WeCloud**Data**

Data Engineering Certificate Program

COURSE INFO & SYLLABUS

CAREER FACTS

WHAT DO DATA ENGINEERS DO?

Data Engineers play an essential role in organizations that collect and manage big data.

In a typical data organization, data engineers gather and collect the data, store it, do batch processing or real-time processing on it, and serve it via the database to a data scientist who can easily query it. In essence, data engineers help manage how data flows in an enterprise's data architecture.

JOB MARKET

The market has seen a surge in demand for data scientists in the past several years and we see almost all universities and colleges offer some kind of data science courses and programs.

However, data engineers are usually harder to train and source because the program needs to be very practical/hands-on and there is not much theory to teach.

JOB TITTLE

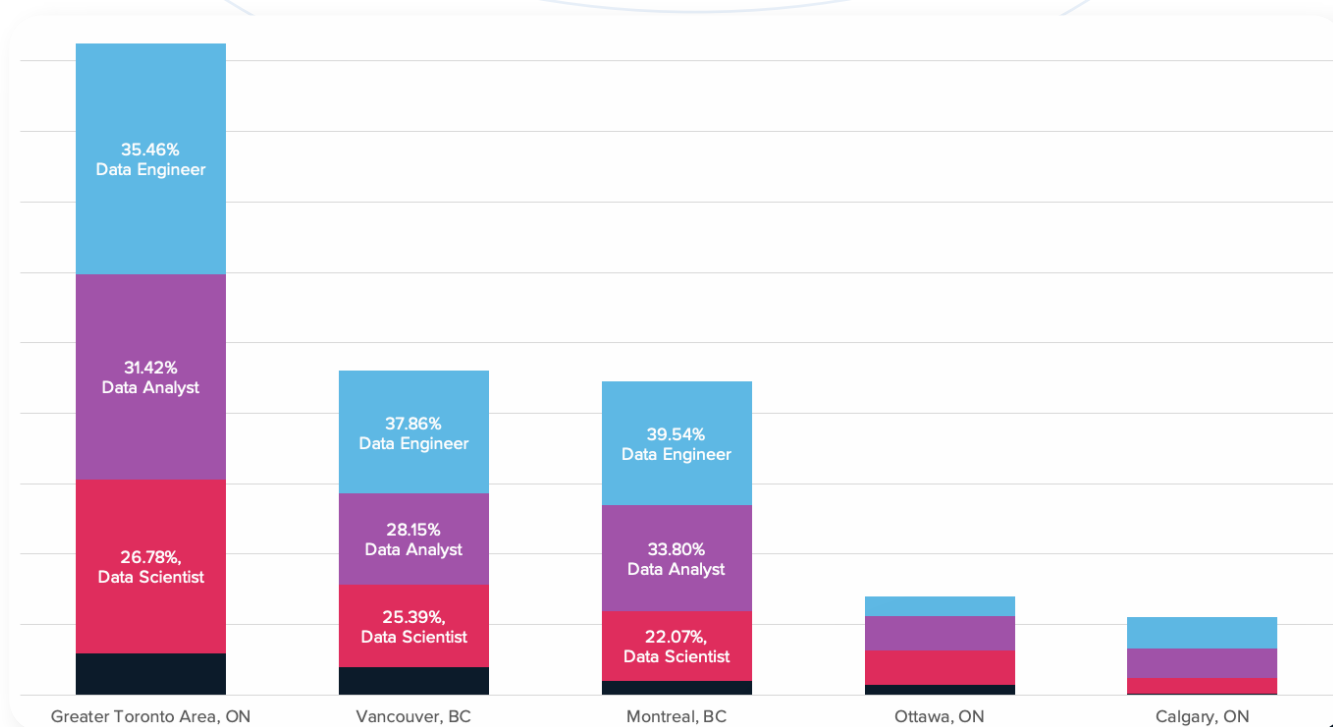| ETL/ELT Engineer	Data-Savvy Software Engineer	BI/Data Warehouse	ML Data Engineer
ETL Developer Data Integration Engineer SQL Developer Big Data Engineer	Computer Science/Engineering Mechanical Engineering Electrical Engineering Data Platform Engineer	Data Architect BI Engineer Big Data Developer Data Warehouse Engineer BI Specialist	Data Engineer Big Data Engineer AI/ML Engineer Model Automation Engineer

