QUALIFICATION FILE-*OEM*

Data Science – IBM

☒Short Term Training (STT) ☐Long Term Training (LTT) ☐Apprenticeship

☒Upskilling ☐ Dual/Flexi Qualification ☐For ToT ☐For ToA

☐General ☐Multi-Skill (MS) ☐ Cross Sectoral (CS) ☐Future Skills ☑ Original Equipment Manufacturer (OEM)

NCrF/NSQF Level: 5

Submitted By:

IT-ITeS SSC nasscom

Table of Contents

Section 1: Basic Details 3
Section 2: Module Summary 5
NOS/s of Qualifications 5
Mandatory NOS/s: 5
Elective NOS/s: 5
Optional NOS/s: 6
Assessment - Minimum Qualifying Percentage 6
Section 3: Training Related 6
Section 4: Assessment Related 6
Section 5: Evidence of the need for the Qualification 7
Section 6: Annexure & Supporting Documents Check List 7
Annexure: Evidence of Level 8
Annexure: Tools and Equipment (Lab Set-Up) 8
Annexure: Industry Validations Summary 9
Annexure: Training & Employment Details 9
Annexure: Blended Learning 10
Annexure: Detailed Assessment Criteria 10
Annexure: Assessment Strategy 11
Annexure: Acronym and Glossary 12

Section 1: Basic Details

<table>
<tr><td>1.</td><td>Qualification Name</td><td colspan="2">Data Science – IBM</td></tr>
<tr><td>2.</td><td>Sector/s</td><td colspan="2">IT-ITES</td></tr>
<tr><td>3.</td><td>Type of Qualification: ☐New ☐ Revised ☐ Has Electives/Options
☒ OEM</td><td>NQR Code & version of existing/previous qualification:</td><td>Qualification Name of existing/previous version:</td></tr>
<tr><td>4.</td><td>a. OEM Name
b. Qualification Name
(Wherever applicable)</td><td colspan="2">IBM Data Science
Data Science – IBM</td></tr>
<tr><td>5.</td><td>National Qualification Register (NQR) Code &Version
(Will be issued after NSQC approval)</td><td>NG-05-IT-02244-2024-V1-NASSCOM</td><td>6. NCrF/NSQF Level: 5</td></tr>
<tr><td>7.</td><td>Award (Certificate/Diploma/Advance Diploma/ Any Other (Wherever applicable specify multiple entry/exits also & provide details in annexure)</td><td colspan="2">IBM Data Science Professional Certificate</td></tr>
<tr><td>8.</td><td>Brief Description of the Qualification</td><td colspan="2">This program provides learners with job-ready tools and skills, including open-source tools and libraries, Python, databases, SQL, data visualization, data analysis, statistical analysis, predictive modelling, and machine learning algorithms. Learners will learn data science through hands-on practice in the IBM Cloud using real data science tools and real-world data sets. Upon completing this course courses, graduates have built a portfolio of data science projects to provide them with the confidence to head into an entry-level profession in data science.</td></tr>
<tr><td>9.</td><td>Eligibility Criteria for Entry for Student/Trainee/Learner/Employee</td><td colspan="2">a. Entry Qualification & Relevant Experience:
<table>
<tr><th>S. No.</th><th>Academic/Skill Qualification (with Specialization - if applicable)</th><th>Required Experience (with Specialization - if applicable)</th></tr>
<tr><td>1</td><td>3rd Year of UG (STEM OR NonSTEM*)</td><td rowspan="3">Basic knowledge of computer literacy and high-school level mathematics, Basic knowledge of calculus and linear algebra and data crunching.</td></tr>
<tr><td>2</td><td>2-year UG Diploma in IT/computer science + 2 years of relevant experience</td></tr>
<tr><td>3</td><td>12th or equivalent (Computer Science)+ 4 years of relevant experience</td></tr>
</table>
b. Age: 18 Years and above</td></tr>
</table>

<table>
<tr><td>10.</td><td>Credits Assigned to this Qualification, Subject to Assessment (as per National Credit Framework (NCrF))</td><td>5.5</td><td>11. Common Cost Norm Category (I/II/III) (wherever applicable):

