Automotive Manufacturing Data Science Specialist

QP Code: ASC/Q6417

NSQF Level: 6

Version: 1.0

Automotive Skills Development Council || 153, Gr Floor, Okhla Industrial Area, Phase – III, Leela Building

New Delhi – 110020

Job Role: Automotive Manufacturing Data Science Specialist

Brief Job Description:

The individual at this job is responsible for designing and developing the visualization platforms for end-to-end visibility, after-sales analytics solutions and services, and determining repair cost estimates and allocating repair timelines, recording and creating daily maintenance logs and progress reports. He/She leads the development activities and guide the team on the technical front in analytics solutions and reviewing of code and supports a range of analytical, visualization, and predictive modelling projects along with project management.

Personal Attributes:

The person should have good technical and analytical skills, should have excellent interpersonal skills, communication, and presentation skills, and a good team leader. The person should have project management skills, and also carry out prioritization of work and mentoring the budding engineers.

Applicable National Occupational Standards (NOS)

Compulsory NOS:
1. Manage work and resources (Manufacturing)
2. Employability NOS (120 Hours)
3. Manage data extraction and collection in automotive manufacturing entities
4. Prepare and analyse data by using analytical tools
5. Develop solutions for complex business problems
6. Analyse assembly line data in automotive manufacturing entity

ASC/N6445: Manage data extraction and collection in automotive manufacturing entities

Description

This NOS unit is about performing tasks related to managing extraction and collection of data from various manufacturing entities.

Scope

The scope covers the following:
- Monitor assessment of project requirements
- Perform and monitor designing of project outline
- Support in selection of data integration platform to integrate the data from various department

Elements and Performance Criteria

Monitor assessment of project requirements

To be competent, the user/individual on the job must be able to:

PC1. Support during evaluation of the project requirements to be catered with either visualization platforms or analytics and predictive modelling solutions

PC2. Monitor and guide team during designing of data architecture for collection and extraction of data from various departments using connectors and platforms

PC3. Identify the people required to execute the business analytics project requirements

PC4. Assess all organizational processes related to the use of data and analytics

PC5. Check that the existing setup is capable or not for data collection and analysis

Perform and monitor designing of project outline

To be competent, the user/individual on the job must be able to:

PC6. Plan and prepare project layout where it should defend the choice of technology and its cost

PC7. Support and prepare the outline of the development process and its requirements for both material and resources

PC8. Define various data attributes and what level of analytics is required to analyse data and deliver value

PC9. Prepare the timeline and resource requirements

PC10. Manage project by using appropriate project tracking tools and task prioritization for all team members

PC11. Obtain the necessary approvals within the organization for data collection and extraction from various departments

Support in selection of data integration platform to integrate the data from various department

To be competent, the user/individual on the job must be able to:

PC12. Select the data integration platform with the capabilities like- data transformation, application connectors, file processing, routing, orchestration, event handling, stream processing, API management, no-vendor lock-in.

PC13. Create and monitor an end-to-end data flow using ETL (Extract-Transform-Load) tool using different connectors for different types of data sources

PC14. Design and create a data warehouse for data acquisition

PC15. Develop data pipelines using connectors to populate the data in the data warehouse

Knowledge and Understanding (KU)

The individual on the job needs to know and understand:

KU1. Organizational policies, procedures, and guidelines that relate to designing and maintaining databases

KU2. Organizational policies and procedures for sharing data

KU3. Organizational policies and procedures for documenting databases architectures and backup mechanisms

KU4. Who to involve while designing and developing the database architecture and pipelines for the solution

KU5. Range of standard platforms and tools available and how to use them

KU6. Database connectors and application connectors for application-cloud communications

KU7. Updated internal and external cybersecurity regulations

KU8. Impacts of network on the environment and human health

KU9. ETL tools like Talend, SQL Server Integration Services (SSIS), etc.

KU10. Basics of SQL

Generic Skills (GS)

User/individual on the job needs to know how to:

GS1. Follow instructions, guidelines, procedures, rules, and service level agreements

