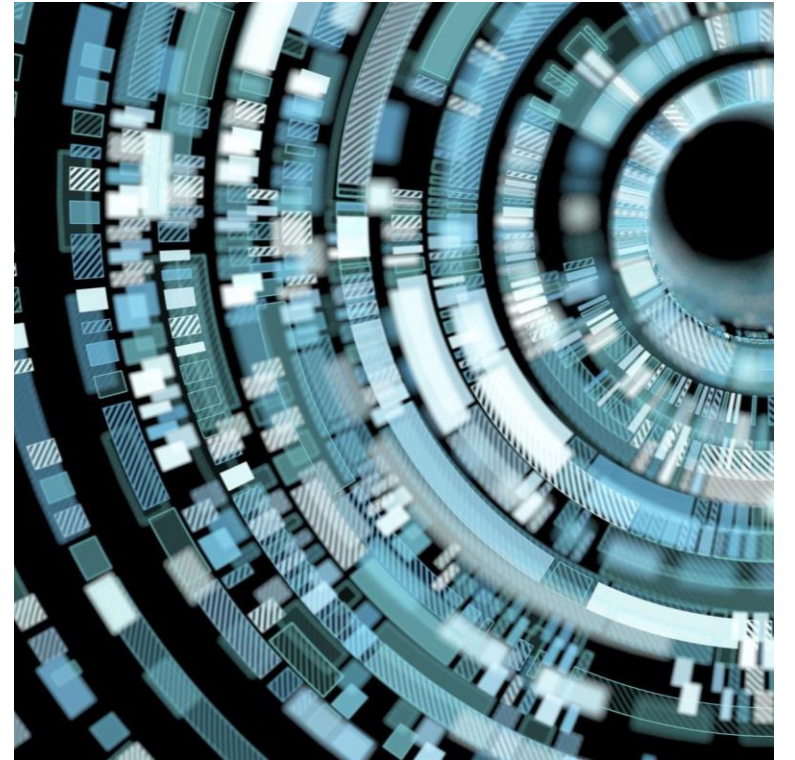


Introduction to Data Engineering

- Overview, Lifecycle, and Applications
- Pawan Kumar Sharma



What is Data Engineering ?