II</td></tr>
<tr><td>12.</td><td>Any Licensing requirements for Undertaking Training on This Qualification (wherever applicable)</td><td colspan="2"></td></tr>
<tr><td>13.</td><td>Training Duration by Modes of Training Delivery (Specify Total Duration as per selected training delivery modes and as per requirement of the qualification)</td><td colspan="2">☐Offline ☐Online ☐Blended
<table>
<tr><th>Training Delivery Modes</th><th>Theory (Hours)</th><th>Practical (Hours)</th><th>OJT Mandatory (Hours)</th><th>OJT Recommended (Hours)</th><th>Total (Hours)</th></tr>
<tr><td>Classroom (offline)</td><td></td><td></td><td></td><td></td><td></td></tr>
<tr><td>Online</td><td>132</td><td>33</td><td></td><td></td><td>165</td></tr>
</table>
(Refer Blended Learning Annexure for details)</td></tr>
<tr><td>14.</td><td>Aligned to NCO/ISCO Code/s (if no code is available mention the same)</td><td colspan="2">NCO-2015/NIL</td></tr>
<tr><td>15.</td><td>Progression path after attaining the qualification (Please show Professional and Academic progression)</td><td colspan="2">Job Roles: Data Scientist, Data Scientist I, Data Scientist II</td></tr>
<tr><td>16.</td><td>Other Indian languages in which the Qualification & Model Curriculum are being submitted</td><td colspan="2">N/A</td></tr>
<tr><td>17.</td><td>Is similar Qualification(s) available on NQR-if yes, justification for this qualification</td><td colspan="2">☐ Yes ☐ No URLs of similar Qualifications:</td></tr>
<tr><td>18.</td><td>Is the Job Role Amenable to Persons with Disability</td><td colspan="2">☒Yes ☐ No
If “Yes”, specify applicable type of Disability: Visual impairment; Hearing impairment; Physical disability</td></tr>
<tr><td>19.</td><td>How Participation of Women will be Encouraged</td><td colspan="2"></td></tr>
<tr><td>20.</td><td>Are Greening/ Environment Sustainability Aspects Covered (Specify the NOS/Module which covers it)</td><td colspan="2">☐ Yes ☐ No</td></tr>
<tr><td>21.</td><td>Is Qualification Suitable to be Offered in Schools/Colleges</td><td colspan="2">Schools ☐ Yes ☐ No Colleges ☒ Yes ☐ No</td></tr>
<tr><td>22.</td><td>Name and Contact Details of Submitting / Awarding Body SPOC
(In case of CS or MS, provide details of both Lead AB & Supporting ABs)</td><td colspan="2">Name: Namrata Kapur
Email: namrata@nasscom.in
Website: https://www.sscnasscom.com/</td></tr>
<tr><td>23.</td><td>Final Approval Date by NSQC: 15th Mar 2024</td><td>24. Validity Duration: 3 years</td><td>25. Next Review Date:15th Mar 2027</td></tr>
</table>

Section 2: Module Summary

NOS/s of Qualifications

(in exceptional cases these could be described as components)

Mandatory NOS/s:

Specify the training duration and assessment criteria at NOS/ Module level. For further details refer curriculum document.

***Th.**-Theory **Pr.**-Practical **OJT**-On the Job **Man.**-Mandatory Training **Rec.**-Recommended **Proj.**-Project*

S. No	NOS/Module Name	NOS/Module Code & Version *(if applicable)*	Core/ Non-Core	NCrF/NSQF Level	Credits as per NCrF	Training Duration (Hours)					Assessment Marks					
						Th.	Pr.	OJT-Man.	OJT-Rec.	Total	Th.	Pr.	Proj.	Viva	Total	Weightage (%) *(if applicable)*
1.	Data Science – IBM	IT-ITeS SSC NASSCOM/N0003/IT/2024	Core	5	5.5	132	33	-	-	165	60	40	-	-	100	

