

DATA ANALYTICS

2020 | Version .03

Unit #1



Python **SQL** **Excel**

The aim of this unit is to set the fundamentals of Data Analysis: How to ask questions; how to interpret questions; how to read, clean, and explore data; what is Machine Learning and how can it help decision making; how to prepare the data for modeling and fit a linear model, how did I validate the model I built and why did I choose it over others (notions of "what does good look like"); what do the results mean; what is the business relevance of this analysis.

Case Study: Healthcare for all

Lab: Customer Analysis

Unit #2



Python **SQL**

The aim of this unit is to learn how to extract and analyze data stored in databases: recognizing two main different types of Databases, identifying the power of SQL, understanding how to query a structured database, manipulating categorical variables, identifying data anomalies and normalizing data, gathering data from various sources, applying window functions, interpreting Entity relations diagrams (ERD) and a brief intro to classification models.

Case Study: Banking

Lab: Movie Marketplace

Unit #3



Python **SQL**

The aim of this unit is to boost SQL skills to query complex databases: joining multiple tables, normalising databases and performing nested subqueries. We will also introduce multi-class classification and the problem of class imbalance.

This unit works as a continuation of the previous unit (2). The output is a classification model built using Python. The students will reinforce their understanding of the whole data process.

Case Study: Banking

Lab: Movie Marketplace

KEY:

Study Weeks

Project Weeks

ML Algorithms:

Linear Regression

Logistic Regression

Unsupervised Learning (K-means)

Soft Skills:

Business

Project Management

Analysis

Presentation

DATA ANALYTICS

2020 | Version .03

Unit #4



The aim of this unit is to expand their comprehension of the Data Analysis toolkit and supervised Machine learning. They are going to expand their data cleaning skills with RegEx, extract features from datasets and tune a KNN approach/methodology considering the bias-variance trade-off.

They will also get introduced to Business Intelligence's main concepts and tools.

Case Study: Healthcare for all

Lab: Customer Analysis

Unit #5



The students will choose their challenge, from both algorithms we learned in the past 4 units. The project briefs are optimized for their portfolio.
For the PT format, unit 5 extends with one extra mini-project that is scheduled right after Unit 1. The aim of this mini project is to foster autonomy, project management skills, and problem-solving at an early stage in the program.

Classification: Real Estate

Regression: Risk Analysis

Part-Time: Fifa Players

Unit #6



The aim of this unit is to boost their Data Analysis skills and consolidate advanced use of the main tech Stack: mapping and understanding the business problem, using advanced functions and detailed expression in their code, optimizing their process using stored procedures in SQL, using excel macros to automate some spreadsheet operations, and finally creating dashboards to visualize results.

Case Study: Banking

Lab: Movie Marketplace

KEY:

Study Weeks

Project Weeks

ML Algorithms:

Linear Regression

Logistic Regression

Unsupervised Learning (K-means)

Soft Skills:

Business

Project Management

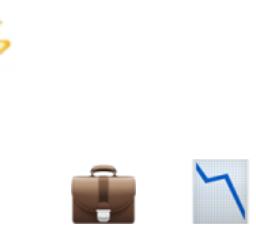
Analysis

Presentation

DATA ANALYTICS

2020 | Version .03

Unit #7



Python

The aim of this unit is to take students' understanding of the logic behind the engineering just one level higher! A lot of the effort in this unit is critical, deep, thinking. We'll be introducing ensemble methods (specifically tree-based algorithms) and deal with high-dimensional datasets. An important inclusion in the ML workflow, is hypothesis testing (including ANOVA). In this unit, the class will find opportunities for debate, critical thinking and deep reflection.

Case Study: Healthcare for all

Lab: Customer Analysis

Unit #8



Python HTML Tableau Presenting

The aim of this unit is to polish their data analytics and engineering skills by performing an end-to-end data product: we will create a program that takes an input from the user and automatically collects data from the internet through web scraping and APIs; then it goes through a clustering model and finally returns an output back to the user. **They will implement agile methodologies to develop the product** and finally they will "sell it" with an engaging presentation

Project: Song Reccomender

Unit #9



Python SQL Tableau Excel Presenting

This is the end of the Bootcamp and the beginning of your career in Data Analytics! Students will complete a full-circle Data Analysis project going through all the process and applying what we've studied in the course.

We'll provide them with datasets and ideas of projects to work with from the main hiring industries, or they can propose their own preference according to their goals.

KEY:

Study Weeks

Project Weeks

ML Algorithms:

Linear Regression

Logistic Regression

Unsupervised Learning (K-means)

Soft Skills:

Business

Project Management

Analysis

Presentation