SOUTH AFRICAN QUALIFICATIONS AUTHORITY

REGISTERED QUALIFICATION THAT HAS PASSED THE END DATE:

Occupational Certificate: Data Science Practitioner

SAQA QUAL ID	QUALIFICATION TITLE			
118708	Occupational Certificate: Data Science Practitioner			
ORIGINATOR				
Development Quality Partner-MICT SETA				
PRIMARY OR DELEGATED QUALITY ASSURANCE FUNCTIONARY			**NQF SUB-FRAMEWORK**	
-			OQSF - Occupational Qualifications Sub-framework	
QUALIFICATION TYPE	**FIELD**		**SUBFIELD**	
Occupational Certificate	Field 10 - Physical, Mathematical, Computer and Life Sciences		Information Technology and Computer Sciences	
ABET BAND	**MINIMUM CREDITS**	**PRE-2009 NQF LEVEL**	**NQF LEVEL**	**QUAL CLASS**
Undefined	185	Not Applicable	NQF Level 05	Regular-ELOAC
REGISTRATION STATUS		**SAQA DECISION NUMBER**	**REGISTRATION START DATE**	**REGISTRATION END DATE**
Passed the End Date - Status was "Registered"		EXCO 0522/24	2022-02-03	2025-12-31
LAST DATE FOR ENROLMENT		**LAST DATE FOR ACHIEVEMENT**		
2026-12-31		2029-12-31		

In all of the tables in this document, both the pre-2009 NQF Level and the NQF Level is shown. In the text (purpose statements, qualification rules, etc), any references to NQF Levels are to the pre-2009 levels unless specifically stated otherwise.

This qualification does not replace any other qualification and is not replaced by any other qualification.

PURPOSE AND RATIONALE OF THE QUALIFICATION

Purpose:
The purpose of this qualification is to prepare a learner to operate as a Data Science Practitioner.

Data Science Practitioners take custody of data and make the data available in a structured form for the Data Scientist to use. They support the data life cycle by collecting, transforming, and analysing data and communicating results to solve elementary business problems. They transform data into robust, comprehensive data sets, aligned with the problem identified in the statement of work and ready for storage.

A qualified learner will be able to:
- Collect large amounts of structured and unstructured data from primary and secondary sources and extract and transform them into a usable format.
- Apply data analysis techniques to uncover patterns and trends in datasets (resultant sets of data that can be viewed as tables or as a "spreadsheet of data") to solve business-related problems.
- Prepare and present descriptive analytic reports on patterns and trends using computer programming languages and explain those patterns and trends through e.g., visualisation, storytelling, etc., using data visualization tools.

Rationale:
The Presidential Commission on 4IR (PC4IR) report states that the key drivers of change in Human Capital and the Future of Work will be ubiquitous high-speed mobile internet, artificial intelligence, widespread adoption of big data analytics and cloud technology. Thus, with the emergence of the '4IR' and the need to properly manage 'Big Data', a new generation of technologies and architectures, designed to economically extract value from very large volumes of a

wide variety of data by enabling high velocity capture, discovery, or analysis, will emerge.

The 4th Industrial Revolution (4IR) is a fusion of advances in artificial intelligence (AI), robotics, process automation, the Internet of Things (IoT), genetic engineering, quantum computing, cyber security, cloud computing and data science.

There is an exponential demand for data analysts, data engineers, data architects and Data Science Practitioners, in response to the proliferation of complex and voluminous data generated by cloud-businesses and social media networks. To meet this demand, many organisations have started to consider developing skills internally by sharing resources, undertaking training programmes and partnering with others in the industry. This plays a crucial role in establishing a data-driven culture and currently available advanced technology to manipulate these big data and complex datasets.