Elective NOS/s

S. No	NOS/Module Name	NOS/Module Code & Version *(if applicable)*	Core/ Non-Core	NCrF/NSQF Level	Credits as per NCrF	Training Duration (Hours)					Assessment Marks					
						Th.	Pr.	OJT-Man.	OJT-Rec.	Total	Th.	Pr.	Proj.	Viva	Total	Weightage (%) *(if applicable)*
1.																
2.																
Duration (in Hours) / Total Marks																

Optional NOS/s:

S. No	NOS/Module Name	NOS/Module Code & Version *(if applicable)*	Core/ Non-Core	NCrF/NSQF Level	Credits as per NCrF	Training Duration (Hours)					Assessment Marks					
						Th.	Pr.	OJT-Man.	OJT-Rec.	Total	Th.	Pr.	Proj.	Viva	Total	Weightage (%) *(if applicable)*
1.																
2.																
Duration (in Hours) / Total Marks																

Assessment - Minimum Qualifying Percentage

*Please specify **any one** of the following:*

Minimum Pass Percentage – Aggregate at qualification level: 70 % *(Every Trainee should score specified minimum aggregate passing percentage at qualification level to successfully clear the assessment.)*

Minimum Pass Percentage – NOS/Module-wise: 70 % *(Every Trainee should score specified minimum passing percentage in each mandatory and selected elective NOS/Module to successfully clear the assessment.)*

Section 3: Training Related

1.	**Trainer's Qualification and experience in the relevant sector (in years)** *(as per NCVET guidelines)*	

2.	Master Trainer's Qualification and experience in the relevant sector (in years) *(as per NCVET guidelines)*	
3.	**Tools and Equipment Required for Training**	☒Yes ☐No *(If "Yes", details to be provided in Annexure)* Annexure: Tools and Equipment
4.	**In Case of Revised Qualification, Details of Any Upskilling Required for Trainer**	

Section 4: Assessment Related

1.	**Assessor's Qualification and experience in relevant sector (in years)** *(as per NCVET guidelines)*	**Online assessment by NCVET approved Assessment Agencies so the assessor qualification & experience is relevant sector of IT-ITeS will be as per the guideline of NCVET or University.**
2.	Proctor's Qualification and experience in relevant sector (in years) *(as per NCVET guidelines)*	
3.	Lead Assessor's/Proctor's Qualification and experience in relevant sector (in years) *(as per NCVET guidelines)*	
4.	**Assessment Mode** *(Specify the assessment mode)*	**Online**
5.	**Tools and Equipment Required for Assessment**	☒ Same as for training ☐ Yes ☐ No *(details to be provided in Annexure-if it is different for Assessment)*

Section 5: Evidence of the need for the Qualification

Provide Annexure/Supporting documents name.

1.	Latest Skill Gap Study (not older than 2 years) (Yes/No):
2.	Latest Market Research Reports or any other source (not older than 2 years) (Yes/No):
3.	Government /Industry initiatives/ requirement (Yes/No):
4.	Number of Industry validation provided:
5.	**Estimated nos. of persons to be trained and employed:**
6.	Evidence of Concurrence/Consultation with Line Ministry/State Departments: If "No", why:

Section 6: Annexure & Supporting Documents Check List

Specify Annexure Name / Supporting document file name

1.	**Annexure:** NCrF/NSQF level justification based on NCrF level/NSQF descriptors *(Mandatory)*	*Annexure: Evidence of Level*
2.	**Annexure:** List of tools and equipment relevant for qualification *(Mandatory, except in case of online course)*	*Annexure: Tools and Equipment*
3.	**Annexure:** Detailed Assessment Criteria *(Mandatory)*	*Annexure: Assessment Criteria*
4.	**Annexure:** Assessment Strategy *(Mandatory)*	
5.	**Annexure:** Blended Learning *(Mandatory, in case selected Mode of delivery is "Blended Learning")*	
6.	**Annexure:** Multiple Entry-Exit Details *(Mandatory, in case qualification has multiple Entry-Exit)*	
7.	**Annexure:** Acronym and Glossary *(Optional)*	
8.	**Supporting Document:** Model Curriculum *(Mandatory – Public view)*	
9.	**Supporting Document:** Career Progression *(Mandatory - Public view)*	
10.	**Supporting Document:** Occupational Map *(Mandatory)*	
11.	**Supporting Document:** Assessment SOP *(Mandatory)*	
12.	**Any other document you wish to submit:**	