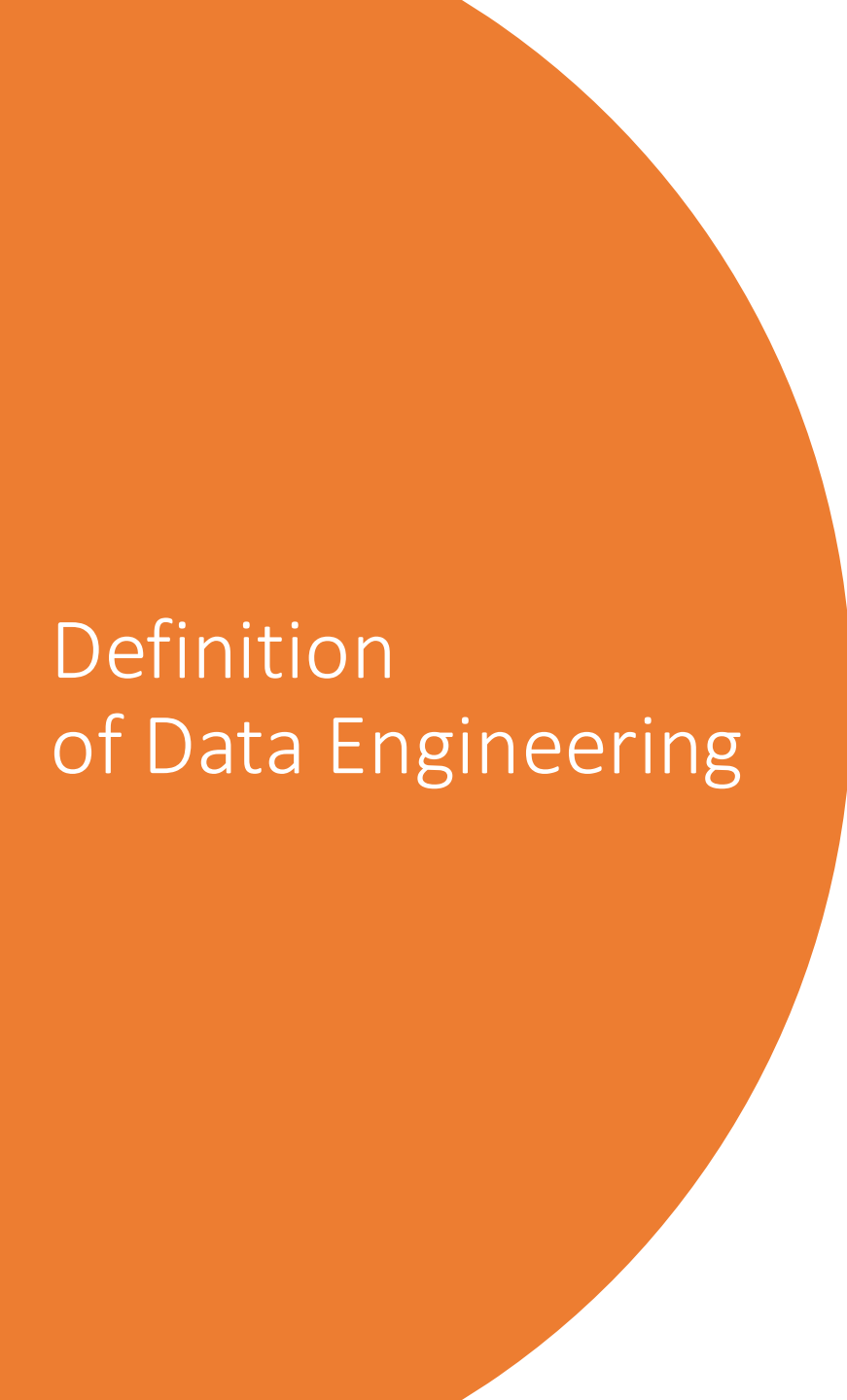
DEFINITION



IMPORTANCE



APPLICATIONS

A large orange circle occupies the left side of the slide, partially cut off by the edge.

Definition of Data Engineering

- The field of software engineering focused on the design, development, and management of systems that handle large volumes of data for Ensuring data is accessible, reliable, and ready for analysis and reporting.





Importance of Data Engineering

- Data engineering is crucial for ensuring that data is accessible, reliable, and ready for analysis. It supports data-driven decision-making, enhances business intelligence, and enables real-time analytics. By managing data pipelines, storage, and processing, data engineering allows organizations to extract valuable insights and maintain data quality and consistency.

Applications of Data Engineering

- Business Intelligence
- Machine Learning
- Data Warehousing
- Real-Time Analytics
- ETL (Extract, Transform, Load) Processes



