IntelliBI
Innovations Technologies

Google Cloud Platform
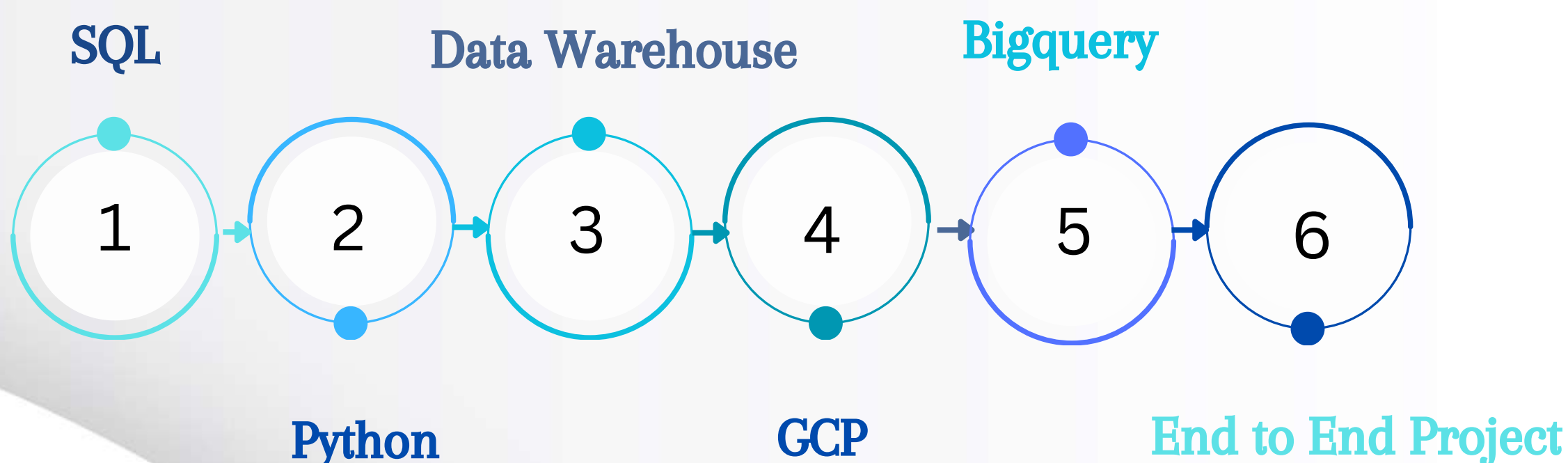
Driving Excellence:

Your BI and AI Journey Starts Here!