GS2. Listen effectively and communicate information accurately

GS3. Follow rule-based decision-making processes

GS4. Make decisions on suitable courses

GS5. Plan and organize the work to achieve targets and meet deadlines

GS6. Apply problem-solving approaches to different situations

GS7. Analyse the business impact and disseminate relevant information to others

GS8. Apply balanced judgments to different situations

GS9. Check the work is complete and free from errors

Assessment Criteria

Assessment Criteria for Outcomes	Theory Marks	Practical Marks	Project Marks	Viva Marks
Monitor assessment of project requirements	**11**	**11**		**5**
PC1. Support during evaluation of the project requirements to be catered with either visualization platforms or analytics and predictive modelling solutions	3	3		1
PC2. Monitor and guide team during designing of data architecture for collection and extraction of data from various departments using connectors and platforms	2	2		1
PC3. Identify the people required to execute the business analytics project requirements	2	2		1
PC4. Assess all organizational processes related to the use of data and analytics	2	2		1
PC5. Check that the existing setup is capable or not for data collection and analysis	2	2		1
Perform and monitor designing of project outline	**20**	**20**		**11**
PC6. Plan and prepare project layout where it should defend the choice of technology and its cost	4	4		2
PC7. Support and prepare the outline of the development process and its requirements for both material and resources	3	3		2
PC8. Define various data attributes and what level of analytics is required to analyse data and deliver value	3	3		2
PC9. Prepare the timeline and resource requirements	3	3		2
PC10. Manage project by using appropriate project tracking tools and task prioritization for all team members	4	4		2
PC11. Obtain the necessary approvals within the organization for data collection and extraction from various departments	3	3		1
Support in selection of data integration platform to integrate the data from various department	**9**	**9**		**4**
PC12. Select the data integration platform with the capabilities like- data transformation, application connectors, file processing, routing, orchestration, event handling, stream processing, API management, no-vendor lock-in.	2	2		1
PC13. Create and monitor an end-to-end data flow using ETL (Extract-Transform-Load) tool using different connectors for different types of data sources	3	3		1

PC14. Design and create a data warehouse for data acquisition	2	2		1
PC15. Develop data pipelines using connectors to populate the data in the data warehouse	2	2		1
NOS Total	**40**	**40**	**-**	**20**

ASC/N6446: Prepare and analyse data by using analytical tools

Description

This NOS unit is about performing exploratory data analysis on the data extracted, deciding which data attributes are required for analytics and analysing the extracted attributes using excel/ open-source python libraires.

Scope

The scope covers the following:

- Identify business goals for which data need to be analysed,
- Support in preparation of data
- Perform statistical analysis of data

Elements and Performance Criteria

Identify business goal for which data need to be analysed

To be competent, the user/individual on the job must be able to:

PC1. Define business problems and business goals which can be achieved using available datasets from the manager

PC2. Select the relevant source of data to define business goal

PC3. Validate the criterion in the business problem with domain person

Support in preparation of data

To be competent, the user/individual on the job must be able to:

PC4. Support and create a set of metadata for the selected dataset

PC5. Identify the attributes or columns in the datasets which are most significant from analysis perspective

PC6. Perform exploratory data analysis to check for missing or duplicate data

Perform statistical analysis of data

To be competent, the user/individual on the job must be able to:

PC7. Perform descriptive statistical analysis on the data by following SOP

PC8. Perform inferential statistics analysis on the data by following SOP

PC9. Prepare list of highly correlated attributes

PC10. Find correlation amongst the selected attributes of the data and plot their heatmap

Knowledge and Understanding (KU)

The individual on the job needs to know and understand:

KU1. Organizational policies, procedures, and guidelines that relate to designing and maintaining databases

KU2. Organizational policies and procedures for sharing data

KU3. Organizational policies and procedures for documenting databases architectures and backup mechanisms

KU4. Descriptive and Inferential statistics for creating charts and predictive analytics modelling

KU5. Types of data wrangling and data cleaning methods to create visualization

KU6. Suitable documentation of the organization for the metadata creation

KU7. Aggregate the charts to create a dashboard to address the business problem

KU8. Addition of filters and chart tips to make the dashboard interactive

Generic Skills (GS)

User/individual on the job needs to know how to:

GS1.   Follow instructions, guidelines, procedures, rules, and service level agreements

GS2.   Listen effectively and communicate information accurately

GS3.   Follow rule-based decision-making processes

GS4.   Make decisions on suitable courses

GS5.   Plan and organize the work to achieve targets and meet deadlines

GS6.   Apply problem-solving approaches to different situations

GS7.   Analyse the business impact and disseminate relevant information to others

GS8.   Apply balanced judgments to different situations

GS9.   Check the work is complete and free from errors

Assessment Criteria

Assessment Criteria for Outcomes	Theory Marks	Practical Marks	Project Marks	Viva Marks
Identify business goal for which data need to be analysed	**11**	**11**		**6**
PC1. Define business problems and business goals which can be achieved using available datasets from the manager	3	3		2
PC2. Select the relevant source of data to define business goal	4	4		2
PC3. Validate the criterion in the business problem with domain person	4	4		2
Support in preparation of data	**16**	**16**		**7**
PC4. Support and create a set of metadata for the selected dataset	6	6		3
PC5. Identify the attributes or columns in the datasets which are most significant from analysis perspective	5	5		2
PC6. Perform exploratory data analysis to check for missing or duplicate data	5	5		2
Perform statistical analysis of data	**13**	**13**		**7**
PC7. Perform descriptive statistical analysis on the data by following SOP	3	3		2
PC8. Perform inferential statistics analysis on the data by following SOP	3	3		2
PC9. Prepare list of highly correlated attributes	3	3		1
PC10. Find correlation amongst the selected attributes of the data and plot their heatmap	4	4		2
NOS Total	**40**	**40**	**-**	**20**

ASC/N6447: Develop solutions for complex business problems

Description

This NOS unit is about developing machine learning models using the extracted data and making predictive analytics solutions for complex business problems. It also involves project management using project tracking tools and task prioritization for all team members.

