NATIONAL COMPETENCY BASED MODULAR CURRICULUM

FOR

DATA MANAGEMENT AND ANALYTICS

KNQF LEVEL: 6

(CYCLE 3)

ISCED PROGRAMME CODE: 0612 554 A

TVET CDACC
P.O BOX 15745-00100
NAIROBI

Council Secretary/CEO
TVET Curriculum Development, Assessment and Certification Council
P.O. Box 15745–00100
Nairobi, Kenya
Email: info@tvetcdacc.go.ke

FOREWORD

The provision of quality education and training is fundamental to the Government's overall strategy for social economic development. Quality education and training will contribute to achievement of Kenya's development blueprint, Vision 2030 and sustainable development goals.

Reforms in the education sector are necessary for the achievement of Kenya Vision 2030 and meeting the provisions of the Constitution of Kenya 2010. The education sector had to be aligned to the Constitution of Kenya 2010 and this resulted to the formulation of the Policy Framework for Reforming Education and Training (Sessional Paper No. 4 of 2016). A key feature of this policy is the radical change in the design and instruction of the TVET training. This policy document requires that training in TVET be competency based, curriculum development be industry led, certification be based on demonstration of competence and mode of instruction allows for multiple entry and exit in TVET programmes.

These reforms demand that Industry takes a leading role in curriculum development to ensure the curriculum addresses its competence needs. It is against this background that these Occupational Standards were developed for the purpose of developing a competency-based curriculum for Data Management and Analytics level 6. These Occupational Standards will also be the basis for assessment of an individual for competence certification.

It is my conviction that these Occupational Standards will play a great role towards development of competent human resources for the ICT sector's growth and development.

PRINCIPAL SECRETARY
STATE DEPARTMENT FOR TVET
MINISTRY OF EDUCATION

PREFACE

Kenya Vision 2030 aims to transform Kenya into a newly industrializing middle-income country, providing high-quality life to all its citizens by the year 2030. Kenya intends to create globally competitive and adaptive human resource base to meet the requirements of a rapidly industrializing economy through lifelong education and training. TVET has a responsibility to facilitate the process of inculcating knowledge, skills, and worker behaviour necessary for catapulting the nation to a globally competitive country, hence the paradigm shift to embrace Competency-Based Education and Training (CBET).

CAP 210A and Sessional Paper No. 1 of 2019 on Reforming Education and Training in Kenya for Sustainable Development emphasized the need to reform curriculum development, assessment, and certification. This called for a shift to CBET to address the mismatch between skills acquired through training and skills needed by industry, as well as increase the global competitiveness of the Kenyan labour force.

This curriculum has been developed in adherence to the Kenya National Qualifications Framework and CBETA standards and guidelines. The curriculum is designed and organized into Units of Learning with Learning Outcomes, suggested delivery methods, learning resources, and methods of assessing the trainee's achievement. In addition, the units of learning have been grouped in modules to concretize the skills acquisition process and streamline upskilling.

I am grateful to all expert trainers and everyone who played a role in translating the Occupational Standards into this competency-based modular curriculum.

CHAIRPERSON
TVET CDACC

ACKNOWLEDGMENT

This curriculum has been designed for competency-based training and has independent units of learning that allow the trainee flexibility in entry and exit. In developing the curriculum, significant involvement and support were received from expert trainers, institutions and organizations.

I recognize with appreciation the role of the ICT National Sector Skills Committee (NSSC) in ensuring that competencies required by the industry are addressed in the curriculum. I also thank all stakeholders in the ICT sector for their valuable input and everyone who participated in developing this curriculum.

I am convinced that this curriculum will go a long way in ensuring that individuals aspiring to work in the ICT Sector acquire competencies to perform their work more efficiently and effectively.

COUNCIL SECRETARY/CEO
TVET CDACC

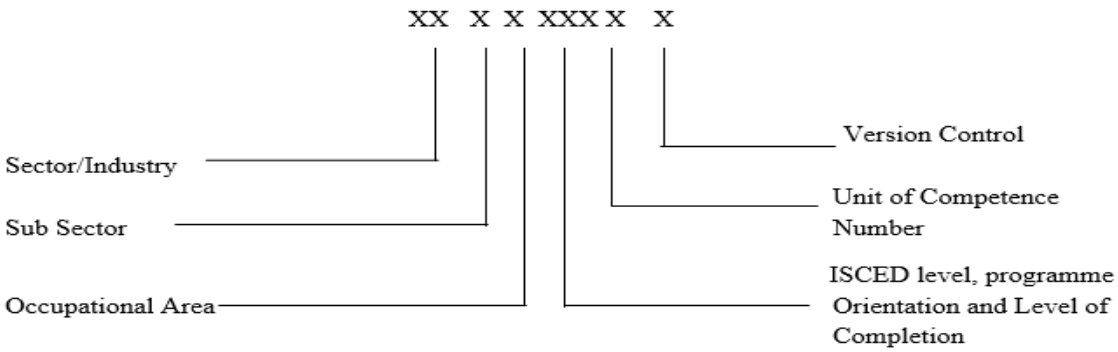
TABLE OF CONTENTS

ABBREVIATIONS AND ACRONYMS

CDACC	Curriculum Development, Assessment and Certification Council
DE	Data Engineer
EMS	Environmental Management Systems
ICT	Information Communication Technology
IEEE	Institute of Electrical and Electronics Engineers
IT	Information Technology
LAN	Local Area Network
MAN	Metropolitan Area Network
OSH	Occupational Health and Safety
PAN	Personal Area Network
RAM	Random Access Memory
POS	Parts of Speech
ROM	Read Only Memory
SQL	Structured Query Language
TVET	Technical and Vocational Education and Training
UML	Unified Modelling Language
WAN	Wide Area Network
KNQA	Kenya National Qualification Authority
KNQF	Kenya National Qualification Framework

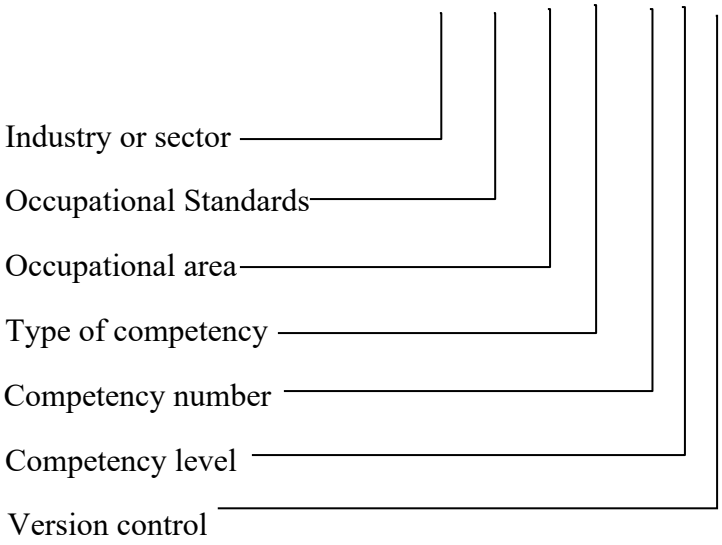
KEY TO UNIT CODE

KEY TO ISCED UNIT CODE



KEY TO TVET CDACC UNIT CODE

IT/CU/DE/BC/01/6/MA



COURSE OVERVIEW

Data Management and Analytics level 6 qualification consists of competencies that an individual must possess to perform various types of data management and analytics in the ICT sector. It involves demonstrating foundational Computer Science Skills, demonstrating Mathematical Skills for data Science, demonstrating Programming Skills Using Python, applying quantitative modelling Skills, applying Python in data Science, designing and developing databases and data warehouses, developing machine Learning application using Python, demonstrating data Mining and analytics skills in big data management, demonstrating project management skills for data science, demonstrating research skills for data science and designing and implementing cloud data base solutions.

Units of learning comprising Data Management and Analytics level 6 qualification include the following units per module:

Summary of Units of Learning

ISCED Unit Code	TVE CDACC Unit Code	Unit of Learning Title	Duration in Hours	Credit Factor
	MODULE I			
0714 551 05A	IT/CU/DE/CC/01/6/MA	Basic Electronics	100	10
0611 551 08A	IT/CU/DE/CR/01/6/MA	Foundational Computer Science Skills	160	16
0613 551 09A	IT/CU/DE/CR/03/6/MA	Programming Skills Using Python	220	22
		SUB TOTAL	**480**	
MODULE II				
0541 551 06A	IT/CU/DE/CR/02/6/MA	Mathematical Skills for Data Science	250	25
0613 551 10A	IT/CU/DE/CR/05/6/MA	Python in Data Science	190	19
		SUB TOTAL	**440**	
MODULE III				
0031 441 02A	IT/CU/DE/BC/01/6/MA	Communication skills	40	4
0542 551 07A	IT/CU/DE/CR/04/6/MA	Quantitative Modelling Skills	180	18

0613 551 14A	IT/CU/DE/CR/09/6/MA	Project Management Skills for Data Science	120	12
0613 551 15A	IT/CU/DE/CR/10/6/MA	Research Skills for Data Science	120	12
		SUB TOTAL	**460**	
MODULE IV				
0417 441 03A	IT/CU/DE/BC/02/6/MA	Work Ethics and Practices	40	4
0413 441 04A	IT/CU/DE/BC/03/6/MA	Entrepreneurial Skills	40	4
0613 551 11A	IT/CU/DE/CR/06/6/MA	Databases and Data Warehouses	190	19
0613 551 16A	IT/CU/DE/CR/11/6/MA	Cloud Database Solutions	160	16
		SUB TOTAL	**430**	
MODULE V				
0613 551 13A	IT/CU/DE/CR/08/6/MA	Data Mining and Analytics Skills in Big Data Management	180	18
0613 551 12A	IT/CU/DE/CR/02/6/MA	Machine Learning Applications Using Python	260	26
		SUB TOTAL	**440**	
		INDUSTRY TRAINING	**480**	48
		GRAND TOTAL	**2730**	**273**

Entry Requirements

An individual entering this course should have any of the following minimum requirements:

a) Kenya Certificate of Secondary Education (K.C.S.E.) with a minimum mean grade of C- (minus)

Or

b) Related level 5 certificate

c) Equivalent qualifications as determined by relevant Regulatory Authority.

Trainer qualification

Qualifications of a trainer for Data management and analytics Level 6 include:

a) Possession of a higher qualification than Data management and analytics Level 6 or in related trade area; and
b) Licence by TVETA.

Industry Training

An individual enrolled in this course will be required to undergo Industry training for a minimum period of 480 hours in Data Management and analytics organization/section in any sector. The industrial training may be taken after completion of all units for those pursuing the full qualification or be distributed equally in each unit for those pursuing part qualification. In the case of dual training model, industrial training shall be as guided by the dual training policy.

Assessment

The course shall be assessed formatively and summatively:

a) During formative assessment all performance criteria shall be assessed based on performance criteria weighting.
b) Number of formative assessments shall minimally be equal to the number of elements in a unit of competency.
c) During summative assessment basic and common units may be integrated in the core units or assessed as discrete units.
d) Theoretical and practical weighting for each unit of learning shall be 40:60.
e) Formative and summative assessments shall be weighted at 60% and 40% respectively in the overall unit of learning score
f) For a candidate to be declared competent in a unit of competency, the candidate must meet the following conditions:
    i) Obtained at least 40% in theory assessment in formative and summative assessments.
    ii) Obtained at least 60% in practical assessment in formative and summative assessment where applicable.

iii) Obtained at least 50% in the weighted results between formative assessment and summative assessment where the former constitutes 60% and the latter 40% of the overall score.

g) Assessment performance rating for each unit of competency shall be as follows:

MARKS	COMPETENCE RATING
80 -100	Attained Mastery
65 - 79	Proficient
50 - 64	Competent
49 and below	Not Yet Competent
Y	Assessment Malpractice/irregularities

h) Assessment for Recognition of Prior Learning (RPL) may lead to award of part and/or full qualification.

