

Data Scientist

Syllabus

Each course lasts two weeks or three weeks and represents approximately 40 hours of study.

9 months

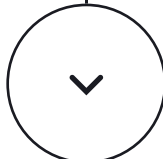
1

Basic Python

Your introduction to the world of data science! Key concepts and basic syntax in Python. Loops, conditions, and functions. The pandas library for data analysis. Your first analytical case study, followed by your first project.

3 weeks,
40 hours

- Chapter 1. Variables, Printing, Data Types, and Arithmetic Operations
- Chapter 2. Strings
- Chapter 3. Lists
- Chapter 4. for Loops
- Chapter 5. Nested Lists
- Chapter 6. Conditions and Loops
- Chapter 7. Creating Functions
- Chapter 8. Dictionaries
- Chapter 9. pandas for Data Analysis
- Chapter 10. Data Preprocessing
- Chapter 11. Analyzing Data and Presenting Results
- Chapter 12. A Quick Overview of the Jupyter Notebook



2

Data Preprocessing

Compensating for less-than-perfect data. Handling missing and duplicate values. Changing data types. Systems thinking for analysts.

- Chapter 1. Introduction to Data Preprocessing
- Chapter 2. Working with Missing and Duplicate Values
- Chapter 3. Changing Data Types
- Chapter 4. Categorizing Data
- Chapter 5. Systems and Critical Thinking for Analysts

+1 project for
your portfolio
3 weeks, 40 hours

3

Exploratory Data Analysis (EDA)

Performing initial scans to detect patterns in data. Building basic graphs and generating your first hypotheses.

- Chapter 1. Introduction to Exploratory Data Analysis (EDA)
- Chapter 2. First Graphs and Conclusions
- Chapter 3. Data Slices
- Chapter 4. Working with Several Data Sources
- Chapter 5. Relationships Between Datasets
- Chapter 6. Validating Results

+1 project for
your portfolio
3 weeks, 40 hours

4

Statistical Data Analysis

Probability theory, the most common distributions, and statistical methods in Python. Sampling and statistical significance. Identifying and handling anomalies.

- Chapter 1. Introduction to Statistical Data Analysis
- Chapter 2. Descriptive Statistics
- Chapter 3. Probability Theory
- Chapter 4. Testing Hypotheses

+1 project for
your portfolio
3 weeks, 40 hours

5

Integrated Project 1

Identify patterns to help you determine whether a given video game will be commercially successful or not.

+1 project for
your portfolio
1 week, 20 hours

1-week break

6

Data Collection and Storage (SQL)

How databases are structured and how to pull data from them using SQL queries. Finding data online.

+1 project for
your portfolio
2 weeks, 40 hours

- Chapter 1. Introduction to Data Collection and Storage (SQL)
- Chapter 2. Retrieving Data from Online Resources
- Chapter 3. SQL as a Tool for Working with Data
- Chapter 4. Advanced SQL Features for Analysis
- Chapter 5. Relationships Between Tables
- Chapter 6. Soft Skills
- Bonus Chapter: PySpark

7

Introduction to Machine Learning

Mastering the basics of machine learning. How the scikit-learn library works and how to apply it in your very first machine learning project.

+1 project for
your portfolio
2 weeks, 40 hours

- Chapter 1. Introduction to Machine Learning
- Chapter 2. Training Your First Model
- Chapter 3. Model Quality
- Chapter 4. Model Improvement
- Chapter 5. Moving on to Regression

8

Supervised Learning

Diving into the most in-demand area of machine learning. How to tune machine learning models, improve metrics, and work with imbalanced data.

+1 project for
your portfolio
2 weeks, 40 hours

- Chapter 1. Introduction to Supervised Learning
- Chapter 2. Feature Preparation
- Chapter 3. Classification Metrics
- Chapter 4. Imbalanced Classification
- Chapter 5. Regression Metrics
- Chapter 6. Soft Skills

9

Machine Learning in Business

Applying what you've learned to business tasks. Discover business metrics, A/B testing, the bootstrapping technique, and data labeling.

- Chapter 1. Course Introduction
- Chapter 2. Business Metrics
- Chapter 3. Implementing New Functionality
- Chapter 4. Data Collection
- Chapter 5. Soft Skills

+1 project for
your portfolio
2 weeks, 40 hours

10

Integrated Project 2

Prepare a prototype of a machine learning model to help a mining company develop efficient solutions.

