



Data Lakes and Data Warehouses Course Summary

Module summary

- Data engineers build data pipelines
- The customers of a data engineer are all the people who make decisions with data

Module summary

- The three primary advantages of doing data engineering in the cloud are:
 - Ability to separate compute and storage
 - Serverless products
 - Not having to manage infrastructure

Module summary

- Difference between a data lake and data warehouse
- Google Cloud Storage as a data lake solution
- Getting your data to Google Cloud

Module summary

- BigQuery as a data warehouse solution
- Differences between ETL, ELT and EL
- Google Cloud reference architectures for ETL, ELT and EL