Certification

A candidate will be issued with a Certificate of Competency upon demonstration of competence in a core Unit of Competency. To be issued with Kenya National TVET Certificate in Data Management and Analytics level 6, the candidate must demonstrate competence in all the Units of Competency as given in the qualification pack. Statement of Attainment certificate may be awarded upon demonstration of competence in certifiable element within a unit.

These certificates will be issued by TVET CDACC

MODULE I

BASIC ELECTRONICS

ISCED UNIT CODE: 0714 551 05A

TVET CDACC UNIT CODE: IT/CU/DE/CC/01/6/MA

Relationship to Occupational Standards

This unit addresses the Unit of Competency: Apply Basic Electronic Skills

Duration of Unit: 100 hours

Unit description

This unit specifies the competencies required to demonstrate basic skills of electronics. It involves identification of electric circuits, electronic components, understand semi-conductor theory, identify and classify memories, apply number systems and identify emerging trends in electronics.

Summary of Learning Outcomes

S/No	ELEMENTS	DURATION (HOURS)
1.	Identify electrical circuits	18
2.	Identify Electronic components	20
3.	Understand Semi-conductor theory	16
4.	Classify computer memory	14
5.	Apply logic gates	12
6.	Perform Boolean Algebraic expressions	20
	TOTAL	**100**

Learning Outcomes, Content and Suggested Assessment Methods

Learning Outcomes	Content	Suggested Assessment Methods
Identify electrical circuits	1.1 Electrical circuit identification 1.1.1 Definition of electrical circuit 1.1.2 Components of electrical circuit 1.2 Electrical quantities and their S.I units' identification	• Practical • Activities • Project work • Demonstration Group discussions • Observation

	1.2.1 Basic electrical quantities and their Units 1.2.1.1 **Emf** in volts 1.2.1.2 Current in Amperes 1.2.1.3 Power in watts 1.2.1.4 Energy in joules 1.2.1.5 Resistance in ohms 1.3 Types of electrical circuits 1.3.1 AC – Alternating Current DC – Direct Current	• Third Party report • Portfolio of • Evidence Written tests
2. Identify Electronic components	2.1 Identification of electronic components 2.1.1 Resistor 2.1.2 Capacitor 2.1.3 Diode 2.1.4 Inductor 2.2 Characteristic of electronic components. 2.3 Application of electronic components. 2.4 Characteristics of integrated circuit	• Practical • Activities • Project work • Demonstration Group discussions • Observation • Third Party report • Portfolio of • Evidence • Written tests •
3. Apply semiconductor theory	3.1 Explanation of semiconductor theory 3.2 Descriptions of structure of matter 3.3 Explanation of Electrons in conductors and semiconductors 3.4 Types of semiconductor materials 3.4.1 Silicon 3.4.2 Germanium 3.5 Explanation of P-type and N-type materials 3.6 Description of P-N junction diodes 3.6.1 Forward biasing	• Practical • Activities • Project work • Demonstration • Group

	3.6.2 Reverse biasing 3.7 Types and operations of transistors 3.7.1 PNP type 3.7.2 NPN type 3.8 Application of Semiconductor theory	
4. Classify computer memory	4.1 Identification of computer memories 4.1.1 Definition of computer memory 4.1.2 Classification of computer memory 4.1.2.1 Primary memory 4.1.2.2 Secondary memory 4.1.3 Types of computer memories 4.1.3.1 RAM 4.1.3.2 4.1.3.2 ROM 4.1.3.3 4.1.3.3 DAM 4.2 Identification of Memory hierarchy speed 4.2.1 Registers 4.2.2 Cache memory 4.2.3 Main memory 4.2.4 Secondary storage 4.2.5 Tertiary storage 4.3 Identification of memory storage levels 4.3.1 Internal 4.3.2 Main 4.3.3 Online 4.3.4 Offline bulk 4.4 Classify computer memories as per the technology used 4.4.1 Semiconductor memory 4.4.2 Magnetic memory	• Practical • Activities • Project work • Demonstration Group discussions • Observation • Third Party report • Portfolio of • Evidence • Written tests

	4.4.3 Optical memory	
5. Apply logic gates	5.1 Identification of Logic gates 5.1.1 Definition of terms 5.1.2 Types of logic gates 5.1.2.1 AND Gate 5.1.2.2 OR Gate 5.1.2.3 NOT Gate 5.1.2.4 NAND Gate 5.1.2.5 NOR Gate 5.1.2.6 XOR Gate 5.1.2.7 XNOR Gate 5.2 Development of Logic circuits 5.3 Simplification of Logic circuits 5.3.1 Logic circuits Simplification Methods 5.3.1.1 Boolean Algebra 5.3.1.2 K-Maps 5.3.1.3 Quine-McCluskey Algorithm 5.3.1.4 Software and CAD Tools 5.4 Application of logic gates in electronic circuits	• Practical • Activities • Project work • Demonstration Group discussions • Observation • Third Party report • Portfolio of • Evidence • Written tests
6. Perform Boolean algebra operations	6.1 Key concepts in Boolean algebra 6.1.1 Boolean variables 6.1.2 Logical operations 6.1.3 Boolean expressions 6.1.4 Laws and rules of Boolean algebra 6.1.5 Truth tables 6.1.6 De Morgan's theorem 6.2 Demonstration of Boolean expressions	• Practical • Activities • Project work • Demonstration Group discussions • Observation • Third Party report

	6.3 Performance of Basic Boolean operations 6.4 Methods of simplifying Boolean expressions 6.5 Illustration of Boolean Laws and Theorems 6.6 Simplification rules for Boolean expressions	• Portfolio of Evidence • Written tests

Suggested Delivery Methods

- Instructor led facilitation using active learning strategies
- Demonstration by trainer
- Practical work by trainee
- Viewing of related videos
- Group discussions
- Direct instructions

Recommended resources for 25 trainees

S/No.	Category/Item	Description/ Specifications	Quantity	Recommended Ratio (Trainee: Item)
A	**Learning Materials**			
1.	Textbooks		5 pcs	5:1
2.	Installation manuals		5 pcs	5:1
3.	Flip Charts		5 pcs	5:1
4.	PowerPoint presentations	For trainer's use		
5.	Magazines/brochures/business cards			
B	**Learning Facilities and infrastructure**			
6.	Lecture/theory room		1	25:1

7.	Laboratory		1	25:1
C	**Consumable materials**			
8.	Printing papers		1 ream	1:20
9.	Foolscaps		1 ream	
10.	Toners		2 pcs	13:1
11.	Assorted colour of whiteboard markers			
D	**Tools and Equipment**			
12.	Computers		25 pcs	1:1
13.	Projector		1 pc	25:1
14.	Printers		2 pcs	13:1
15.	Whiteboard		1 pc	25:1
16.	Ohmmeter		5	5:1
17.	Ammeter		5	5:1
18.	Digital Multi meter		5	5:1
19.	Power supplies		5	5:1
20.	Circuits		5	5:1
21.	Semiconductor materials		10	3:1
22.	Conductors e.g., copper, gold, silver		25	1.1 1:1
23.	Insulators		5	5:1
24.	Screw Drivers		5	5:1
25.	Resistors		5	5:1
26.	Capacitors		5	5:1
27.	Logic gates		5	5:1

28.	Inductors		5	5:1
29.	Transistors		5	5:1
30.	Transformers batteries, power supplies		5	5:1
31.	Conducting wires		5	5:1

FOUNDATIONAL COMPUTER SCIENCE SKILLS

ISCED UNIT CODE: 0611 551 08A

TVET CDACC UNIT CODE: IT/CU/DE/CR/01/6/MA

Relationship to Occupational Standards

This unit addresses the Unit of Competency: Demonstrate Foundational Computer Science Skills

Duration of Unit: 160 hours

This unit covers the competencies required to demonstrate foundational computer science skills. It involves identifying computer components, performing computer arithmetic, solving digital logic, demonstrating basic networking skills, demonstrating spreadsheet skills using MS Excel and demonstrating presentation skills using MS PowerPoint

Summary of Learning Outcomes

S/No	ELEMENTS	DURATION (HOURS)
1.	Identify computer components	28
2.	Perform computer arithmetic	20
3.	Solve Digital Logic Problems	26
4.	Demonstrate networking skills	30
5.	Demonstrate spreadsheet skills using MS Excel	28
6.	Demonstrate presentation skills using MS PowerPoint	28
	TOTAL	**160**

Learning Outcomes, Content and Suggested Assessment Methods

Learning Outcome	Content	Suggested Assessment Methods
1. Identify computer components	1.1 Definition of computer components 1.1.1 Computer hardware 1.1.2 Computer software 1.2 Types of software 1.2.1 Application software 1.2.2 Systems software 1.2.3 Utility software 1.3 Components of computer hardware 1.3.1 Input devices 1.3.2 Output devices 1.3.3 CPU 1.3.4 Memory 1.3.5 Secondary storage 1.4 Functions of computer hardware components 1.5 Functions of computer software types 1.6 Installing and operating different Operating Systems 1.6.1 Windows 1.6.2 Linux 1.6.3 Unix based 1.7 Troubleshooting common I/O devices 1.7.1 Resolving computer problems using task manager 1.7.2 Resolving computer startup problems	• Practical exercises • Oral tests • Written tests • Observation

	1.7.3 Troubleshooting disk storage 1.7.4 Removing computer malware	
2. Perform computer arithmetic	2.1 Number systems 2.1.1 Types 2.1.2 Operations 2.1.3 Conversion 2.2 IEEE-based integer and floating-point representations 2.3 Integer and floating-point arithmetic 2.3.1 Addition 2.3.2 Subtraction 2.3.3 Multiplication	• Oral tests • Written tests • Observation
3. Solve Digital Logic Problems	3.1 Boolean algebra 3.1.1 Definition of Boolean algebra 3.1.2 Uses of Boolean algebra 3.1.3 Key Terminology 3.2 Boolean operations 3.2.1 AND 3.2.2 OR 3.2.3 NOT 3.2.4 NAND 3.2.5 NOR 3.2.6 EX-OR 3.2.7 EX-NOR 3.3 Writing Boolean Expressions 3.3.1 Order of basic operations 3.3.2 Symbols 3.4 Methods of simplifying Boolean expressions 3.4.1 Using algebraic functions 3.4.2 Using Truth tables	• Oral tests • Written tests • Observation

