a data cleaning technique?
A.   Remove duplicates.
B.   Remove irrelevant data.
C.   Standardise capitalisation.
D.   Convert data from text to numeric

ANSWER: D

QUESTION EIGHT
Part of the Cross-Industry Standard Process for data mining requires business understanding. Which of the following is the **CORRECT** description of business understanding?
A.   Defining specific goals and requirements for data mining
B.   Obtain an overview of the available data and their quality
C.   Data preparation for further validation
D.   Data is reviewed in the context of the business

ANSWER: A

QUESTION NINE
Which of the following questions is **NOT** relevant in the understanding of data as per the Cross-Industry Standard Process for data mining?
A.   Is the data complete (does it cover all the cases required)?
B.   Does the data contain errors?
C.   Are there missing values in the data?
D.   Is the data current?

ANSWER: D

QUESTION TEN
Data analytics can be analysed at four levels:
1.   What happened?
2.   What could happen?
3.   What should happen?
4.   Why it happened?

Which of the following is the alternative description of Level 4?
A.   Predictive
B.   Descriptive
C.   Diagnostic
D.   Prescriptive

ANSWER: C

QUESTION ELEVEN
Mhasibu is using data analytics to make management decisions. Some decisions and types of analytics are as follows:
1.   Prepare variance analysis report (Descriptive Analytics)
2.   Prepare flexible budgets evaluating various scenarios (Prescriptive)

Which of the following statement(s) is (are) **TRUE**?
A.   Statement One only
B.   Statement Two only
C.   Both Statements are True
D.   None of the statements are True

ANSWER: D

QUESTION TWELVE
Which of the following software does **NOT** focus only on data visualisation tools?
A.   Google Charts
B.   Python
C.   Grafana
D.   Tableu

ANSWER: B

QUESTION THIRTEEN
A pie chart will likely be classified as what type of data visualisation?
A.   Relationship
B.   Distribution
C.   Composition
D.   Comparison

ANSWER: B

QUESTION FOURTEEN
Under data ethics, when algorithms are used to predict past decisions by human beings, this may lead to which unethical behaviour?
A.   Human bias
B.   Discrimination
C.   Intrusion to privacy
D.   Lack of transparency

ANSWER: A

QUESTION FIFTEEN
While data protection deals solely with personal data, data security is directed to protect all sorts of information, regardless of whether it's digital or analog, comes from the natural person or not.

Which of the above is **CORRECT**?
A.   The definition of data protection only
B.   The definition of data security only
C.   Both definitions
D.   None of the definitions

ANSWER: C

QUESTION SIXTEEN
Which one of the following concepts does **NOT** apply in data protection of digital platforms?
A.   Encryption of personal data.
B.   Anonymization of personal data.
C.   Pseudonymization of personal data.
D.   Biometric access

ANSWER: D

QUESTION SEVENTEEN
Which of the following will **NOT** lead to a challenge in data privacy in an organisation?
A.   Data Privacy regulations
B.   Poor data visibility
C.   Too much data
D.   Too many devices

ANSWER: A

QUESTION EIGHTEEN
Companies must do everything they can to use customer data ethically, preserve customers' privacy, keep them informed of how their data is being used, provide consumers with options to opt in or out, and walk the fine line between serving up relevant, targeted content and overstepping boundaries.
This definition will likely be?
A.   Data confidentiality
B.   Data Security
C.   Data Privacy
D.   Data protection

ANSWER: A

QUESTION NINETEEN
Which of the following is **NOT** a benefit of data analytics?
A.   Customer retention
B.   Strategic analysis
C.   Profitability
D.   Risk identification

ANSWER: C

QUESTION TWENTY
Despite the appreciation and adoption of big data and analytics by organisations, which of the following is a risk of poor use of big data and analytics?
A.   Cost optimisation
B.   Staff dependence
C.   Infrastructure upgrade
D.   Poor heuristics

ANSWER: D

....

CPA ADVANCED LEVEL

BUSINESS DATA ANALYTICS (PRACTICAL PAPER)

THURSDAY: 27 April 2023. Morning Paper.   **Time allowed: 3 hours.**

Answer ALL questions in SECTION I and any THREE (3) questions in SECTION II. SECTION I has twenty (20) Multiple Choice Questions each carrying two (2) marks. SECTION II has five (5) practical questions each carrying twenty (20) marks.

Under SECTION II, you are required to create Ms Excel Worksheets with the name of the entity in each question and input your workings and solutions. You may use the Excel template within the question.

SECTION I (40 MARKS)

Question One

The acronym "CRISP" stands for:
A.   Cross-industry standard process for data mining
B.   Cross-industry statistical process for data mining
C.   Cross-industry standard practice for data mining
D.   Cross-industry statistical package for data mining

Question Two

You are given the following formulas for computing variance and standard deviation of a population and sample:
1.   =VAR.S( )
2.   =VAR.P( )
3.   =STDEV.P( )
4.   =STDEV.S( )

Which of the following statements in **CORRECT?**
A.   Statements 1 and 3 contain the formulas for variance and standard deviation of a population
B.   Statements 2 and 4 contain the formulas for variance and standard deviation of a sample
C.   Statements 1 and 4 contain the formulas for variance and standard deviation of a sample
D.   Statements 2 and 3 contain the formulas for variance and standard deviation of a sample

Question Three

Which of the following focuses on the discovery of previously unknown properties on the data?
A.   Data mining
B.   Big data
C.   Data wrangling
D.   Data archiving

Question Four

A data analyst would like to determine the number of times revenues have exceeded Sh.10 million over the past 10 years. If the revenues are listed vertically in column A (From Cell A2 to Cell A11) of Excel, which of the following formulas will provide the correct output? :
A.   COUNTIF(A2:A11,"=10")
B.   COUNTIF(A2:A11,">10")
C.   COUNTIF(A2:J11, ">10")
D.   COUNTIF(">10", A2:A11)

