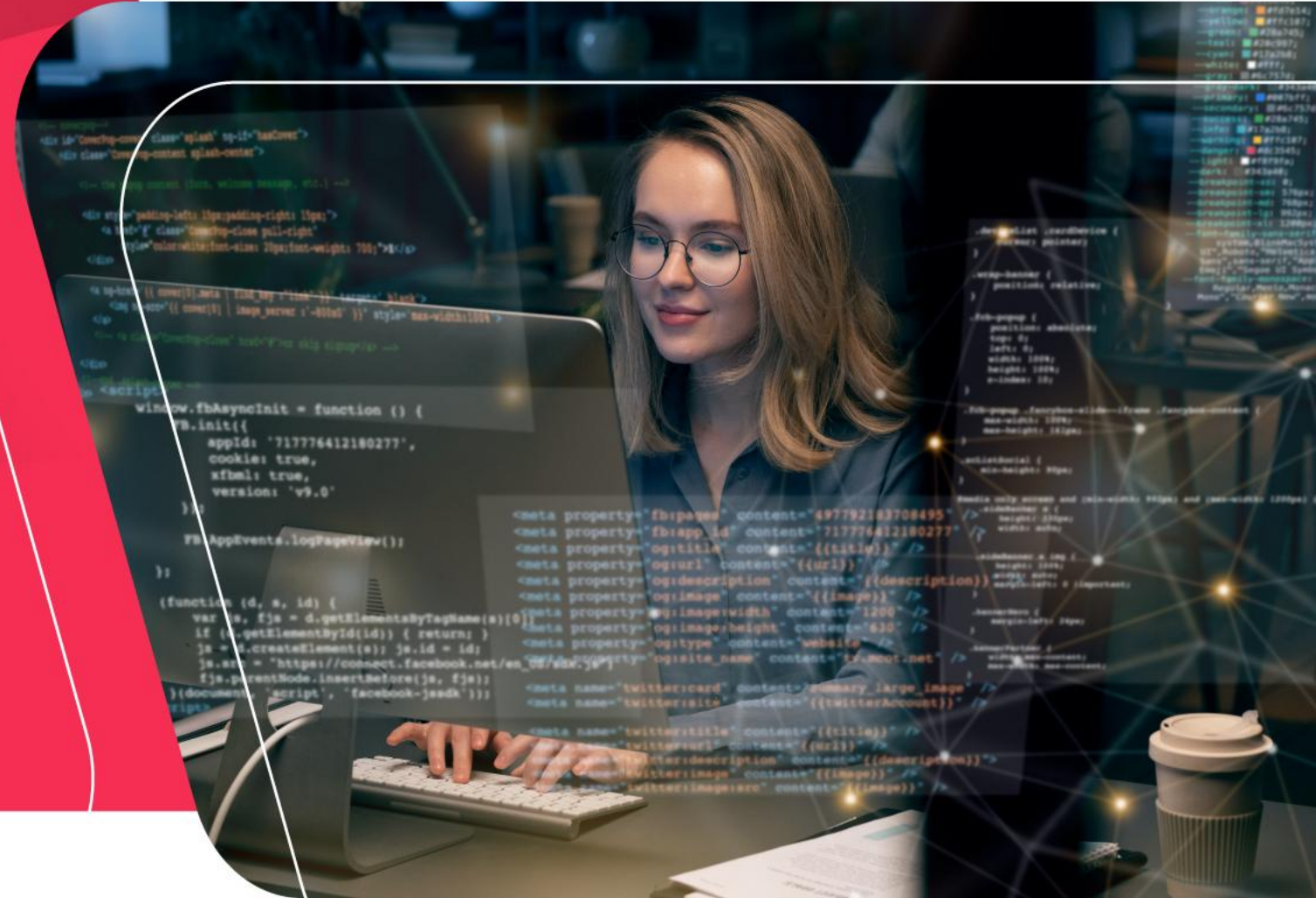# Python for Data Science

Day 01