	3.4.3 Using Karnaugh Maps	
4. Demonstrate networking skills	4.1 Computer networking terminologies 4.2 Computer network components 4.3 Types of networks 4.3.1 LAN 4.3.2 MAN 4.3.3 WAN 4.3.4 PAN 4.4 Illustration of network topologies 4.4.1 Star 4.4.2 Ring 4.4.3 Mesh 4.4.4 Bus 4.5 Internet Protocols 4.5.1 HTTP 4.5.2 TCP 4.5.3 FTP 4.5.4 UDP 4.6 Network troubleshooting tools 4.6.1 Function of various network troubleshooting tools 4.6.2 Demonstration of network troubleshooting tools as per IEEE standard	• Practical exercises • Oral tests • Written tests • Observation
5. Demonstrate spreadsheet skills using MS Excel	5.1 Spreadsheets 5.1.1 Definition 5.1.2 Functions 5.1.3 Examples of spreadsheets 5.2 Creating worksheets 5.2.1 Excel interface 5.2.2 Types of data	• Practical exercises • Oral tests • Written tests • Observation

	5.2.3 Formatting of worksheet data 5.2.4 Managing a workbook 5.3 Importing and linking 5.3.1 External data sources 5.3.2 Updating source data 5.3.3 Refreshing excel worksheet 5.4 Formulas and functions in MS excel 5.4.1 Definitions 5.4.2 Applying formulas on a data set 5.4.3 demonstrating various types of excel functions 5.4.4 Using name manager 5.4.5 Applying functions on a data set 5.4.6 Using formula auditing 5.5 Excel data tools 5.5.1 Applying Data validation 5.5.2 Sorting and filtering 5.5.3 Removing duplicates 5.5.4 Consolidating worksheets 5.5.5 Performing What-if analysis 5.6 Analysing and summarising data using Pivot tables 5.6.1 Defining fields	

	5.6.2 Setting filters 5.6.3 Creating slicers and timelines 5.7 Visualising data using charts based on different sets of data 5.7.1 Customizing and formatting charts 5.7.2 Spark lines 5.7.3 Pivot charts 5.8 Exporting data in MS Excel	
6. Demonstrate presentation skills using MS PowerPoint	6.1 MS PowerPoint 6.1.1 Definition of MS PowerPoint 6.1.2 Uses of PowerPoint 6.1.3 Characteristics of effective PowerPoint presentations 6.2 Developing a PowerPoint presentation 6.2.1 Gathering content requirements 6.2.2 Choice of design elements 6.2.3 Insertion of PowerPoint slides 6.2.4 Inserting slide elements 6.2.5 Adding transitions and animations 6.3 Using different presentation views 6.3.1 Normal view 6.3.2 Slide sorter view 6.3.3 Notes page view 6.3.4 Slide show view 6.4 Creating presentation masters 6.4.1 Slide masters 6.4.2 Notes master 6.4.3 Handout master	• Practical exercises • Oral tests • Written tests • Observation

	6.5 Data importation from Microsoft word and Microsoft excel 6.6 Customizing external templates 6.7 Creating a presentation using a set of requirements	

Suggested Delivery Methods

- Instructor led facilitation using active learning strategies
- Demonstration by trainer
- Practical work by trainee
- Viewing of related videos
- Group discussions
- Direct instructions

Recommended resources for 25 trainees

S/No.	Category/Item	Description/ Specifications	Quantity	Recommended Ratio (Trainee: Item)
A	**Learning Materials**			
1.	Textbooks		5 pcs	5:1
2.	Installation manuals		5 pcs	5:1
3.	Flip Charts		5 pcs	5:1
4.	PowerPoint presentations	For trainer's use		
5.	Magazines/brochures/business cards			
B	**Learning Facilities & infrastructure**			

6.	Lecture/theory room		1	25:1
7.	Computer Laboratory		1	25:1
C	**Consumable materials**			
8.	Printing papers		1 ream	1:20
9.	Foolscaps		1 ream	
10.	Toners		2 pcs	13:1
11.	Assorted colour of whiteboard markers			
D	**Tools and Equipment**			
12.	Computers		25 pcs	1:1
13.	Projector		1 pc	25:1
14.	Printers		2 pcs	13:1
15.	Whiteboard		1 pc	25:1
16.	Ms Office		Per computer	
17.	Internet		Per computer	

PROGRAMMING SKILLS USING PYTHON

ISCED UNIT CODE: 0613 551 09A

TVET CDACC UNIT CODE: IT/CU/DE/CR/02/6/MA

Relationship to Occupational Standards

This unit addresses the Unit of Competency: Demonstrate Programming Skills Using Python

Duration of Unit: 220 hours

Unit description

This unit covers the competencies required to demonstrate programming skills using python. It involves identifying programming building blocks, working in the python environment, performing data operations, using control structures, applying functions for problem solving, demonstrating Object Oriented programming, handling errors in a program, working with files, demonstrating unit testing and demonstrating git version control

Summary of Learning Outcomes

S/No	ELEMENTS	DURATION (HOURS)
1.	Identify programming building blocks	24
2.	Work in the Python environment	20
3.	Perform data operations	26
4.	Use Control Structures	24
5.	Apply functions for problem solving	26
6.	Demonstrate Object Oriented Programming	28
7.	Handle errors in a program	18
8.	Work with files	18
9.	Demonstrate unit testing skills	18
10.	Demonstrate git version control skills	18
	TOTAL	**220**

Learning Outcomes, Content and Suggested Assessment Methods

Learning Outcome	Content	Suggested Assessment Methods

1. Identify programming building blocks	1.1 Definition of programming 1.2 Phases of program development 1.3 Key terms used in programming 1.4 Types of code 1.5 Translators used in programming 1.6 Designing program specification	• Practical tests • Oral tests • Written tests
2. Work in the Python environment	2.1 Installation of python 6.7.1 Downloading python setup 6.7.2 Running python setup 2.2 Python Programming environment 2.1.1 Using Python command line interpreter 2.1.2 Navigating the Python IDE 2.2 Features of python 2.3 Python syntax 2.3.1 The zen of python 2.3.2 Applying Python enhancement proposal 8 (PEP 8) 2.3.3 Declaring variables 2.3.4 Inserting comments 2.3.5 Using Python key words	• Practical tests • Oral tests • Written tests
3 Perform data operations	3.1 Python Data Types 3.2 Illustrating Python statements 3.3 Creating an using variables and constants 3.4 Illustrating Python Data operations 3.5 Creation of program to perform specified operation	• Practical tests • Oral tests • Written tests
4 Use Control Structures	4.1 Python Control structures 4.1.1 Decision making statements 4.1.2 Looping statements 4.2 Demonstrating uses of different control structures in python	• Practical tests • Oral tests • Written tests

	4.3 Creation of programs using control structures	
5 Apply functions for problem solving	5.1 Python functions 5.1.1 Definition 5.1.2 Structure 5.2 Types of functions 5.3 Demonstration of functions 5.3.1 Creating functions 5.3.2 Functions calling 5.3.3 Using command line arguments 5.4 Creation of programs to implement functions and their return values	• Practical tests • Oral tests • Written tests
6 Demonstrate Object Oriented Programming	6.1 Object oriented programming concepts (OOP) 6.1.1 Definition of OOP 6.1.2 Classes 6.1.3 Objects 6.1.4 Inheritance 6.1.5 Encapsulation 6.1.6 Abstraction 6.1.7 Polymorphism 6.2 Demonstrating Classes and Objects 6.2.1 Declaring attributes 6.2.2 Creating Methods 6.2.3 Creating objects 6.2.4 Calling methods 6.3 Creation of programs to implement inheritance	• Practical tests • Oral tests • Written tests
7 Handle errors in a program	7.1 Types of errors 7.2 Catching and raising errors 7.3 Creation of programs using error handling	• Practical tests • Oral tests • Written tests

8 Work with files	8.1 Open files for writing 8.2 Writing to a file 8.3 Closing a file 8.4 Reading a file	• Practical tests • Oral tests • Written tests
9 Demonstrate unit testing skills	9.1 Unit testing 9.1.1 Definition 9.1.2 Concepts 9.1.3 Steps of unit testing 9.2 Identifying test modules 9.3 Building test cases	• Practical tests • Oral tests • Written tests
10 Demonstrate git version control skills	10.1 Git version control 10.1.1 Definition 10.1.2 Installing git version control 10.1.3 Git commands 10.2 Initializing a git version control repository 10.2.1 Adding files to git version control 10.2.2 Committing files to git version control 10.2.3 Pushing files to a remote git repository	• Practical tests • Oral tests • Written tests

Suggested Delivery Methods

- Instructor led facilitation using active learning strategies
- Demonstration by trainer
- Practical work by trainee
- Viewing of related videos
- Group discussions
- Direct instructions

Recommended resources for 25 trainees

S/No.	Category/Item	Description/ Specifications	Quantity	Recommended Ratio (Trainee: Item)
A	**Learning Materials**			
1.	Textbooks		5 pcs	5:1
2.	Installation manuals		5 pcs	5:1
3.	Flip Charts		5 pcs	5:1
4.	PowerPoint presentations	For trainer's use		
5.	Magazines/brochures/business cards			
B	**Learning Facilities and infrastructure**			
6.	Lecture/theory room		1	25:1
7.	Computer Laboratory		1	25:1
C	Consumable materials			
8.	Printing papers		1 ream	1:20
9.	Foolscaps		1 ream	
10.	Toners		2 pcs	13:1
11.	Assorted colour of whiteboard markers			
D	**Tools and Equipment**			
12.	Computers		25 pcs	1:1
13.	Projector		1 pc	25:1
14.	Printers		2 pcs	13:1
15.	Whiteboard		1 pc	25:1
16.	Ms Office		Per computer	

17.	Internet		Per computer	
18.	Python IDE			
19.	git			
20.	GitHub			

MODULE II

MATHEMATICAL SKILLS FOR DATA SCIENCE

ISCED UNIT CODE: 0541 551 06A

TVET CDACC UNIT CODE: IT/CU/DE/CC/02/6/MA

Relationship to Occupational Standards
This unit addresses the Unit of Competency: Demonstrate Mathematical Skills for Data Science

Duration of Unit: 250 hours

Unit Description:
This unit covers the competencies required to demonstrate mathematical skills for data science. It involves performing Calculus Operations, performing Linear Algebra operations, analysing events using probability theory and analysing data using statistics.