DEMAND FOR DATA ENGINEERING

35.46% Data Engineer	
31.42% Data Analyst	
26.78%, Data Scientist	

37.86% Data Engineer

28.15% Data Analyst

25.39%, Data Scientist

39.54% Data Engineer

33.80% Data Analyst

22.07%, Data Scientist

Greater Toronto Area, ON	Vancouver, BC	Montreal, BC	Ottawa, ON	Calgary, ON

DATA ENGINEER SALARIES IN CANADA

Average in Canada

$83,662 per year

Salary Distribution

$65,000 $229,000

TECHNICAL SKILLS

Java

Hive

Kafka

AWS

NoSQL

Python

SQL

Hadoop

R

Azure

Spark

Scala

Redshift
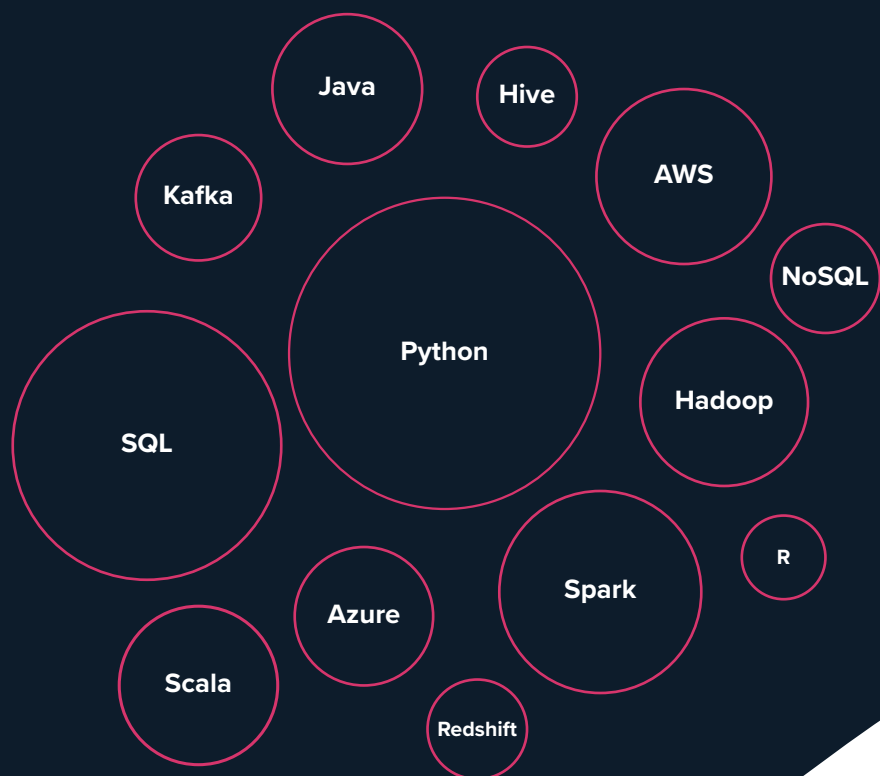
REQUIRED YEARS OF EXPERIENCE

| Years of experience	Percentage
1	11.76%
2	7.06%
3	31.47%
4	8.43%
5	31.57%
6	1.37%
8	8.33%
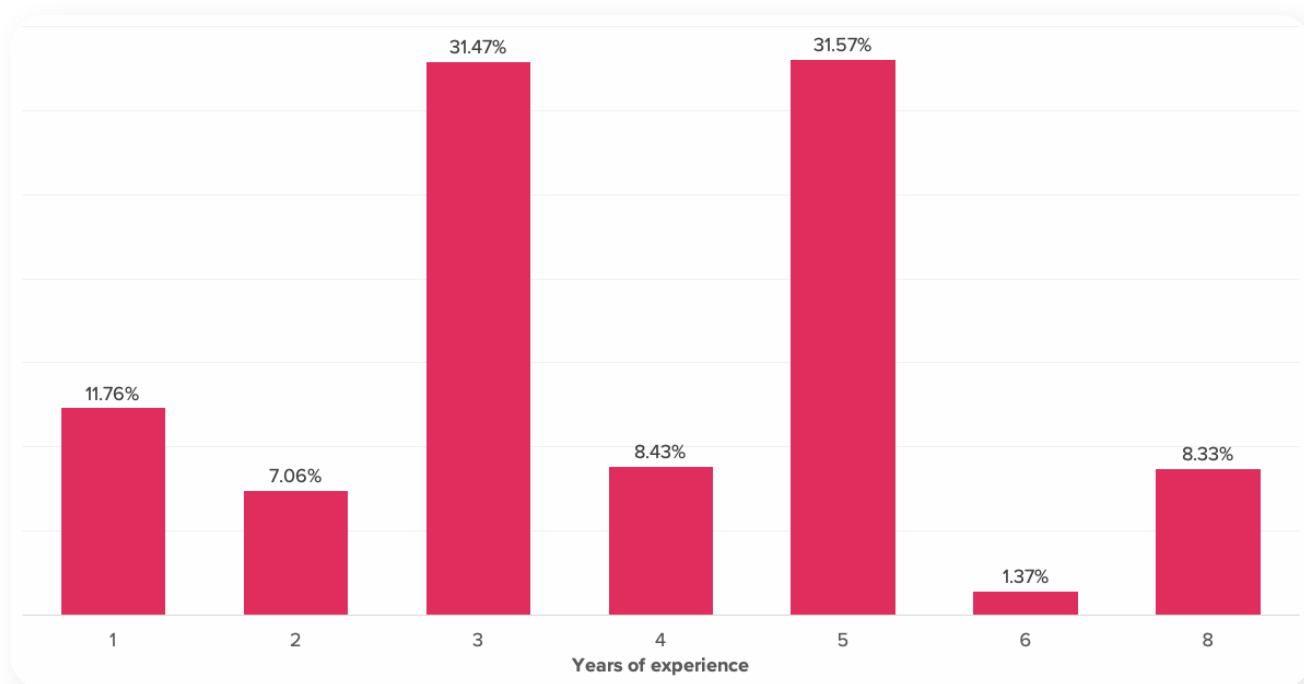
ABOUT THE PROGRAM

OVERVIEW

WeCloudData's Data Engineering certificate program focuses on helping students to acquire the essential data engineering skills, gain professional experience and prepare for data engineer careers.

WeCloudData is a Toronto-based data skill training academy. We've partnered with many Canadian corporations on upskilling their employees. No. 1 best data science bootcamp in 2020-2021 by SwitchUp.
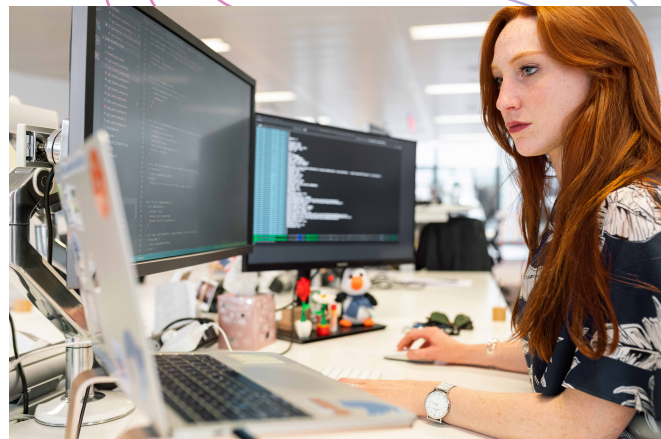
WHO IS IT FOR?

This weekend program is suitable for students who want to become a Machine Learning Engineer or AI Engineer

Students accepted into the program should have acquired fundamental knowledge and skills of Python programming and machine learning already

