

# Sela.

GCPFDIGC

## From Data to Insights with Google Cloud

college@sela.co.il

03-6176666





# From Data to Insights with Google Cloud

GCPFDIGC - Version: 1

## 3 days course

### Description:

Explore ways to derive insights from data at scale using BigQuery, Google Cloud's serverless, highly scalable, and cost-effective cloud data warehouse. This course uses lectures, demos, and hands-on labs to teach you the fundamentals of BigQuery, including how to create a data transformation pipeline, build a BI dashboard, ingest new datasets, and design schemas at scale.

### Intended Audience:

Data Analysts, Business Analysts, Business Intelligence professionals

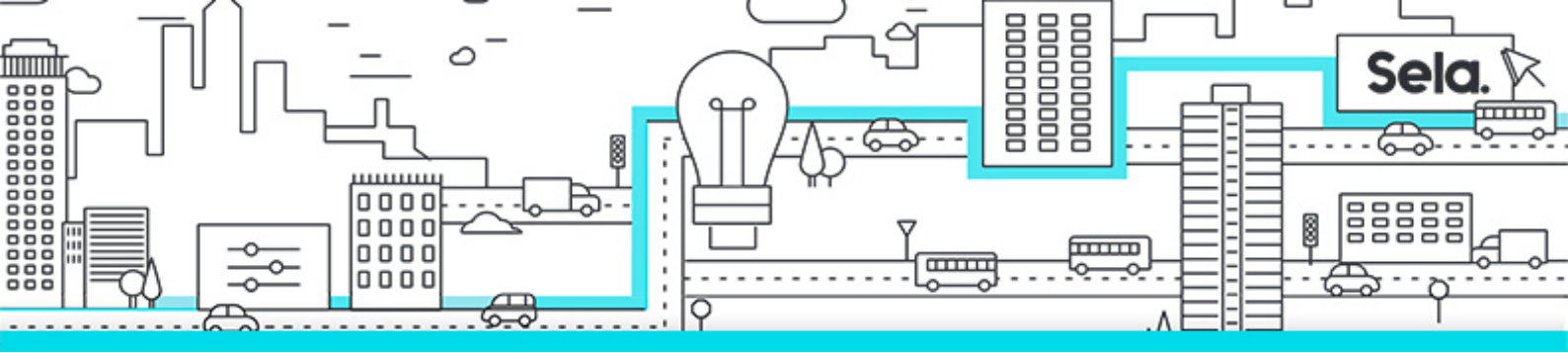
Cloud Data Engineers who will be partnering with Data Analysts to build scalable data solutions on Google Cloud Platform

### Prerequisites:

- To get the most out of this specialization, we recommend participants have some proficiency with ANSI SQL.

### Objectives:

- • Derive insights from data using the analysis and visualization tools
- on Google Cloud
- • Load, clean, and transform data at scale with Dataprep



- Explore and Visualize data using Google Data Studio
- Troubleshoot, optimize, and write high performance queries
- Practice with pre-built ML APIs for image and text understanding
- Train classification and forecasting ML models using SQL with
- BigQuery ML

## Topics:

### • Introduction to Data on Google Cloud

- Analytics Challenges Faced by Data Analysts
- Big Data On-premise Versus on the Cloud
- Real-world Use Cases of Companies Transformed Through Analytics on the Cloud
- Google Cloud Project Basics

### • Analyzing Large Datasets with BigQuery

- Data Analyst Tasks, Challenges, and Google Cloud Data Tools
- Fundamental BigQuery Features
- Google Cloud Tools for Analysts, Data Scientists, and Data Engineers

### • Exploring your Public Dataset with SQL

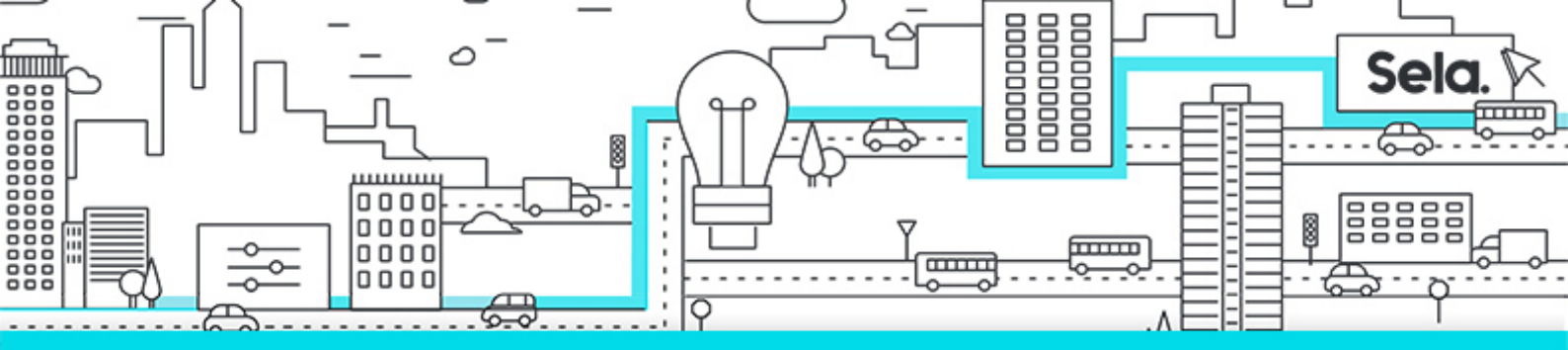
- Common Data Exploration Techniques
- Use SQL to Query Public Datasets