Summary of Learning Outcomes

S/No	ELEMENTS	DURATION (HOURS)
1.	Perform Calculus Operations	80
2.	Perform Linear Algebra Operations	70
3.	Analyse events using Probability Theory	50
4.	Analyse data using statistics	50
	TOTAL	**250**

Learning Outcomes, Content and Suggested Assessment Methods

Learning Outcome	Content	Suggested Assessment Methods
1. Perform Calculus Operations	1.1. Calculus 1.1.1. Definition of calculus 1.1.2. Types of calculus 1.1.3. Functions 1.2. Applications of calculus in data science 1.3. Graphing of functions	• Practical tests • Oral tests • Written tests

	1.3.1 Intercepts 1.3.2 Limits 1.3.3 Graphing of functions using a graphing calculator 1.4. Differential calculus 1.4.1 Computing rate of change 1.4.2 Applying rules of derivatives 1.4.3 Optimizing derivative functions 1.4.4 Solving first and second order differential equations 1.5. Integral calculus 1.5.1 Illustrating definite and indefinite integrals 1.5.2 Solving integration problems using integration by reserve chain rule and substitution	
2. Perform Linear Algebra Operations	2.1 Definition of Linear algebra 2.2 Applications of linear algebra in data science 2.3 Solving linear equations 2.3.1 Methods of solving 2.3.2 Formation 2.4 Vectors	• Practical tests • Oral tests • Written tests

	2.4.1 Definition 2.4.2 Types 2.5 Illustrating vector operations 2.5.1 Addition 2.5.2 Subtraction 2.5.3 Multiplication 2.5.4 Scalar 2.5.5 Dot product 2.6 Matrices 2.6.1 Definition 2.6.2 Types 2.6.3 Determinant 2.7 Illustrating matrix operations 2.7.1 Addition 2.7.2 Scalar multiplication 2.7.3 Transposition 2.8 Illustrating inverse of square matrix	
3. Analyse events using Probability Theory	3.1 Key terminologies in probability 3.1.1 Samples spaces 3.1.2 events 3.1.3 sets 3.1.4 outcomes 3.2 Applications of probability theory in data science 3.3 Illustrating probability axioms and counting problems 3.4 Illustrating Permutations and combinations 3.5 Illustrating conditional probability and multiplication rule	• Practical tests • Oral tests • Written tests

4. Analyse data using statistics	4.1 Key terminologies in statistics 4.2 Applications of statistics in data science 4.3 Illustrating distribution in statistics 4.3.1 Binomial 4.3.2 Normal 4.3.3 Poison 4.4 Illustrating data representation techniques 4.4.1 Histogram 4.4.2 Pie charts 4.4.3 Scatter plot 4.4.4 Bar graph 4.5 Descriptive statistics 4.5.1 Measures of central tendency 4.5.2 Measures of spread 4.6 Illustrating measures of central tendency 4.7 Illustrating measures of spread 4.8 Inferential statistics 4.8.1 Linear regression 4.8.2 Correlation 4.8.3 Analysis of Variance 4.9 Illustration of linear regression and correlation	• Practical tests • Oral tests • Written tests

Suggested Delivery Methods

- Instructor led facilitation using active learning strategies
- Demonstration by trainer

- Practical work by trainee
- Viewing of related videos
- Group discussions
- Direct instructions

Recommended resources for 25 trainees

S/No.	Category/Item	Description/ Specifications	Quantity	Recommended Ratio (Trainee: Item)
A	**Learning Materials**			
1.	Textbooks		5 pcs	5:1
2.	Installation manuals		5 pcs	5:1
3.	Flip Charts		5 pcs	5:1
4.	PowerPoint presentations	For trainer's use		
5.	Calculator			
6.	Magazines/brochures/business cards			
B	**Learning Facilities & infrastructure**			
7.	Lecture/theory room		1	25:1
8.	Computer Laboratory		1	25:1
C	**Consumable materials**			
9.	Printing papers		1 ream	1:20
10.	Foolscaps		1 ream	
11.	Toners		2 pcs	13:1

12.	Assorted colour of whiteboard markers			
D	**Tools and Equipment**			
13.	Computers		25 pcs	1:1
14.	Projector		1 pc	25:1
15.	Printers		2 pcs	13:1
16.	Whiteboard		1 pc	25:1
17.	Ms Office		Per computer	
18.	Internet		Per computer	

PYTHON IN DATA SCIENCE

ISCED UNIT CODE: 0613 551 10A

TVET CDACC UNIT CODE: IT/CU/DE/CR/03/6/MA

Relationship to Occupational Standards

This unit addresses the Unit of Competency: Apply Python in Data Science

Duration of Unit: 190 hours

Unit Description:

This unit covers the competencies required to apply Python in Data Science. It involves identifying data science concepts, performing python data processing, performing python data visualization and performing statistical data analysis.

Summary of Learning Outcomes

S/No	ELEMENTS	DURATION (HOURS)
1.	Identify key data science concepts	80
2.	Perform Python data processing	70
3.	Perform Python data visualization	50
4.	Perform statistical data analysis	50
	TOTAL	**250**

Learning Outcomes, Content and Suggested Assessment Methods

Learning Outcome	Content	Suggested Assessment Methods
1. Identify key data science concepts	1.1. Definition of data science 1.2. Key terms used in data science 1.3. Foundations of data science 1.4. Data science libraries 1.5. Data requirements in data science	• Oral tests • Written tests • Practical tests

	1.6. Preparation of a data set from a specified requirement	
2. Perform Python data processing	2.1. Installation of Python scientific libraries 2.1.1. Installing Pandas 2.1.2. Installing Numpy 2.2. Choosing scientific libraries in processing python data 2.3. Importation of data formats using pandas 2.4. Demonstrating Exploratory Data Analysis 2.5. Demonstrating Data formatting and data type conversions 2.6. Demonstrating Data Cleaning 2.7. Demonstrating pandas operations	1. Oral tests 2. Written tests 3. Practical tests
3. Perform Python data visualization	3.1. Visualizations of python data 3.1.1. Using Matplotlib 3.1.2. Using Seaborn 3.2. Demonstrating various types of data visualisation from given data 3.3. Creating sub plots from given data 3.4. Adding visualisation elements 3.5. Creation of data visualization using a given dataset	4. Practical tests 5. Oral tests 6. Written tests

4. Perform statistical data analysis	4.1. Types of Statistics 4.1.1. Descriptive 4.1.2. Inferential 4.2. Demonstrating Descriptive Statistics 4.2.1. Measures of Dispersion 4.2.2. Measures of Central Tendency 4.3. Demonstrating Inferential statistics 4.3.1. Regression 4.3.2. Correlation 4.3.3. Confidence intervals 4.4. Publication of data science blog	7. Practical tests 8. Oral 9. Written tests

Suggested Delivery Methods

- Instructor led facilitation using active learning strategies
- Demonstration by trainer
- Practical work by trainee
- Viewing of related videos
- Group discussions
- Direct instructions

Recommended resources for 25 trainees

S/No.	Category/Item	Description/ Specifications	Quantity	Recommended Ratio (Trainee: Item)
A	**Learning Materials**			
1.	Textbooks		5 pcs	5:1
2.	Installation manuals		5 pcs	5:1

3.	Flip Charts		5 pcs	5:1
4.	PowerPoint presentations	For trainer's use		
5.	Jupyter Notebook			
6.	Magazines/brochures/business cards			
B	**Learning Facilities & infrastructure**			
7.	Lecture/theory room		1	25:1
8.	Computer Laboratory		1	25:1
C	**Consumable materials**			
9.	Printing papers		1 ream	1:20
10.	Foolscaps		1 ream	
11.	Toners		2 pcs	13:1
12.	Assorted colour of whiteboard markers			
D	**Tools and Equipment**			
13.	Computers		25 pcs	1:1
14.	Projector		1 pc	25:1
15.	Printers		2 pcs	13:1
16.	Whiteboard		1 pc	25:1
17.	Ms Office		Per computer	
18.	Internet		Per computer	

MODULE III

COMMUNICATION SKILLS

ISCED UNIT CODE: 0031 441 02A

TVET CDACC UNIT CODE: IT/CU/DE/BC/02/6/MA

Relationship to Occupational Standards
This unit addresses the Unit of Competency: Apply Communication Skills

Duration of Unit: 40 hours

Unit Description
This unit covers the competencies required to demonstrate communication skills. It involves applying communication channels, written, non-verbal, oral, and group communication skills and job entry techniques.

Summary of Learning Outcomes

S/No	ELEMENTS	DURATION (HOURS)
1.	Apply communication channels	5
2.	Apply written communication skills	8
3.	Apply non-verbal communication skills	5
4.	Apply oral communication skills	8
5.	Apply group discussion skills	7
6.	Apply job entry techniques	7
	TOTAL	**40**

Learning Outcomes, Content, and Suggested Assessment Methods

Learning Outcome	Content	Suggested Assessment Methods
1. Apply communication channels	1.1. Communication process 1.2. Principles of effective communication 1.3. Channels/medium/modes of communication 1.4. Factors to consider when selecting a channel of	• Oral questions • Written assessment • Observation • Portfolio of Evidence

Learning Outcome	Content	Suggested Assessment Methods
	communication 1.5. Barriers to effective communication 1.6. Flow/patterns of communication 1.7. Sources of information 1.8. Organizational policies	• Practical assessment • Third party report
2. Apply written communication skills	2.1 Types of written communication 2.2 Elements of communication 2.3 Organization requirements for written communication	• Oral assessment • Written assessment • Observation • Portfolio of Evidence • Practical assessment • Third party report
3. Apply non-verbal communication skills	3.1 Utilize body language and gestures 3.2 Apply body posture 3.3 Apply workplace dressing code	• Oral assessment • Written assessment • Observation • Portfolio of Evidence • Practical assessment • Third party report
4. Apply oral communication skills	4.1 Types of oral communication pathways 4.2 Effective questioning techniques 4.3 Workplace etiquette	• Oral assessment • Written assessment • Observation • Portfolio of Evidence

Learning Outcome	Content	Suggested Assessment Methods
	4.4 Active listening	• Practical assessment • Third party report
5. Apply group discussion skills	4.1 Establishing rapport 4.2 Facilitating resolution of issues 4.3 Developing action plans 4.4 Group organization techniques 4.5 Turn-taking techniques 4.6 Conflict resolution techniques 4.7 Team-work	• Oral assessemnt • Written assessment • Observation • Portfolio of Evidence • Practical assessment
6. Apply job entry techniques	6.1 Types of job opportunities 6.1.1 Self employment 6.1.2 Service provision 6.1.3 product development 6.1.4 salaried employment 6.2 Sources of job opportunities 6.3 Resume/ curriculum vitae 6.3.1 What is a CV 6.3.2 How long should a CV be 6.3.3 What to include in a AC 6.3.4 Format of CV 6.3.5 How to write a good CV 6.3.6 Don’ts of writing a CV 6.4 Job application letter	5. Observation 6. Oral assessment 7. Portfolio of evidence 8. Third party report 9. Written assessment

Learning Outcome	Content	Suggested Assessment Methods
	6.4.1 What to include 6.4.1.1 Addressing a cover letter 6.4.1.2 Signing off a cover letter 6.5 Portfolio of Evidence 6.5.1 Academic credentials 6.5.2 Letters of commendations 6.5.3 Certification of participations 6.5.4 Awards and decorations 6.6 Interview skills 6.6.1 Listening skills 6.6.2 Grooming 6.6.3 Language command 6.6.4 Articulation of issues 6.6.5 Body language 6.6.6 Time management 6.6.7 Honesty 6.7 Generally knowledgeable in current affairs and technical area	

Suggested Methods of Instruction

- Discussion
- Roleplaying
- Simulation
- Direct instruction
- Demonstration
- Field trips

Recommended Resources for 25 trainees

S/No.	Category/Item	Description/ Specifications	Quantity	Recommended Ratio (Trainee: Item)
A	**Learning Materials**			
1.	Textbooks		5 pcs	5:1
2.	Flip Charts		5 pcs	5:1
3.	PowerPoint presentations	For trainer's use		
4.	Magazines/brochures/business cards			
B	**Learning Facilities & infrastructure**			
5.	Lecture/theory room		1	25:1
6.	Computer Laboratory		1	25:1
C	Consumable materials			
7.	Printing papers		1 ream	1:20
8.	Foolscaps		1 ream	
9.	Toners		2 pcs	13:1
10.	Assorted colour of whiteboard markers			
D	**Tools and Equipment**			
11.	Computers		25 pcs	1:1
12.	Projector		1 pc	25:1
13.	Printers		2 pcs	13:1
14.	Whiteboard		1 pc	25:1
15.	Ms Office		Per computer	
16.	Internet		Per computer	

QUANTITATIVE MODELLING SKILLS

ISCED UNIT CODE: 0542 551 07A

TVET CDACC UNIT CODE: IT/CU/DE/CC/03/6/MA

Relationship to Occupational Standards

This unit addresses the Unit of Competency: Apply Quantitative Modelling Skills Structures

Duration of Unit: 180 hours

Unit Description

This unit covers the competencies required to apply quantitative modelling skills. It involves identifying key quantitative modelling concepts, performing regression modelling, performing linear programming, applying simulation modelling techniques and performing statistical quality control.

