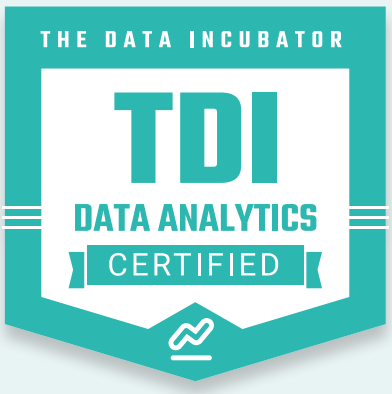

Data Analytics Fellowship

FROM THE DATA INCUBATOR

The Data Analytics Fellowship from The Data Incubator is an intensive, hands-on, business-focused bootcamp that gives working professionals and students the skills they need to fill critical skills gaps in major organizations. This 20-week, part-time, online program was developed to give students the skills and tools they need to help any organization better utilize and understand its data. Students will learn how to wrangle data using Python, perform processes using SQL, effectively communicate and present data to key stakeholders, and will be introduced to machine learning.

OUR CURRICULUM

The Data Incubator uses open source tools, popular programs like Tableau, and provides each student with a Jupyter server—hosted by DigitalOcean—for the duration of the course for coursework completion. Jupyter notebooks are used so students can follow along in lectures and see and edit the raw code that's being run.

Learn more about our *Data Analytics Fellowship* by visiting **thedataincubator.com**.

The Data Incubator
A Pragmatic Institute Company

Data Analytics Fellowship

PROGRAM SCHEDULE

SECTION 1

Data Wrangling with Python

Between 60-73% of an organization's data is untouched for analysis, leaving valuable opportunities on the table. Students will learn how to use Python and the Python ecosystem to begin processing data, distill key insights and trends, and turn that information into clear, business-focused visualizations. Students will work with NumPy, Pandas and matplotlib packages, and common data formats like CSV, JSON and XML.

Topics include:

- Python data structures
- Python control structures
- Common file and data formats
- Object-oriented programming (OOP) in Python
- NumPy and arrays
- Ingesting, filtering and merging data with Pandas
- Aggregating data with Pandas
- Handling time series data with Pandas
- Visualizing data with Pandas and matplotlib
- String processing with regular expressions

SECTION 2

SQL

Companies collect data from a number of different places, but they don't all arrive in the same way. Students will learn how to use SQL for accessing data from a variety of sources to process data more quickly and efficiently. Starting with the basic syntax of the language, students will learn how to filter, aggregate and join data, and integrate SQL queries into a Python analysis.

Topics include:

- Basic queries and filtering data
- Joining tables
- Aggregation functions
- Creating tables and ingesting data
- Use of temp tables and views
- Interacting with SQL from Python
- Uses of SQL outside of RDS

SECTION 3

Data Communication

Many business professionals lack the data literacy skills to interpret and understand the insights drawn by data analysts, and aren't comfortable accessing or using the data at their disposal. Students will learn to communicate and present data insights using Python plotting tools and business intelligence platforms to a general audience. Best practices for presenting clear visuals with a compelling story will allow students to product reports that stand out.

Topics include:

- Components of visualization
- The power of different visual cues
- Exploratory and explanatory visualizations
- Visualization tools using Python
- Loading data into Tableau
- Visualizations with Tableau
- Choosing the right problems to solve
- Metrics and levels
- Communicating results

SECTION 4

Machine Learning

Every business leader wishes they could predict the future of their market and their business. With machine learning, you can do just that. Students will learn the basics of machine learning with a focus on linear and logistic regressions. Students will understand the power and limitations of machine learning, and will be well-positioned to move on to more advanced machine learning models from their experience with Scikit-learn.

Topics include:

- Machine learning basics
- Scikit-learn for machine learning
- Techniques and metrics of regression
- Linear regression
- Overfitting and cross-validation
- Techniques and metrics of classification
- Logistic regression
- Dimensionality reduction with PCA
- Clustering
- Feature engineering with transformers

Learn more about our *Data Analytics Fellowship* by visiting **thedataincubator.com**.

The Data Incubator
A Pragmatic Institute Company