Role of a Data Engineer



Responsibilities



Skills Required



Tools Used

Responsibilities of a Data Engineer



Data Pipeline
Development



Data Integration



Database
Management



Data Quality
Assurance



Collaboration

Skills Required for a Data Engineer



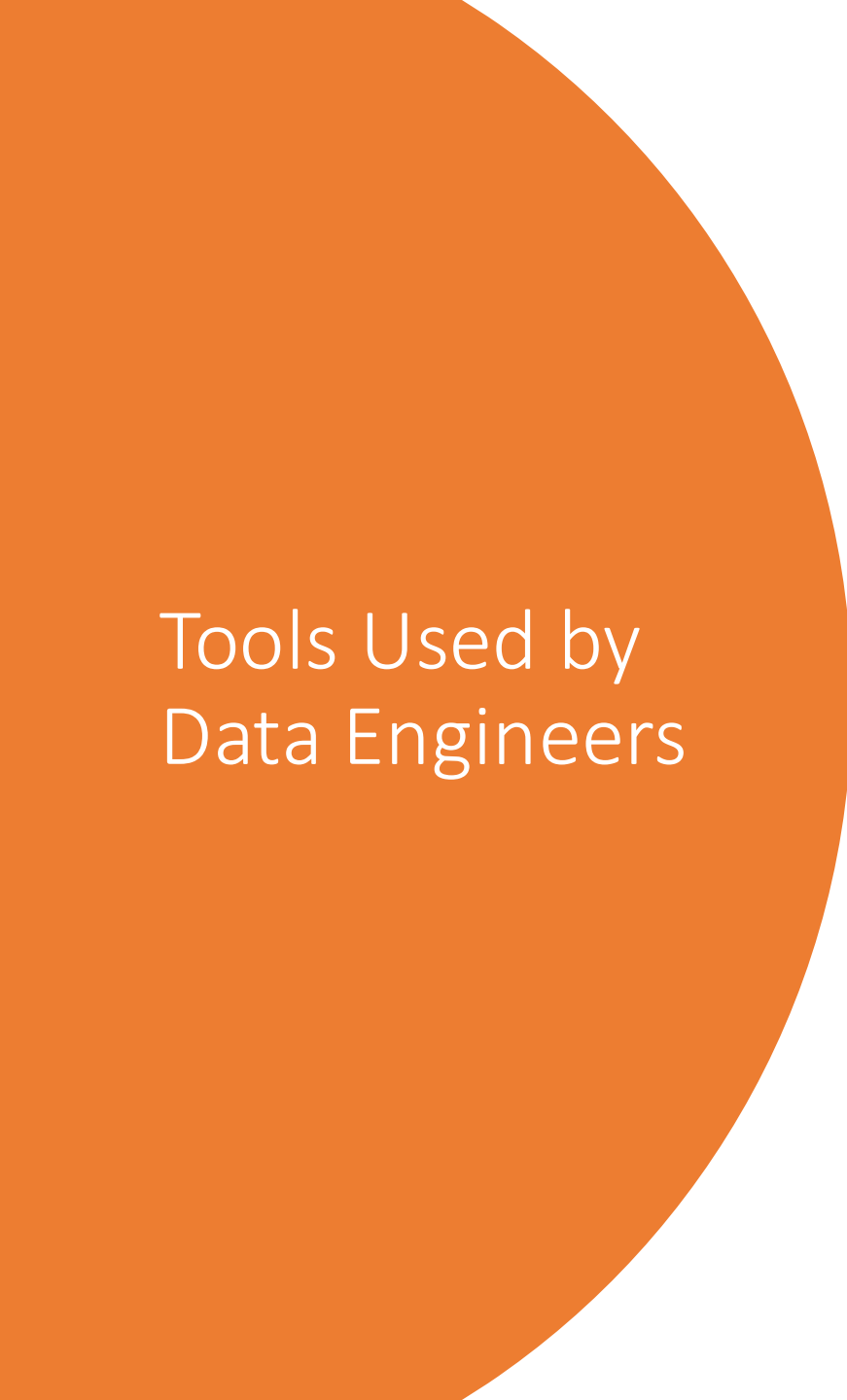
Proficiency in programming languages: Python, Java, Scala.




Knowledge of SQL and NoSQL databases.



Experience with ETL (Extract, Transform, Load) processes.

A large orange circle on the left side of the slide, partially cut off by the edge.

Tools Used by Data Engineers

- SQL Databases: MySQL, PostgreSQL, Oracle.
 - NoSQL Databases: MongoDB, Cassandra, DynamoDB.
 - ETL Tools(SSIS)
 - Data pipeline tools(Airflow)
- 
- A yellow dashed line in the bottom right corner, consisting of several short, curved segments.