Summary of Learning Outcomes

S/No	ELEMENTS	DURATION (HOURS)
1.	Identify key quantitative modelling concepts	30
2.	Perform regression modelling	40
3.	Perform linear programming	38
4.	Apply simulation modelling techniques	38
5.	Perform statistical quality control	34
	TOTAL	**180**

Learning Outcomes, Content and Suggested Assessment Methods

Learning Outcome	Content	Suggested Assessment Methods

1. Identify key Quantitative Modelling Concepts	1.1. Definition of quantitative modelling 1.2. Key terms in quantitative modelling 1.3. Quantitative modelling techniques 1.3.1 Definition 1.3.2 Application	• Written tests • Oral tests
2 Perform Regression Modelling	2.3 Types of regression models 2.3.1 Linear 2.3.2 Logistic 2.3.3 Polynomial 2.3.4 Ridge 2.3.5 Lasso 2.4 Key assumptions in regression 2.5 Illustration of linear regression modelling 2.5.1 Identifying the variables 2.5.2 Defining the regression equation 2.5.3 Solving the equation 2.5.4 Interpreting model 2.6 Evaluation of the regression model 2.6.1 R^2 2.6.2 F-test 2.7 Creating linear regression models to solve real world problems	• Written tests • Oral tests • Practical tests
3 Perform Linear Programming	3.1 Illustration of Linear programming using graphical method 3.1.1 Steps of graphing 3.1.2 Finding the optimal solution	• Written tests • Oral tests • Practical tests

	3.2 Illustration of Linear programming modelling using simplex method 3.2.1 Simplex method steps 3.2.2 Finding the optimal solution 3.3 Creating linear programming models to solve a real world problems.	
4 Apply simulation modelling techniques	4.1 Types of simulation models 4.1.1 Monte Carlo 4.1.2 Agent-based 4.1.3 Discrete events 4.1.4 System dynamics 4.2 Illustration of Monte Carlo Simulation model 4.2.1 Steps of Monte Carlo simulation modelling 4.2.2 Interpreting a Monte Carlo model 4.3 Creation of a Monte Carlo financial simulation model	• Written tests • Oral tests • Practical tests
5. Perform statistical quality control	5.1 Statistical quality control (SQC) 5.1.1 Definition 5.1.2 Purpose of statistical quality control 5.2 Key terms in statistical quality control 5.3 Specification of typical quality problems relating to manufacturing	• Written tests • Oral tests • Practical tests

	5.4 Selection and application of statistical quality control tools on given data 5.5 Application of Statistical quality controls are in the Lean Six Sigma Process 5.5.1 Definition of Six Sigma 5.5.2 Application of SQC tools in the Six Sigma DMAIC phases	

Suggested Delivery Methods

- Instructor led facilitation using active learning strategies
- Demonstration by trainer
- Practical work by trainee
- Viewing of related videos
- Group discussions
- Direct instructions

Recommended Resources for 25 trainees

S/No.	Category/Item	Description/ Specifications	Quantity	Recommended Ratio (Trainee: Item)
A	**Learning Materials**			
1.	Textbooks		5 pcs	5:1
2.	Flip Charts		5 pcs	5:1
3.	PowerPoint presentations	For trainer's use		
4.	Magazines/brochures/business cards			

5.	Calculator			
6.	Mathematical tables			
B	**Learning Facilities & infrastructure**			
7.	Lecture/theory room		1	25:1
8.	Computer Laboratory		1	25:1
C	**Consumable materials**			
9.	Printing papers		1 ream	1:20
10.	Foolscaps		1 ream	
11.	Toners		2 pcs	13:1
12.	Assorted colour of whiteboard markers			
D	**Tools and Equipment**			
13.	Computers		25 pcs	1:1
14.	Projector		1 pc	25:1
15.	Printers		2 pcs	13:1
16.	Whiteboard		1 pc	25:1
17.	Internet		Per compute r	

PROJECT MANAGEMENT SKILLS FOR DATA SCIENCE

ISCED UNIT CODE: 0613 551 14A

TVET CDACC UNIT CODE: IT/CU/DE/CR/07/6/MA

Relationship to Occupational Standards
This unit addresses the Unit of Competency: Demonstrate Project Management Skills for Data Science

Duration of Unit: 120 hours

Unit Description

This unit covers the competencies required to demonstrate project management skills for data science. It involves identifying concepts in project management, demonstrating business understanding using CRISP-DM, demonstrating data understanding using CRISP-DM, demonstrating data preparation skills using CRISP-DM, demonstrating data modelling and evaluation skills using CRISP-DM and deploying data mining model using CRISP-DM.

Summary of Learning Outcomes

S/No	ELEMENTS	DURATION (HOURS)
1.	Identify concepts in project management	18
2.	Demonstrate business understanding using CRISP-DM	18
3.	Demonstrate data understanding using CRISP-DM	18
4.	Demonstrate data preparation skills using CRISP-DM	20
5.	Demonstrate data modelling and evaluation skills using CRISP-DM	26
6.	Deploy data mining model using CRISP-DM	20
	TOTAL	**120**

Learning Outcomes, Content and Suggested Assessment Methods

Learning Outcome	Content	Suggested Assessment Methods
1. Identify concepts in project management	1.1. Project Management 1.1.1. Definition 1.1.2. Key terms 1.2. Project Management Methodologies 1.3. Comparing Project Management Software features 1.4. Selecting appropriate project management methodology and software	1. Oral tests 2. Written tests 3. Practical tests
2. Demonstrate business understanding using CRISP-DM	2.1 Identifying project output 2.1.1 Setting objectives 2.1.2 Specifying business success criteria 2.2 Assessing the current situation 2.2.1 Inventory of resources 2.2.2 Requirements, assumptions and constraints 2.2.3 Risk and contingencies 2.2.4 Cost and benefits 2.3 Identifying data miming goals 2.3.1 Business criteria 2.3.2 Specifying data mining success criteria 2.4 Creating a project plan using selected project management software	4. Oral tests 5. Written tests 6. Practical tests
3. Demonstrate data understanding using CRISP-DM	3.1 Performing data collection 3.1.1 Identifying data sources	7. Oral tests 8. Written tests 9. Practical tests

	3.1.2 Selecting collection methods 3.2 Describing data 3.2.1 Data format 3.2.2 Quantity of data 3.3 Exploring data 3.3.1 Distribution 3.3.2 Relationships between pairs 3.3.3 Simple aggregations 3.3.4 Properties of significant sub-populations 3.3.5 Simple statistical analysis 3.4 Performing data verification 3.5 Preparation of data quality report	
4. Demonstrate data preparation skills using CRISP-DM	4.1 Selecting data selection 4.1.1 Row data selection 4.1.2 Column data selection 4.2 Performing data cleaning 4.2.1 Selection of clean subset of data 4.2.2 Insertion of suitable defaults 4.2.3 Estimation of missing data 4.3 Demonstration of data construction 4.3.1 Derived attributes 4.3.2 Generated records 4.4 Demonstration of data integration 4.4.1 Merging data 4.4.2 Aggregating data	10. Oral tests 11. Written tests 12. Practical tests
5. Demonstrate data modelling and	5.1 Selecting modelling techniques	13. Oral tests 14. Written tests 15. Practical tests

evaluation skills using CRISP-DM	5.1.1 Documenting the modelling technique 5.1.2 Stating modelling assumptions 5.2 Identifying test metrics 5.2.1 Regression test metrics 5.2.2 Classification test metrics 5.2.3 Clustering test metrics 5.3 Demonstrate model building 5.3.1 Setting parameters 5.3.2 Fitting the model 5.3.3 Creating a model report 5.4 Assessing a model 5.4.1 Summarizing results 5.4.2 Revising parameter settings 5.4.3 Approving model based on business criteria 5.4.4 Review Process 5.4.5 Identifying Activities to be repeated 5.4.6 Locating missed activities 5.5 Determination of next steps 5.5.1 Listing possible actions 5.5.2 Decision making	

6. Deploy data mining model using CRISP-DM	6.1 Creating model deployment plan 6.1.1 Selecting Deployment strategies 6.1.2 Setting timelines 6.2 Creating a monitoring and maintenance plan 6.3 Creating final project report and reviewing project 6.3.1 Documenting Project successes 6.3.2 Documenting unexpected events 6.3.3 Documenting lessons learnt 6.3.4 Documenting project performance	16. Oral tests 17. Written tests 18. Practical tests

Suggested Delivery Methods

- Instructor led facilitation using active learning strategies
- Demonstration by trainer
- Practical work by trainee
- Viewing of related videos
- Group discussions
- Direct instructions

Recommended Resources for 25 trainees

S/No.	Category/Item	Description/ Specifications	Quantity	Recommended Ratio (Trainee: Item)
A	**Learning Materials**			
1.	Textbooks		5 pcs	5:1
2.	Flip Charts		5 pcs	5:1

3.	PowerPoint presentations	For trainer's use		
4.	Magazines/brochures/business cards			
5.	Appropriate project management software			
6.	Appropriate data mining tools			
B	**Learning Facilities & infrastructure**			
7.	Lecture/theory room		1	25:1
8.	Computer Laboratory		1	25:1
C	Consumable materials			
9.	Printing papers		1 ream	1:20
10.	Foolscaps		1 ream	
11.	Toners		2 pcs	13:1
12.	Assorted colour of whiteboard markers			
D	**Tools and Equipment**			
13.	Computers		25 pcs	1:1
14.	Projector		1 pc	25:1
15.	Printers		2 pcs	13:1
16.	Whiteboard		1 pc	25:1
17.	Internet		Per computer	

RESEARCH SKILLS FOR DATA SCIENCE

ISCED UNIT CODE: 0613 551 15A

TVET CDACC UNIT CODE: IT/CU/DE/CR/08/6/MA

Relationship to Occupational Standards

This unit addresses the unit of competency**:** Demonstrate research skills for data science

Duration of Unit: 120 hours

Unit Description:

This unit covers the competencies required demonstrate research skills for data science. It involves identifying foundational research concepts, selecting and using data collection methods, organizing collected data using a statistical software tool, analyzing research data using a statistical software tool and presenting and reporting of research findings.

