Data Science Bootcamp

Full-Time

GENERAL ASSEMBLY

Last updated on 06-02-2025

Overview

General Assembly's Data Science Bootcamp (DSB) is a transformative course that prepares students to break into data careers.

Created to transform your career in just 12 weeks, this full-time on-campus bootcamp will equip you with the tools and techniques needed to go from novice to competent and develop a data career full of opportunities.

You'll get skilled in everything from conducting an exploratory data analysis to running supervised and unsupervised machine learning models, before progressing on to learning about different types of neural networks and understanding how large language models work.

Once you complete this bootcamp, you'll present potential employers with an impressive resume highlighting your capabilities with the most in-demand tools they are looking for. You'll also have sought-after soft skills that will make you stand out.

Throughout this expert-designed course, you'll:

- Collect, extract, query, clean, and aggregate data for analysis.
- Gather, store, and organize data using SQL, APIs, and manage code using Git and GitHub.
- Perform visual and statistical analysis on data using Python and its associated libraries and tools.
- Craft and share compelling narratives through data visualization.
- Build and implement appropriate machine learning models and algorithms to evaluate data science problems spanning finance, public policy, and more.

- Clearly and effectively communicate to technical and nontechnical audiences through creating reports, meeting with stakeholders, and presenting results.
- Apply question, modeling, and validation problem-solving processes to data sets from various industries to provide insight into real-world problems and solutions.
- Prepare for the world of work, compiling a professional-grade portfolio of projects.
- Learn the principles of data engineering and how to manage data, from creating databases and pipelines to data governance.
- Learn how large language models (LLMs) work, and how to apply pretrained models to new data.
- Get an introduction to working with images in machine learning models through computer vision.
- Understand what generative AI is and techniques for effective prompt engineering.

Prerequisites

The course assumes no prior knowledge of either Python programming or data science. However, we recommend that students arrive with a mathematical foundation and familiarity with Python and programming fundamentals. Some students have a technical background, such as a degree in mathematics or computer science or work experience in research or analysis. Other students engage in self-learning to build a foundation ahead of class. Suitable work experience would be one or more years as a data analyst before starting the course.

Our Admissions team can discuss your background and learning goals to advise if this bootcamp is a good fit for you.

What to expect

Pre-course learning paths

Set yourself up for success in this bootcamp with up to 12 hours of preparatory lessons covering essential concepts in Python programming and applied math for data science. Designed to introduce you to foundational data skills and context, Data Science Fundamentals is a self-paced online learning path that will help you hit the ground running on day one of class.

The in-class experience

Engage in project-based learning that's designed to inspire a lifetime of discovery. As a DSB student, you'll:

- Explore new concepts and tools through expert-led lectures and discussions.
- Complete hands-on programming and modeling exercises to reinforce newly learned skills in a supervised learning environment.
- Develop fluency in industry-essential topics and techniques via independent, pair, and group labs.
- Receive individualized feedback and support from your expert instructional team.
- Build out a professional portfolio to showcase your job-ready data science skills to potential employers and collaborators.

Dedicated coaching hours

As a bootcamp student, you'll receive dedicated support from career coaches who will help you set goals, build connections with employers, and stay on track with your job search. Throughout the course, you'll:

- Get an inside look at the industry through day-in-the-life talks, hiring panel discussions, guest speakers, and more.
- Cultivate a competitive candidate mindset, learning to assess your skill set against job descriptions, track progress, and recognize opportunities.
- Develop your professional brand; polish your online and in-person presence.
- Practice technical challenges and whiteboarding skills to set yourself apart in interviews.
- Tap into an exclusive global network of experts, influencers, and peers, plus learn strategies for leveraging your existing connections, whether in person or online.
- Become an active contributor to the data science community.

After graduation, you'll also gain access to resources to help fuel a lifetime of learning. Dive into new topics or continue honing your data science skills.

What you'll learn

Unit 1 Python and Data Science Fundamentals

Learn how to approach data science challenges and get acquainted with essential data science tools and techniques by learning Python, SQL, Git and GitHub.

- Learn about the data science lifecycle.
- Sharpen your presentation skills for both technical and nontechnical audiences.
- Learn the base Python you will need for data science to navigate data sources and collections.
- Run SQL in BigQuery.
- Learn to track changes and iterations using Git version control from your terminal.
- Share your code with others via GitHub.

Unit 2 Exploratory Data Analysis

Learn what exploratory data analysis (EDA) is and how to do it, including collecting and connecting data programmatically via APIs to add nuance to your models.

- Perform exploratory data analysis using pandas to turn a messy dataset into one that can be analyzed or used in a machine learning model.
- Generate visual and statistical analyses, using Python and its associated libraries and tools such as Matplotlib and seaborn.
- Collect and connect external data using APIs, in order to add nuance to your models.

Unit 3 Machine Learning Fundamentals

Build machine learning models. Implement regression and classification supervised learning models, and learn how to optimize them.

- Learn how to shape a data science problem.
- Use the scikit-learn and statsmodels libraries to build linear regression and logistic regression supervised machine learning models.
- Articulate the bias-variance trade-off as part of evaluating a model.
- Apply optimization and regularization for fitting and tuning models.
- Understand the principles of working with time series data.
- Optimize your code by leveraging generative AI tools.

Unit 4 Advanced Modeling

Explore the differences between supervised and unsupervised learning via clustering and dimensionality reduction.