Recent Grads

from CS and Engineering who want to gain advantage in the job market by gaining practical ML engineering skills

Data Scientists

who want to build more advanced skills in MLOps and model deployment

Entrepreneurs

who want to build ML systems and platforms for her/his next startup

Career Switchers

who have a full-time job and would like to switch career

PROGRAM DETAILS

24 WEEKS
Wednesday evening & Saturday full day

10+ GUEST SPEAKERS
to make connection with

2 END-TO-END PROJECTS
to build your portfolio

TA OFFICE HOUR
Mon-Thu 6PM-8PM

4 : 1 RATIO
students vs instructors

1 BUSINESS PROJECT
from a real client
(full-time students only)

LEARNING OUTCOMES

- Solid understanding of major big data and cloud platforms such as Hadoop, Spark, Databricks and AWS
- Able to build and deploy data pipelines on cloud platforms using workflow orchestration tools such as Apache Airflow and DBT
- Comfortable with building real-time and batch data ingestion
- Deep understanding of the pros and cons of different database systems
- Hands-on experience with data warehouse modeling and ETL
- Hands-on experience with DataOps tools such as Git, Docker, Kubernetes
- Familiar with all things data engineering for data scientists: maintaining pipelines, building feature stores, scheduling model training, monitoring, and deployment
- Solid understanding of data governance and managing data engineering projects