Summary of Learning Outcomes

S/No	ELEMENTS	DURATION (HOURS)
1.	Identify foundational research concepts	20
2.	Select and use data collection methods	22
3.	Organize collected data using a statistical software tool	20
4.	Analyse research data using a statistical software tool	32
5.	Presentation and reporting of research findings	26
	TOTAL	**120**

Learning Outcomes, Content and Suggested Assessment Methods

Learning Outcome	Content	Suggested Assessment Method
1. Identify foundational research concepts	1.1 Foundational research concepts for data science 1.1.1. Definition 1.1.2. Types of research 1.1.3. Statistics in data science research	• Practical tests • Written tests • Oral tests

	1.2 Types of data in data science 1.2.1 Qualitative 1.2.2 Quantitative 1.3 Levels of measurements for data science 1.3.1 Nominal 1.3.2 Ordinal 1.3.3 Interval 1.3.4 ratio 1.4 Sources of data 1.4.1 Primary source 1.4.2 Secondary source 1.5 Choice of sources of data for research assignments	
2. Select and use data collection methods	2.1 Methods of data collection 2.1.1 Observation 2.1.2 Interviews 2.1.3 Questionnaires 2.2 Developing an interview instrument 2.3 Developing a questionnaire instrument 2.4 Demonstrating data collection from a research assignment	• Practical tests • Written tests • Oral tests
3. Organize collected data using a statistical software tool	3.1 Types of data in data Organisation 3.1.1 Structured 3.1.2 Un-structured 3.2 Demonstrating methods of data organization 3.2.1 By Location 3.2.2 Alphabetically 3.2.3 By Time 3.2.4 By Hierarchy 3.2.5 By Category 3.3 Comparing features of statistical software tools	• Practical tests • Written tests • Oral tests

	3.4 Development of codebook using R 3.5 Data Entry using Codebook 3.6 Data cleaning using R	
4. Analyse research data using a statistical software tool	4.1 Summarizing data using descriptive statistics in R 4.2 Summarizing data using Inferential statistics in R 4.3 Creating visualizations using R	• Practical tests • Written tests • Oral tests
5. Presentation and reporting of research findings	5.1 Methods of Data Presentation 5.1.1 Textual 5.1.2 Tabular 5.1.3 Graphical 5.2 Presentation of data using various methods 5.3 Creating research reports using R outputs	• Practical tests • Written tests • Oral tests

Suggested Delivery Methods

- Instructor led facilitation using active learning strategies
- Demonstration by trainer
- Practical work by trainee
- Viewing of related videos
- Group discussions
- Direct instructions

Recommended Resources for 25 trainees

S/No.	Category/Item	Description/ Specifications	Quantity	Recommended Ratio (Trainee: Item)
A	**Learning Materials**			
1.	Textbooks		5 pcs	5:1

2.	Flip Charts		5 pcs	5:1
3.	PowerPoint presentations	For trainer's use		
4.	Magazines/brochures/business cards			
5.	R studio			
6.	Other statistical software			
B	**Learning Facilities & infrastructure**			
7.	Lecture/theory room		1	25:1
8.	Computer Laboratory		1	25:1
C	**Consumable materials**			
9.	Printing papers		1 ream	1:20
10.	Foolscaps		1 ream	
11.	Toners		2 pcs	13:1
12.	Assorted colour of whiteboard markers			
D	**Tools and Equipment**			
13.	Computers		25 pcs	1:1
14.	Projector		1 pc	25:1
15.	Printers		2 pcs	13:1
16.	Whiteboard		1 pc	25:1
17.	Internet		Per computer	

MODULE IV

WORK ETHICS AND PRACTICES

ISCED UNIT CODE: 0417 441 03A

TVET CDACC UNIT CODE: IT/CU/DE/BC/03/6/MA

Relationship to Occupational Standards

This unit addresses the Unit of Competency: Apply work ethics and practices.

Duration of Unit: 40 hours

Unit Description

This unit covers competencies required to demonstrate employability skills. It involves the ability to: conduct self-management, promote ethical work practices and values, promote teamwork, manage workplace conflicts, maintain professional and personal development, apply problem-solving, and promote customer care.

Summary of Learning Outcomes

S/No	ELEMENTS	DURATION (HOURS)
1.	Apply self-management skills	7
2.	Promote ethical work practices and values	7
3.	Promote Team work	6
4.	Maintain professional and personal development	6
5.	Apply Problem solving skills	7
6.	Promote Customer Care	7
	TOTAL	**40**

Learning Outcomes, Content, and Suggested Assessment Methods

Learning Outcome	Content	Suggested Assessment Methods
1. Apply self-management skills	1.1. Self-awareness 1.2. Formulating personal vision, mission, and goals	● Observation ● Written assessment ● Oral assessment ● Third party reports

Learning Outcome	Content	Suggested Assessment Methods
	1.3. Healthy lifestyle practices 1.4. Strategies for overcoming work challenges 1.5. Emotional intelligence Coping with Work Stress. Assertiveness versus aggressiveness and passiveness 1.5.1 Developing and maintaining high self-esteem 1.5.2 Developing and maintaining positive self-image 1.6. Time management 1.7. Setting performance targets 1.8. Monitoring and evaluating performance targets	● Portfolio of evidence ● Project ● Practical
2. Promote ethical work practices and values	2.1 Integrity 2.2 Core Values, ethics and beliefs 2.3 Patriotism 2.4 Professionalism 2.5 Organizational codes of conduct 2.6 Industry policies and procedures	● Observation ● Written assessment ● Oral assessment ● Third party reports ● Portfolio of evidence ● Project ● Practical
3. Promote Teamwork	3.1 Types of teams Team building Individual responsibilities in a team	● Observation ● Written assessment ● Oral assessment ● Third party reports

Learning Outcome	Content	Suggested Assessment Methods
	3.2 Determination of team roles and objectives 3.3 Team parameters and relationships 3.3.1 Benefits of teamwork 3.3.2 Qualities of a team player 3.3.3 Leading a team 3.4 Team performance and evaluation 3.5 Conflicts and conflict resolution 3.6 Gender and diversity mainstreaming 3.7 Developing Healthy workplace relationships 3.8 Adaptability and flexibility 3.9 Coaching and mentoring skills	● Portfolio of evidence ● Project ● Practical
4. Maintain professional and personal development	4.1 Personal vs professional development and growth 4.2 Avenues for professional growth 4.3 Recognizing career advancement 4.4 Training and career opportunities 4.5 Assessing training needs 4.6 Mobilizing training resources 4.7 Licenses and certifications for professional growth and development	● Observation ● Written assessment ● Oral assessment ● Third party reports ● Portfolio of evidence ● Project ● Practical

Learning Outcome	Content	Suggested Assessment Methods
	4.8 Pursuing personal and organizational goals 4.9 Managing work priorities and commitments 4.10 Dynamism and on-the-job learning	
5. Apply Problem-solving skills	5.1 Causes of problems 5.2 Methods of solving problems 5.3 Problem-solving process 5.4 Decision making 5.5 Creative thinking and critical thinking process in development of innovative and practical solutions	• Observation • Written assessment • Oral assessment • Third party reports • Portfolio of evidence • Project • Practical
6. Promote Customer Care	6.1 Identifying customer needs 6.2 Qualities of good customer service 6.3 Customer feedback methods 6.4 Resolving customer concerns 6.5 Customer outreach programs 6.6 Customer retention	• Observation • Written assessment • Oral assessment • Third party reports • Portfolio of evidence • Project • Practical

Suggested Delivery Methods

- Instructor led facilitation using active learning strategies
- Demonstration by trainer
- Practical work by trainee
- Viewing of related videos

- Group discussions
- Direct instructions

Recommended Resources for 25 trainees

S/No.	Category/Item	Description/ Specifications	Quantity	Recommended Ratio (Trainee: Item)
A	**Learning Materials**			
1.	Textbooks		5 pcs	5:1
2.	Flip Charts		5 pcs	5:1
3.	PowerPoint presentations	For trainer's use		
4.	Magazines/brochures/business cards			
B	**Learning Facilities & infrastructure**			
5.	Lecture/theory room		1	25:1
6.	Computer Laboratory		1	25:1
C	**Consumable materials**			
7.	Printing papers		1 ream	1:20
8.	Foolscaps		1 ream	
9.	Toners		2 pcs	13:1
10.	Assorted colour of whiteboard markers			
D	**Tools and Equipment**			
11.	Computers		25 pcs	1:1
12.	Projector		1 pc	25:1
13.	Printers		2 pcs	13:1
14.	Whiteboard		1 pc	25:1

15.	Internet		Per computer	

ENTREPRENEURIAL SKILLS

ISCED UNIT CODE: 0413 441 04A

TVET CDACC UNIT CODE: IT/CU/DE/BC/04/6/MA

Relationship to occupational standards

This unit addresses the unit of competency: Apply Entrepreneurial skills.

Duration of unit: 60 hours

Unit Description:

This unit covers the competencies required to demonstrate an understanding of entrepreneurship. It involves demonstrating an understanding of financial literacy, applying entrepreneurial concepts identifying entrepreneurship opportunities, applying business legal aspects, and developing business innovative strategies and business plans.

Summary of Learning Outcomes

S/No	ELEMENTS	DURATION (HOURS)
1.	Apply Financial Literacy	8
2.	Apply entrepreneurial concept	6
3.	Identify entrepreneurial opportunities	6
4.	Apply business legal aspects	6
5.	Innovate Business strategies	7
6.	Develop Business Plan	7
	TOTAL	**40**

Learning Outcomes, Content and Suggested Assessment Methods

Learning Outcome	Content	Suggested Assessment Methods
1. Apply financial literacy	1.1.Personal finance management 1.2.Balancing between needs	• Observation • Project

Learning Outcome	Content	Suggested Assessment Methods
	and wants 1.3. Budget Preparation 1.4.Saving management 1.5. Factors to consider when deciding where to save 1.6.Debt management 1.7.Factors to consider before taking a loan 1.8.Investment decisions 1.9.Types of investments 1.10. Factors to consider when investing money 1.11. Insurance services 1.12. Insurance products available in the market 1.13. Insurable risks	• Written assessment • Oral assessment • Third party report • Interviews
2.Apply entrepreneurial concept	2.1 Difference between Entrepreneurs and Business persons 2.2 Types of entrepreneurs 2.3 Ways of becoming an entrepreneur 2.4 Characteristics of Entrepreneurs 2.5 salaried employment and self-employment 2.6 Requirements for entry into self-employment 2.7 Roles of an Entrepreneur in an enterprise 2.8 Contributions of Entrepreneurship	• Observation • Project • Written assessment • Oral assessment • Third party report
3.Identify entrepreneurship opportunities	3.1 Sources of business ideas	• Observation • Project • Written

Learning Outcome	Content	Suggested Assessment Methods
	3.2 Factors to consider when evaluating business opportunity 3.3 Business life cycle	assessment • Oral assessment • Third party report
4.Apply business legal aspects	4.1 Forms of business ownership 4.2 Business registration and licensing processing 4.3 Types of contracts and agreements 4.4 Employment laws 4.5 Taxation laws	• Observation • Project • Written assessment • Oral assessment • Third party report
5.Innovate business Strategies	5.1 Creativity in business 5.2 Innovative business strategies 5.3 Entrepreneurial Linkages 5.4 ICT in business growth and development	• Observation • Project • Written assessment • Oral assessment • Third party report
6.Develop Business Plan	6.1 Business description 6.2 Marketing plan 6.3 Organizational/Management 6.4 plan 6.5 Production/operation plan 6.6 Financial plan 6.7 Executive summary 6.8 Business plan presentation 6.9 Business idea incubation	• Observation • Written assessment • Project • Oral assessment • Third party report

Suggested Delivery Methods

- Instructor led facilitation using active learning strategies
- Demonstration by trainer

- Practical work by trainee
- Viewing of related videos
- Group discussions
- Direct instructions

Recommended Resources for 25 trainees

S/No.	Category/Item	Description/ Specifications	Quantity	Recommended Ratio (Trainee: Item)
A	**Learning Materials**			
1.	Textbooks		5 pcs	5:1
2.	Flip Charts		5 pcs	5:1
3.	PowerPoint presentations	For trainer's use		
4.	Magazines/brochures/business cards			
B	**Learning Facilities & infrastructure**			
5.	Lecture/theory room		1	25:1
6.	Computer Laboratory		1	25:1
C	**Consumable materials**			
7.	Printing papers		1 ream	1:20
8.	Foolscaps		1 ream	
9.	Toners		2 pcs	13:1
10.	Assorted colour of whiteboard markers			
D	**Tools and Equipment**			