Our Adventure Path and Destination
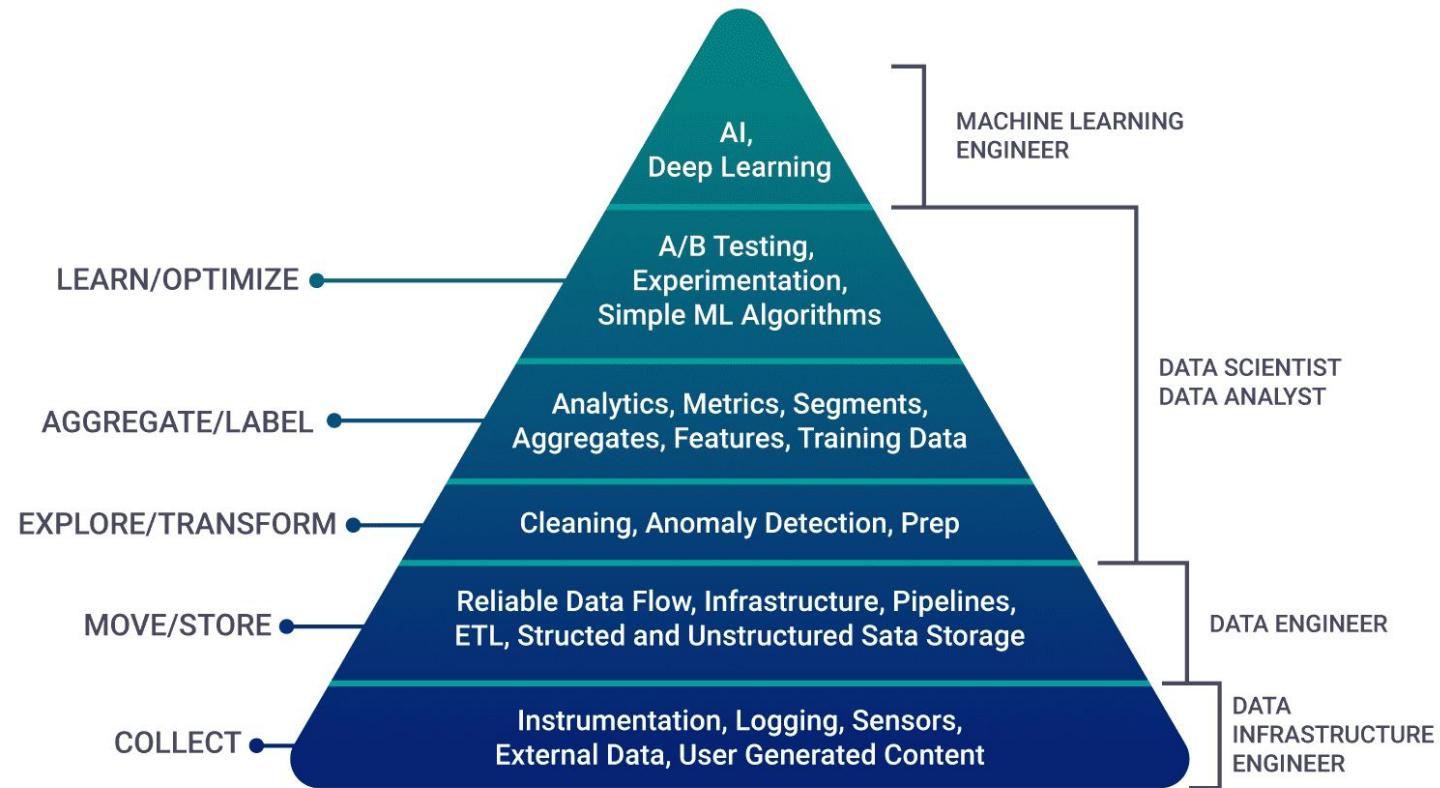
Adventure Path:

- Python Programming
- Data Warehouse Modelling
- SQL Programming
- Data Pipeline
- Distributed Computing
- Change Data Capture
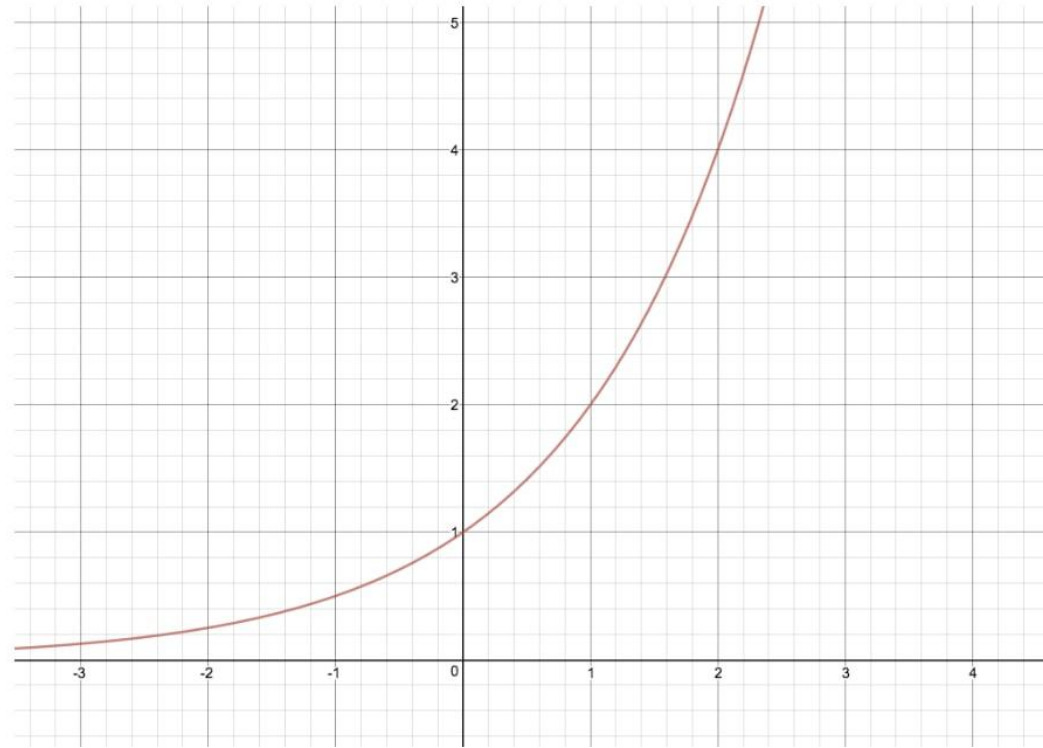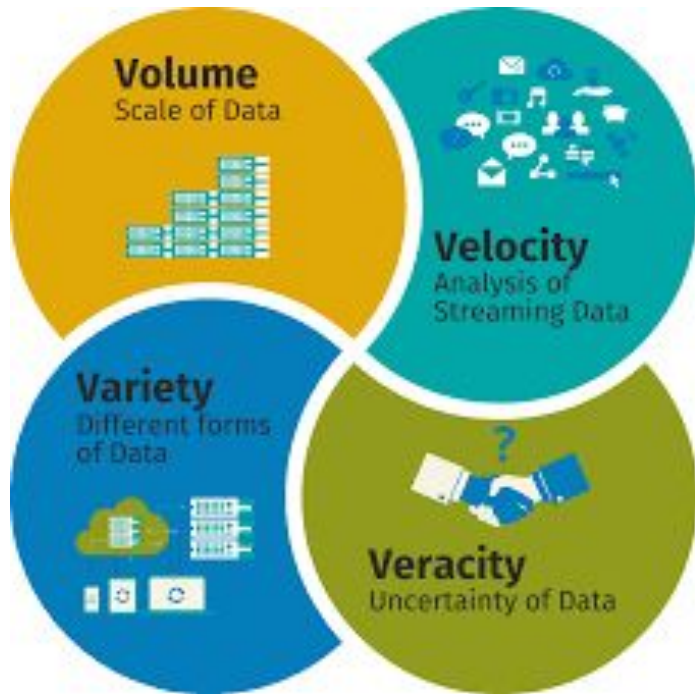- Data Quality

Our Destination:

- Design Data Warehouse and analyse according business requirements
- Understand ETL tasks from any data source into a data lake
- Put DWH design and ETL tasks together to create a data pipeline
- Implement and Monitor ETL tasks on production level
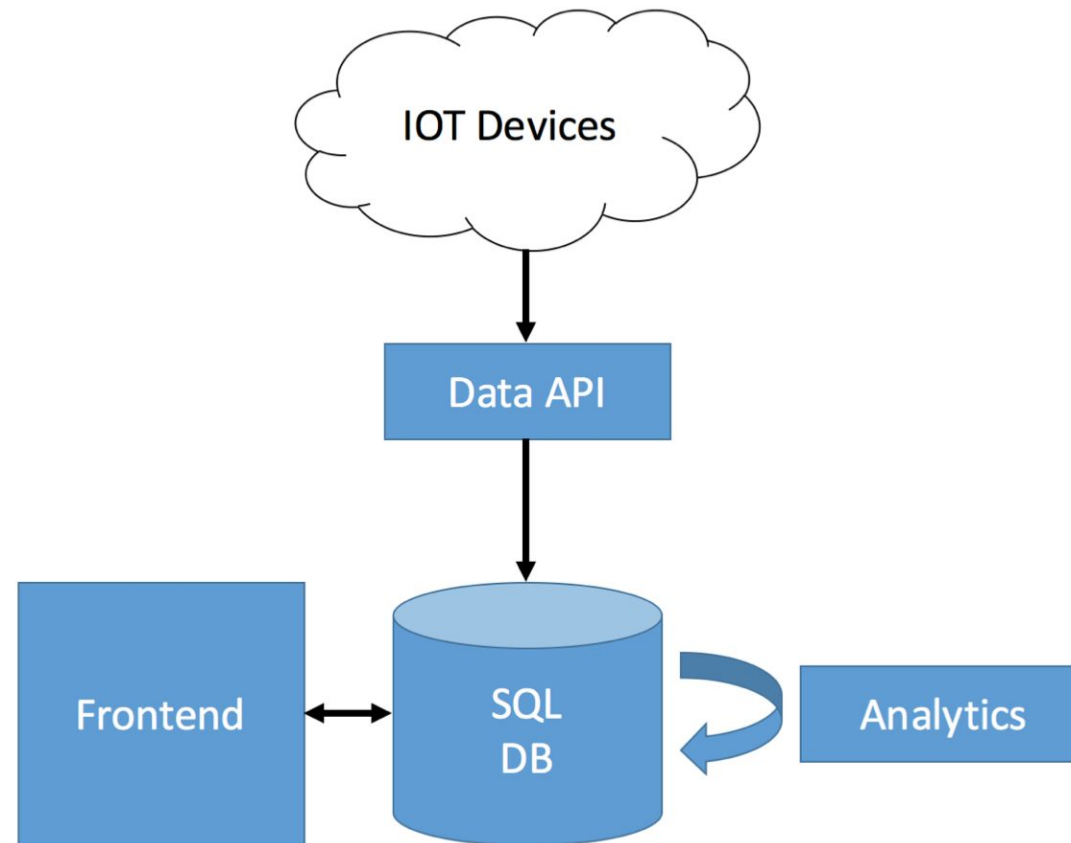
Introduction to Data Engineer