The demand for qualified big data analysts is exceeding supply to the point where it can take many months to fill vacancies. The root problem of this is that big data analytics is a new field and the existing workforce skill sets must be adjusted to be able to work with large, sophisticated datasets. This shortage is acute and is growing exponentially. Recent research indicated that in 2020 the shortage of data scientists can best be summarised as follows:
- Year-on-year there is a growth of 37% in job listings for data scientists.
- Data scientist ranked 3rd amongst top jobs for 2020.
- The average annual salary increase of data scientists is 14%.

Data science is an inter-disciplinary field that uses scientific methods, processes, algorithms, and systems to extract knowledge and insights from many structured and unstructured data. Data science is related to data mining, machine learning and big data. The data science practitioner's duties can include developing strategies for analysing data, preparing data for analysis, exploring, analysing, and visualizing data, building models with data using programming languages and deploying models into applications. This qualification covers the collection and transformation of data, solving business-related problems through the analysis of data to uncover patterns and trends and the preparation and presentation of descriptive analytic reports using programming techniques, mathematics, and statistics.

The above information confirms the growing need for the Occupational Certificate: Data Science Practitioner. There is a plethora of similar qualifications registered on the NQF. None of these qualifications are at NQF 5 and they are not occupational qualifications.

Data science will bring many benefits to society, touching a wide range of aspects in the daily life of the individual. Scientists can now develop algorithms that can help predict infections based on data analysis, hours before physical symptoms appear. Big data is key to the success of healthcare organizations. They can deliver immunizations, healthcare, and water to some of the world's poorest populations by analysing big data. Companies use the data they collect from the individual to determine what kind of product - whether music, movies, or consumables - to produce.

The target group for this qualification is school leavers, graduates from TVET colleges, new entrants into the sector and existing employees who have experience in this field, but without formal recognition of skills and competencies. No professional registration or licencing is expected for Data Science Practitioner to seek employment in the sector. Data Science Practitioners can find employment as Data Analyst Assistants, Junior Data Analysts, Data Miners, Data Modellers, Data Custodians or Management Information Analysts.

LEARNING ASSUMED TO BE IN PLACE AND RECOGNITION OF PRIOR LEARNING

Recognition of Prior Learning (RPL):

RPL for Access to the External Integrated Summative Assessment
Accredited providers and approved workplaces must apply the internal assessment criteria specified in the related curriculum document to establish and confirm prior learning. Accredited providers and workplaces must confirm prior learning by issuing a statement of result.

RPL for Access to the Qualification
- Learners will gain access to the qualification through RPL for Access as provided for in the QCTO RPL Policy. RPL for access is conducted by accredited education institution, skills development provider or workplace accredited to offer that specific qualification/part qualification.
- Learners who have acquired competencies of the modules of a qualification or part qualification will be credited for modules through RPL.

RPL for access to the external integrated summative assessment
Accredited providers and approved workplaces must apply the internal assessment criteria specified in the related curriculum document to establish and confirm prior learning. Accredited providers and workplaces must confirm prior learning by issuing a statement of result.

Entry Requirements:
The minimum entry requirement for this qualification is:
- NQF Level 4 with Mathematics.

RECOGNISE PREVIOUS LEARNING?

Y

QUALIFICATION RULES

This qualification is made up of compulsory Knowledge, Practical Skill and Work Experience Modules:

Knowledge Modules
- 251102-001-00-KM-01 Introduction to Data Science and Data Analysis, Level 4, 6 Credits.
- 251102-001-00-KM-02 Logical Thinking and Basic Calculations: Refresher, Level 4, 4 Credits.
- 251102-001-00-KM-03 Computers and Computing Systems, Level 4, 4 Credits.
- 251102-001-00-KM-04 Computing Theory, Level 4, 2 Credits.
- 251102-001-00-KM-05 Basic Statistics for Data Analytics, Level 4, 10 Credits.
- 251102-001-00-KM-06 Statistics Essentials for Data Analytics, Level 5, 4 Credits.
- 251102-001-00-KM-07 Data Science and Data Analysis, Level 5, 12 Credits.
- 251102-001-00-KM-08 Data Analysis and Visualisation, Level 5, 16 Credits.
- 251102-001-00-KM-09 Introduction to Governance, Legislation and Ethics, Level 4, 3 Credits.
- 251102-001-00-KM-10 Fundamentals of Design Thinking and Innovation, Level 4, 4 Credits.
- 251102-001-00-KM-11 4IR and Future Skills, Level 4, 1 Credit.

