



Xtreme Talent Accelerator Program (XTAP)

Data Science Bootcamp

OVERVIEW



12 Weeks



60 Hours/Week



Data Science Certification

Being a data scientist requires so much more than just knowing how to read datasets. From SQL and Python to Machine Learning and beyond, each module of our immersive, industry-driven curriculum is designed to equip you with the skills, knowledge and confidence needed to succeed and grow in the field of data science.

WHO IS THIS FOR?

You may be a recent graduate or someone looking for a career pivot. No tech background? No problem. Previous experience in data science, data analysis or programming is *not* required. All that is needed to be successful in this program is your willingness to work hard.

ELIGIBILITY

We invite you to apply if:

- You're unemployed, or
- You're working, but your job is unstable, inconsistent, uncertain or requires fewer than 20 hours per week, or
- Your family is living with low household income, and
- You have the time (60+ hours/week) to dedicate to participating in the program
- You'd like to build the skills needed to secure a longer-term career, and
- You're a resident of Ontario, and
- You're 18 years or older, and
- You're a Canadian Citizen, Permanent Resident or Protected Person

This program requires a commitment of 60+ hours per week for the 12-week duration. You will spend two hours each day, from Monday to Friday, on live lectures, and 8–10 hours on self-paced studies. It is estimated that you will likely need to spend 20–30 hours on projects and studies over weekends, as well.

WHAT YOU WILL LEARN

Designed and delivered by Lighthouse Labs, the bootcamp covers a myriad of industry-relevant technologies vetted by employers, subject matter experts and members of the broader tech community. In addition, career services offered by XTAP will help you develop important soft skills to get noticed and get ahead.

The Data Science Bootcamp

The curriculum covers:

- **Data Science Environment**

Set up an environment on your own machine using Anaconda, Git and Virtual Environments.

- **Coding**

Learn the basics of programming and computer science, including Python, Jupyter Notebook and VSCode.

- **Math Foundations**

Dive into the basics of probability theory, statistics and linear algebra using NumPy and Statsmodels.

- **Data Wrangling**

Master the art of data manipulation and preparation for data science using Pandas, APIs and Regex.

- **Databases**

Discover relational and non-relational databases such as SQL, NoSQL (MongoDB, Redshift), SQLite, Postgres, SQLAlchemy, Turbodbc and PyODBC.

- **Machine Learning**

Using sklearn, Spacy, NLTK, Gensim, Tensorflow and Keras, you'll dig into the elements of Machine Learning including Supervised Learning, Unsupervised Learning, Deep Learning, NLP, Time Series and Recommenders.

- **Visualization**

Learn the principles of graphical integrity and the guidelines for visualization context and guidelines with tools like Matplotlib, Seaborn, Plotly, Mpl_toolkits, Geopandas and Ipywidgets.

- **Development Process**

Master the development process, using AWS, Flash, Docker, and Spark to move through design, experimentation (prototyping), production code and deployment.

Your learning will prepare you to:

- Use and navigate popular data science tools, libraries, environments and workflows.
- Demonstrate and apply various data-wrangling techniques and methodologies with a multitude of data types and representations.
- Differentiate data visualization concepts and the applicable types of graphs to tell a story from an existing dataset.
- Explain the main function and purpose of experimental design and how to apply it to deploy machine-learning models.
- Recognize and describe the fundamental concepts, best practices, applications and algorithms of machine learning.
- Solve common machine-learning problems using Python and associated libraries to implement machine-learning algorithms.
- Summarize and compare deep-learning architectures and the opportunities and limitations of each.
- Design and deploy your own machine-learning model and develop data engineering pipelines.
- Explain and present data science concepts, best practices and applications, processes and findings to both technical and non-technical audiences.

EMPLOYMENT PREPARATION, ON-THE-JOB SKILLS, AND ETHICS AND EMPLOYEE ADVOCACY

These additional career services provided by XTAP cover:

- Personality Assessments
- The Anatomy of a Strong Resume
- Virtual Interviewing Skills for Job Seekers
- Job Search Strategies
- Tech, EQ and You: Exploring and Understanding Emotional Intelligence in Tech
- Effective Communication in the Workplace
- Designing Your Professional Digital Presence
- Time Management
- Mental Health on the Job
- Diversity, Equity, and Inclusion — Call It Out: Racism, Racial Discrimination, and Human Rights
- Employee Rights in the Workplace

CAREER OUTLOOK

Successful graduates will be equipped to walk into any one of these roles:

- Machine Learning Engineer
- DataOps Engineer
- Data Science Generalist
- Data Analyst
- Senior Data Analyst

YOUR LEARNING PATH

The Data Science Bootcamp is offered through a mix of daily live instructor-led lectures and self-paced work such as reading, group projects, assignments and quizzes. You will have access to tech mentors on-demand to get help with your specific questions and needs as you move through the independent and group work.

