DATA ANALYTICS INFORMATION SESSION

Cynthia V. Marcello DM, MS, MBA

OpenHub

Innovation | Collaboration | Education

WELCOME

What is data analytics?

Why is data analytics needed?

What are the career opportunities?

What does this course cover?

Who is a good fit for a data analytics career?

DATA ANALYTICS – WHY?

A <u>data analyst</u> – someone who uses data to generate insights for an organization. What story does the data tell?

Descriptive – uses historical data to determine what happened or what is happening now

Diagnostic – uses historical data to determine what happened and why

Predictive – uses historical data to forecast possible outcome based on statistical models and machine learning

Prescriptive – uses historical data to make a recommendation for a course of action

OpenHub
Innovation | Collaboration | Education

CAREER OPPORTUNITIES

Every industry needs data analysts!

Add data analytics to your current core discipline!

Data analysts need less training and cost less than data scientists.

By 2022, 85% of companies will have adopted big data and analytics technologies and 96% of companies seeking to hire new permanent staff! (World Economic Forum)

Skills needed:

Math, Statistics

Pattern recognition

Organized

Communication skills

Excel and SQL

Python, R, Tableau, Power BI

OpenHub
Innovation | Collaboration | Education

BOOTCAMP COURSE TOPICS

Course 1 – Data Sourcing, Feature Engineering, and Preprocessing

Course 2 – Exploratory Data Analysis

Course 3 – Python

Course 4 – R

Course 5 – Python and R

Course 6 – Data Visualization

The Data Analytics Bootcamp is an intensive, hands-on learning experience that focuses on providing exposure to highly desirable skills sought by the job market. Over a 6-month period, students build their project portfolio on Github with the goal of being "resume-ready" for their next job interview or for seeking a promotion.

Timeframe: 26 weeks, 3-hour lecture per week, 4-6 hours per week on quizzes and ongoing real-world projects for homework.

OpenHub

Innovation | Collaboration | Education

DATA SOURCING, FEATURE ENGINEERING, AND PREPROCESSING

In this course, students will learn the various methods used to source and load data sets selected for analysis. The features and attributes of data will be reviewed after a data quality assessment has been conducted. Using a variety of data cleaning techniques, students will normalize the data in preparation for feature aggregation, sampling, dimensionality reduction and one hot encoding.

At the end of the course, student will know:

1. How to perform a data quality assessment.
2. How to clean a data set in preparation for data analysis.
3. How to identify different statistical data types ideal for different analytical algorithms.
4. How to perform feature aggregation and sampling for analytical efficiency.

OpenHub
Innovation | Collaboration | Education

EXPLORATORY DATA ANALYSIS

In this course, students will learn about the three main rules of data analysis and the associated statistical concepts relevant to performing different analytical methods. Using insight gained from the data preprocessing step, students will learn how to identify a simple model with the ideal number of predictor variables.

At the end of the course, student will know:

1. How use the three main rules of data analysis to inform model selection.
2. How to run statistical analysis on data sets and interpret the results.
3. How to determine which variables align with the research questions posed.

OpenHub
Innovation | Collaboration | Education

PYTHON

In this course, students will learn how to load a data set into Python for analysis. The Python language will be covered to allow the student to become familiar with data analysis techniques accomplished. Specific algorithms and related libraries will be used to develop machine learning models used for prediction.

At the end of the course, student will know:

1. How to use Python to perform analysis on different size data sets.
2. How to import data sets from Github.
3. How to identify and import libraries into Python IDE.
4. How to create a workflow to perform data analytics in Python.

OpenHub

Innovation | Collaboration | Education

R

In this course, students will learn how to load a data set into R for analysis. The R language will be covered to allow the student to become familiar with data analysis techniques accomplished. Specific algorithms and related packages will be used to develop machine learning models used for prediction.

At the end of the course, student will know:

1. How to use R to perform analysis on different size data sets.
2. How to work with packages in R to streamline project efficiency.
3. How to use the R Studio IDE to run different algorithms used for data analytics projects.

4. How to create a workflow to perform data analytics in R.

PYTHON AND R

In this course, students will expand on skills acquired in the prior courses by learning how to combine R and Python capability into one project using R Studio and Anaconda. Students will develop a working Notebook using R Markdown and transfer results into an interactive web-based dashboard using Shiny.

At the end of the course, student will know:

1. How to use Python and R together to efficiently perform data analysis.
2. How to import data sets from different repositories.
3. How to create a workflow using R Markdown to maintain data analytics projects and communicate results to stakeholders.
4. How to develop an interactive web-based dashboard using Shiny.

OpenHub
Innovation | Collaboration | Education

DATA VISUALIZATION

In this course, students will learn how to put data into a chart, graph, or other visual format that helps inform analysis and interpretation using Excel, Tableau, R Studio, and PowerBI.

At the end of the course, student will know:

1. How to select appropriate visualization options based on the type of analysis being conducted.
2. How to effectively communicate data analytics results using visualizations that facilitate insight for business owners.
3. How to interpret and report analysis results to various stakeholders.

OpenHub
Innovation | Collaboration | Education