Google Cloud Platform(GCP)
Data Engineer

IntelliBI
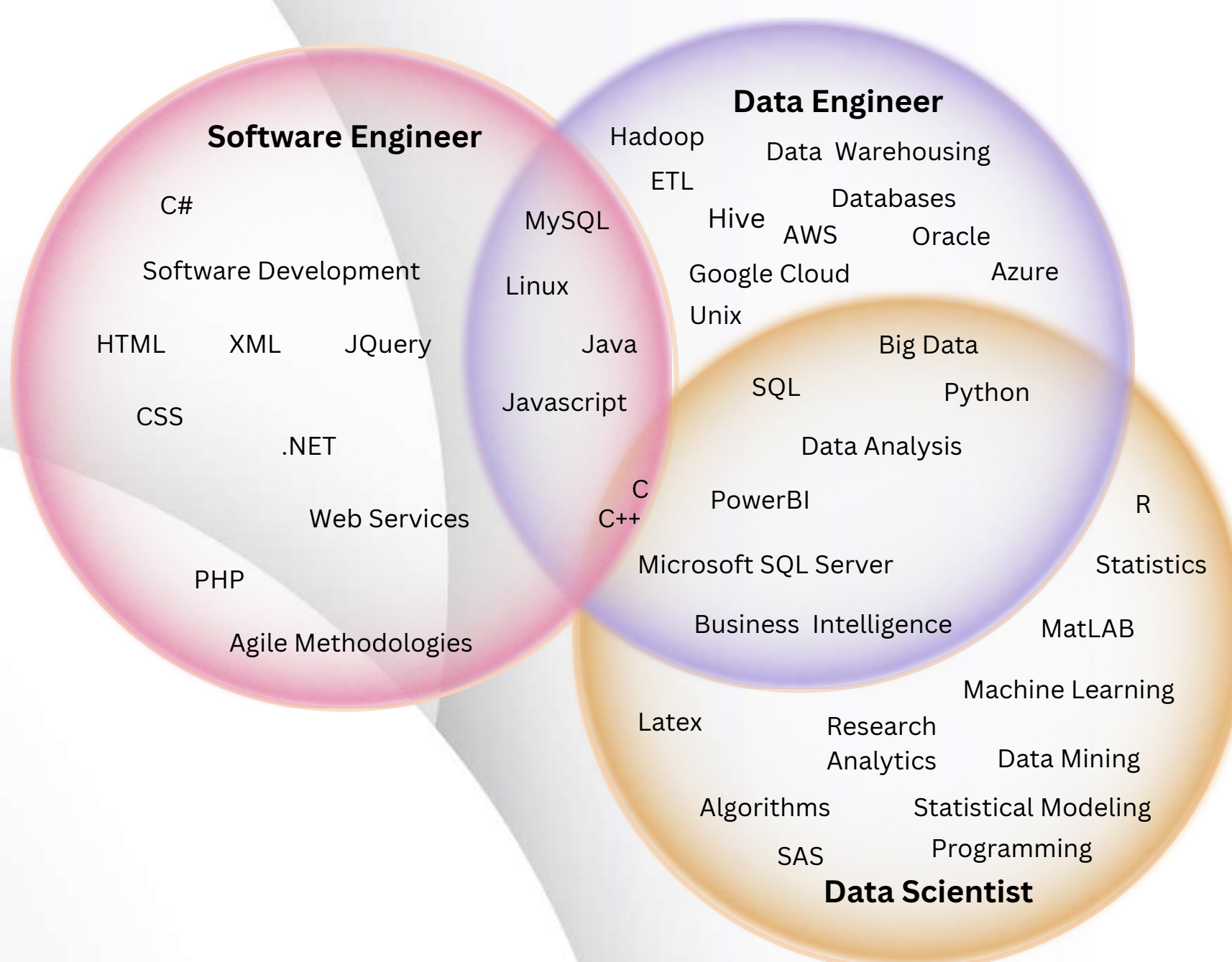## Innovations Technologies

Course Duration 5.5 Months	Batches Weekends/ Daily	Mode Online(Zoom)

Technologies

SQL		Data Warehouse		Bigquery	
1	**2**	**3**	**4**	**5**	**6**
	Python		GCP		End to End Project

Skills And Qualifications

Software Engineer

C#

Software Development

HTML   XML   JQuery

CSS

.NET

Web Services

PHP

Agile Methodologies

Data Engineer

Hadoop   Data Warehousing

ETL   Databases

MySQL   Hive   AWS   Oracle

Linux   Google Cloud   Azure

Java   Unix

Javascript   Big Data

SQL   Python

Data Analysis

C   PowerBI   R

C++

Microsoft SQL Server   Statistics

Business Intelligence   MatLAB

Machine Learning

Latex   Research   Data Mining
Analytics

Algorithms   Statistical Modeling

SAS   Programming

Data Scientist

Introduction

SQL Type

- DDL
- DML
- TCL
- DCL
- DQL

Constraints

- Primary Key
- Unique
- Foreign Key
- Check
- Default
- NOT Null

Clause

- DISTINCT
- WHERE
- ORDER BY
- GROUP BY
- HAVING

Operators

- IS NULL
- LIKE
- BETWEEN
- IN

Functions

- String
- Date
- Numeric
- Conversion
- Aggregate

Joins

- Inner Join
- Left Outer Join
- Right Outer Join
- Full Outer Join
- Cross Join
- Self Join

Subqueries

- Single-row subqueries
- Multiple-row subqueries
- Correlated subqueries

Set Operator

- Union All
- Union
- Intersect
- Minus

IntelliBI
Innovations Technologies

Views

- Simple View
- Complex View
- Materialized views

Indexes

- B-Tree Indexes
- Bitmap Indexes
- Function Based Indexes

Analytical Functions

- RANK
- DENSE_RANK
- ROW_NUMBER
- FIRST_VALUE,
- LAST_VALUE
- LAG
- LEAD

Advanced SQL

- PIVOT
- UNPIVOT
- PARTITIONS

PLSQL

- Stored Procedure
- Stored Function

Python

- Setup Python
- "Hello world"
- Datatypes-Numbers, Strings, Boolean
- String Functions: Indexing, Slicing
- Print Formating with Strings
- Variables
- Operators
- List
- Dictionary
- Tuples
- Sets in Python
- If, Elif and Else
- For Loops
- While Loops (including break and continue, Pass)
- Functions and Recursions
- Local and Global variable
- Error and Exception Handling

