

- Batch → Process data in chunks
- Stream → Processes data in real-time as it arrives

16. Cloud vs. On-premise Databases

- cloud databases are scalable, managed by providers, while on-premise ones require in-house infrastructure and maintenance.

17. Data Governance

- Framework for managing data access, quality, policies and compliance

18. Challenges in large-scale Databases

- Scalability
- Backup & recovery
- Performance tuning
- Data consistency
- Security

19. Data Integrity & Quality

- Integrity → Accuracy and ~~data consistency~~ ^{consistency} via rules
- Quality → Completeness, validity, timeliness

20. Popular Database platforms

- MySQL/PostgreSQL
- Oracle
- SQL Server
- Mongo DB
- Snowflake/BigQuery
- Amazon Redshift