11.	Computers		25 pcs	1:1
12.	Projector		1 pc	25:1
13.	Printers		2 pcs	13:1
14.	Whiteboard		1 pc	25:1
15.	Internet		Per computer	

DATABASES AND DATA WAREHOUSES

ISCED UNIT CODE: 0613 551 11A

TVET CDACC UNIT CODE: IT/CU/DE/CR/04/6/MA

Relationship to Occupational Standards

This unit addresses the Unit of Competency: Design and Develop Databases and Data Warehouses

Duration of Unit: 190 hours

Unit Description

This unit covers the competencies required to demonstrate designing and development of databases and data warehouses. It involves identifying key database concepts, designing a relational database from given requirements, using Structured Query Language to implement a database design, designing a data warehouse and implementing a data warehouse

Summary of Learning Outcomes

S/No	ELEMENTS	DURATION (HOURS)
1.	Identify key Database concepts	28
2.	Design a relational Database from given requirements	38
3.	Use Structured Query Language to implement database design	40
4.	Design a data warehouse	42
5.	Implement a Data warehouse Design	42
	TOTAL	**190**

Learning Outcomes, Content and Suggested Assessment Methods

Learning Outcome	Content	Suggested Assessment Methods
1. Identify key database concepts	1.1. Database 1.1.1. Definition of database 1.1.2. Database terminologies 1.1.3. Reasons of using databases 1.1.4. Definition of relational model 1.2. Relational Modelling Concepts 1.2.1 Relations/tables 1.2.2 Attributes/Columns 1.2.3 Domain 1.2.4 Tuples/Rows 1.2.5 Primary Key 1.2.6 Foreign Key 1.3. Comparison of RDBMS products 1.3.1 Oracle 1.3.2 MS SQL server 1.3.3 MySQL 1.3.4 Ms Access 1.3.5 Mongo DB 1.3.6 Casandra 1.4. Installation of MS SQL server 1.5. MS SQL server interface 1.6. Properties of MS SQL server Database 1.6.1 Deleting a database 1.6.2 Deleting data or log files	• Oral questioning • Written tests • Practical tests

	1.6.3 Increasing database size 1.6.4 Shrinking database 1.6.5 Renaming database 1.6.6 Importing a database 1.6.7 Exporting a database 1.7. Prescription of RDBMS product for a simulated environment 1.8. Illustration of Object Relational Mappers	
2. Design a relational Database from given requirements	2.2 Phases of database Design Conceptual database design (ERM Modelling) 2.2.1 Logical database design 2.2.2 Physical database design 2.3 Illustrating Entity modelling Components 2.4 Designing Entity Model using UML (Unified Modelling Language) 2.5 Demonstrating Normalisation 2.6 Definition 2.7 Demonstration of normalisation 2.8 Validating model according to the requirements / specified transactions (CRUD matrix)	• Written tests • Oral questioning • Practical tests

3. Use Structured Query Language to implement database design	3.1 SQL 3.1.1 Definition 3.1.2 Characteristics 3.1.3 Components 3.2 Data definition queries 3.2.1 CREATE 3.2.2 DROP 3.2.3 ALTER 3.3 Demonstration of CREATE TABLE statement 3.4 Demonstration of CREATE TABLE constraints 3.5 Editing table schema using SQL ALTER statement 3.5.1 Adding an attribute 3.5.2 Dropping an attribute 3.5.3 Modifying attribute domain 3.6 Dropping table using SQL DROP TABLE statement 3.7 Using data manipulation query statements 3.7.1 INSERT 3.7.2 SELECT 3.7.3 UPDATE 3.7.4 DELETE 3.8 Data Manipulation Query Statements 3.8.1 Retrieving records using SELECT statement 3.8.2 Insertion of records using INSERT INTO statements 3.8.3 Deleting records using DELETE statement	• Practical • Oral questioning • Observation • Written tests

	3.8.4 Updating records using UPDATE. SET statement 3.9 SQL Joins 3.9.1 Definition of a join 3.9.2 Types of joins 3.10 Creating and querying a database from a validated ER model. 3.11 Creating a simple join from a database	
4. Design a data warehouse	4.1 Data warehouse 4.1.1 Definition of Data warehouse 4.1.2 Terminologies used in Data warehousing 4.1.3 Types of Data warehouse 4.2 Online analytical processing (OLAP) 4.2.1 Definition of OLAP 4.2.2 Illustration of OLAP 4.3 Online Transaction Processing (OLTP) 4.3.1 Definition of OLTP 4.3.2 Illustration of Relational-based OLAP 4.4 Designing Data warehouse schemas	• Practical tests • Oral tests • Written tests
5. Implement a Data warehouse Design	5.1 Data Mining Query Language (DMQL) 5.1.1 Definition of DMQL 5.1.2 Terminologies in DMQL	• Practical tests • Oral tests • Written tests

	5.1.3 Illustrating syntax of DMQL commands 5.2 Creating cubes and dimensional tables using schema specifications 5.3 Performing Extract Transform Load (ETL) operations	

Suggested Delivery Methods

- Instructor led facilitation using active learning strategies
- Demonstration by trainer
- Practical work by trainee
- Viewing of related videos
- Group discussions
- Direct instructions

Recommended Resources for 25 trainees

S/No.	Category/Item	Description/ Specifications	Quantity	Recommended Ratio (Trainee: Item)
A	**Learning Materials**			
1.	Textbooks		5 pcs	5:1
2.	Flip Charts		5 pcs	5:1
3.	PowerPoint presentations	For trainer's use		
4.	Magazines/brochures/business cards			

B	**Learning Facilities & infrastructure**			
5.	Lecture/theory room		1	25:1
6.	Computer Laboratory		1	25:1
C	Consumable materials			
7.	Printing papers		1 ream	1:20
8.	Foolscaps		1 ream	
9.	Toners		2 pcs	13:1
10.	Assorted colour of whiteboard markers			
D	**Tools and Equipment**			
11.	Computers		25 pcs	1:1
12.	Projector		1 pc	25:1
13.	Printers		2 pcs	13:1
14.	Whiteboard		1 pc	25:1
15.	Internet		Per compute r	

CLOUD DATABASE SOLUTIONS

ISCED UNIT CODE: **0613 551 16A**

TVET CDACC UNIT CODE: IT/CU/DE/CR/09/6/MA

Relationship to Occupational Standards

This unit addresses the Unit of Competency: Designing and Implementing Cloud Database Solutions

Duration of Unit: 160 hours

Unit description

This unit covers the competencies required to design and implement cloud database solutions. It involves identifying key concepts of cloud computing, designing and implementing database solutions for SQL Server, monitoring and troubleshooting database implementation in Azure.

Summary of Learning Outcomes

S/No	ELEMENTS	DURATION (HOURS)
1.	Identify key concepts of cloud computing	28
2.	Design and implement database solutions for SQL Server and Microsoft Azure	48
3.	Manage, design and implement database security and privacy	44
4.	Monitor and Troubleshoot Database implementation in Azure	40
	TOTAL	**160**

Learning Outcomes, Content and Suggested Assessment Methods

Learning Outcome	Content	Suggested Assessment Methods
1. Identify key concepts of cloud computing	1.1. Cloud computing 1.1.1. Definition 1.1.2. Characteristics 1.1.3. Benefits and challenges 1.2. Cloud computing architecture 1.3. Illustration of cloud computing technologies 1.4. Illustration of cloud computing deployment models 1.5. Illustration of Cloud computing service models 1.6. Cloud service providers for IaaS, PaaS and SaaS 1.7. Advantages and disadvantages of cloud solutions. 1.8. Prescribing deployment and service model for a simulated organization	• Written tests • Oral tests • Practical tests
2. Design and implement database solutions for SQL Server and Microsoft Azure	2.1 Microsoft Windows Azure 2.1.1 Definition 2.1.2 Available Azure services Services (Paas and IaaS) 2.2 Creating an account on the Azure portal 2.3 Illustrating Azure Components 2.3.1 Data Management 2.3.2 Identity and Access 2.3.3 Big Data and Big data Compute 2.3.4 Mobile Service 2.3.5 Back Up	• Written tests • Oral tests • Practical tests

	2.4 Development of SQL Server database solution on Azure 2.3.1 Creating a single database using the Azure portal 2.3.2 Configuring firewall rules for the database 2.3.3 Importing SQL server databases using the Data Migration Assistant (DMA) 2.5 Implementing SQL Server on Azure Virtual machines 2.5.1 Benefits of SQL Server Azure VMs 2.5.2 Demonstration of procedure for creating SQL server virtual machine in Azure portal 2.5.3 Other virtual machine types 2.5.4 Ubuntu 2.5.5 Redhat 2.5.6 Centos	
3. Manage, design and implement database security and privacy	3.1 SQL server database security issues: Specifying security authentication and authorization requirements for database and server 3.2 Implementing SQL Authentication and Authorization 3.2.1 Comparing and demonstrating Windows authentication and mixed mode authentication	• Written tests • Oral tests • Practical tests

	3.2.2 Configuring server and database roles 3.3 Azure SQL Database security capabilities 3.4 Implementing Azure SQL server Database security capabilities 3.4.1 Database security best practices in line standards including ISO 27001/27002 3.4.2 Enabling Azure SQL databases security for best practice compliance	
4. Monitor and Troubleshoot Database implementation in Azure	4.1 Identification of resources that need monitoring on Azure SQL server databases 4.2 Diagnosing database performance problems 4.2.1 Running related problems 4.2.2 Waiting related problems 4.3 Selecting and configuring monitoring tools 4.4 Using Azure portal monitoring chart 4.5 using Query Performance Insight 4.6 Enabling Automatic tuning	• Written tests • Oral tests • Practical tests

Suggested Delivery Methods

- Instructor led facilitation using active learning strategies
- Demonstration by trainer
- Practical work by trainee
- Viewing of related videos
- Group discussions

- Direct instructions

Recommended Resources for 25 trainees

S/No.	Category/Item	Description/ Specifications	Quantity	Recommended Ratio (Trainee: Item)
A	**Learning Materials**			
1.	Textbooks		5 pcs	5:1
2.	Flip Charts		5 pcs	5:1
3.	PowerPoint presentations	For trainer's use		
4.	Magazines/brochures/business cards			
B	**Learning Facilities & infrastructure**			
5.	Lecture/theory room		1	25:1
6.	Computer Laboratory		1	25:1
C	Consumable materials			
7.	Printing papers		1 ream	1:20
8.	Foolscaps		1 ream	
9.	Toners		2 pcs	13:1
10.	Assorted colour of whiteboard markers			
D	**Tools and Equipment**			
11.	Computers		25 pcs	1:1

12.	Projector		1 pc	25:1
13.	Printers		2 pcs	13:1
14.	Whiteboard		1 pc	25:1
15.	Internet		Per computer	
16.	Microsoft SQL Server			
17.	Microsoft Azure			
18.	Ubuntu			
19.	Redhat			
20.	Centos			
21.	Network devices			

MODULE V

DATA MINING AND ANALYTICS SKILLS IN BIG DATA MANAGEMENT

ISCED UNIT CODE: 0613 551 13A

TVET CDACC UNIT CODE: IT/CU/DE/CR/06/6/MA

Relationship to Occupational Standards

This unit addresses the Unit of Competency: Demonstrate Data Mining and Analytics Skills in Big Data Management

Duration of Unit: 180 hours

Unit Description

This unit covers the competencies required to demonstrate data mining and analytics skills in big data management. It involves identifying key concepts of data mining and big data, applying data mining techniques, visualizing real world big data problems, managing big data using Hadoop and recognizing ethical, social and legal issues in computing and big data.