Annexure: Evidence of Level

NCrF/NSQF Level Descriptors	Key requirements of the job role/ outcome of the qualification	How the job role/ outcomes relate to the NCrF/NSQF level descriptor	NCrF/NSQF Level
Professional Theoretical Knowledge/Process	This Professional Certificate is suitable for those with and without college degrees. It does not require any prior programming or cloud skills. It does require basic understanding	Learners will understand the basics of data mining and management using Python, SQL and other data engines and industry tools.	5

	of computing environments and computer literacy such as familiarity working with files and folders in an operating system, launching, using, and switching between applications, running commands, copying and pasting text and files, taking screenshots.		
Professional and Technical Skills/ Expertise/ Professional Knowledge	Experience basic skills to develop as a Data Scientist	Learners will apply Python programming language and Linux/UNIX shell scripts to extract, transform and load (ETL) data. Learners will also work with Relational Databases (RDBMS) and query data using SQL statements and use NoSQL databases as well as unstructured data.	5
Employment Readiness & Entrepreneurship Skills & Mind-set/Professional Skill	Experience basic skills to develop as a Data Scientist	Skills are focused on core programming and data mining concepts, industry tools used in data engineering, critical thinking, and ability to convert data into business intelligence insights.	5
Broad Learning Outcomes/Core Skill	Data Analysis, Database Management, Python Programming, Data Warehousing and Management	Learners will design, deploy, and manage an end-to-end data engineering platform.	5
Responsibility	Ability to develop big data architecture that helps analyze and process data in the way the organization needs it.	Can work on improving an organization's data infrastructure, reliability, efficiency and quality	5

Annexure: Tools and Equipment (Lab Set-Up)

List of Tools and Equipment

Batch Size:

S. No.	Tool / Equipment Name	Specification	Quantity for specified Batch size
1.	Computer	i3/i5 with 16GB Ram	one per learner
2.	Installed Browser	Modern Browser	installed on the computer
3.	Internet	Stable internet with at least 40 mbps speed	internet access for the learners

Classroom Aids

The aids required to conduct sessions in the classroom are:

1.
2.

Annexure: Industry Validations Summary

Provide the summary information of all the industry validations in table. This is not required for OEM qualifications.

S. No	Organization Name	Representative Name	Designation	Contact Address	Contact Phone No	E-mail ID	LinkedIn Profile *(if available)*

Annexure: Training & Employment Details

Training and Employment Projections:

Year	Total Candidates		Women		People with Disability	
	Estimated Training #	Estimated Employment Opportunities	Estimated Training #	Estimated Employment Opportunities	Estimated Training #	Estimated Employment Opportunities
2024	422,330		43,926			
2025	506,795		52,711			
2026	557,475		57,982			

Data to be provided year-wise for next 3 years

Training, Assessment, Certification, and Placement Data for previous versions of qualifications:

Qualification Version	Year	Total Candidates				Women				People with Disability			
		Trained	Assessed	Certified	Placed	Trained	Assessed	Certified	Placed	Trained	Assessed	Certified	Placed

Applicable for revised qualifications only, data to be provided year-wise for past 3 years.

List Schemes in which the previous version of Qualification was implemented:

1.
2.