Question Five

"Alteryx" is an example of a _____.
A.   Data management tool
B.   Data cleaning tool
C.   Data visualisation tool
D.   Data analysis tool

Question Six

Which of the following approaches to data collection will require **SIGNIFICANT** data cleansing?
A.   Online administered questionnaire
B.   Email administered questionnaire
C.   Physically administered questionnaire
D.   All of the above

Question Seven

Which of the following function can be used to find data with unique codes arranged in the top-most row of the dataset in MS Excel.
A.   MATCH
B.   HLOOKUP
C.   VLOOKUP
D.   SET UP

Question Eight

Which of the following reasons will make an organisation **AVOID** cloud computing as a means for data management?
A.   Data Costs
B.   Data Scalability
C.   Data Integrity
D.   Data Safety

Question Nine

Correlation analysis is an example of:
A.   Predictive analytics
B.   Prescriptive analytics
C.   Descriptive analytics
D.   Exploratory analytics

Question Ten

The following are the main examples of data visualisation:
1.   Comparison
2.   Composition
3.   Relationship
4.   Distribution

Which of the following summarises the order of the examples from simple to complex?
A.   Distribution, comparison, composition, relationship
B.   Distribution, composition, comparison, relationship
C.   Comparison, composition, distribution, relationship
D.   Relationship, composition, distribution, comparison

Question Eleven

One or more of the following activities is/are not from Phase 1 of Data Science Life Cycle.
A.   Learning the target domain
B.   Developing initial hypothesis
C.   Visualise initial hypothesis
D.   All of the above

Question Twelve

The following statements relate to the 'Vs' of big data:
1.   Variability is the evolving nature of data sources
2.   Variability is the different types of data structures

Which of the following statements is **CORRECT?**
A.   Only statement 1 is correct
B.   Only statement 2 is correct
C.   Both statements are correct
D.   Both statements are wrong

Question Thirteen

The following is considered by many to be the most important language for Data Science:
A.   Ruby
B.   R
C.   Java
D.   MS Excel 2010

Question Fourteen

Which of the following choices best represents the correct flow of data models:
A.   Conceptual, logical and physical
B.   Physical, logical and conceptual
C.   Logical, physical and conceptual
D.   None of the above

Question Fifteen

Choose the correct keyword for this definition: A graphical representation of a data set:
A.   Data Set
B.   Investigative Cycle
C.   Visualisation
D.   Data Plot

Question Seven

The_____ data model gives the data analyst the chance to gain an overview of the system to be analysed without being concerned with the details of how it will be analysed.

A.   Conceptual
B.   Logical
C.   Physical
D.   Rational

Question Seventeen

A bank collected data on visitors' viewing habits at the bank's website. Which technique can be best used to identify pages commonly viewed during the same visit to the website?
A.   Clustering
B.   Classification
C.   Association rules
D.   Panel analysis

Question Eighteen
The following statements apply to data mining:
1.   Predictive data mining is a type of analysis that extracts data that may be helpful in determining an outcome.
2.   Description data mining is a type of analysis that informs users of that data of a given outcome.

Which of the following is **CORRECT?**
A.   Only statement 1 is true
B.   Only statement 2 is true
C.   Both statements 1 and 2 are true
D.   Both statements are not true

Question Nineteen
Which of the following steps is performed by a data scientist after collecting the data?
A.   Data integration
B.   Data replication
C.   Data cleansing
D.   Data manipulation

Question Twenty
Which of the following best describes the work of a data architect?
A.   Utilise large sets of data to gather information that meets their company's needs
B.   Work with businesses to determine the best usage of the information yielded from data
C.   Develop data solutions that are optimised for performance and design applications
D.   Evaluate data to reach at logical conclusions
……………..………………………………………………………….

CPA POST ADVANCED LEVEL

BUSINESS DATA ANALYTICS (PRACTICAL PAPER)

THURSDAY: 24 August 2023. Morning Paper.   **Time allowed: 3 hours.**

Answer ALL questions in SECTION I and SECTION II. Answer any one (1) question out of the three (3) questions in section III. SECTION I has twenty (20) Multiple Choice Questions each carrying one (1) mark. SECTION II has two (2) practical questions each carrying thirty (30) marks. SECTION III has three (3) practical questions each carrying twenty (20) marks.

Under SECTION II and SECTION III, you are required to create Ms Excel Worksheets with the name of the entity in each question and input your workings and solutions. You may use the Excel template within the question.

RATES OF TAX (Including wife's employment, self-employment and professional income rates of tax).

Year of income 2022.

Monthly taxable pay (Sh.)		Annual taxable pay (Sh.)		Rate of tax % in each Sh.
1 - 24,000		1 - 288,000		10%
24,001 - 32,333		288,001 - 388,000		25%
Excess over - 32,333		Excess over - 388,000		30%

Personal relief Sh.2,400 per month (Sh.28,800 per annum).

Investment allowance: Capital expenditure incurred on:	Rate of investment allowance	Residual value (25% per year on equal instalments)	Prescribed benefit rates of motor vehicles provided by employer (i) Saloons, Hatch Backs and Estates	Monthly rates (Sh.)	Annual rates (Sh.)
(a) **Buildings:**					
• Hotel building	50% in the first year of use	25%	Up to 1200 cc	3,600	43,200
• Building used for manufacture	50% in the first year of use	25%	1201 - 1500 cc	4,200	50,400
• Hospital buildings	50% in the first year of use	25%	1501 - 1750 cc	5,800	69,600
• Petroleum or gas storage facilities	50% in the first year of use	25%	1751 - 2000 cc	7,200	86,400
• Educational/hostels building	10% per year on straight line basis		2001 - 3000 cc	8,600	103,200
• Commercial building	10% per year on straight line basis		Over - 3000 cc	14,400	172,800
(b) **Machinery:**			(ii) Pick-ups, Panel Vans (unconverted)		
• Machinery used for manufacture	50% in the first year of use	25%	Up to - 1750 cc	3,600	43,200
• Hospital equipment	50% in the first year of use	25%	Over - 1750 cc	4,200	50,400
• Ships or aircraft	50% in the first year of use	25%			
• Motor vehicles and heavy earth moving equipment	25% per year on straight line basis				
• Computer software, calculators, copiers and duplicating machines	25% per year on straight line basis				
• Furniture and fittings	10% per year on straight line basis				
• Telecommunication equipment	10% per year on straight line basis				
• Film equipment by a local producer	25% per year on straight line basis				
• Machinery used to undertake operations under prospecting rights and exploration under mining rights	50% in the first year of use	25%			
• Other machinery	10% per year on straight line basis				
(c) Purchase/acquisition of right to use fibre optic cable by telecommunication operation	10% per year on straight line basis		(iii) Land Rovers/Cruisers	7,200	86,400
(d) Farm works	50% in the first year of use	25%			