Total number of credits for Knowledge Modules: 66

Practical Skill Modules
- 251102-001-00-PM-01 Apply Logical Thinking and Maths Refresher, Level 4, 3 Credits.
- 251102-001-00-PM-02 Apply Code to Use a Software Toolkit/Platform in the Field of Study or Employment, Level 4, 4 Credits.
- 251102-001-00-PM-03 Use Spreadsheets to Analyse and Visualise Data, Level 4, 3 Credits.
- 251102-001-00-PM-04 Use a Visual Analytics Platform to Analyse and Visualise Data, Level 5, 4 Credits.
- 251102-001-00-PM-05 Apply Statistical Tools and Techniques, Level 5, 4 Credits.
- 251102-001-00-PM-06 Collect and Pre-Process Large Amounts of Unruly Data, Level 5, 12 Credits.
- 251102-001-00-PM-07 Apply Data Analysis Techniques to Uncover Patterns and Trends in Datasets, Level 5, 12 Credits.
- 251102-001-00-PM-08 Prepare and Present Descriptive Analytic Reports for Decision Making, Level 5, 12 Credits.
- 251102-001-00-PM-09 Participate in a Design Thinking for Innovation Workshop, Level 5, 3 Credits.
- 251102-001-00-PM-10 Collaborate Ethically and Effectively in the Workplace, Level 5, 2 Credits.

Total number of credits for Practical Skill Modules: 59

Work Experience Modules
- 251102-001-00-WM-01 Data Collection and Pre-processing Processes, Level 5, 16 Credits.
- 251102-001-00-WM-02 Statistical Data Analysis Processes, Level 5, 16 Credits.
- 251102-001-00-WM-03 Data visualisation and Reporting Processes, level 5, 16 Credits.
- 251102-001-00-WM-04 Capstone Project Using an Appropriate Toolkit, Level 5, 12 Credits.

Total number of credits for Work Experience Modules: 60

EXIT LEVEL OUTCOMES

1. Collect large amounts of structured and unstructured data from primary and secondary sources and extract and transform them into a usable format.
2. Apply data analysis techniques to uncover patterns and trends in datasets (resultant sets of data that can be viewed as tables or as a "spreadsheet of data") to solve business-related problems.
3. Prepare and present descriptive analytics reports on patterns and trends using computer programming languages and explain those patterns and trends through e.g., visualization and storytelling etc., using data visualisation tools.

ASSOCIATED ASSESSMENT CRITERIA

Associated Assessment Criteria for Exit Level Outcome 1:
- Frame the problem to exactly define what is needed or expected and well-defined goals are set.
- Undertake data collection from multiple sources, including primary, secondary, internal, and external sources to generate relevant data for analysis.
- Retrieve and extract data from multiple sources into a scratch table using suitable platforms to make data visible before analysis.
- Clean, wrangle, scrub and reorganise data in a format that can easily be read by either human or machine.

Associated Assessment Criteria for Exit Level Outcome 2:
- Design and use a query to retrieve data which fit the criteria from a database.
- Apply a combination of analytical skills, problem-solving skills, logical reasoning and communication skills to define tendencies in data sets.
- Apply statistical principles, methods, techniques and platforms to analyse and interpret data sets (sets of groupings of datasets or data sources), paying particular attention to trends and patterns that could be valuable for diagnostic and predictive analytics efforts.
- Test the system to ensure it provides accurate results or performs the required tasks accurately.

Associated Assessment Criteria for Exit Level Outcome 3:
- Confirm visual elements to ensure they accurately represent the results of the data analysis.
- Create visualizations, including dashboards, flowcharts, and graphs to communicate business concepts through visuals.