Scope

The scope covers the following:

- Execute training phase in the machine learning project lifecycle
- Execute testing phase in the machine learning project lifecycle
- Deployment of the developed analytics model solution into production
- Deployment of the analytics model solution into production line

Elements and Performance Criteria

Execute training phase in the machine learning project lifecycle

To be competent, the user/individual on the job must be able to:

PC1. Select and install relevant libraries and tools for model making

PC2. Split and prepare the dataset into training, validation and testing sets

PC3. Configure hyperparameters for the selected model, establish the training pipelines and execute the training phase

PC4. Store the model and network parameters to be used in the testing phase

PC5. Prevent underfitting and overfitting of the model

PC6. Solve the imbalanced dataset problem when the samples from minority class are very few

PC7. Evaluate the training performance of the machine learning model for training and validation accuracy

Execute testing phase in the machine learning project lifecycle

To be competent, the user/individual on the job must be able to:

PC8. Test the models with testing datasets

PC9. Ensure the inference time per sample is as per the business requirement

PC10. Evaluate the testing performance of the machine learning model for testing accuracy

Deployment of the developed analytics model solution into production

To be competent, the user/individual on the job must be able to:

PC11. Develop a front-end application to fetch inputs from the user and consume developed model for inference

PC12. Verify the production performance of the machine learning model

PC13. Analyse performance of the machine learning model and prepare feedback on the wrong predictions

PC14. Implement the feedback back to the training phase and retrain the machine learning model

Deployment of the analytics model solution into production line

To be competent, the user/individual on the job must be able to:

PC15. Select relevant libraries and machine learning operations (MLOPS) tools and packages for deploying the analytics model solution into production line

PC16. Install the selected libraries and tools for machine learning operations tasks

PC17. Setup the continuous training, continuous integration, and continuous delivery pipelines for the developed machine learning models

PC18. Monitor the analytics model solution performance in the deployment phase

PC19. Carry out commissioning of the end-to-end system

Knowledge and Understanding (KU)

The individual on the job needs to know and understand:

KU1. Organizational policies, procedures, and guidelines that relate to designing and maintaining databases

KU2. Organizational policies and procedures for sharing data

KU3. Organizational policies and procedures for documenting databases architectures and backup mechanisms

KU4. Who to involve while developing the different stages in the machine learning lifecycle

KU5. Descriptive and Inferential statistics for creating charts and predictive analytics modelling

KU6. Types of data wrangling and data cleaning methods to create visualization

KU7. Suitable documentation of the organization for the metadata creation

KU8. Updated internal and external network regulations

KU9. How to make an API interface for the developed machine learning model in Python using REST API framework

KU10. How to diagnose and resolve underfitting, overfitting and imbalanced dataset issues

KU11. How to use different machine learning algorithms for specific functions like regression, classification and clustering

KU12. How to use python programming constructs for developing machine learning models using open-source libraries like for example, scikit-learn

KU13. How to develop necessary front end to consume the developed analytics solution

Generic Skills (GS)

User/individual on the job needs to know how to:

GS1. Follow instructions, guidelines, procedures, rules, and service level agreements