Commissioner's prescribed benefit rates: Services	Monthly rates (Sh.)	Annual rates (Sh.)
(i) Electricity (Communal or from a generator)	1,500	18,000
(ii) Water (Communal or from a borehole)	500	6,000
Agriculture employees: Reduced rates of benefits		
(i) Water	200	2,400
(ii) Electricity	900	10,800

Answer ALL questions in this section.

QUESTION ONE

How fast data is processed and accessed is represented by which of the following V's of Big Data?

A.   Velocity
B.   Validity
C.   Virality
D.   Veracity

ANSWER: A

QUESTION TWO

Data modelling is a process that helps one to understand the structures, form, and relationships of information. It requires one to first think of the model, then identify the model details and then how the model will be implemented. Which of the following is the correct name of data model implementation?

A.   Conceptual
B.   Logical
C.   Physical
D.   Rational

ANSWER: C

QUESTION THREE

The following activities are included in the Cross-Industry Standard Process for data mining:

1.   Data Preparation
2.   Data Understanding
3.   Data Deployment
4.   Business Understanding
5.   Data Evaluation
6.   Data modelling

Which of the following is the proper sequence of these activities?

A.   2,1,3,5,6,4
B.   1,3,5,6,4,2
C.   1,6,5,2,3,4
D.   4,2,1,6,5,3

ANSWER: D

QUESTION FOUR

Which of the following is **NOT** an activity in the modeling stage in the Cross-Industry Standard Process for Data Mining?

A.   Build
B.   Select
C.   Design
D.   Evaluate

ANSWER: D

QUESTION FIVE

For data visualisation, a bar chart will likely be classified as a_____.

A.   Distribution
B.   Relationship
C.   Comparison
D.   Composition

ANSWER: C

QUESTION SIX

A data analyst would like to present data in the form of a distribution. He will likely use all the following **EXCEPT**?

A.   Line chart
B.   Histogram
C.   Scatter graph
D.   Pie Chart

ANSWER: D

QUESTION SEVEN

Data Ethics are the norms of behaviour that promote appropriate judgements and accountability when acquiring, managing, or using data, with the goals of protecting civil liberties, minimising risks to individuals and society, and maximising the public good. All the following promote data ethics **EXCEPT**?

A. Complying with the law
B. Promoting transparency
C. Respecting privacy and confidentiality
D. Fairness

ANSWER: A

QUESTION EIGHT

Which of the following is **NOT** a concept of data integrity?

A. Data availability
B. Data currency
C. Data accuracy
D. Data validity

ANSWER: A

QUESTION NINE

With the high volume of data available for businesses, collecting meaningful data is a big challenge. To enable handling large volumes of data, data collection tools are critical. Which of the following will likely be a data collection software?

A. Excel
B. Python
C. R
D. Zoho

ANSWER: D

QUESTION TEN

A data analyst makes the following comments about functions in excel:

1. To determine the variance of a list of 10 numbers we use =VAR.P(A1:A10)
2. To determine the standard deviation of 10 numbers we use = VAR.P(A1:A10)
3. To determine the variance of a list of 10 numbers we use = VAR.S(A1:A10)

Which of the following statements is **TRUE**?

A. Only statement 1 is correct
B. Only statement 2 is correct
C. Only statement 3 is correct
D. All the statements are correct

ANSWER: C

QUESTION ELEVEN

The following are the various phases of data models:

1. Physical
2. Logical
3. Conceptual

Which of the following represents the correct order of the models?

A. 1,2,3
B. 2,3,1
C. 3,1,2
D. 3,2,1

ANSWER: D

QUESTION TWELVE

Which of the following techniques is **LEAST** effective in obtaining data for easy storage and analysis?

A. Interviews
B. Self-administered email questionnaire
C. Online questionnaire such as on google forms
D. Recorded interviews

ANSWER: A

QUESTION THIRTEEN

According to cross-industry standard process for data mining, data preparation involves:

A. Obtaining data and information from different sources, processing and storing for future reference.
B. Fixing or removing incorrect, corrupted, incorrectly formatted data and information.
C. Collecting data and information about business requirements from stakeholders and end users.
D. Creating a visual representation of either a whole information system or parts of it to communicate connections between data points and structures.

ANSWER: B

QUESTION FOURTEEN

The following statements are associated with big data:

(i)   Big data is a collection of data that is huge in volume, yet growing exponentially with time.
(ii) Big data that can be stored, accessed and processed in the form of fixed format is termed as structured data.
(iii) Big data that is heterogeneous and contains a combination of simple text files and images is semi structured.

Which of the following choices is correct?

A. Statements (i) and (ii) are correct
B. Statements (i) and (iii) are correct
C. Statements (ii) and (iii) are correct
D. All the statements are correct

ANSWER: A

QUESTION FIFTEEN

Which of the following is **NOT** the correct definition of one of the 'V's of Big Data?

A. Velocity: The speed at which big data is collected and processed
B. Value: The usefulness of the information obtained from analysing big data
C. Veracity: Multifactor, unstructured and dynamic nature of big data
D. Volume: The size of big data

ANSWER: C

QUESTION SIXTEEN

The following are the different categories of data analytics:

1. Diagnostic
2. Predictive
3. Descriptive
4. Prescriptive

Which of the following is the correct order of value?