Summary of Learning Outcomes

S/No	ELEMENTS	DURATION (HOURS)
1.	Identify key concepts of data mining and Big data	28
2.	Apply data mining techniques	44
3.	Visualize real world big data problems	28
4.	Manage Big data using Hadoop	44
5.	Recognize ethical, social and legal issues in computing and Big Data	36
	TOTAL	**180**

Learning Outcomes, Content and Suggested Assessment Methods

Learning Outcome	Content	Suggested Assessment Methods
1. Identify key concepts of data mining and Big data	1.1 Data mining 1.1.1. Definition of Data Mining 1.1.2. Terminologies used in data mining 1.1.3. Purpose for data mining 1.1.4. Mineable data 1.1.5. Techniques of data mining 1.1.6. Technologies used in data mining 1.2 Architecture of data mining 1.3 Illustrating data mining Process 1.4 Applications of data mining 1.5 Demonstration of data pre-processing from provided data 1.6 Big data 1.6.1 Definition of big data 1.6.2 The five Vs of big data 1.6.3 Big data sources 1.7 Prescribing potential big data analytics areas	• Oral tests • Written tests • Practical tests
2. Apply data mining techniques	2.1 Selecting mining technique 2.2 Selecting data mining software mining tool 2.3 Applying classification techniques 2.3.1 Performing predictions 2.3.2 Evaluating classifier 2.4 Applying regression techniques 2.4.1 Performing predictions	• Oral tests • Written tests • Practical tests

	2.4.2 Evaluating regressor 2.5 Applying clustering techniques 2.5.1 Performing predictions 2.5.2 Evaluating cluster	
3. Visualize real world big data problems	3.1 Comparison of big data visualization tools 3.2 Visualising given data using Ms Excel 3.2.1 Tables 3.2.2 Pivot Tables 3.2.3 Sparklines for trends 3.2.4 Charts 3.2.5 Pivot Charts 3.2.6 Slicers 3.2.7 Dashboards 3.2.8 Powerpivot 3.3 Visualising given data using Ms PowerBi 3.3.1 Components of PowerBi 3.3.2 Importing Data 3.3.3 Refreshing PowerBi Data 3.3.4 Building interactive visualization	• Oral tests • Written tests • Practical tests
4. Manage Big data using Hadoop	4.1 Hadoop 4.1.1 Definition of Hadoop 4.1.2 Hadoop architecture 4.2 Hadoop environment 4.2.1 Demonstrating hadoop download process 4.2.2 Hadoop operation modes 4.2.3 Installing appropriate Hadoop operation mode	• Oral tests • Written tests • Practical tests

	4.2.4 Verifying installation mode 4.2.5 Demonstrating Hadoop launching 4.3 Hadoop file system (HDFS) 4.3.1 Features 4.3.2 Illustrating HDFS architecture 4.3.3 Demonstration of HDFS operations 4.3.4 Installing snakebite package 4.4 MapReduce 4.4.1 Definition of MapReduce 4.4.2 MapReduce Algorithm 4.4.3 Demonstrating Hadoop streaming with python	
5. Recognise ethical, social and legal issues in computing and Big Data	5.1 Definition of Ethics 5.2 Legal and ethical issues in computing 5.2.1 Legal issues 5.2.2 Ethical issues 5.3 Social issues and emerging trends in computing 5.4 Big data ethical concerns and principles	• Oral tests • Written tests

Suggested Delivery Methods

- Instructor led facilitation using active learning strategies
- Demonstration by trainer
- Practical work by trainee
- Viewing of related videos
- Group discussions

- Direct instructions

Recommended Resources for 25 trainees

S/No.	Category/Item	Description/ Specifications	Quantity	Recommended Ratio (Trainee: Item)
A	**Learning Materials**			
1.	Textbooks		5 pcs	5:1
2.	Flip Charts		5 pcs	5:1
3.	PowerPoint presentations	For trainer's use		
4.	Magazines/brochures/business cards			
B	**Learning Facilities & infrastructure**			
5.	Lecture/theory room		1	25:1
6.	Computer Laboratory		1	25:1
C	**Consumable materials**			
7.	Printing papers		1 ream	1:20
8.	Foolscaps		1 ream	
9.	Toners		2 pcs	13:1
10.	Assorted colour of whiteboard markers			
D	**Tools and Equipment**			
11.	Computers		25 pcs	1:1
12.	Projector		1 pc	25:1
13.	Printers		2 pcs	13:1

14.	Whiteboard		1 pc	25:1
15.	Internet		Per computer	
16.	Microsoft SQL Server			
17.	Python IDE			
18.	Ms Excel			
19.	Ms PowerBi			
20.	Hadoop			

MACHINE LEARNING APPLICATIONS USING PYTHON

ISCED UNIT CODE: 0613 551 12A

TVET CDACC UNIT CODE: IT/CU/DE/CR/05/6/MA

Relationship to Occupational Standards

This unit addresses the Unit of Competency: Develop Machine Learning Applications Using Python

Duration of Unit: 260 hours

Unit description:

This unit covers the competencies required to develop machine learning applications. It involves identifying concepts of machine learning, developing classification-based applications, developing regression-based applications, developing clustering-based applications, applying ensemble learning techniques, using cross validation to optimize machine learning methods and demonstrating model deployment skills

Summary of Learning Outcomes

S/No	ELEMENTS	DURATION (HOURS)
1.	Identify key concepts of machine learning	36
2.	Develop classification-based applications	44
3.	Develop regression-based applications	38
4.	Develop clustering-based applications	36
5.	Apply ensemble learning techniques	36
6.	Use cross validation to optimize machine learning methods	34
7.	Demonstrate model deployment skills	36
8.	**TOTAL**	**260**

Learning Outcomes, Content and Suggested Assessment Methods

Learning Outcome	Content	Suggested Assessment Methods
1. Identify key concepts of machine learning	1.1. Machine Learning 1.1.1. Definition 1.1.2. Machine learning techniques 1.2. Key Terms in Machine Learning 1.3. Importance of data preparation 1.4. Foundations of Machine Learning 1.5. Types of Machine Learning 1.5.1 Supervised 1.5.2 Unsupervised 1.6. Applications of Machine Learning 1.7. Scikit-learn 1.7.1 Installation of Scikit-learn 1.7.2 Libraries in Scikit-learn 1.8. Prescription of Machine learning in real life problems	• Written tests • Oral tests • Practical tests
2 Develop classification-based applications	2.1 Classification 2.1.1 Definition 2.1.2 Classification algorithms 2.1.3 How classification algorithms work 2.2 Identifying features and targets from given data 2.3 Demonstrating data preprocessing 2.3.1 Label encoding 2.3.2 Scaling	• Written tests • Oral tests • Practical tests

	2.3.3 Cleaning data 2.3.4 Feature extraction 2.4 Demonstrating fitting a classifier 2.5 Demonstrating predictions 2.6 Demonstrating classifier Evaluation 2.6.1 Evaluation on the training set 2.6.2 Evaluation on the test set 2.7 Saving the classifier 2.8 Creation of classifiers using given data	
3 Develop regression based applications	3.1 Regression 3.1.1 Definition 3.1.2 Regression algorithms 3.1.3 How regression algorithms work 3.2 Identifying features and targets from given data 3.3 Demonstrating data preprocessing 3.3.1 Label encoding 3.3.2 Scaling 3.3.3 Cleaning data 3.3.4 Feature extraction 3.4 Demonstrating fitting a regressor 3.5 Demonstrating predictions 3.6 Demonstrating regressor evaluation 3.6.1 Evaluation on the training set 3.6.2 Evaluation on the test set 3.7 Saving the regressor	• Written tests • Oral tests • Practical tests

	3.8 Creation of regressors using given data	
4 Develop clustering based applications	4.1 Clustering 4.1.1 Definition 4.1.2 Clustering algorithms 4.1.3 How clustering algorithms work 4.2 Identifying features and targets from given data 4.3 Demonstrating data preprocessing 4.3.1 Label encoding 4.3.2 Scaling 4.3.3 Cleaning data 4.3.4 Feature extraction 4.4 Demonstrating fitting a cluster 4.5 Demonstrating predictions 4.6 Demonstrating cluster evaluation 4.6.1 Evaluation on the training set 4.6.2 Evaluation on the test set 4.7 Saving the cluster 4.8 Creation of clusters using given data	• Written tests • Oral tests • Practical tests
5 Apply ensemble learning techniques	5.1 Ensemble learning 5.1.1 Definition 5.1.2 How it works 5.1.3 Ensemble learning Algorithms 5.2 Comparing ensemble learning methods 5.2.1 Gradient boosting 5.2.2 Bagging	• Written tests • Oral tests • Practical tests

	5.3 Advantages and disadvantages of ensemble methods 5.4 Demonstrating ensemble learning in regression and classification using given data 5.5 Demonstrating evaluation of the ensemble learning regressor and classifier	
6 Use cross validation to optimize machine learning methods	6.1 Cross Validation 6.1.1 Define cross validation 6.1.2 How cross validation works Illustrating Cross validation techniques 6.2.1 K-fold 6.2.2 Stratified K-fold 6.2.3 Leave one out cross validation (LOOCV) 6.2.4 Demonstration of Cross Validation using given data 6.2.5 Hyper parameter tuning using GridSearchCV	• Written tests • Oral tests • Practical tests
7 Demonstrate model deployment skills	7.1 Installation of flask HTTP methods and JSON 7.2.1 GET 7.2.2 PUT 7.2.3 POST 7.2.4 DELETE Serialization and deserialization 7.3.1 Definition 7.3.2 Using Flask-Mashmallow Connecting to a database 7.4.1 Using SQLAlchemy Demonstration of model deployment	• Written tests • Oral tests • Practical tests

Suggested Delivery Methods

- Instructor led facilitation using active learning strategies
- Demonstration by trainer
- Practical work by trainee
- Viewing of related videos
- Group discussions
- Direct instructions

Recommended Resources for 25 trainees

S/No.	Category/Item	Description/ Specifications	Quantity	Recommended Ratio (Trainee: Item)
A	**Learning Materials**			
1.	Textbooks		5 pcs	5:1
2.	Flip Charts		5 pcs	5:1
3.	PowerPoint presentations	For trainer's use		
4.	Magazines/brochures/business cards			
B	**Learning Facilities & infrastructure**			
5.	Lecture/theory room		1	25:1
6.	Computer Laboratory		1	25:1
C	**Consumable materials**			
7.	Printing papers		1 ream	1:20
8.	Foolscaps		1 ream	
9.	Toners		2 pcs	13:1
10.	Assorted colour of whiteboard markers			

D	Tools and Equipment			
11.	Computers		25 pcs	1:1
12.	Projector		1 pc	25:1
13.	Printers		2 pcs	13:1
14.	Whiteboard		1 pc	25:1
15.	Internet		Per computer	
16.	Microsoft SQL Server			
17.	Python			