- Check information to ensure it answers the questions or explains solutions to the problems defined at the start of the data analysis life cycle.
- Use the results of the data analysis to perform predictive analysis.
- Discover meaning in data, create insights and storytelling or use other techniques to present data.
- Identify new process improvement opportunities.

Integrated Assessment:
Integrated Formative Assessment
The skills development provider will use the curriculum to guide them on the stipulated internal assessment criteria and weighting. They will also apply the scope of practical skills and applied knowledge as stipulated by the internal assessment criteria. This formative assessment together with work experience leads to entrance in the integrated external summative assessment.

Integrated summative assessment:
An external integrated summative assessment, conducted through the relevant QCTO Assessment Quality Partner is required for the issuing of this qualification. The external integrated summative assessment will focus on the exit level outcomes and associated assessment criteria.

The external integrated summative assessment will be conducted through a theoretical assessment and the evaluation of practical tasks at decentralised approved assessment sites in a simulated environment and conducted by an assessor(s) registered with the relevant AQP.

INTERNATIONAL COMPARABILITY

The Occupational Certificate: Data Science Practitioner was compared with the training provided in India and Canada since they seem to have globally accepted qualifications in this field, and in the case of Toronto School of Management, it is a member of the Global University Systems group of companies.

Canada
The Occupational Certificate: Data Science Practitioner was compared with the Diploma in Data Analytics Co-op presented by the Toronto School of Management, Canada. This Diploma is presented over a period of 52 weeks, including scheduled breaks. 24 Weeks are in class and 12 weeks are on-the-job training which provides the opportunity to integrate academic studies with related on-the-job training experiences. The remainder of the 52 weeks are regarded as self-study, etc. It also includes a Capstone project, which is a final assignment executed in the workplace and provides the learners with the opportunity to build a portfolio with references. Mentoring is provided with the capstone project and work experience. Toronto School of Management has partnerships with over 70 co-op partners across a wide range of industries to provide on-the-job training. Entry requirements are specified as having an Ontario Secondary School Diploma or Equivalent and a pass in the Wonderlic test.

The overall objective of this Diploma is to analyse data using cutting-edge technology or traditional methods to drive proactive decision-making, optimize business performance and to foster ability to interpret and transform large sets of data into actionable insights, whereby increasing business efficiencies. The program is Amazon Web Services (AWS) and Tableau certified.

Qualification Outcomes are:
- Enhancing the quality and usefulness of data analytics and theoretical concepts by handling and designing data.
- Applying automated data collection and traditional methods of data collection to enable the development of methodologically sound approaches.
- Understanding theoretical concepts and practical applications of data auditing, handling and collecting.
- Handling and analysing data to gain informative and useful insights.
- Accessing and manipulating data using analysis software such as SQL and SAS as well as the accurate tools for this and for effective decision making.
- Using appropriate graphs and charts as well as specialist data visualization packages and tools to visualize data and impact the decision-making process.
- Developing presentation skills to facilitate the understanding of findings and to make informed decisions while seeking to understand and monitor processes to drive efficiency.
- Displaying professional conduct and developing and implementing strategies to promote professional competence.

The overall focus of the Capstone project is evaluating and providing a solution that is supported by the following outcomes:
- Identifying opportunities and/or problems.
- Evaluating existing processes.
- Assessing workflow.
- Determining requirements to address opportunities/problems, evaluating alternatives, and recommending an appropriate solution.
- Developing a plan for implementing, maintaining, and supporting the solution.

Similarities
Both qualifications include knowledge, practical (utilising software platforms) and work experience components, and the duration is a full year. Both qualifications specify matric as entry requirements. The content of both qualifications is similar and similar learning outcomes are achieved. Both qualifications include a Capstone project as well as soft skills related to professionalism.