A. 3,1,4,2
B. 3,4,2,1
C. 2,3,4,1
D. 1,3,2,4

ANSWER: A

QUESTION SEVENTEEN

David Mongare is the audit manager of Mkaguzi Auditors. He is carrying out analytical reviews at two stages: at planning and at substantive testing. Which type of analytics relates to the two stages?

A. Predictive analytics for planning and diagnostic analytics for substantive testing
B. Diagnostic analytics for planning and predictive analytics for substantive testing
C. Descriptive analytics for planning and diagnostic analytics for substantive testing
D. Predictive analytics for planning and descriptive analytics for substantive testing

ANSWER: D

QUESTION EIGHTEEN

Which of the following is **NOT** a strategy that the data analyst will use in handling missing data to enable data analysis:

A. Drop observations that have missing values
B. Estimate the missing value using other observations
C. Assign an integer that will enable the values to be processed as null values
D. Highlighting the observation with a tag

ANSWER: D

QUESTION NINETEEN

You are given the following categories of data visualisation:
1.   Composition
2.   Distribution
3.   Comparison
4.   Relationship
5.   Trends

Which of the following choices represents the order from simple to complex in terms of analysing data for visualisation?
A.   5,3,2,4,1
B.   4,3,2,1,5
C.   2,3,1,4,5
D.   1,5,4,3,2

ANSWER: A

QUESTION TWENTY

Which of the following data visualisation charts will likely present more than two variables over time effectively?
A.   Scatter Graph
B.   Bubble Chart
C.   Column Chart
D.   Line Chart

ANSWER: D

....

CPA ADVANCED LEVEL

BUSINESS DATA ANALYTICS (PRACTICAL PAPER)

WEDNESDAY: 14 December 2022. Morning Paper.   **Time allowed: 3 hours.**

Answer ALL questions in SECTION I and any THREE (3) questions in SECTION II. SECTION I has twenty (20) Multiple Choice Questions each carrying two (2) marks. SECTION II has five (5) practical questions each carrying twenty (20) marks.

Under SECTION II, you are required to create Ms Excel Worksheets with the name of the entity in each question and input your workings and solutions. You may use the Excel template within the question.

SECTION I - TOTAL 40 MARKS

Question One
Data models evolve from conceptual (that is; a quick, high-level view of the business requirement) to logical (where the entities involved are expanded and include more detail) and finally the physical data model, which can be implemented with a specific database provider (like Oracle and SQL Server).

Which of the following choices is correct?
A.   The entire statement is true
B.   Only the statement on conceptual data model is true
C.   Only the statement on logical data model is true
D.   Only the statement on physical data model is true   (2 marks)

Question Two
Which of the following would be more appropriate to replace the question mark in the following diagram?
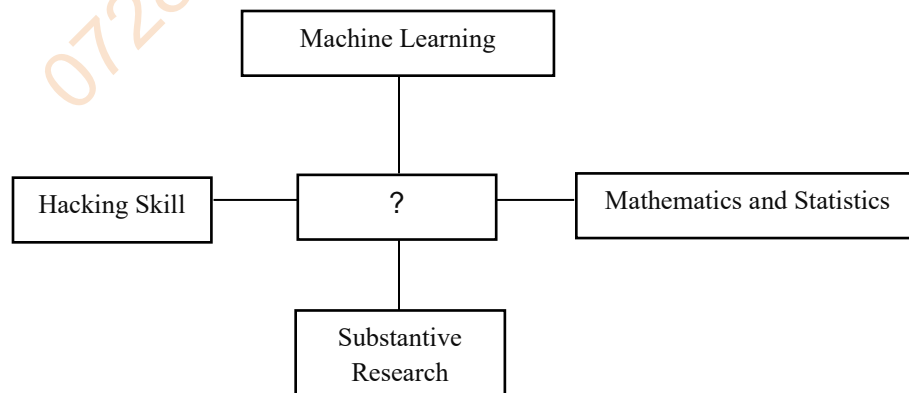
A.   Data analysis
B.   Data science
C.   Statistical inference
D.   Predictive modelling   (2 marks)

Question Three
According to cross-industry standard process for data mining, data modelling involves:
A. Obtaining data and information from different sources, processing and storing for future reference
B. Fixing or removing incorrect, corrupted, incorrectly formatted data and information
C. Collecting data and information about business requirements from stakeholders and end users
D. Creating a visual representation of either a whole information system or parts of it to communicate connections between data points and structures   (2 marks)

Question Four
Read the statements below and answer the question that follows:

(i) Data mining relates to turning raw data into useful information.

(ii) Data mining using built-in algorithms should guarantee a result.

Which of the following choices apply?
A. The two statements are true
B. Only the first statement is true
C. Only the second statement is true
D. None of the statements is true   (2 marks)

Question Five
Which of the following 'Vs' of data describes data as multifactor, unstructured and dynamic?
A. Veracity
B. Value
C. Variability
D. Variety
                    (2 marks)

Question Six
Which of the following is an example of discrete data?
A. Number of children
B. Height of children
C. Behaviour of children
D. Test scores of children   (2 marks)

Question Seven
Ms Dare Mongare is the Chief Finance Officer of Modern Company Limited. She is using data analytics in estimating future risks that the company is facing and also cash budgeting, with scenario analysis.
By carrying out risk management and cash budgeting, she is applying:
A. Predictive analytics for risk management and cash budgeting
B. Predictive analytics for risk management and prescriptive analytics for cash budgeting
C. Predictive analytics for cash budgeting and prescriptive analytics for risk Management
D. Prescriptive analytics for risk management and cash budgeting   (2 marks)

Question Eight
Based on the principles in the Unified Ethical Frame for Big Data Analytics, which of the following applies to the principle of Fairness?
A. Thinking through the potential impacts of our data use on all interested parties
B. Sustainability of the data over time
C. Transparency and inclusivity of the data
D. Data benefiting both the business and customers   (2 marks)

Question Nine
Which of the following applications will likely NOT be used for cloud computing?
A. Azure
B. AWS
C. SQL
D. Alibaba Cl   (2 marks)

Question Ten

In data science, a relationship between two entities is called _____.

