WEB AGE SOLUTIONS

Providing Technology Training and Mentoring For Modern Technology Adoption

PROGRAM SYLLABUS

Data Engineering

Overview

Learn to design data pipelines and API's in the cloud, perform analytics in the cloud and automate this complete process flow. At the end of the program, demonstrate your mastery by finishing a capstone project that combines all the critical concepts leant.

You need to have intermediate SQL and Python programming skills. If you do not have the required prerequisites, the following courses will get you ready:

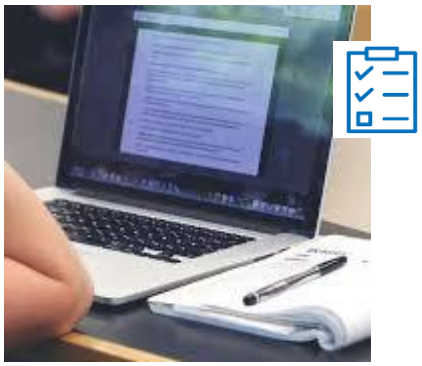
- Introduction to Python Development
- Introduction to SQL

Educational Objectives:

You will learn to

- Create scalable data processing solutions using Hive and Spark
- Build data pipelines with Storm and Kafka
- Create scalable and efficient data warehouses
- Create Data Services and API's
- Work with the analytics solutions in the cloud and automate the process flow
- Build an end to end solution to deliver data to your clients who are creating analytics solutions

Estimated Time: 4 to 5 weeks (fifth week is optional)



Prerequisites: Intermediate Python & SQL

Flexible Learning: We can deliver virtually or in-person; in half-days, 2 evenings a week, or on a 9-5 schedule.

Need Help? Discuss this program with a Web Age advisor.

Week 1: Spark and Data Pipelines Week

In this week, you will learn to create scalable data processing solutions using Hive and Spark. You will also build data pipelines with Storm and Kafka.

Course Project

For this project you will use an existing Order Management System, a Microservices based application, and implement the functionality to integrate Kafka. The expectation is that you will be able to make the necessary changes to an existing project to decouple the Microservices using Kafka

	LEARNING OUTCOMES	
LESSON ONE	**Parallel Computing (Hadoop/Hive/Spark)**	• start using the Hadoop platform • create scalable data processing solutions using Apache Hive and Apache Spark • compare suitable use cases for Hive and Spark
LESSON TWO	**Data Pipeline (Storm, Kafka)**	• appreciate the value proposition of Storm and Kafka messaging platforms • build data pipelines with Storm and Kafka

Week 2: SQL and APIs in the Cloud Week

In this week, you will learn about the AWS services. You will also learn about how to create Data Services and API's.

Course Project

For this project you will use an existing MS SQL Database and export the data into JSON files. The expectation is to stream the JSON output through Storm or Kafka to an AWS Aurora RDS and select the table data with Athena.

	LEARNING OUTCOMES	
LESSON ONE	**Introduction to AWS**	• recognize AWS capabilities and its services • match practical use cases with AWS services • appreciate and communicate the benefits of managed services
LESSON TWO	**Data Services and API's (Docker/K8s/Microservices)**	• appreciate the value of API management solutions for data processing needs • take advantage of data services using APIs
LESSON THREE	**Data formats (CSV, JSON, XML, Avro, Parquet)**	• identify a variety of data formats and their specific use cases • make informed decision about where to use which format

LESSON FOUR **Cloud-based RDS servers**

- effectively use the Aurora RDS service
- appreciate the benefits of AWS managed RDS services

Week 3: Data Analytics in the Cloud Week

In this week, you'll learn about the architecture of the analytics solutions in the cloud and work with those. You will also automate these solutions truly in the infrastructure-as-a-service fashion.

Course Project

For this project, you will use the Amazon Web Services (AWS), which includes cloud services that can be used to support all stages of data analysis from storage to analytics. For this project, you will use AWS services to store, process, query and visualize a given data set. The expectation is you will utilize AWS CDK, AWS Glue, AWS Athena and AWS QuickSight to accomplish these tasks. You are given tasks to perform and are expected to use what you have learned in course to fill in the steps.

