

Data Science Professional Course Planner

Course Information

Duration
6 Months

Mode of Delivery
Online

General Information

Description

The curriculum, learning content, and projects are designed and created by Careerera faculties in coordination with leading Industry experts. This professional program helps you to build extra skill sets in Data Science and Machine Learning thereby making you a future-ready technology professional.

Learning Goals

- Students will develop relevant programming abilities.
- Students will demonstrate proficiency with statistical analysis of data.
- Students will develop the ability to build and assess data-based models.
- Students will execute statistical analyses with professional statistical software.
- Students will demonstrate skills in data management.

MODULE 01 : PRE-PROGRAM CONTENT

1. INTRODUCTION TO ANACONDA ENVIRONMENT AND INSTALLATION

WEEK 0
1 Hour

2. WORKING ENVIRONMENTS

- ✦ Jupyter Notebook
- ✦ Kaggle
- ✦ Google Colab
- ✦ MySQL Workbench

WEEK 0
2 Hours

3. INTRODUCTION TO:

- ✦ Data Science
- ✦ Machine Learning
- ✦ Artificial Intelligence
- ✦ Deep Learning

WEEK 0

2 Hours

4. THE CRISP-DM FRAMEWORK

- ✦ Business and Data Understanding
- ✦ Data Preparation
- ✦ Modelling
- ✦ Evaluation
- ✦ Deployment

WEEK 0

2 Hours

MODULE 02 : PYTHON & ITS LIBRARIES FOR DATA SCIENCE

1. GETTING STARTED WITH PYTHON

- ✦ Syntax, variables and Data types
- ✦ Data structures in python – List, Tuple, Set, Dictionary
- ✦ Conditional statements – if, Else, Elif
- ✦ Loops – While & For Loops
- ✦ Functions – lambda, recursions, map, filter& reduce functions
- ✦ OOP – Class, Objects, Inheritance, Polymorphism, Encapsulation, Abstraction, Generators, Iterators
- ✦ Programming in Python – Questions Practice
- ✦ Doubts Clearing/Assessment

WEEK 1-4

24 Hours

PYTHON PROGRAMMING ASSIGNMENT

- ✦ Problem Statement
- ✦ Final Submission
- ✦ Solution

2. PYTHON FOR DATA SCIENCE

- ✦ Introduction to NumPy
- ✦ Introduction to Pandas
- ✦ Introduction to Matplotlib
- ✦ Introduction to Seaborn
- ✦ Data Visualization using Seaborn & Matplotlib
- ✦ Doubts Clearing/Assessment

WEEK 5-7

18 Hours

3. EXPLORATORY DATA ANALYSIS

WEEK 8-9

- ✦ Data Sourcing
- ✦ Data Cleaning – null and infinite values, outliers, capping, sanity checks, data formatting, etc.
- ✦ Types of Variables – Categorical & Continuous
- ✦ Univariate Analysis
- ✦ Bivariate Analysis
- ✦ Multivariate Analysis
- ✦ Doubts Clearing/Assessment

12 Hours

EDA CASE STUDY ASSIGNMENT

- ✦ Problem Statement
- ✦ Final Submission
- ✦ Solution

MODULE 03 : DATA ANALYSIS USING SQL

4. DATA ANALYSIS USING SQL

WEEK 10-12

18 Hours

- ✦ Introduction to DBMS & RDBMS
- ✦ OLAP vs OLTP
- ✦ Database Design
- ✦ Database creation in MySQLWorkbench
- ✦ DDL and DML statements
- ✦ MySQL Data Types &Clauses
- ✦ Retrieve Data from Single Tables
- ✦ Retrieve and Transform Data From Multiple Tables Using JOINS and SET Operations
- ✦ Grouping and Summarizing Records
- ✦ Creating Views with Aggregate Functions
- ✦ Writing Subqueries
- ✦ Doubt Clearing/Assessment