A.   Binary
B.   Quartenary
C.   Unary
D.   None of the above   (2 marks)

Question Eleven

Which of the following data visualisation tools will likely present a relationship of more than two variables effectively?

A.   Scatter Graph
B.   Bubble Chart
C.   Column Chart
D.   Line Chart   (2 marks)

Question Twelve

_____ graph displays information as a series of data points connected by straight line segments.

A.   Line
B.   Bar
C.   Scatter
D.   Histogram   (2 marks)

Question Thirteen

A data breach occurs when the organisation's data suffers a security incident resulting in a breach of confidentiality, availability or integrity. According to the applicable data protection law, in the case of a data breach, the organisation should:

A.   Notify the supervisory authority within 48 hours of the incident whether or not it poses a risk to the organisation and affected individuals
B.   Notify the supervisory authority within 48 hours of the incident, only if it poses a risk to the organisation and affected individuals
C.   Notify the supervisory authority within 72 hours of the incident whether or not it poses a risk to the organisation and affected individuals
D.   Notify the supervisory authority within 72 hours of the incident, only if it poses a risk to the organisation and individuals   (2 marks)

Question Fourteen

Data fishing is sometimes referred to as

A.   Data bagging
B.   Data dredging
C.   Data merging
D.   Data pooling   (2 marks)

Question Fifteen

One of the challenges of big data is the fact that there is so much data, so many techniques and models to analyse the data and several ways to interpret the findings and results. The data analyst should therefore be more sceptical in the following aspects **EXCEPT** on the:

A.   Sources of data
B.   Data collection
C.   Data analysis
D.   Assumptions   (2 marks)

Question Sixteen

Which of the following is NOT a principle of data protection as provided by the data protection laws in various countries?

A.   Process data lawfully
B.   Maximise data collection
C.   Ensure data quality
D.   Limit data processing   (2 marks)

Question Seventeen

Two challenges of big data and analytics is inaccessibility of data and low speed of access to data. Which one of the following actions by an organisation will address the two challenges?

A. Providing access to all data to specified employees
B. Providing access to selected data to all employees
C. Restricting access to all data for specified employees
D. A database management system   (2 marks)

Question Eighteen

Which of the following is an open source revision/version control system?

A. Numpy
B. Git
C. Scipy
D. Loft   (2 marks)

Question Nineteen

Alteryx is an example of a _____.

A. Data management tool
B. Data cleaning tool
C. Data visualisation tool
D. Data presentation tool   (2 marks)

Question Twenty

Which of the following formulas in Ms Excel will provide a subtotal of variables provided in a list of vertically listed cells?

A. =sum (A1:H20)

B. =sum (A1:A20)

C. =sum (A1;H20)

D. =sum (A;A20)

SECTION II – TOTAL 60 MARKS

Question Twenty one

You are provided with the following extracts of the statement of profit or loss for Sepetuka Limited:

Sepetuka Limited
Statement of profit or loss extract for the year ended 30 September:

Year	2019 Sh."000"	2020 Sh."000"	2021 Sh."000"	2022 Sh."000"
Sales	54,000	64,800	81,000	95,580
Cost of sales	(32,400)	(32,400)	(32,400)	(38,232)
Gross profit	21,600	32,400	48,600	57,348
Operating expenses	(10,800)	(10,125)	(21,094)	(14,934)
Operating profit	10,800	22,275	27,506	42,414
Depreciation	(600)	(800)	(750)	(900)
Profit before interest and tax	10,200	21,475	26,756	41,514
Finance costs	(5,000)	(7,000)	(9,000)	(8,000)
Profit before tax	5,200	14,475	17,756	33,514
Income tax expense	(1,560)	(4,343)	(5,327)	(10,054)
Profit after tax	3,640	10,132	12,429	23,460

Required:

(a)   Calculate and interpret the following ratios:

       (i)   Annual revenue growth rates for years 2020, 2021 and 2022.   (3 marks)

       (ii)   Three years cumulative average growth rate (CAGR) for year 2022.   (3 marks)

       (iii)   Effective tax rate for the period 2019 to 2022.   (2 marks)

(b)   Now assume the following for Sepetuka Limited:

      1.   Revenue growth rates are forecast under three scenarios namely base case, optimistic case and pessimistic case. The base case growth rate for the first year forecast is the 2022 CAGR. This is expected to reduce by 2% annually until the last year of the forecast subject to a minimum of 15%. The optimistic case is 20% above the base case while the pessimistic case is 10% below the base case in all forecast periods.

      2.   Gross profit margin for the first year of the forecast is the 3-year average for the period from year 2020 to year 2022. This is expected to reduce by 2% annually until the last year of the forecast subject to a minimum of 50%.

      3.   Operating expense ratios are modelled as 3-year averages for the period from 2020 to 2022. These are assumed to remain constant over the forecast period.

      4.   Depreciation to revenue ratio is assumed to remain constant as the 3-year average for the period 2020 to 2022.

      5.   Finance costs are expected to reduce steadily as the loans are repaid. Use the reduction rate in year 2022 over the forecast period.

      6.   Income tax expense is calculated as the historical effective tax rate.

     Required:

     Prepare five-year forecast statements of profit or loss for Sepetuka Limited from year 2023 to year 2027.

                                                                    (12 marks)
                                                                (Total: 20 marks)

Question Twenty two

Mrs Jane Wakwa is the Marketing Director of Vuma Limited, a company that makes and sells electronic devices.

The company is considering the launch of a new mobile phone model branded "Trex". The available data is not fully reliable though Jane still feels that she can make a recommendation on whether or not to launch "Trex".