### • Cleaning and Transforming your Data with Dataprep

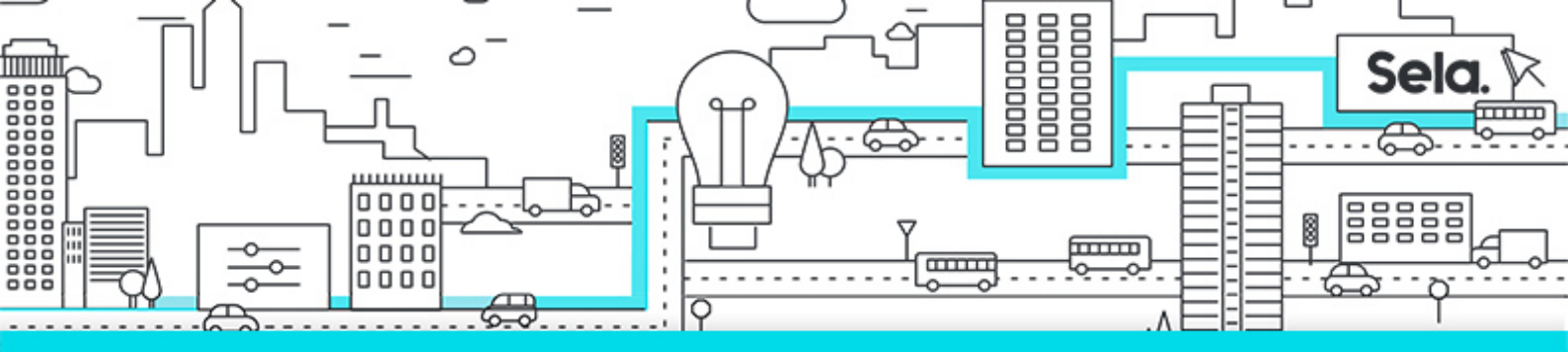
- 5 Principles of Dataset Integrity
- Dataset Shape and Skew
- Clean and Transform Data using SQL
- Introducing Dataprep by Trifacta

### • Visualizing Insights and Creating Scheduled Queries

- Data Visualization Principles



- • Common Data Visualization Pitfalls
- • Google Data Studio
- **Storing and Ingesting New Datasets**
  - • Permanent Versus Temporary Data Tables
  - • Ingesting New Datasets
- **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
  - • JOIN Examples and Pitfalls
- **Advanced Features and Partitioning your Queries and Tables for Advanced Insights**
  - • Advanced Functions (Statistical, Analytic, User-defined)
  - • Date-Partitioned Tables
- **Designing Schemas that Scale: Arrays and Structs in BigQuery**
  - • BigQuery Versus Traditional Relational Data Architecture
  - • ARRAY and STRUCT Syntax
  - • BigQuery Architecture
- **Optimizing Queries for Performance**
  - • BigQuery Performance Pitfalls
  - • Prevent Data Hotspots
  - • Diagnose Performance Issues with the Query Explanation Map
- **Controlling Access with Data Security**
  - • Hashing Columns
  - • Authorized Views
  - • IAM and BigQuery Dataset Roles



- • Access Pitfalls

- **Predicting Visitor Return Purchases with BigQuery ML**

- • Machine Learning on Structured Data
- • Scenario: Predicting Customer Lifetime Value
- • Choosing the Right Model Type
- • Creating ML models with SQL

- **Deriving Insights From Unstructured Data Using Machine Learning**

- • ML Drives Business Value
- • How does ML on unstructured data work?
- • Choosing the Right ML Approach
- • Pre-built AI Building Blocks
- • Customizing Pre-built Models with AutoML
- • Building a Custom Model