

Data Science Immersive

 **GENERAL ASSEMBLY**



Overview

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

Designed to get you hired, this 12-week data science bootcamp features best-in-class instruction, career coaching, and professional connections to top employers. Join us on campus or online in our Remote classroom.

As a graduate, you'll leave poised to succeed in a variety of data science and advanced analytics roles, creating predictive models that drive decision-making and strategy throughout organizations of all kinds.

Throughout the expert-designed course, you'll:

- Collect, extract, query, clean, and aggregate data for analysis.
- Gather, store and organize data using SQL and Git.
- 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.
- Compile clear stakeholder reports to communicate the nuances of your analyses.
- 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 solo, group, and client projects.



Prerequisites

This is an intermediate-level course with some prerequisites. 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.

Our [Admissions team](#) can discuss your background and learning goals to advise if this Immersive is a good fit for you.



What To Expect

Pre-Course Learning Paths

Set yourself up for success in this Immersive 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 full-time, project-based learning that's designed to inspire a lifetime of discovery. As a DSI 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.
- Develop fluency in industry-essential topics and techniques via independent, pair, and group labs and real-world client work.
- 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 Career Coaching

As an Immersive 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, in person and 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 with discounts on a suite of tools, passes and packages to premier events, and more. You can also apply tuition discounts to future GA courses, classes, and workshops, both on campus and online.

General Assembly's Connected Classroom

Experience the best of online and in-person education in our Connected Classroom.* Our innovative format seamlessly blends instruction, technology, and collaboration into a proven approach to career transformation, preparing students to succeed on the job.

- Learn from GA's top instructors from around the globe who broadcast live to your classroom.
- Get in-person, individualized training from local pros who are on hand to offer hands-on support and lead discussion.
- Collaborate and connect with a network of alumni and peers with diverse professional backgrounds across major cities.

**Data Science Immersive is run in a Connected Classroom format at all GA campuses except for London.*





What You'll Learn

Pre-Work

Data Science Fundamentals

Dive into a series of self-paced lessons on the essentials of Python programming and applied math for data science before the course begins.

- Explore fundamental Python programming concepts, including variables, lists, loops, dictionaries, and data sets.
- Leverage programming tools like GitHub and the command line interface to manage data science projects.
- Practice solving coding challenges similar to the questions used in task-based data science interviews.
- Write and run Python functions using multiple arguments.
- Discover how key math concepts like statistical significance and probability distribution are applied throughout data science.

Unit 1 Fundamentals

Get acquainted with essential data science tools and techniques, working in a programming environment to gather, organize, and share projects and data with Git and UNIX.

- Demonstrate familiarity with introductory programming concepts using Python and NumPy to navigate data sources and collections.
- Utilize UNIX commands to navigate file systems and modify files.
- Learn to track changes and iterations using Git version control from your terminal.
- Define and apply descriptive statistical fundamentals to sample data sets.
- Practice plotting and visualizing data using Python libraries like Matplotlib and Seaborn.

Project: Apply NumPy and Python programming skills to answer questions based on a clean data set.



Unit 2 Exploratory Data Analysis

Perform exploratory data analysis. Generate visual and statistical analyses, using Python and its associated libraries and tools to approach problems in fields like finance, marketing, and public policy.

- Design an experimental study with a well-thought-out problem statement and data framework
- Use Pandas to read, clean, parse, and plot data, extracting and rearranging data through indexing, grouping, and JOINing.
- Review statistical testing concepts (p values, confidence intervals, lambda functions, correlation/causation) with SciPy and StatsModels.
- Learn to scrape website data using popular scraping tools.
- Explore bootstrapping, Resampling and building inferences about your data.

Project: Leverage Pandas to apply advanced NumPy and Python skills cleaning, analyzing, and testing data from multiple messy data sets.

Unit 3 Classical Statistical Modeling

Explore effective study design and model evaluation and optimization, implementing linear and logistic regression, and classification models. Collect and connect external data to add nuance to your models using web scraping and APIs.

- Use scikit-learn and StatsModels to run linear and logistic regression models and learn to evaluate model fit.
- Begin to look at classification models by implementing the k-nearest neighbors (kNN) algorithm.
- Articulate the bias-variance trade-off as you practice evaluating classical statistical models.
- Use feature selection to deepen your knowledge of study design and model evaluation.
- Learn to apply optimization and regularization for fitting and tuning models.
- Dive into the math and theory behind how gradient descent helps to optimize loss functions for machine learning models.

Project: Explore, clean, and model data based on a provided data set, outlining your strategy and explaining your results.