Additional information:

1.   Trex is estimated to have a shelf life of five years commencing year 2023.

2.   Trex will require the purchase of a machine at a cost of Sh.100 million at the end of year 2022, after which the machine will be sold for Sh.20 million at the end of the fifth year.

3.   The selling price and cost structures of Trex (for the first year 2023) with expected inflation factors are as follows:

	Sh. (Per unit)	Inflation rate (%) - from year 2024 onwards
Selling price	5,000	2%
Material costs	2,000	4%
Direct labour costs	1,000	5%
Incremental fixed cost (excludes depreciation)	500	10%

4.   The company is eligible for capital allowances (depreciation for tax purposes) at the rate of 25% on reducing balance.

5.   At the end of the project when the machine is sold, any gain or loss on disposal will be considered for tax.

6.   The tax rate on income and capital allowances is at the rate of 30% per annum. Assume that the tax for a given period is paid in the same year.

7.   The project will require an initial investment in working capital of Sh.20 million which will be increasing by Sh.5 million at the end of each year to cater for general inflation. The whole amount together with the periodic increase will, however, revert at the end of the project.

8.   Experience has shown that demand for new products is not exactly known in year one but tends to be stable thereafter. Jane has come up with the following estimates of demand for year 2023.

Probability	Expected sales (Units)
30%	40,000
40%	30,000
30%	10,000

Jane expects an initial increase in demand in year 2024 of 25% then a decline of 50% in year 2025. This level will remain the same till the end of the project.

9. Vuma Limited has a real weighted average cost of capital (WACC) of 8% and general inflation is expected to be at 4%.
   Due to the risk of the project, Jane feels that the relevant nominal WACC should be increased by 3%.

Required:
Compute the following:

(a) The weighted average cost of capital to be used to evaluate the project. (2 marks)

(b) The relevant cash flows over the project period. (15 marks)

(c) The net present value (NPV) of the project. Advise on the viability of the project. (3 marks)
(Total: 20 marks)

Question Twenty three
You are evaluating a four-year project with an initial investment of Sh.10,000,000 on 1 January 2023 and the following cash flow characteristics:

Date	Cash flow (Sh.)
1 January 2023	(10,000,000)
30 June 2023	2,750,000
31 July 2024	4,250,000
30 September 2025	3,250,000
31 December 2026	2,750,000

The discount rate is given as 8%.

Required:
(a) (i) Calculate the NPV and XNPV of the project and hence determine whether the project is viable. (6 marks)

(ii) Highlighting the cause of the difference between NPV and XNPV, explain which one you would use in your analysis and why. (4 marks)

(b) (i) Calculate the IRR and XIRR of the project and hence determine whether the project is viable. (6 marks)

(ii) Highlighting the cause of the difference between IRR and XIRR, explain which one you would use in your analysis and why. (4 marks)
(Total: 20 marks)

Question Twenty four

Bamuda Limited presented the following financial statements for the years ended 30 June 2021 and 30 June 2022:

Statement of profit or loss for the year ended 30 June 2022:

	Sh."million"	Sh."million"
Revenue		473
Cost of sales		(229)
Gross profit		244
Gain on financial assets at fair value		5
Investment income		6
		255
Other expenses		
Administration expenses	48	
Distribution costs	76	
Finance costs	17	(141)
Profit before tax		114
Income tax expense		(47)
Profit after tax for the year		67

Statement of financial position as at 30 June:

	Sh."million" 2022		Sh."million" 2021	
Assets:				
Non-current assets:				
Property, plant and equipment		327		264
Intangible assets		40		50
Financial assets through other comprehensive income (OCI)		22		10
		389		324
Current assets:				
Inventory	123		176	
Trade receivables	95		87	
Financial assets at fair value	65		30	
Cash and cash equivalents	29	312	0	293
Total assets		701		617
Equity and liabilities:				
Equity:				
Ordinary share capital (Sh.10 each)		230		150
Share premium		30		0
Revaluation reserve		36		67
Finacial assets through OCI - equity reserve		2		0
Retained profits		121		91
		419		308
Non-current liabilities:				
10% convertible loan stock		100		150
Current liabilities:				
Bank overdraft			22	
Trade payables	156		100	
Interest payable	7		3	
Income tax payable	19	182	34	159
Total liabilities		282		309
Total equity and liabilities		701		617

Additional information:

1. Property, plant and equipment held by Bamuda Limited are items of plant and equipment and freehold premises. During the year ended 30 June 2022, items of plant and equipment which originally cost Sh.40 million were disposed of resulting in a loss of Sh. 6 million charged in administrative expenses. These items had a net book value of Sh.28 million as at the date of disposal.
2. Depreciation charge for the year ended 30 June 2022 was Sh.43 million.

3. Sh.50 million of convertible loan stock was converted to Sh.50 million ordinary share capital at par during the year ended 30 June 2022.

Required:
Prepare the statement of cash flows for Bamuda Limited for the year ended 30 June 2022 as per IAS 7 "Statement of Cash Flows".

(Total 20 marks)

Question Twenty five
Farmgate Company Limited produces and sells shovels.

The company provides the following data for quantities of shovels produced and the related production cost for the calendar year 2021:

Month (year 2021)	Quantity produced (units)	Production cost Sh.
January	150,000	18,000,000
February	120,000	14,000,000
March	200,000	23,000,000
April	170,000	19,000,000
May	120,000	16,000,000
June	250,000	30,000,000
July	220,000	27,000,000
August	90,000	11,000,000
September	180,000	24,000,000
October	300,000	32,000,000
November	280,000	29,000,000
December	350,000	36,000,000

Required:

(a)   Determine the variable cost per unit and the total variable costs using regression analysis.   (6 marks)

(b)   Estimate the total production cost if the following number of units are produced:

      (i)   100,000 units

      (ii)   150,000 units

      (iii)   200,000 units

      Include the lower and upper boundaries at 95% confidence intervals.

(12 marks)

(c)   Prepare a chart plotting the predicted Y versus the actual Y from the regression function determined in (a) above. Interpret your answer.   (2 marks)
(Total 20 marks)
....