GS2. Listen effectively and communicate information accurately

GS3. Follow rule-based decision-making processes

GS4. Make decisions on suitable courses

GS5. Plan and organize the work to achieve targets and meet deadlines

GS6. Apply problem-solving approaches to different situations

GS7. Analyse the business impact and disseminate relevant information to others

GS8. Apply balanced judgments to different situations

GS9. Check the work is complete and free from errors

Assessment Criteria

Assessment Criteria for Outcomes	Theory Marks	Practical Marks	Project Marks	Viva Marks
Execute training phase in the machine learning project lifecycle	**16**	**16**		**7**
PC1. Select and install relevant libraries and tools for model making	2	2		1
PC2. Split and prepare the dataset into training, validation and testing sets	3	3		1
PC3. Configure hyperparameters for the selected model, establish the training pipelines and execute the training phase	3	3		1
PC4. Store the model and network parameters to be used in the testing phase	2	2		1
PC5. Prevent underfitting and overfitting of the model	2	2		1
PC6. Solve the imbalanced dataset problem when the samples from minority class are very few	2	2		1
PC7. Evaluate the training performance of the machine learning model for training and validation accuracy	2	2		1
Execute testing phase in the machine learning project lifecycle	**5**	**5**		**3**
PC8. Test the models with testing datasets	2	2		1
PC9. Ensure the inference time per sample is as per the business requirement	1	1		1
PC10. Evaluate the testing performance of the machine learning model for testing accuracy	2	2		1
Deployment of the developed analytics model solution into production	**8**	**8**		**4**
PC11. Develop a front-end application to fetch inputs from the user and consume developed model for inference	2	2		1
PC12. Verify the production performance of the machine learning model	2	2		1
PC13. Analyse performance of the machine learning model and prepare feedback on the wrong predictions	2	2		1
PC14. Implement the feedback back to the training phase and retrain the machine learning model	2	2		1
Deployment of the analytics model solution into production line	**11**	**11**		**6**
PC15. Select relevant libraries and machine learning operations (MLOPS) tools and packages for	2	2		1

deploying the analytics model solution into production line				
PC16. Install the selected libraries and tools for machine learning operations tasks	2	2		1
PC17. Setup the continuous training, continuous integration, and continuous delivery pipelines for the developed machine learning models	3	3		2
PC18. Monitor the analytics model solution performance in the deployment phase	2	2		1
PC19. Carry out commissioning of the end-to-end system	2	2		1
NOS Total	**40**	**40**	**-**	**20**

ASC/N6448: Analyse Assembly line data in automotive manufacturing entity

Description:

This NOS unit is about performing analysis projects related to production process in assembly line using analytics and Business Intelligence tools.

Scope

The scope covers the following:
- Optimization of production processes to improve output and efficiency
- Identification of data needs expert analysis
- Deployment of the projects in local server or on the cloud

Elements and Performance Criteria

Optimization of production processes to improve output and efficiency

To be competent, the user/individual on the job must be able to:

PC1. Create demand forecast of the project to ensure the right spare parts mix in assembly line

PC2. Analyse safety and quality data to reduce risk in assembly line

PC3. Conduct predictive health maintenance of assembly line machines

Identification of data needs expert analysis

To be competent, the user/individual on the job must be able to:

PC4. Support in identification and verification of sensor level data sources in assembly line

PC5. Validate the data which needs expert analysis

Deployment of the projects in local server or on the cloud

To be competent, the user/individual on the job must be able to:

PC6. Support in deployment of the project on the local server or cloud

PC7. Monitor and verify the compatibility of dashboard on different devices

PC8. Monitor the alert system in real time dashboard as per requirement

Knowledge and Understanding (KU)

The individual on the job needs to know and understand:

KU1. Organizational policies, procedures, and guidelines that relate to designing and maintaining networks

KU2. Organizational policies and procedures for sharing data

KU3. Organizational policies and procedures for documenting network designs and fall-back mechanisms

KU4. Different types of visualizations charts Bar Graph, Line Graph, Stacked Bar Graph, Pie Chart, Scatter Plot Chart, etc.

KU5. Different types and categories of data variables qualitative, quantitative, nominal, ordinal, discrete, continuous, etc.

KU6. Different types of visualizations tools like Microsoft Power BI Desktop, Tableau Public

KU7. Local machine server architecture

KU8. Python based on tools like Anaconda, Jupyter, VS Code, etc.

Generic skills (GS)

User/individual on the job needs to know how to:

GS1. Follow instructions, guidelines, procedures, rules, and service level agreements

GS2. Listen effectively and communicate information accurately

GS3. Follow rule-based decision-making processes

GS4. Make decisions on suitable courses

GS5. Plan and organize the work to achieve targets and meet deadlines

GS6. Refer anomalies to the supervisor

GS7. Ask for clarification and advice from appropriate people

GS8. Analyse the business impact and disseminate relevant information to others

GS9. Apply balanced judgments to different situations

GS10. Check the work is complete and free from errors

GS11. Work independently and collaboratively

Assessment Criteria

Assessment Criteria for Outcomes	Theory Marks	Practical Marks	Project Marks	Viva Marks
Optimization of production processes to improve output and efficiency	**16**	**16**		**8**
PC1. Create demand forecast of the project to ensure the right spare parts mix in assembly line	5	5		3
PC2. Analyse safety and quality data to reduce risk in assembly line	5	5		2
PC3. Conduct predictive health maintenance of assembly line machines	6	6		3
Identification of data needs expert analysis	**11**	**11**		**5**
PC4. Support in identification and verification of sensor level data sources in assembly line	6	6		3
PC5. Validate the data which needs expert analysis	5	5		2
Deployment of the projects in local server or on the cloud	**13**	**13**		**7**
PC6. Support in deployment of the project on the local server or cloud	4	4		2
PC7. Monitor and verify the compatibility of dashboard on different devices	5	5		3
PC8. Monitor the alert system in real time dashboard as per requirement	4	4		2
NOS Total	**40**	**40**	**-**	**20**