PREREQUISITES

- Know the basics of Python and SQL (We will also provide online Python and SQL courses for you)
- An operating system with at least 8G of RAM (16G is preferred), 4 cores of CPU

GUEST SPEAKERS

THOMSON REUTERS | Microsoft | LAYER 6 AI | rubikloud | DESSA | RBC | TORONTO

adeptmind | CANACCEDE FINANCIAL GROUP | amazon | ELEMENT AI | payright

TOP-NOTCH CURRICULUM

The only academy that provides the most comprehensive data engineering curriculum

Topics	Content	Topics	Content
Programming	Linux SQL Python/Scala	Data Pipelines	Apache Airflow CI/CD Basics
Cloud Computing	Amazon Web Service	Docker Kubernetes	Containerization Container Orchestration
Big Data	Hadoop, MapReduce, Hive Apache Spark NoSQL Database Presto/Druid	Data Governance	Data Quality Data Cataloguing Metadata Management Data Privacy
Data Warehouse	Snowflake, Redshift, BigQuery Data Modelling Database Sharding, Replication	AI for Data Engineers	Machine Learning Processes Building Feature Stores Model Deployment
ETL/ETL	Data Ingestion Data Integration Streaming Data Processing	Data Project Management	Project Management Time Management Communications

FLEXIBLE SCHEDULE

PART-TIME

$9,100

- ✓ Learning Schedule:
  - Wednesday Evening
  - Saturday Full Day
- ✓ 6-month Career Support after graduation
- ✓ Hands-on projects
- Full-time TA (additional projects)
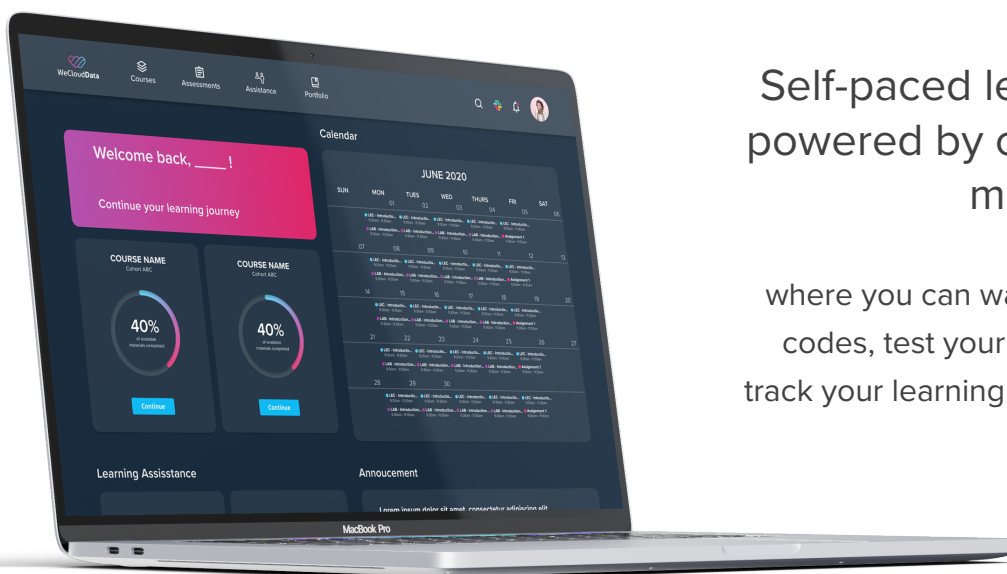- Real client project

FULL-TIME

$13,000

- ✓ Learning Schedule:
  - Wednesday Evening
  - Saturday Full Day
- ✓ 6-month Career Support after graduation
- ✓ Hands-on projects
- ✓ Full-time TA (additional projects)
- ⭐ **Real client project**

CUSTOMIZED LEARNING PLATFORM

Self-paced learning guided and powered by our unique learning management system

where you can watch lectures, work on your codes, test your progress with live quizzes, track your learning journey, and communicate with instructors and TA's.