CPA ADVANCED LEVEL

BUSINESS DATA ANALYTICS (PRACTICAL PAPER)

PILOT PAPER

December 2022.   Time allowed: 3 hours.

Answer ALL questions. This paper has TWO sections. SECTION I has twenty (20) Multiple Choice Questions of two (2) marks each. SECTION II has three (3) practical questions of twenty (20) marks each.

Under SECTION II, you are required to create Ms Excel Worksheets with the name of the entity in each question and input your workings and solutions.

SECTION I – TOTAL 40 MARKS

Question One
Which of the following choices provides the correct definition of a data model?
A.   A visual representation of data objects
B.   An organisation of data objects
C.   A conceptual representation of data objects
D.   A graphical representation of data objects   (2 marks)

Question Two
The Cross Industry Standard Process for Data Mining (CRISP) requires a data analyst to understand the business as the first step. Which of the following components does **NOT** form part of this activity?
A.   Specific goals and requirements for data mining
B.   An overview of the available data and their quality
C.   Inventory and resources available
D.   Cost benefit analysis   (2 marks)

Question Three
Which of the following is the most common goal of statistical modelling?
A.   Inference
B.   Descriptive analysis
C.   Sub-setting
D.   Risk analysis   (2 marks)

Question Four
A collection of interesting and useful patterns in database is called _____.
A.   Data Set
B.   Information
C.   Knowledge
D.   Analytics   (2 marks)

Question Five
Which of the following is included in a data cleaning report?
A.   Data structure
B.   Data content
C.   Data cleaning activities
D.   Data models   (2 marks)

Question Six

Which of the following short-key functions will an analyst use in excel to paste contents into a cell?

A.   Control + F
B.   Control + V
C.   Control + C
D.   Control + U   (2 marks)

Question Seven

The following statements relate to the formula for computing the standard deviation in Ms Excel:

(i)   STDEV.P = Computes the standard deviation of an entire sample
(ii)   STDEV.S = Computes the standard deviation of an entire sample
(iii)   STDEVA = Computes the standard deviation of a selected sample

Which of the following choices is correct?

A.   Statements (i) and (ii) are correct
B.   Statements (i) and (iii) are correct
C.   Statements (ii) and (iii) are correct
D.   All the statements are correct   (2 marks)

Question Eight

Which one is not an array function in Ms Excel?

A.   MMULT
B.   TRANSPOSE
C.   VLOOKUP
D.   None of the above   (2 marks)

Question Nine

_____ is the basic data structure that can be thought of in an SQL table or a spreadsheet data representation.

A.   Series
B.   List
C.   Dataframe
D.   Points   (2 marks)

Question Ten

Which of the following components does **NOT** fall within the realm of "Velocity" in big data?

A.   Processes
B.   Streams
C.   Real-Time
D.   Terabytes   (2 marks)

Question Eleven

Which of the following 'Vs' of data provides for data as being authentic, trustworthy and original?

A.   Veracity
B.   Value
C.   Variability
D.   Variety   (2 marks)

Question Twelve

The following statements relate to the definition of the types of data analytics:

(i)   Descriptive analytics provide historical data and information
(ii)   Prescriptive analytics provide data and information for decision making
(iii)   Predictive analytics provide data and information to explain deviation from expected

Which of the following is the correct answer?

A.   Only statement (i) is correct
B.   Only statement (ii) is correct
C.   Only statement (iii) is correct
D.   Only statements (i) and (ii) are correct   (2 marks)

Question Thirteen
Which one ranks highest among the data analytics tools?
A.   Prescriptive analytics
B.   Predictive analytics
C.   Descriptive analytics
D.   Summative analytics   (2 marks)

Question Fourteen
Which of the following is **NOT** a data cleaning strategy?
A.   Removing incorrect data
B.   Removing duplicate data
C.   Fixing data with structural errors
D.   Filling in for missing data   (2 marks)

Question Fifteen
Which of the following software will a data analyst use to clean data?
A.   Data Stage
B.   SQL
C.   Tableu
D.   Microstrategy   (2 marks)

Question Sixteen
The following activities apply to data visualisation:
1.   The format
2.   The content
3.   The structure
4.   The purpose

Which of the following is the correct order of the activities?
A.   3,4,2 and 1
B.   4,3,1 and 2
C.   3,2,4, and 1
D.   4, 2,3 and 1   (2 marks)

Question Seventeen
Which of the following graph will likely present better a relationship between variables in data visualisation?
A.   Exploded Pie Chart
B.   Stacked Bar Chart
C.   Progress Line graph
D.   Radar Chart   (2 marks)

Question Eighteen
Which of the following correctly defines composition as a data visualisation technique?
A.   The method of graphical representation of data and information
B.   The method that shows the differences or similarities between values, with and without axes
C.   The method of presenting a part-to-whole relationship of data and information
D.   The method that shows relationships and connections between the data and information   (2 marks)

Question Nineteen
Which of the following people in an organisation will **LEAST** likely be involved in implementing the data protection laws in a country?
A.   Chief Risk Officer
B.   Chief Legal Officer
C.   Chief Information Officer
D.   Chief Executive Officer   (2 marks)

Question Twenty

Under the data protection laws, organisations will have to take steps to demonstrate they know the data they hold, where it is stored, and who it is shared with. This is referred to as:

A.   Data Portability
B.   Data Inventory
C.   Data Traceability
D.   Data Management   (2 marks)

SECTION II – TOTAL 60 MARKS

Question Twenty One

Amani Pharmaceuticals Limited is a mid-sized firm exploring prospects of future growth. The management of the company has prepared the statement of profit or loss for the year ended 30 June 2022 with forecasts of expected growth rates in the various line items for the next three years.