Pandas And Numpy

- Create Pandas DataFrame
- Numpy Array Operations
- Pandas DataFrame Operations

Datawarehouse

- Why Data warehousing?
- What is Data warehouse?
- DW Architecture
- OLTP vs OLAP
- ETL Basics
- Dimensional Modeling (Fact, Dimensions)
- Star Schema
- Snowflake Schema
- Slowly Changing Dimension

IntelliBI
Innovations Technologies

Introduction

- Explore the role of a data engineer.
- Analyze data engineering challenges.
- Intro to BigQuery.
- Data Lakes and Data Warehouses.
- Transactional Databases vs Data Warehouses.
- Partner effectively with other data teams.
- Manage data access and governance.
- Build production-ready pipelines.
- Review GCP customer case study.

Building a Data Lake

- Introduction to Data Lakes.
- Data Storage and ETL options on GCP.
- Building a Data Lake using Cloud Storage.
- Securing Cloud Storage.
- Storing All Sorts of Data Types.
- Cloud SQL as a relational Data Lake.

Building a Data Warehouse

- The modern data warehouse.
- Intro to BigQuery.
- Getting Started.
- Loading Data.

- Exploring Schemas.
- Nested and Repeated Fields.
- Optimizing with Partitioning and Clustering.
- Schema Design.
- Lab: Loading Data into BigQuery.

Introduction to Building Batch Data Pipelines, EL, ELT and ETL

- Quality considerations.
- How to carry out operations in BigQuery.
- Shortcomings.
- ETL to solve data quality issues.

Executing Spark on Cloud Dataproc

- The Hadoop ecosystem.
- Running Hadoop on Cloud Dataproc.
- GCS instead of HDFS.
- Optimizing Dataproc.

Serverless Data Processing with Cloud Dataflow

- Cloud Dataflow
- Why customers value Dataflow.
- Dataflow Pipelines.
- Dataflow Templates.
- Dataflow SQL.

Manage Data Pipelines with Cloud Data Fusion and Cloud Composer

- Building Batch Data Pipelines visually with Cloud Data Fusion.
- Components.
- UI Overview.
- Building a Pipeline.
- Exploring Data using Wrangler.
- Orchestrating work between GCP services with Cloud Composer.
- Apache Airflow Environment.
- DAGs and Operators.
- Workflow Scheduling.
- Monitoring and Logging.

Introduction to Processing Streaming Data

- Processing Streaming Data.

Serverless Messaging with Cloud Pub/Sub

- Cloud Pub/Sub.
- Publish Streaming Data into Pub/Sub.

Cloud Dataflow Streaming Features

- Cloud Dataflow Streaming Features.
- Streaming Data Pipelines.

High-Throughput BigQuery and Bigtable Streaming Features

- BigQuery Streaming Features.
- Cloud Bigtable.
- Streaming Analytics and Dashboards.

Advanced BigQuery Functionality and Performance

- Analytic Window Functions.
- Using With Clauses.
- GIS Functions.
- Performance Considerations.
- Optimizing your BigQuery Queries for Performance.

Introduction to Analytics and AI

- What is AI?.
- From Ad-hoc Data Analysis to Data Driven Decisions.
- Options for ML models on GCP.

Prebuilt ML model APIs for Unstructured Data

- Unstructured Data is Hard.
- ML APIs for Enriching Data.
- Using the Natural Language API to Classify Unstructured Text.

IntelliBI
Innovations Technologies

Big Data Analytics with Cloud AI Platform Notebooks

- What is a Notebook?
- BigQuery Magic and Ties to Pandas.

Production ML Pipelines with Kubeflow

- Ways to do ML on GCP.
- Kubeflow.
- AI Hub.

Custom Model building with SQL in BigQuery ML

- BigQuery ML for Quick Model Building.
- Supported Models.

Custom Model building with Cloud AutoML

- Why Auto ML
- Auto ML Vision.
- Auto ML NLP.
- Auto ML Tables.

Our Alumni Work in Top Companies