Differences

The modules on soft skills of the Occupational Certificate: Data Science Practitioner are more comprehensive than those of the Canadian Diploma and include aspects such as ethics. The Occupational Certificate: Data Science Practitioner includes refresher and introductory modules, mathematics, and programming. The Occupational Certificate: Data Science Practitioner carries no vendor specific certification and is vendor agnostic, while the Diploma in Data Analytics Co-op is Amazon Web Services (AWS) and Tableau certified.

India
The Occupational Certificate: Data Science Practitioner was compared with the foundational course "Data Science with Python" presented by Analytixlabs, one of India's top ranked data science institutes. This course is offered as fully interactive, live, online video-based and classroom. The duration of the course is stated as 45 hours of classes, 21 hours e-learning, video-based (which is self-paced) training, 64 hours self-study plus practice including 10 assignments and projects. The target group is specified as candidates from different technical or quantitative backgrounds like Engineering, Finance, Maths, Statistics and Business Management who wish to start their career in Data Science and Machine Learning skills. For aspirants from a non-technical background, it is recommended to have prior knowledge of basic data analytics tools like Excel/ SQL/ Tableau.

Qualification objectives include:
- Introduction to basic statistics.
- Introduction to mathematical foundations.
- Introduction to analytics and data science.
- Visualising data.
- Numerical Python.
- Python essentials.
- Data structure.
- Cleansing data.
- Data Analysis with Python.
- Data Visualization with Python.
- Visualisation tools.
- Statistical methods and hypothesis testing.
- Linear Regression.
- Logistic Regression.

Similarities:
The content of both qualifications is comparable and similar learning outcomes are achieved. Both qualifications include a Capstone project or final assignments as well as knowledge and practical components.

Differences
The Data Science with Python course is shorter since the entry requirements include competency in mathematics and statistics. The Data Science with Python course does not include work experience. The Occupational Certificate: Data Science Practitioner includes modules on soft skills and ethics, which is not the case with the Data Science with Python course. The Occupational Certificate: Data Science Practitioner carries no vendor specific certification and is vendor agnostic, while the Data Science with Python course is IBM certified. The qualifications differ in terms of entry requirements and target groups.

The target group for the Data Science with Python course is specified as candidates from different technical or quantitative backgrounds like Engineering, Finance, Maths, Statistics and Business Management who wish to start their career in Data Science and Machine Learning skills. For aspirants from a non-technical background, it is recommended to have prior knowledge of basic data analytics tools like Excel/ SQL/ Tableau.

Conclusion
The South African qualification compares favourably with the competencies covered in international qualifications.

ARTICULATION OPTIONS

This qualification provides opportunities for horizontal and vertical articulation options.

Horizontal Articulation:
- National Certificate: Business Analysis Support Practice, NQF Level 5.

Vertical Articulation:
- National Certificate: Business Analysis, NQF Level 6.

NOTES

Qualifying for External Assessment:
To qualify for an external assessment, learners must provide proof of completion of all required knowledge and practical modules by means of statements of results and a record of completed work experience.

Additional Legal or Physical Entry Requirements:
- None.

Criteria for the accreditation of providers
Accreditation of providers will be done against the criteria as reflected in the relevant curriculum on the QCTO website.

The curriculum title and code are: Data Science Practitioner: 251102-001-00.

Encompassed Trade:
This qualification encompasses the following trades as recorded on the NLRD:
- This is not a trade qualification.

Assessment Quality Partner (AQP)
- MICT SETA.

LEARNING PROGRAMMES RECORDED AGAINST THIS QUALIFICATION:

NONE

PROVIDERS CURRENTLY ACCREDITED TO OFFER THIS QUALIFICATION:

This information shows the current accreditations (i.e. those not past their accreditation end dates), and is the most complete record available to SAQA as of today. Some Primary or Delegated Quality Assurance Functionaries have a lag in their recording systems for provider accreditation, in turn leading to a lag in notifying SAQA of all the providers that they have accredited to offer qualifications and unit standards, as well as any extensions to accreditation end dates. The relevant Primary or Delegated Quality Assurance Functionary should be notified if a record appears to be missing from here.

NONE