Content availability for previous versions of qualifications:

☐ Participant Handbook ☐ Facilitator Guide ☐ Digital Content ☐Qualification Handbook ☐Any Other:

Languages in which Content is available:

Annexure: Blended Learning

Blended Learning Estimated Ratio & Recommended Tools:

Refer NCVET "Guidelines for Blended Learning for Vocational Education, Training & Skilling" available on:
https://ncvet.gov.in/sites/default/files/Guidelines%20for%20Blended%20Learning%20for%20Vocational%20Education,%20Training%20&%20Skilling.pdf

S. No.	Select the Components of the Qualification	List Recommended Tools – for all Selected Components	Offline : Online Ratio
1	☐Theory/ Lectures - Imparting theoretical and conceptual knowledge	NA	
2	☐Imparting Soft Skills, Life Skills, and Employability Skills /Mentorship to Learners	NA	
3	☐Showing Practical Demonstrations to the learners	NA	
4	☐Imparting Practical Hands-on Skills/ Lab Work/ workshop/ shop floor training	NA	
5	☐Tutorials/ Assignments/ Drill/ Practice	NA	
6	☐Proctored Monitoring/ Assessment/ Evaluation/ Examinations	NA	
7	☐On the Job Training (OJT)/ Project Work Internship/ Apprenticeship Training	NA	

Annexure: Detailed Assessment Criteria

Detailed assessment criteria for each NOS/Module are as follows:

NOS/Module Name	Assessment Criteria for Performance Criteria/Learning Outcomes	Theory Marks	Practical Marks	Project Marks	Viva Marks
What is Data Science?	• Define data science and its importance in today's data-driven world. • Describe the various paths that can lead to a career in data science. • Summarize advice given by seasoned data science professionals to data scientists who are just starting out. • Explain why data science is considered the most in-demand job in the 21st century.	60	40		
	Total Marks	60	40		
Tools for Data Science	• Describe the Data Scientist's tool kit which includes: Libraries & Packages, Data sets, Machine learning models, and Big Data tools • Utilize languages commonly used by data scientists like Python, R, and SQL • Demonstrate working knowledge of tools such as Jupyter notebooks and RStudio and utilize their various features • Create and manage source code for data science using Git repositories and GitHub. .	60	40		
	Total Marks	60	40		
Data Science Methodology	• Describe what a data science methodology is and why data scientists need a methodology. • Apply the six stages in the Cross-Industry Process for Data Mining (CRISP-DM) methodology to analyze a case study. • Evaluate which analytic model is appropriate among predictive, descriptive, and classification models used to analyze a case study. • Determine appropriate data sources or your data science analysis methodology.	60	40		
	Total Marks	60	40		
Python for Data Science, AI & Development	• Describe Python Basics including Data Types, Expressions, Variables, and Data Structures. • Apply Python programming logic using Branching, Loops, Functions, Objects & Classes.	60	40		

	• Demonstrate proficiency in using Python libraries such as Pandas, Numpy, and Beautiful Soup. • Access web data using APIs and web scraping from Python in Jupyter Notebooks. .				
	Total Marks	60	40		
Python Project for Data Science	• Play the role of a Data Scientist / Data Analyst working on a real project. • Demonstrate your Skills in Python - the language of choice for Data Science and Data Analysis. • Apply Python fundamentals, Python data structures, and working with data in Python. • Build a dashboard using Python and libraries like Pandas, Beautiful Soup and Plotly using Jupyter notebook.	60	40		
	Total Marks	60	40		
Databases and SQL for Data Science with Python	• Analyze data within a database using SQL and Python. • Create a relational database and work with multiple tables using DDL commands. • Construct basic to intermediate level SQL queries using DML commands. • Compose more powerful queries with advanced SQL techniques like views, transactions, stored procedures, and joins.	60	40		
	Total Marks	60	40		
Data Analysis with Python	• Develop Python code for cleaning and preparing data for analysis - including handling missing values, formatting, normalizing, and binning data • Perform exploratory data analysis and apply analytical techniques to real-word datasets using libraries such as Pandas, Numpy and Scipy • Manipulate data using dataframes, summarize data, understand data distribution, perform correlation and create data pipelines • Build and evaluate regression models using machine learning scikit-learn library and use them for prediction and decision making	60	40		
	Total Marks	60	40		