This information is presented below:

Amani Pharmaceuticals Limited
Statement of profit or loss for the year ended 30 June 2022:

		Growth Rates: Scenario		
	Sh."000"	Worst %	Base %	Best %
Revenue	4,500.00	2	5	10
Cost of goods sold	(3,200.00)	2	5	10
Gross Profit	1,300.00			
Other incomes - Investment income	15.00	2	3	4
Total Income	1,315.00			
Administration expenses	(210.00)	5	5	5
Distribution expenses	(210.00)	2	5	10
Earnings before interest, tax, depreciation and amortisation	895.00			
Depreciation and amortisation	(150.00)	5	10	10
Earnings before interest and tax	745.00			
Interest expense	(120.00)	10	10	10
Earnings before tax	625.00			
Income tax expense	(160.00)			
Profit after tax (for the period)	465.00			
Dividends paid	(280.00)	10.00	10.00	10.00
Retained Profit for the year	185.00			
Retained profit b/f	550.00			
Retained Profit c/f	735.00			
Total equity	2,200.00			
Total assets	4,500.00	5%	10%	10%

Additional information:
1.   Assume a corporation tax rate of 30%.
2.   The growth in equity is only attributable to changes in retained profits.
       Hint: In applying the growth rates, ignore changes in opening and closing balances.
3.   Total equity will grow only as a result of retained profits.

Required:
(a)   Prepare a spreadsheet showing the forecast statement of profit or loss for the next three years under each of the three
        scenarios.   (10 marks)

(b)   Compute the following metrics for evaluating performance under each scenario:

        (i)   Net profit margin.   (2 marks)

        (ii)   Return on assets.   (2 marks)

        (iii)   Return on equity.   (2 marks)

(c)   Present a summarised dash board with suitable graphs to highlight the revenues and net profit under each scenario.
                                                                                                                          (4 marks)
                                                                                                              (Total: 20 marks)

Question Twenty Two

You are provided with the following extracts of the statement of profit or loss and statement of financial position for Darubini Limited for the years ended 30 June 2020, 2021 and 2022:

Darubini Limited

Statement of profit or loss extract for the year ended 30 June (figures in Sh."000"):

	2020	2021	2022
Sales	54,000	64,800	81,000
Cost of sales	(32,400)	(32,400)	(32,400)
Gross profit	**21,600**	**32,400**	**48,600**
Operating expenses	(10,800)	(10,125)	(21,094)
Operating profit	**10,800**	**22,275**	**27,506**
Depreciation	(600)	(800)	(750)
Profit before interest and tax	**10,200**	**21,475**	**26,756**
Finance costs	(5,000)	(7,000)	(9,000)
Profit before tax	**5,200**	**14,475**	**17,756**
Income tax expense	(1,560)	(4,343)	(5,327)
Net profit	**3,640**	**10,133**	**12,429**

Statement of financial position:

Assets:

Non-current assets:

	2020	2021	2022
Property, plant and equipment	25,800	45,200	46,700
Current assets:			
Inventory	15,000	21,750	27,488
Receivables	4,500	8,750	15,460
Cash and cash equivalents	0	0	3,823
Total current assets	**19,500**	**30,500**	**46,771**
Total assets	**45,300**	**75,700**	**93,471**
Equity and Liabilities:			
Long-term liabilities:			
Long term borrowings	33,333	46,667	60,000
Current liabilities:			
Accounts payable	3,750	6,658	7,989
Bank overdraft	5,297	9,322	0
Sub-total	**9,047**	**15,980**	**7,989**
Equity:			
Paid up capital	10,000	10,000	10,000
Retained earnings	(7,080)	3,053	15,482
Total equity	**2,920**	**13,053**	**25,482**
Total equity and liabilities	**45,300**	**75,700**	**93,471**

Required:

(a) Prepare the statement of cash flows for the company for the two years ended 30 June 2021 and 30 June 2022.

(8 marks)

(b) Calculate and comment on the following key ratios:

(i) Revenue growth rates for the years ended 30 June 2021 and 30 June 2022. (2 marks)

(ii) Two years' cumulative average revenue growth rate for the year ended 30 June 2022. (2 marks)

(iii) Return on equity (ROE) for each of the three years. Discuss your results. (2 marks)

(c) Using the two-years cumulative average revenue growth rate calculated in (b) (ii) above, prepare a one year forecast of the statement of profit or loss for the year ending 30 June 2023. Assume that the vertical ratios for year ended 30 June 2022 apply for the 2023 forecast except that depreciation and finance costs remain constant. Tax is assumed to be at 30%. (6 marks)

(Total: 20 marks)

Question Twenty Three

(a) You are a financial management consultant in an investment firm. A potential client requests you to analyse an investment with the following features:

- The cost of the investment is Sh.10,000,000.

- The investment pays out a sum of Sh."X" at the end of the first year; this payout grows at the rate of 10 percent per year for 11 years.

Required:

(i) If your discount rate is 15%, calculate the minimum value of "X" that should attract the investor to make the investment. (3 marks)

(ii) The respective "NPVs" of the investment if the discount rates were to range between 10% and 15%. (3 marks)

(iii) Assume that you have three project scenarios as outlined below:

Scenario	Discount rate	Initial cost Sh.	Starting cash flow Sh.	Growth in initial cash flow
Scenario 1	15%	10,000,000	1,500,000	10%
Scenario 2	10%	15,000,000	2,000,000	8%
Scenario 3	13%	20,000,000	4,000,000	7%

Prepare detailed outcomes of the project under each of the above scenarios. (8 marks)

(b) Assume that an investor takes an investment loan of Sh.150,000,000. The loan is for 48 months at an annual interest rate of 15 percent (which the bank translates to a monthly rate of 15%/12 = 1.25%). The payments are to be made at the end of each of the next 48 months.

Required:

(i) Calculate the monthly payment on the loan. (2 marks)

(ii) Assume that you have the following options to negotiate with the bank:

Option	Annual interest rate (%)	Term
Option 1	15	48 months
Option 2	18	60 months
Option 3	13	36 months

Use a two-way data table to determine which option would be the most favourable to the investor. (4 marks)

(Total: 20 marks)

……………………………………………………………………………