## [Sf=ir] Institute

Google Cloud | GCP200DA

## From Data to Insights with Google Cloud Platform

3jours / 21h

## Target audience

Data Analysts, Business Analysts, Business Intelligence professionals Cloud Data Engineers who will be partnering with Data Analysts to build scalable data solutions on Google Cloud Platform

## Course Outline

The course includes presentations, demonstrations, and hands-on labs.

Module 1: Intro to Google Cloud Platform

Highlight Analytics Challenges Faced by Data Analysts

Compare Big Data On-Premises vs on the Cloud

Learn from Real-World Use Cases of Companies Transformed through Analytics on the Cloud

Navigate Google Cloud Platform Project Basics

Module 2: Analyzing Large Datasets with BigQuery

Walkthrough Data Analyst Tasks, Challenges, and Introduce Google Cloud Platform Data Tools

Demo: Analyze 10 Billion Records with Google BigQuery

Explore 9 Fundamental Google BigQuery Features

Compare GCP Tools for Analysts, Data Scientists, and Data Engineers

Lab: BigQuery Basics

Module 3: Exploring your Public Dataset with SQL

Compare Common Data Exploration Techniques

Learn How to Code High Quality Standard SQL

Explore Google BigQuery Public Datasets

Visualization Preview: Google Data Studio

Lab: Explore your Ecommerce Dataset with SQL in Google BigQuery

Module 4: Cleaning and Transforming your Data with Cloud Dataprep

Examine the 5 Principles of Dataset Integrity

Characterize Dataset Shape and Skew

Clean and Transform Data using SQL

Clean and Transform Data using a new UI: Introducing Cloud Dataprep

Lab: Creating a Data Transformation Pipeline with Cloud Dataprep

Module 5: Visualizing Insights and Creating Scheduled Queries

Overview of Data Visualization Principles

Exploratory vs Explanatory Analysis Approaches

Demo: Google Data Studio UI

Connect Google Data Studio to Google BigQuery

Lab: How to Build a BI Dashboard Using Google Data Studio and BigQuery

Module 6: Storing and Ingesting new Datasets

Compare Permanent vs Temporary Tables

Save and Export Query Results

Performance Preview: Query Cache

Lab: Ingesting New Datasets into BigQuery

Module 7: Enriching your Data Warehouse with JOINs

Merge Historical Data Tables with UNION

Introduce Table Wildcards for Easy Merges

Review Data Schemas: Linking Data Across Multiple Tables

Walkthrough JOIN Examples and Pitfalls

Lab: Troubleshooting and Solving Data Join Pitfalls

Module 8: Partitioning your Queries and Tables for Advanced Insights

Review SQL Case Statements

Introduce Analytical Window Functions

Safeguard Data with One-Way Field Encryption

Discuss Effective Sub-query and CTE design

Compare SQL and Javascript UDFs

Lab: Creating Date-Partitioned Tables in BigQuery

Module 9: Designing Schemas that Scale: Arrays and Structs in BigQuery

Compare Google BigQuery vs Traditional RDBMS Data Architecture

Normalization vs Denormalization: Performance Tradeoffs

Schema Review: The Good, The Bad, and The Ugly

Arrays and Nested Data in Google BigQuery

Lab: Querying Nested and Repeated Data

Lab: Schema Design for Performance: Arrays and Structs in BigQuery

Module 10: Optimizing Queries for Performance

Walkthrough of a BigQuery Job

Calculate BigQuery Pricing: Storage, Querying, and Streaming Costs

Optimize Queries for Cost

Module 11: Controlling Access with Data Security Best Practices

**Data Security Best Practices** 

Controlling Access with Authorized Views

Module 12: Predicting Visitor Return Purchases with BigQuery ML

Intro to ML

Feature Selection

Model Types

Machine Learning in BigQuery

Lab: Predict Visitor Purchases with a Classification Model with BigQuery ML

Module 13: Deriving Insights from Unstructured Data using Machine Learning

Structured vs Unstructured ML

Prebuilt ML models

Lab: Extract, Analyze, and Translate Text from Images with the Cloud ML APIs

Lab: Training with Pre-built ML Models using Cloud Vision API and AutoML

Module 14: Completion

