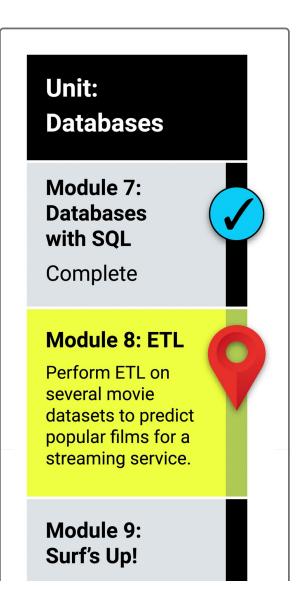# 8.0.2   Module 8 Roadmap

Looking Ahead

In this module, you'll learn how to use the Extract, Transform, Load (ETL) process to create data pipelines. A data pipeline moves data from a source to a destination, and the ETL process creates data pipelines that also transform the data along the way. Analysis is impossible without access to good data, so creating data pipelines is often the first step before any analysis can be performed. Therefore, understanding ETL is an essential skill for data analysis.

What You Will Learn

By the end of this module, you will be able to:

Unit:
Databases

Module 7:
Databases
with SQL

Complete ✓

Module 8: ETL

Perform ETL on several movie datasets to predict popular films for a streaming service.

Module 9:
Surf's Up!

- Create an ETL pipeline from raw data to a SQL database.

- Extract data from disparate sources using Python.

- Clean and transform data using Pandas.

- Use regular expressions to parse data and to transform text into numbers.

- Load data with PostgreSQL.

Planning Your Schedule

Here's a quick look at the lessons and assignments you'll cover in this module. You can use the time estimates to help pace your learning and plan your schedule.

- Introduction (15 minutes)

- Getting Started with ETL (15 minutes)

- Overview of the ETL Process (30 minutes)

- Extract the Data (45 minutes)

- Transform: Clean Individual Datasets (7 hours)

- Transform: Merge Datasets (2 hours)

- Load (1 hour, 30 minutes)

- Application (5 hours)