CURRICULUM
Semester 1 | 12 weeks

WEEK 1
Linux

- Linux System basic
- Linux System Configuration
- Linux Text Editors
- Linux Commands
- Linux Shell Scripting
- Git
- Work on A Mini Project

WEEK 2
Docker

- Docker Container
- Docker Compose
- Docker Image Creation
- Docker in CICD
- Work on A Mini Project

WEEK 3
Data Warehouse I

- Advanced SQL
- Data Modeling
- Snowflake

WEEK 4
Data Warehouse II

- ETL design
- ETL SQL Scripting
- ETL Shell Scripting
- Work on A Mini Project

WEEK 5
Python in Cloud Projects I

- Useful Python Packages
- The Best Practice in Python projects
- Project Setup on Cloud (AWS)
    - AWS EC2
    - AWS S3
    - AWS IAM

WEEK 6
Python in Cloud Projects II

- AWS Lambda
- A Cloud Python project

WEEK 7
Big Data I

- Big Data Theory
- Presto and Hadoop Review
- Apache Spark Theory
- Spark Programing I

WEEK 8
Big Data II

- Spark Programing II
- Best Practice in Spark Projects
- Work on A Mini Project

WEEK 9
Big Data III

- AWS EMR
- Spark in EMR
- CICD in Spark EMR projects
- Databricks
- Spark Cluster Optimization

WEEK 10
Data Pipeline: Apache Airflow

- Airflow Setup
- Airflow DAGs
- Operators and Hooks
- Airflow Operation
- Work on A Mini Project

WEEK 11
Data Lake I

- Data Lake Catalog
    - AWS Glue
- Analysis on Data Lake
    - AWS Athena
    - Apache Superset

WEEK 12/13
Mid-Term Project Preparation

- Advanced Spark Project
- Advanced Data Pipeline Setup
  - EC2, S3, Lambda, EMR, Glue, Athena
  - Apache Airflow and Superset
  - Docker and Airbytes

WEEK 14/15
Mid-Term Project

- Mid-term Project

WEEK 16
No-SQL Database: DynamoDB

- DynamoDB Theory and Setup
- DynamoDB Syntax
- Best Practice in Real Project
- Work on A Mini Project

WEEK 17
No-SQL Database: Elasticsearch

- Elasticsearch Theory and Setup
- ELK Query
- Best Practice in Real Project
- Work on A Mini Project

WEEK 18
Data Lake II: Data Ingestion — Nifi

- Nifi Theory and Setup
- NIfi Operators
- Best Practice in Real Project
- Work on A Mini Project

WEEK 19
Data Lake II: Data Ingestion — Kafka

- Kafka Theory and Setup
- Kafka Structures: Topics, Producer, Consumer
- Data Ingestion with Kafka
- Work on A Mini Project

WEEK 20
Data Lake II: Lakehouse — Apache Hudi

- The Theory of Lakehouse
- The Architecture of Apache Hudi
- Hudi Lakehouse on S3
- Hudi Best Practice
- Work on A Mini Project

WEEK 21
Data Steaming: Spark Streaming

- Data Streaming Theory
- Spark Streaming Programing
- NIFI in Streaming

WEEK 22
Data Steaming: Streaming Project

- Spark Streaming Project Setup
- Streaming best practice

WEEK 23/24
Capstone Project

- Capstone Project

ADMISSION PROCESS

Book a Counseling Meeting

Our program advisor will have a 1-on-1 meeting with you to see if the program is a good fit for you. There is no application fee. It takes about 30 minutes

Pass the Challenge

There will be a technical test and an interview. Applicants spend up to 2 hours on the challenge.

Reserve the Spot & Sign the Contract

Our admission officer will work with you directly to help you fill out a contract, pay $500 deposit and assist you to apply any kind of grant and finance options.

Start your Pre-bootcamp Learning

Starting ahead will gain you more experience and competence. Research shows that preview and preparation account for 73.7% success in academic achievement of university students.

TUITION, GRANTS AND FINANCE OPTIONS

$9,100
PART-TIME

$13,000
FULL-TIME

As an Ontario registered private career college, you can apply for student line of credit from BMO with lower interest rate.

DO YOU KNOW?

With Ontario Second Career grant you may be eligible for up to $28,000 for costs including: tuition, books, manuals, transportation, basic living allowance, child care.

Contact our program advisor, Amir, for more information
amir.asadian@weclouddata.com