- Build and evaluate ensemble models using decision trees, random forests, bagging, and boosting.
- Define clustering and its advantages and disadvantages as compared to classification models.
- Use Principal Component Analysis (PCA) to carry out dimensionality reduction.
- Dive into the math and theory for how gradient descent helps to optimize loss functions for machine learning models.
- Build a basic recommender system.

Unit 5 Introduction to Data Engineering and Machine Learning Operations (MLOps)

Explore the infrastructure used to support effective data science. As well as an introduction to data modeling, data pipelines, and data storage, learn how machine learning models are deployed within organizations. Practice sound data governance.

- Create a data model and a database.
- Explain the forms and benefits of data normalization.
- Differentiate between Extract, Transform, Load (ETL) and Extract, Load, Transform (ELT) processes when managing data pipelines.
- Describe the tools used for data orchestration and the use cases for batch vs real-time processing.
- Articulate the differences between data lakes and data warehouses, their architectures, use cases, and best practices for optimizing data storage.
- Learn data governance principles that ensure security, integrity, and compliance.
- Explain how machine learning models are deployed and managed in organizations using machine learning operations (MLOps).

Unit 6 Deep Learning and Large Language Models (LLMs)

Dive into neural networks, comparing and contrasting different types of neural networks, and learn about transformers and large language models (LLMs): how they work and how to apply pretrained models to new data.

- Describe how neural networks work and common applications of deep learning.
- Use common deep learning toolkits: PyTorch, Hugging Face, and torchvision.
- Understand neural network architectures: feedforward networks, convolutional neural networks (CNNs), and recurrent neural networks (RNNs).
- Conduct transfer learning by fine-tuning pre-trained models for specific tasks (e.g., classification, object detection).
- Use a CNN to classify images (computer vision).
- Get acquainted with natural language processing (NLP)-the basis of large language models (LLMs)-to translate human language into a form that computers can use and apply to solve problems.
- Dive into transformer architecture and apply select models from the transformer family to solve a real-world problem.
- Develop generative AI skills through prompt engineering and chaining via LangChain.

Frequently asked questions

Why is this course relevant today?

Data science is highly relevant today due to rapid technological advancements that are transforming industries and tasks across sectors. According to the World Economic Forum's Future of Jobs Report 2025, data science is ranked among the top 15 fastest-growing jobs through 2030, with many new roles emerging that will involve creating algorithms and machine learning applications. This shift emphasizes the growing demand for skilled professionals in data science who can develop and implement these technologies. As traditional roles evolve, reskilling and upskilling in data science become crucial for workers transitioning into new jobs, ensuring they remain competitive in the workforce. Overall, data science is essential for navigating the interconnected trends shaping the global labor market.

Will I earn a certificate?

Yes! Upon passing this course, you will receive a signed certificate of completion. Thousands of GA alumni use their course certificates to demonstrate skills to potential employers—including our 19K+ hiring partners—along with their LinkedIn networks. GA's data programs are well-regarded by many top employers, who contribute to our curriculum and partner with us to train their own teams.

What are the professional backgrounds of data science students?

DSB students come from all walks of life but share one common mission: They are passionate about launching a career in data science or advanced analytics. We see people developing from data analyst roles to career-changers from diverse professional backgrounds, including engineers and recent STEM graduates, mid-career marketing and financial analysts, and business strategists, as well as those from less directly related fields like sales and law. It is suitable for graduates of General Assembly's Data Analytics Bootcamp (DAB), though we would recommend obtaining some work experience between taking the two programs.

Who is eligible for tuition sponsorship by Tamkeen and GA? And what does tuition sponsorship cover?

You must be a Bahraini citizen to receive tuition sponsorship. Tuition sponsorship covers:

- Expert instruction in the skills you need to enter the workforce.
- Self-paced pre-work to explore software engineering fundamentals that helps you hit the ground running on day one of class.
- Robust coursework, including expert-vetted lesson decks, project toolkits, and more.
- A professional-grade portfolio of projects taken from concept to completion—each mirroring real problems that engineers face—that allows you to showcase the breadth of your technical skills to employers.
- Individual feedback and guidance from instructors and TAs during office hours.
- Dedicated Career Services to help you navigate your personal job search experience from technical challenges to salary negotiation and more.
- Technical interview prep, including resume reviews, mock interviews, and whiteboarding practice.
- A GA course certificate to showcase your new skill set on LinkedIn.

What projects will I work on during this course?

For your capstone project, you'll tackle a real-world data problem from end to end. Develop a pitch and problem statement, source and collect relevant data, conduct an exploratory data analysis, and build a predictive model. You'll document and share your findings through a presentation, technical report, and nontechnical summary.

Throughout this bootcamp, you'll also compile a portfolio of projects designed to reinforce what you've learned in each unit. Gain hands-on experience with statistical and machine learning models, Python programming tools, recommender systems, neural networks, and more.

How does this course relate to GA's other data programs?

This Data Science Bootcamp course is meant for students who are committed to changing careers and provides the most direct pathway to data science and other advanced analytics roles.

For professionals who already work with data and want to perform more complex analysis involving computation, we also offer a Data Science Short Course.

If you're searching for a more entry-level course, Data Analytics teaches beginners how to perform rigorous analysis with Excel, SQL, and Tableau.

Where are the GA & Tamkeen courses delivered?

The GA and Tamkeen courses are delivered in-person at the Bahrain Institute of Banking and Finance (BIBF).

Take the next step

Have questions about our Data Science Bootcamp course? Our Admissions team is here to help and can advise on if this program is right for you and your learning goals. You can also:

- Apply for the course here.
- Explore upcoming classes and workshops