	LEARNING OUTCOMES	
LESSON ONE	**Cloud - AWS Analytics Solutions and Architecture (Lambda, S3, Glue, Redshift) [Or Azure Data Factory(HDInsights) /GCP equivalent]**	• differentiate between a variety of AWS analytics tools and services, including Lambda, Glue, and Redshift • make informed decisions about applicable use cases for each tool and service presented in the module • identify suitable data processing patterns in the cloud
LESSON TWO	**JavaScript (Typescript) - working with AWS CDK [or Azure/GCP equivalent]**	• appreciate the benefits of using AWS CDK • create and deploy re-usable CDK scripts

Week 4 (Optional): Coding Best Practices, EDA and Data Security - "Thinking-out-of-the-Box" Week

In this optional week, you'll learn about coding best practices, explore visualization and EDA for data quality and learn the best practices of data security.

	LEARNING OUTCOMES	
LESSON ONE	**Coding best practices**	• recognize coding best practices • write better code • reduce the number of defects in their programs
LESSON TWO	**Visualization and EDA (for data quality)**	• effectively use EDA in data analysis • identify suitable graphics to get their message across • apply visualization in practical situations
LESSON THREE	**Data Security**	• identify cloud security domains • distinguish various security concerns • make informed decisions regarding data security

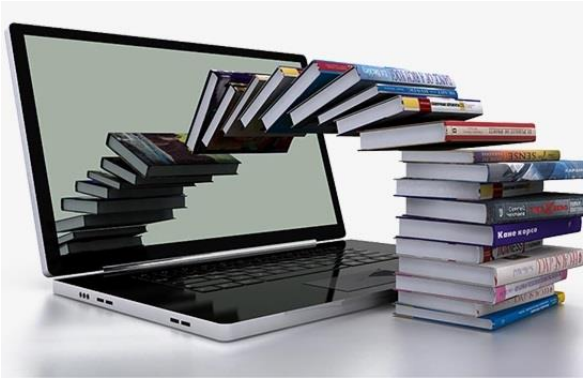
Capstone Project

Apply the key concepts learned throughout the program to create a typical Data Engineering solution.

Course Project
Data Engineering
Capstone

The Python language can be used to support all stages of exploratory data analysis (EDA) from storage to analytics. For this project, you will use Python to store, clean, process, and query a given data set. Further to this, you will use Python modules like pandas for visualization. It challenges you to move the cleansed data via a serverless application to AWS for visualization.

Our Classroom Experience



REAL-WORLD PROJECTS

Build your skills through projects that exercise your skills in real-life situations. Get constant feedback from our trainers throughout the program.

KNOWLEDGE

Work with instructors with recent hands-on experience in the industry.

STUDENT PORTAL

All information about the program is maintained in one central student portal.

CLOUD BASED SETUP

We will provide cloud based setup that you can access 24x7 throughout the program.

QUIZZES

Weekly quizzes verify your understanding.

CLIENT-SPECIFIC CUSTOMIZATION

Every program is customized to the client need. The projects can also be customized

WEEKLY PROGRESS REPORT

Weekly touch base meetings with the stakeholders along with progress reports ensures that the client goals are met.

Why Choose Web Age Solutions

We are a leading provider of Instructor Led Technical Training, collaborative mentoring, and custom curriculum development to many of the world's largest organizations. We earned our "Preferred Vendor" status with the Fortune 100 global leaders as a result of our highly progressive approach to education and project-based mentoring. We use industry-standard methodologies to help our clients by providing comprehensive, focused solutions

Two-thirds of Fortune 100's have chosen Web Age as a trusted partner

350+ Proprietary Courses, with cutting edge services

Customized courses to optimize your in-class time

Experts in all mediums of delivery including instructor-led live, virtual and eLearning