You will follow a vertical, module-based learning process covering the following topics:

1

Data Foundations:

- | | | |
|---|--|---|
| <input checked="" type="checkbox"/> Fundamentals <ul style="list-style-type: none">• Development Environment• Git• Unix• Bash• Assertions and Unit Testing | <input checked="" type="checkbox"/> Python Fundamentals <ul style="list-style-type: none">• Functions• Loops• Conditionals• Importing Modules• Dictionaries• Lists• Tuples• Numpy Arrays | <input checked="" type="checkbox"/> Probability and Statistics |
|---|--|---|

2

Data Wrangling:

- | | | |
|---|---|---|
| <input checked="" type="checkbox"/> Pandas | <input checked="" type="checkbox"/> ETL | <input checked="" type="checkbox"/> Data Types <ul style="list-style-type: none">• HTML• Text• XML• JSON• Images |
| <input checked="" type="checkbox"/> Mini Project | <input checked="" type="checkbox"/> APIs | |
| <input checked="" type="checkbox"/> SQL | | |

3

Data Visualization:

- | | | |
|--|--|--|
| <input checked="" type="checkbox"/> Concepts | <input checked="" type="checkbox"/> Complex Plots (Scatter, Density, 3D, Contour, Marginal Distribution, Pairs) | <input checked="" type="checkbox"/> DataViz in Python <ul style="list-style-type: none">• Matplotlib• Seaborn• Plotly |
| <input checked="" type="checkbox"/> Overview of Tools | | |

4

Machine Learning:

- Data Preparation**
- Feature Engineering**
- Sampling and Dimensionality Reduction**
- Introduction to Machine Learning Modelling**
- Modelling Techniques**
- Optimization With Gradient Descent (and Stochastic Gradient Descent)**
- Training and Evaluation**
- Unsupervised:**
 - Hard Clustering Methods (K-means, Medoids, Modes)
 - Hierarchical Methods
 - Soft (probabilistic) Clustering Methods: GMM, HDBSCAN
- Supervised:**
 - Logistic Regression
 - Decision Trees
 - Random Forests
 - Boosting
 - Support Vector Machines

5

Mid-Term Project (54 hours):

You will have an opportunity to practise your skills, from kick-off to demo, with an end-to-end group project that includes:

- Problem understanding**
- Design of the solution**
- Combination of data from different sources**
- Prototyping**
- Evaluation**
- Presentation of the results**

6

Data Engineering and Dev-Ops (40 hours):

- Pipelines**
- Deployment of Machine Learning Models**
- SPARK**
- Docker and Containerization**
- Flask**
- AWS**

7

Deep Learning (30 hours):

- Introduction to Neural Networks
- Architecture – FNN, RNN (LSTM, GRU), CNN
- Back-propagation With Gradient Descent
- Tensorflow Core 2.0 (Keras API)

8

Applications of Machine Learning (80 hours):

- Recommender Engines:**
 - Matrix Factorization
 - Collaborative Filtering
 - Collaborative Filtering With DL
- Time Series Analysis**
- NLP:**
 - Applying FNN, RNN and CNN
 - Word2Vec, Word/Sentence Embeddings, Cosine Distance
 - Topic Modelling
 - TSNE/UMAP
 - Sentiment Analysis
- A Brief Overview of Advanced Applications (Reinforcement Learning, DL With Attention, Online Learning, Graphical Models)**

9

Final Project:

You'll put all your skills to the test with a final group project including the following elements:

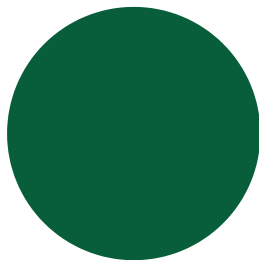
- Problem understanding**
- Design of the solution**
- Combination of data from different sources**
- Prototyping**
- Evaluation**
- Deployment**
- Presentation of the results**

DATES AND APPLICATION DEADLINES

September 13 – December 3, 2021
Application deadline: July 23, 2021

October 25, 2021 – January 28, 2022
Application deadline: September 25, 2021

November 2021 – February 2022
Application deadline: TBD





Ready to advance your career?

An application takes only 5–10 minutes to complete. Plus, you will be able to save it and return to finish it later.

[Apply Now](#)



investottawa.ca/xtreme-talent-accelerator



The Xtreme Talent Accelerator is made possible through the [SkillsAdvance Ontario \(SAO\) pilot project](#), with funding provided by the Government of Ontario, Ministry of Labour, Training and Skills Development.