



Data Engineering on
Google Cloud Platform
Streaming and
Performance
Course Summary

Hello and welcome to our course on data engineering.
My name is ...

Course Summary

- Streaming is data processing for unbounded data sets
- Cloud Pub/Sub is a messaging service used as a starting point to process streaming data
- Cloud Dataflow provides a serverless service for processing batch and streaming data
- BigQuery allows you to stream records into a table
- Cloud BigTable can handle NoSQL queries over large datasets (petabytes) in milliseconds

Course Summary

- Use BigQuery's analytic window functions for advanced analysis
- BigQuery has built in GIS functionality
- Use BigQuery's WITH clause to help modularize your queries
- There are many ways to optimize your query performance