Unit 4 **Machine Learning Models**

Build machine learning models. Explore the differences between supervised and unsupervised learning via clustering, natural language processing, and neural networks.

- Define clustering and its advantages and disadvantages as compared to classification models.
- Build and evaluate ensemble models using decision trees, random forests, bagging, and boosting.
- Get acquainted with natural language processing (NLP) through sentiment analysis of scraped website data.
- Learn how Naive Bayes can simplify the process of analyzing data for supervised learning algorithms.
- Explore the history and use of Hadoop, as well as the advantages and disadvantages of using parallel or distributed systems to store, access, and analyze big data.
- Understand how Hive interacts with Hadoop and discover Spark's advantages through big data case studies.
- Analyze and model time series data using the ARIMA model.

Project: Students will scrape and model their own data using multiple methods, outlining their approach and evaluating any risks or limitations.

Unit 5 **Advanced Topics and Trends**

Dive deeper into recommender systems, neural networks, and computer vision models, implementing what you've learned to productize models.

- Compare and contrast different types of neural networks and demonstrate how they are fit with back propagation.
- Build and apply basic recommender systems in order to predict on sample user data.
- Work with career coaches to create and polish your professional portfolio.
- Practice with data science case studies to prepare for job interviews.

Project: Choose a data set to explore and model, providing detailed notebook of your technical approach and a public presentation on your findings.



Frequently Asked Questions

Why is this course relevant today?

It's ranked among the top fields in LinkedIn's [Emerging Jobs Report](#) for three years running and, according to the latest numbers, data scientist roles are experiencing 37% annual growth. Across industries from information technology and software to financial services and higher education, organizations are moving quickly to build robust in-house teams of data scientists and advanced analysts, and there's not enough talent to go around.

For those with quantitative backgrounds or existing analytical skills, building out this specialized skill set can spell ample opportunity to secure a high-paying job in data science, advanced analytics, business intelligence, and more.

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 certificate 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?

DSI 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 career-changers from diverse professional backgrounds, including engineers and recent STEM graduates, mid-career marketing and financial analysts, and business strategists, as well as more those from more far-flung fields like sales and the law.

What does my tuition cover?

Here are just some of the benefits full-time students can expect at GA:

- Expert instruction in the skills you need to successfully transition into a data science career.
- Self-paced pre-work to explore data science fundamentals and prepare to hit the ground running on day one of class.
- Robust coursework, including expert-vetted lesson decks, project toolkits, and more. Refresh and refine your knowledge throughout your professional journey as needed.
- A professional-grade portfolio to showcase your ability to solve real-world data problems to potential employers and collaborators.
- Individual feedback and guidance from instructors and TAs during office hours. Stay motivated and make the most of your experience with the help of GA's dedicated team.
- 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.
- Exclusive access to alumni discounts, networking events, and career workshops.
- A GA course certificate to showcase your new skill set on LinkedIn.



- Connections with a professional network of instructors and peers that lasts well beyond the course. The global GA community can help you navigate and succeed in the field.

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 non-technical summary.

Throughout this Immersive, 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 Immersive 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 an evening and weekend [Data Science](#) 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.

Which format should I take this data science bootcamp in – on campus or online?

It's up to you! Our Remote courses offer a learning experience that mirrors GA's on-campus offerings but allow you to learn from the comfort of home. If you don't live near a GA campus, have a busy travel schedule, or just want to save yourself the commute, a Remote course could be a good option for you, if available in your market. You'll still get access to the expert instruction, learning resources, and support network that GA is known for.

If you prefer to learn alongside your peers and can make it to campus, our in-person courses allow you to take advantage of our beautiful classrooms and workspaces.

Our [Admissions team](#) can advise you on the best format for your personal circumstances and learning style.

What do career services look like for Remote students?

The same as our on-campus experience! We approach our Remote Outcomes programming with the same philosophy, promise to, and expectations of our students. Get an insider's look at the industry through virtual sessions like day-in-the-life talks, hiring panel discussions, guest speakers, and more. You'll work individually with your career coach to understand your local job market, find opportunities, and connect with the local data community.

Our [Admissions team](#) can provide more details on the dedicated support you'll receive on the path to landing a data science role.



Take the Next Step

Have questions about our Data Science Immersive 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:

- Attend an info session [online](#) or at your [local campus](#).
- Explore your [financing options](#).
- [Apply](#) to enroll in the course.*

**Course modality options vary by location, pending market availability. Please contact our Admissions team to discuss what version is available in your location.*