#### Module 1: Python

Python is the most important and necessary topic that every data scientist should have knowledge about. In this section, our instructors will take you through the basics of Python and areas where it can be used. You will learn how to use some of the current tools such as Numpy, Pandas, and Matplotlib. Therefore, module 1 includes –

* Environment set-up
* Jupyter overview
* Python Numpy
* Python Pandas
* Python Matplotlib

#### Module 2: R

Used for statistical and data analysis, R programming language is one of the advanced statistical languages used in data science. This module teaches you how to explore data sets using R. Here you will learn –

* An introduction to R
* Data structures in R
* Data visualization with R
* Data analysis with R

#### Module 3: Statistics

When working with data, the knowledge of statistics is necessary and an important skill set that you must have. In this module, you will learn –

* Important statistical concepts used in data science
* Difference between population and sample
* Types of variables
* Measures of central tendency
* Measures of variability
* Coefficient of variance
* Skewness and Kurtosis

#### Module 4: Inferential statistics

Inferential statistics is used to make generalizations of populations, from which samples are drawn. This is a new branch of statistics, which helps you learn to analyze representative samples of large data sets. In this module, you will learn –

* Normal distribution
* Test hypotheses
* Central limit theorem
* Confidence interval
* T-test
* Type I and II errors
* Student’s T distribution

#### Module 5: Regression and Anova

This lesson will help you understand how to establish a relationship between two or more objects. ANOVA or analysis of variance is used to analyze the differences among sample sets. Here you will learn –

* Regression
* ANOVA
* R square
* Correlation and causation

#### Module 6: Exploratory data analysis

In this lesson you will learn –

* Data visualization
* Missing value analysis
* The correction matrix
* Outlier detection analysis

#### Module 7: Supervised machine learning

This is a comprehensive module to help you understand how to make machines or computers interpret human language. You will learn –

* Python Scikit tool
* Neural networks
* Support vector machine
* Logistic and linear regression
* Decision tree classifier

#### Module 8: Tableau

Tableau is a sophisticated business intelligence tool used for data visualization. In this lesson, you will learn –

* Working with Tableau
* Deep diving with data and connection
* Creating charts
* Mapping data in Tableau
* Dashboards and stories

#### Module 9: Machine learning on cloud

In this lesson, you will learn –

* ML on cloud platform
* ML on AWS
* ML on Microsoft Azure

Some of the tools used in Crampete’s data science course are –

* Jupyter
* Python Pandas
* Python Numpy
* Python Matplotlib
* Statistical tools like T-test and ANOVA
* SAS
* Exploratory data analysis tools like Apache Spark, Tableau, etc.
* Excel
* Machine learning tools
* Python Ski-kit tool
* Software to learn about neural networks and fuzzy logic
* Natural Language Toolkit (NLTK), which are used by computers to interpret human languages.