

# PROGRAM CURRICULUM

## Introduction to Data Science & Analytics Techniques

Gain foundational data science skills to prepare for a career or further ... of data science across fields, and learn how data analysis can help you make data ... Senior Developer Advocate with IBM Center for Open Data and AI Technologies.

### Key Learning Objectives

- ✓ Introduction to Data Science
- ✓ Introduction to Python
- ✓ Data Analysis Pipeline
- ✓ What is Data Extraction

### Types of Data

- ✓ Raw and Processed Data
- ✓ Data Wrangling
- ✓ Overview of the Analytics Techniques

### Analytics

- ✓ Business Intelligence
- ✓ Business Analytics
- ✓ Industry Examples

## SQL for Data Science

SQL (or Structured Query Language) is a powerful programming language that is used for communicating with and extracting various data types from databases. A working knowledge of databases and SQL is necessary to advance as a data scientist or a machine learning specialist.

### **Basic SQL**

- ✓ Introduction to SQL
- ✓ DDL & DML Statement
- ✓ SELECT Statement AGGREGATE functions
- ✓ WHERE, ORDER BY, DISTINCT, GROUP BY, LIKE, AND & OR clause
- ✓ UPDATE & DELETE query

### **Advanced SQL**

- ✓ JOINS
- ✓ UNION, UNION ALL, INTERSECT
- ✓ Using VIEWS & INDEXES
- ✓ Sub Queries
- ✓ NULL values & DATE function

NumPy is a library for the Python programming language, adding support for large, multi-dimensional arrays and matrices, along with a large collection of high-level mathematical functions to operate on these arrays

### Numpy

- ✓ Indexing
- ✓ ndarray
- ✓ Array Creation
- ✓ Data Type Objects
- ✓ Data type Object (dtype) in NumPy
- ✓ Basic Slicing and ALinear Algebra
- ✓ Sorting, Searching and Counting

#### **Set 1 (Introduction)advanced Indexing**

- ✓ Iterating Over Array
- ✓ Binary Operations
- ✓ Mathematical Function
- ✓ String Operations

#### **Set 2 (Advanced)**

- ✓ Multiplication of two Matrices in Single line using Numpy in Python

Pandas is an open source, BSD-licensed library providing high-performance, easy-to-use data structures and data analysis tools for the Python programming.

## **Pandas DataFrame**

- ✓ Creating a Pandas DataFrame
- ✓ Dealing with Rows and Columns in Pandas DataFrame
- ✓ Indexing and Selecting Data with Pandas
- ✓ Boolean Indexing in Pandas
- ✓ Conversion Functions in Pandas DataFrame
- ✓ Iterating over rows and columns in Pandas DataFrame
- ✓ Working with Missing Data in Pandas
- ✓ Working With Text Data
- ✓ Working with Dates and Times
- ✓ Merging, Joining and Concatenating

## **Data Analysis**

- ✓ Data visualization using Bokeh
- ✓ Exploratory Data Analysis in Python
- ✓ Data visualization with different Charts in Python
- ✓ Data Analysis and Visualization with Python
- ✓ Math operations for Data analysis

## Introduction to NLP

Natural Language Processing (NLP) is a branch of artificial intelligence dealing with the interaction between humans and computers using a natural language.

- ✓ Syntactical Parsing
- ✓ Text Preprocessing
- ✓ Text to Features
- ✓ (Feature Engineering on text data)
- ✓ Noise Removal
- ✓ Lexicon Normalization
- ✓ Lemmatization
- ✓ Stemming
- ✓ Object Standardization
- ✓ Dependency Grammar
- ✓ Part of Speech Tagging
- ✓ Entity Parsing
- ✓ Phrase Detection
- ✓ Named Entity Recognition
- ✓ Topic Modelling
- ✓ N-Grams
- ✓ Statistical features
- ✓ TF – IDF
- ✓ Frequency / Density Features
- ✓ Readability Features
- ✓ Word Embeddings
- ✓ Important tasks of NLP
- ✓ Text Classification
- ✓ Text Matching
- ✓ Levenshtein Distance
- ✓ Phonetic Matching
- ✓ Flexible String Matching
- ✓ Important NLP libraries

# Machine Learning with Python

Machine Learning with Python course dives into the basics of machine learning using an approachable, and well-known, programming language.

## Introduction to Machine Learning

- ✓ Introduction to Machine Learning
- ✓ Python for Machine Learning
- ✓ Supervised vs Unsupervised

## Regression

- ✓ Introduction to Regression
- ✓ Simple Linear Regression
- ✓ Model Evaluation in Regression Models
- ✓ Evaluation Metrics in Regression Models
- ✓ Multiple Linear Regression
- ✓ Non-Linear Regression

## Classification

- ✓ Introduction to Classification
- ✓ K-Nearest Neighbours
- ✓ Evaluation Metrics in Classification
- ✓ Introduction to Decision Trees
- ✓ Building Decision Trees
- ✓ Intro to Logistic Regression
- ✓ Logistic regression vs Linear regression
- ✓ Logistic Regression Training
- ✓ Support Vector Machine

## Clustering

- ✓ Intro to Clustering
- ✓ Intro to k-Means
- ✓ More on k-Means
- ✓ Intro to Hierarchical Clustering
- ✓ More on Hierarchical Clustering
- ✓ DBSCAN

## Recommender Systems

- ✓ Intro to Recommender Systems
- ✓ Content-based Recommender Systems
- ✓ Collaborative Filtering

## Data Analysis & Wrangling

Data wrangling is the practice of converting and then plotting data from one "raw" form into another. The aim is to make it ready for downstream analytics.

- ✓ Data Analysis Pipeline
- ✓ What is Data Extraction
- ✓ Types of Data
- ✓ Raw and Processed Data
- ✓ Data Wrangling
- ✓ Exploratory Data Analysis

## Statistics & Probability About Data

Probability and Statistics form the basis of Data Science.

### **About Data**

- ✓ Data definition
- ✓ Raw and Processed data
- ✓ Data Types (NOIR)

### **Descriptive Stats**

- ✓ Measure of Central Tendency
- ✓ Measure of Dispersion
- ✓ Measure of Association

## **Probability**

- ✓ Basic terminology
- ✓ Rules and Events
- ✓ Conditional probability and Bayes theorem

## **Data Distribution**

- ✓ Skewness
- ✓ t-Distribution
- ✓ Uniform Distribution
- ✓ Binomial Distribution
- ✓ Poisson Distribution
- ✓ Geometric Distribution
- ✓ Gaussian Distribution
- ✓ Standard Normal Distribution
- ✓ Central Limit Theorem

## **Inferential Stats**

- ✓ Estimation technique
- ✓ Hypothesis Testing (t-statistic calculations)

## **Sampling techniques**

- ✓ Random Sampling,
- ✓ Stratified Sampling

## **Statistical Tests**

- ✓ ANOVA
- ✓ Chi-Square

# 6 Month Live Internship

- ✓ Live Projects in an Internship Company with access to Virtual Cloud based Linux System.

## TOOLS & MORE



## EXPERIENTIAL LEARNING



Capstone Project



Assignments



Case Studies



Hackathons



## **Bengaluru**

Bygrad Education Pvt. Ltd.  
7th Cross Rd, HAL 2nd Stage, Indiranagar, Bengaluru,  
Karnataka 560038  
Phone No: +91 9582889465

## **Noida**

Bygrad Education Pvt Ltd.  
C-76, C Block, Sector 63 Noida 201301  
Call us at: +91 9599691368