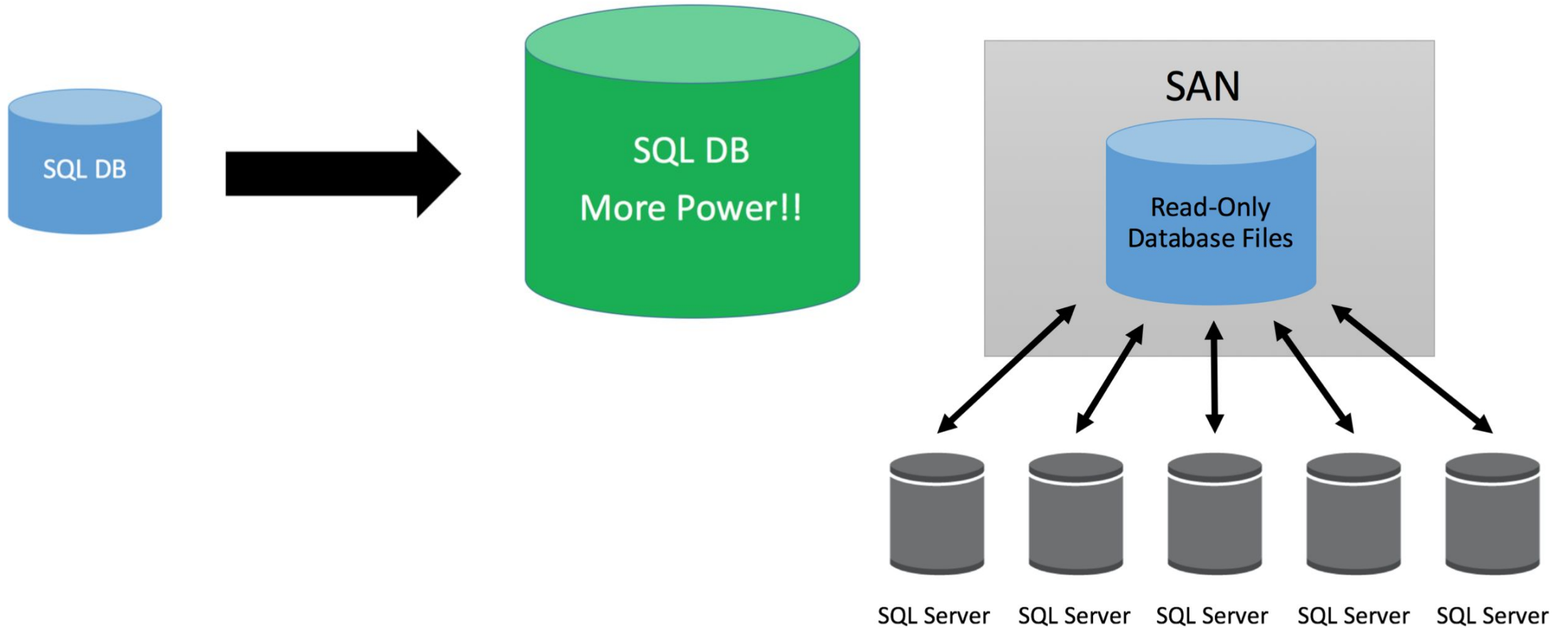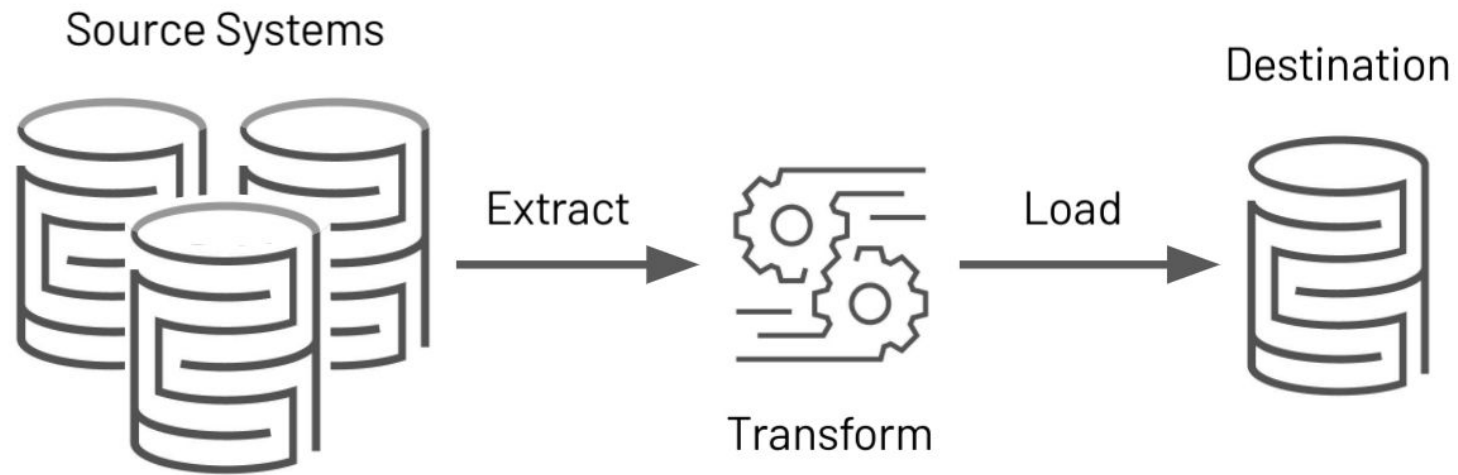
Introduction to Data Engineer

Big Data

Big Data

Big Data - Scaling

Data Ingestion - ETL

Stream Processing

Three methods of streaming:

1. At Least Once (Batch)

2. At Most Once (Semi Real-Time)

3. Exactly Once (Real-Time)