SQL ASSIGNMENT

- ✦ Problem Statement
- ✦ Final Submission
- ✦ Solution

MODULE 04 : QUANTITATIVE RESEARCH & STATISTICS FOR DATA SCIENCE

5. STATISTICS

WEEK 13-14

12 Hours

- ✦ Introduction to Statistics
 - ✦ Basics of Probability
 - ✦ Discrete Probability Distribution
 - ✦ Continuous Probability Distribution
 - ✦ Normal Distribution
 - ✦ Poisson's Distribution
 - ✦ Bayes' Theorem
 - ✦ Central Limit Theorem
 - ✦ Doubt Clearing/Assessment
-

STATISTICS ASSIGNMENT

- ✦ Problem Statement
 - ✦ Final Submission
 - ✦ Solution
-

6. HYPOTHESIS TESTING

WEEK 15-16

12 Hours

- ✦ Concepts of Hypothesis Testing: Null and Alternate Hypothesis
 - ✦ Making a Decision and Critical Value Method
 - ✦ p-Value Method and Types of Errors
 - ✦ One-Sample T-Test, Two Sample T-Test
 - ✦ Z-Test, ANOVA, Chi-Square, A/B Testing
 - ✦ Pearson Co-Relation, Co-Variance, Chebyshev-Inequality Formula
 - ✦ Doubt Clearing/Assessment
-

HYPOTHESIS TESTING ASSIGNMENT

- ✦ Problem Statement
- ✦ Final Submission
- ✦ Solution

MODULE 05 : MACHINE LEARNING(ML) FOR DATA SCIENCE & ANALYTICS

1. SUPERVISED LEARNING - REGRESSION

WEEK 17-18

15 Hours

- ✦ Simple Linear Regression
 - ✦ Multiple Linear Regression
 - ✦ Performance Metrics: Accuracy Scores, R^2 And Adjusted R^2 To Compare the Model with Different Numbers of Independent Variables.
 - ✦ Approaches to Feature Selection: Univariate Selection, Feature Importance, RFE
 - ✦ Parameter Tuning and Model Evaluation
 - ✦ Data Transformation and Normalization
 - ✦ Ridge & Lasso Regression (L1 & L2)
 - ✦ Doubts Clearing/Assessment
-

LINEAR REGRESSION ASSIGNMENT

- ✦ Problem Statement
 - ✦ Final Submission
 - ✦ Solution
-

2. LOGISTIC REGRESSION

WEEK 19

6 Hours

- ✦ Concept of Logistic Regression
 - ✦ Univariate Logistic Regression
 - ✦ Multivariate Logistic Regression: Model Building and Evaluation
 - ✦ Dealing with Categorical Independent Variable – One Hot Encoding Vs Dummy Variable
 - ✦ Doubt Clearing/Assignment
-

LOGISTIC REGRESSION ASSIGNMENT

- ✦ Problem Statement
 - ✦ Final Submission
 - ✦ Solution
-

3. DECISION TREES

WEEK 20

6 Hours

- ✦ Concept of Decision Trees
- ✦ Importance and Usage of Gini And Entropy
- ✦ Information Gain
- ✦ Visualizing Decision Trees Nodes and Splits
- ✦ Working of The Decision Tree Algorithm
- ✦ Evaluating Decision Trees Models: Accuracy, Precision, Recall, Confusion Matrix
- ✦ Doubt Clearing/Assignment

DECISION TREES ASSIGNMENT

- ✦ Problem Statement
 - ✦ Final Submission
 - ✦ Solution
-

4. CLASSIFICATION TECHNIQUES – CART

WEEK 21

6 Hours

- ✦ Extending Decision Trees to Regressing Problems
 - ✦ Advantages of Using CART
 - ✦ The Bayes Theorem
 - ✦ KNN Classifier
 - ✦ The Gaussian Naïve's Bayes Classifier – Assumptions of The Naïve Bayes Classifier, Functioning of The Naïve's Bayes Algorithm
 - ✦ Doubts Clearing/Assessment
-

