



Soha Khalid

BWT – Data Engineering

Task 01

Try to understand the data domains and core responsibilities of every role.

1. Data Analyst:

- Collect, organize, and analyse data to answer business questions
- Create reports, visualizations, and dashboards to present findings
- Identify trends, patterns, and correlations in data

2. Data Scientist:

- Develop and apply advanced statistical and machine learning models
- Design experiments, collect data, and analyse results
- Communicate insights and recommendations to stakeholders

3. Data Engineer:

- Design, build, and maintain large-scale data systems
- Ensure data quality, security, and scalability
- Collaborate with data scientists and analysts to support data workflows

4. Data Architect:

- Design and implement data management systems and architectures
- Ensure data integration, security, and compliance
- Oversee data governance and quality

5. Business Analyst (Data Focus):

- Work with stakeholders to understand business needs and requirements
- Analyse data to inform business decisions and drive growth

- Develop and implement data-driven solutions

6. Data Quality Specialist:

- Ensure data accuracy, completeness, and consistency
- Develop and implement data validation rules and checks
- Monitor and report on data quality metrics

7. Data Governance Manager:

- Develop and implement data governance policies and procedures
- Ensure data compliance and regulatory adherence
- Oversee data security and access controls

Scenario based questions

A company XYZ has a data engineer, a data analyst, and a data scientist working there. Business users were writing queries on the transactional database of company and notice that same queries are being executed slowly, whom should they contact to debug?

- Data Scientist
- **Data Engineer ✓** If there is a proper data ware house, A data engineer will take care of it. And if transactional DB exists on production level backend engineers will take care of the database.
- Data Analyst
- It's not either's job

A business user identified the app is not recommending products based on the history of the user's recent favourites but showing random recommendations for that user. While for another user the recommendations are correct. Whom should she contact? Select one or more.

- Data Analyst
- **Data Engineer ✓** The first thing to check is the source of the data which the model is using. A data engineer is providing data to data scientists. There might be some malfunctioning in data which is causing the model to act incorrectly. A model can overfit but that's a rare chance. This problem will be debugged on the end of data engineers. If there is no issue in the data then the problem will be catered by a data scientist.
- Data Scientist
- Database Administrator
- Backend Engineer

A business wants to track how many customers are being dropped from subscriptions every day to analyse the potential cause of this, the CEO wants to see a moving dashboard on Power BI for quick visuals. Who will contribute to make this happen? Select one or more.

- Data Engineer & Data Scientist
- **Data Engineer & Data Analyst ✓**
- Data Analyst & Data Scientist
- **Data Engineer & Power BI Developer ✓**
- Data Analyst & Power BI Developer
- Power BI Developer & Data Scientist

This analysis won't be possible if a data engineer doesn't provide data to the analyst. A data analyst will then analyse the data and a Power BI developer will develop a dashboard.