

Data Engineering for Everyone

Discover how data engineers lay the groundwork that makes data science possible. No coding involved!

2 hours



Hadrien Lacroix

Curriculum Manager at DataCamp

2

Python Programming

Introduction to Data Engineering

Learn about the world of data engineering with an overview of all its relevant topics and tools!

4 hours

3



Vincent Vankrunkelsven

Data and Software Engineer @DataCamp

4

Streamlined Data Ingestion with pandas

Learn to acquire data from common file formats and systems such as CSV files, spreadsheets, JSON, SQL databases, and APIs.

4 hours



Amany Mahfouz

Data scientist via spatial analytics and geography.

5

Writing Efficient Python Code

Learn to write efficient code that executes quickly and allocates resources skillfully to avoid unnecessary overhead.

4 hours



Logan Thomas

Scientific Software Technical Trainer, Enthought

Writing Functions in Python

6

Learn to use best practices to write maintainable, reusable, complex functions with good documentation.

4 hours



Shayne Miel

Director of Software Engineering @ American Efficient

7

Introduction to Shell

The Unix command line helps users combine existing programs in new ways, automate repetitive tasks, and run programs on clusters and clouds.

4 hours

DataCamp Content Creator
Course Instructor

8

Data Processing in Shell

Learn powerful command-line skills to download, process, and transform data, including machine learning pipeline.

4 hours



Susan Sun
Data Person

Introduction to Bash Scripting

Bash scripting allows you to build analytics pipelines in the cloud and work with data stored across multiple files.

4 hours



Alex Scriven
Data Scientist @ New South Wales Government

Unit Testing for Data Science in Python

Learn how to write unit tests for your Data Science projects in Python using pytest.

4 hours



Dibya Chakravorty
Senior Python Developer, TECH-5

11

Object-Oriented Programming in Python

Dive in and learn how to create classes and leverage inheritance and polymorphism to reuse and optimize code.

4 hours



Alex Yarosh

Curriculum Developer @ Cockroach Labs

12

Introduction to Airflow in Python

Learn how to to implement and schedule data engineering workflows.

4 hours



Mike Metzger

Data Engineer Consultant @ Flexible Creations

Introduction to PySpark

13

Learn to implement distributed data management and machine learning in Spark using the PySpark package.

4 hours



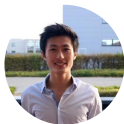
Nick Solomon
Data Scientist

14

Building Data Engineering Pipelines in Python

Learn how to build data engineering pipelines in Python.

4 hours



Kai Zhang
Data Engineer at Data Minded

15

Introduction to AWS Boto in Python

Learn about AWS Boto and harnessing cloud technology to optimize your data workflow.

4 hours



Maksim Pecherskiy
Data Engineer

16

Data Analysis in SQL (PostgreSQL)

Introduction to Relational Databases in SQL

Learn how to create one of the most efficient ways of storing data - relational databases!

4 hours

17



Timo Grossenbacher
Project Lead Automated Journalism at Tamedia

Database Design

Learn to design databases in SQL.

4 hours

18



Lis Sulmont
Workspace Architect at DataCamp

Introduction to Scala

Begin your journey with Scala, a popular language for scalable applications and data engineering infrastructure.

3 hours



David Venturi
Data Science Educator

Big Data Fundamentals with PySpark

Learn the fundamentals of working with big data with PySpark.

4 hours



Upendra Kumar Devisetty
Science Analyst at CyVerse

21

Cleaning Data with PySpark

Learn how to clean data with Apache Spark in Python.

4 hours



Mike Metzger

Data Engineer Consultant @ Flexible Creations

22

Introduction to Spark SQL in Python

Learn how to manipulate data and create machine learning feature sets in Spark using SQL in Python.

4 hours



Mark Plutowski

Big Data Architect & Scientist @ Flipboard

Cleaning Data in SQL Server Databases

Develop the skills you need to clean raw data and transform it into accurate insights.

4 hours



Miriam Antona
Software Engineer

Transactions and Error Handling in SQL Server

Learn to write scripts that will catch and handle errors and control for multiple operations happening at once.

4 hours



Miriam Antona
Software Engineer

Building and Optimizing Triggers in SQL Server

Learn how to design and implement triggers in SQL Server using real-world examples.

4 hours



Florin Angelescu
Database Developer

26

Improving Query Performance in SQL Server

In this course, students will learn to write queries that are both efficient and easy to read and understand.

4 hours



Dean Smith

Founder, Atamai Analytics

27

Introduction to MongoDB in Python

Learn to manipulate and analyze flexibly structured data with MongoDB.

4 hours



Donny Winston

Donny is a computer systems engineer at Lawrence Berkeley National Lab.