5. SUPPORT VECTOR MACHINE(SVM)

WEEK 22

6 Hours

- ✦ What is Support Vector Machine?
 - ✦ How does SVM Work?
 - ✦ Different Types of SVM
 - ✦ SVM Kernels
 - ✦ SVM Use Cases
 - ✦ How to Implement SVM?
 - ✦ Doubts Clearing/Assessment
-

SVM ASSIGNMENT

- ✦ Problem Statement
 - ✦ Final Submission
 - ✦ Solution
-

6. UNSUPERVISED LEARNING: CLUSTERING

WEEK 23

6 Hours

- ✦ Introduction to Clustering
 - ✦ K-Means Clustering: Linkage, Use of Elbow Curve & Silhouette Score
 - ✦ Hierarchical Clustering – Agglomerative & Divisive, Distance Matrix, Dendrogram
 - ✦ Doubts Clearing/Assessment
-

CLUSTERING ASSIGNMENT

- ✦ Problem Statement
- ✦ Final Submission

- ✦ Solution
-

7. PRINCIPAL COMPONENT ANALYSIS(PCA)

WEEK 24

6 Hours

- ✦ Noise in The Data and Dimensionality Reduction
 - ✦ Capturing Variance – The Concept of Principal Components
 - ✦ Assumptions in Using PCA
 - ✦ Eigenvectors and Orthogonality of Principal Components
 - ✦ What Is the Complexity Curve?
 - ✦ Advantages of Using PCA
 - ✦ The Working of The PCA Algorithm
 - ✦ Singular Value Decomposition
 - ✦ Doubts Clearing/Assessment
-

PCA ASSIGNMENT

- ✦ Problem Statement
 - ✦ Final Submission
 - ✦ Solution
-

8. ENSEMBLE MODELLING

WEEK 25-26

12 Hours

- ✦ Popular Ensembles
 - ✦ Bagging
 - ✦ Bootstrap
 - ✦ Introduction to Random Forests
 - ✦ Feature Importance in Random Forests
 - ✦ Boosting
 - ✦ How Boosting Algorithm Works?
 - ✦ Adaboost
 - ✦ Doubts Clearing/Assessments
-

ENSEMBLING MODELLING ASSIGNMENT

- ✦ Problem Statement
 - ✦ Final Submission
 - ✦ Solution
-

ML CASE STUDY ASSIGNMENT

- ✦ Overview of The Domain and Associated Concepts
- ✦ Problem Statement
- ✦ Evaluation Rubric

- ✦ Final Submission

MODULE 06 : BUSINESS ANALYSIS WITH TABLEAU

TABLEAU BASICS

WEEK 27

6 Hours

- ✦ Introduction to Data Visualization and Tableau
- ✦ Tableau Products Suite
- ✦ Data Types
- ✦ Tableau File Types
- ✦ Connecting to Data Sources
- ✦ Handling R Data
- ✦ Connecting to MS Access Database
- ✦ Loading & Reshaping Data Aggregation
- ✦ Working with Continuous and Discrete Data Using Filters
- ✦ Working with Dates
- ✦ Using Calculated Fields & Parameters
- ✦ Creating Tables and Charts
- ✦ Building Dash Boards and Storyboards
- ✦ Sharing Your Work and Publishing for A Wider Audience

TABLEAU ASSIGNMENT

- ✦ Problem Statement
- ✦ Final Submission
- ✦ Solution

CAPSTONE PROJECT PHASE

- ✦ Overview of The Domain and Associated Concepts
- ✦ Problem Statement
- ✦ Evaluation Rubric
- ✦ Final Submission

EMPLOYABILITY ENHANCEMENT PROGRAM

WEEK28

6 Hours

- ✦ Training Need Analysis
- ✦ Communication Skills
- ✦ Placement Skills
- ✦ Behavioral Science
- ✦ Resume Building
- ✦ Corporate Etiquette
- ✦ Personality Enhancement, Etc.

DISCLAIMER:

Modules may change in order to keep up with the demand and developing issues in the relevant discipline or to reflect the research interests of teaching staff in current affairs.