Data Visualization with Python	• Implement data visualization techniques and plots using Python libraries, such as Matplotlib, Seaborn, and Folium to tell a stimulating story • Create different types of charts and plots such as line, area, histograms, bar, pie, box, scatter, and bubble • Create advanced visualizations such as waffle charts, word clouds, regression plots, maps with markers, & choropleth maps • Generate interactive dashboards containing scatter, line, bar, bubble, pie, and sunburst charts using the Dash framework and Plotly library	60	40		
	Total Marks	60	40		
Machine Learning with Python	• Describe the various types of Machine Learning algorithms and when to use them • Compare and contrast linear classification methods including multiclass prediction, support vector machines, and logistic regression • Write Python code that implements various classification techniques including K-Nearest neighbors (KNN), decision trees, and regression trees • Evaluate the results from simple linear, non-linear, and multiple regression on a data set using evaluation metrics	60	40		
	Total Marks	60	40		
Applied Data Science Capstone	• Demonstrate proficiency in data science and machine learning techniques using a real-world data set and prepare a report for stakeholders • Apply your skills to perform data collection, data wrangling, exploratory data analysis, data visualization model development, and model evaluation • Write Python code to create machine learning models including support vector machines, decision tree classifiers, and k-nearest neighbors • Evaluate the results of machine learning models for predictive analysis, compare their strengths and weaknesses and identify the optimal model .	60	40		
	Total Marks	60	40		

Annexure: Assessment Strategy

This section includes the processes involved in identifying, gathering, and interpreting information to evaluate the Candidate on the required competencies of the program. Mention the detailed assessment strategy in the provided template.

<1. Assessment System Overview:

- Batches assigned to the assessment agencies for conducting the assessment on SIP or email ● Assessment agencies send the assessment confirmation to VTP/TC looping SSC
- Assessment agency deploys the ToA certified Assessor for executing the assessment ● SSC monitors the assessment process & records

2. Testing Environment:

- Check the Assessment location, date and time
- If the batch size is more than 30, then there should be 2 Assessors.
- Check that the allotted time to the candidates to complete Theory & Practical Assessment is correct. 3. Assessment Quality Assurance levels/Framework:
- Question bank is created by the Subject Matter Experts (SME) are verified by the other SME ● Questions are mapped to the specified assessment criteria
- Assessor must be ToA certified & trainer must be ToT Certified 4. Types of evidence or evidence-gathering protocol:
- Time-stamped & geotagged reporting of the assessor from assessment location ● Centre photographs with signboards and scheme specific branding

5. Method of verification or validation:

- Surprise visit to the assessment location

6. Method for assessment documentation, archiving, and access ● Hard copies of the documents are stored

On the Job:

1. Each module (which covers the job profile of Automotive Service Assistant Technician) will be assessed separately. 2. The candidate must score 60% in each module to successfully complete the OJT.
3. Tools of Assessment that will be used for assessing whether the candidate is having desired skills and etiquette of dealing with customers, understanding needs & requirements, assessing the customer and perform Soft Skills effectively:

• Videos of Trainees during OJT •

4. Assessment of each Module will ensure that the candidate is able to: • Effective engagement with the customers
- Understand the working of various tools and equipment • .....>

Annexure: Acronym and Glossary

Acronym

Acronym	Description
AA	Assessment Agency
AB	Awarding Body

ISCO	International Standard Classification of Occupations
NCO	National Classification of Occupations
NCrF	National Credit Framework
NOS	National Occupational Standard(s)
NQR	National Qualification Register
NSQF	National Skills Qualifications Framework
OJT	On the Job Training

Glossary

Term	Description
National Occupational Standards (NOS)	NOS define the measurable performance outcomes required from an individual engaged in a particular task. They list down what an individual performing that task should know and also do.
Qualification	A formal outcome of an assessment and validation process which is obtained when a competent body determines that an individual has achieved learning outcomes to given standards
Qualification File	A Qualification File is a template designed to capture necessary information of a Qualification from the perspective of NSQF compliance. The Qualification File will be normally submitted by the awarding body for the qualification.
Sector	A grouping of professional activities on the basis of their main economic function, product, service or technology.
Long Term Training	Long-term skilling means any vocational training program undertaken for a year and above. https://ncvet.gov.in/sites/default/files/NCVET.pdf