Projects and Requirements

These are the real-world projects you will complete during this program. Each project has a recommended deadline.

Data Modeling with Apache Cassandra

Students will model event data to create a non-relational database and ETL pipeline for a music streaming app. They will define queries and tables for a database built using Apache Cassandra.

Data Warehouse

In this project, you'll build an ETL pipeline that extracts data from S3, stages data in Redshift, and transforms data into a set of dimensional tables for an analytics team.

STEDI Human Balance Analytics

In this project, you'll work with sensor data that trains a machine learning model. You'll load S3 JSON data from a data lake into Athena tables using Spark and AWS Glue.

Data Pipelines

Students work on a music streaming company's data infrastructure by creating and automating a set of data pipelines with Airflow, monitoring and debugging production pipelines

Improve Your LinkedIn Profile

Find your next job or connect with industry peers on LinkedIn. Ensure your profile attracts relevant leads that will grow your professional network.

Optimize Your GitHub Profile

Other professionals are collaborating on GitHub and growing their network. Submit your profile to ensure your profile is on par with leaders in your field.