+1 project for
your portfolio
1 week, 20 hours

1-week break

11

Linear Algebra

Taking a deeper look at some algorithms you've already studied and understanding how to apply them. Key concepts in linear algebra: vectors, matrices, and linear regression.

- Chapter 1. Course Introduction
- Chapter 2. Vectors and Vector Operations
- Chapter 3. Distance Between Vectors
- Chapter 4. Matrices and Matrix Operations
- Chapter 5. Linear Regression from the Inside

+1 project for
your portfolio
2 weeks, 40 hours

12

Numerical Methods

Analyzing a number of algorithms that use numerical methods and applying them to practical tasks. Gradient descent, gradient boosting, and neural networks.

- Chapter 1. Course Introduction
- Chapter 2. Algorithm Analysis
- Chapter 3. Gradient Descent
- Chapter 4. Gradient Descent Training
- Chapter 5. Gradient Boosting
- Chapter 6. Soft Skills

+1 project for
your portfolio
2 weeks, 40 hours

13

Time Series

Exploring the time series. Understanding trends, seasonality, and feature creation.

- Chapter 1. Course Introduction
- Chapter 2. Time Series Analysis
- Chapter 3. Time Series Forecasting

+1 project for
your portfolio
2 weeks, 40 hours

14

Machine Learning for Texts

Applying machine learning to text data. Finding out how to convert text into numbers and how to use bag-of-words, TF-IDF, as well as embeddings and BERT.

- Chapter 1. Course Introduction
- Chapter 2. Text Vectorization
- Chapter 3. Language Representations

+1 project for
your portfolio
2 weeks, 40 hours

15

Computer Vision

How to handle simple computer vision tasks using premade neural networks and the Keras library. A quick look at deep learning.

- Chapter 1. Course Introduction
- Chapter 2. Fully Connected Networks
- Chapter 3. Convolutional Neural Networks
- Chapter 4. Soft Skills

+1 project for
your portfolio
2 weeks, 40 hours

16

Unsupervised Learning

Figuring out what to do when you have no target features. Handling clustering tasks and looking for anomalies.

- Chapter 1. Course Introduction
- Chapter 2. Clustering
- Chapter 3. Search for Anomalies

2 weeks, 40 hours

17

Final Project

Apply everything you've learned in a two-week bootcamp that simulates the experience of working as a junior data scientist.

+1 project for
your portfolio
2 weeks, 40 hours

Career help

In addition to the main educational course, our career help is divided into three parts: the Career Prep Course, Career Acceleration Program, and the Apiary projects.

Career Prep Course

This is a course devoted to preparing for life after Practicum. During this course, you will learn how to create a resume, a LinkedIn profile, and a GitHub account, along with improving networking and interviewing skills. This course is self-paced and ends with a final task. We'll also perform a review of your career artifacts.

40 hours
+ Resume, LinkedIn
and Github profiles

1 Resume

Learn how to write an eye-catching resume, transform your non-tech experience into a strength.

- Compile a ready-to-use resume
- Gain access to a resume improvement tool

2 Creating an Online Presence

Assemble your GitHub portfolio and ensure your LinkedIn looks professional and informative.

- Produce a production-ready portfolio
- Launch your LinkedIn profile

3 Being a Networking Ninja

Learn how to become a networking professional, and how to write the perfect cover letter.

- Unlock a networking roadmap
- Prepare a cover letter template

4

The Job Search

Learn where to find a job and prepare for the search!

- Access job searching resources & application tracker tool
- Produce a target job list

5

An Interview Masterclass

Familiarize yourself with different interview types, common questions you might face, and practice tech assignments.

- Learn interview do's & don'ts
- Master the STAR technique and sound more professional
- Get tech interview help

Career Acceleration Program

Prepare for real-world interviews and gain experience through authentic practice. This program is designed to help you find a real job and also provides some work with technical skills.

Up to 6 months
after graduation

- Attend mock interviews
- Receive 1:1 career coaching
- Write technical articles and demonstrate your knowledge
- Produce demo videos of your work
- Participate in extracurricular activities
- Join the Slack community

Apiary Projects

You'll gain confidence solving work tasks that use a real company's data to provide them with valuable insights. Learn to communicate with clients, meet their expectations, exchange peer reviews with colleagues, and present results to the company. The Apiary projects become available for participants sometime between the 8-10th Sprint, depending on the project. They are also available after graduation.

+ 1- ∞ real projects
for your portfolio,
5-6 weeks

- Assemble a portfolio project based on actual data
- Get a recommendation on LinkedIn by a real company
- Gain experience with freelance project workflow