

PROGRAMMING ESSENTIALS AND VISUALIZATION FOR DATA SCIENCE

PYTHON ESSENTIALS

Module 1

Introduction to Python IDE's, Concept of Packages, Data Types & Data objects, Basic Operations, Control flow & conditional statements, Python Built-in Functions

DATA CLEANING

Module 2

Sub Setting / Filtering / Slicing Data, Using indexing, Using functions, Mutation of table, Binning data, Renaming columns or rows, Type conversions, Setting index, Handling duplicates /missing/Outliers, Creating dummies from categorical data, Data manipulation tools

DATA SCIENCE OPERATIONS

Module 3

What is NumPy?, Overview of functions & methods in NumPy, Data structures in NumPy, Creating arrays and initializing, Reading arrays, Slicing and indexing, Combining arrays, What are pandas, Pandas Data Structures (Series & Data Frames), Functions & methods.

DATA VISUALIZATION

Module 4

Introduction to Matplotlib, Basic Plotting with Matplotlib, Line Plots. Simple to complex visualizations Advanced Visualization Tools

DATA SCRAPPING, WRANGLING

Module 5

Data Scrapping, Finding data across sources, Querying an API directly Stocks, Weathers etc., Browser-based Scrapping, Scrapping tables such as Wikipedia/IMDB, Data Wrangling, Integrating data

APPLIED STATISTICS FOR DATA SCIENCE USING R/PYTHON

STATISTICS PRIMER

Module 6

Measures of central tendencies, Measures of variance, Measures of frequency, Measures of Rank, Basics of Probability, distributions, Conditional Probability (Bayes Theorem)

STATISTICAL METHODS, HYPOTHESIS TESTING

Module 7

Descriptive vs. Inferential Statistics, Discrete & Continuous distributions, Concept of Sampling & types of Sampling, Hypothesis Testing and Applications, Statistical Methods - Z/t-tests (One sample, independent, paired), ANOVA, Correlation and Chi- square.

STATISTICAL DATA ANALYSIS

Module 8

Exploratory data analysis, Descriptive statistics, Frequency Tables and summarization, Uni-variate Analysis (Distribution of data & Graphical Analysis), Bi-Variate Analysis(Cross Tabs, Distributions & Relationships, Graphical Analysis)

APPLIED ECONOMETRICS USING PYTHON

Module 9

Working with Financial Datasets and Time series, Webscrapping financial information using Python, Performing regression, PCA application for multivariate financial datasets, Forecasting models using GARCH, Financial Time Series models

MACHINE LEARNING & PREDICTIVE MODELLING FOR DATA SCIENCE

INTRODUCTION TO PREDICTIVE MODELLING

Module 10

Concept of model in analytics and how it is used, Common terminology used in modeling process, Types of Business problems - Mapping of Algorithms, Different Phases of Predictive Modeling, Data Exploration for modeling, Exploring the data and identifying any problems with the data, Identify missing/Outliers in the data, Visualize the data trends and patterns

SUPERVISED LEARNING REGRESSION PROBLEM

Module 11

Linear Regression, Polynomial Regression, Multivariate Regression, Support Vector Machines, Ridge Regression, Lasso Regression, Problems with Regression

CLASSIFICATION PROBLEMS

Module 12

Logistic Regression, K-Nearest Neighbor, Naïve Bayes Classifier, Decision Trees, Ensemble Learning - Bagging, Random Forest, Adaboost, Gradient Boost, XGBoost, Support Vector Classifier

NEURAL NETWORKS

Module 13

Multi Layer Perceptrons, Convolutional Neural Networks, Recurrent Neural Network, Auto Encoder, Generative Adversarial Network, Graph Neural Networks, Applications

APPLIED BUSINESS ANALYTICS

APPLIED MARKETING ANALYTICS

Module 14

You will learn how to analyze Marketing campaigns with Python. You will learn how to analyze Social media data in Python. You will perform advanced Marketing analytics such as Market Basket Analysis using Python, Customer Segmentation Analysis, Customer Analytics and Customer Churn Analytics.

APPLIED FINANCIAL ANALYTICS

Module 15

You will learn how to work with open-source Financial Datasets and Time series. You will learn how to web-scrape historical time series, fundamental data using Python You would learn Portfolio Modeling with Python, Portfolio Optimization, Portfolio Analytics using Python, Risk Analytics, Value at Risk (VaR) calculations, Monte Carlo Simulations for Portfolio Modeling, Simulating using Python.

SOCIAL MEDIA ANALYTICS

Module 16

Learn to analyse the unstructured textual data to derive meaningful insights. - Text Mining and Natural Language Processing (NLP). Word Clouds, Sentiment Analysis, Semantic network, Clustering, Extract user reviews of the product/services from Amazon, Snapdeal and trip advisor, Install Libraries from Shell, Extraction and text analytics in Python, LDA / Latent Dirichlet Allocation, Topic Modelling, Sentiment Extraction, Lexicons & Emotion Mining.

BUSINESS INTELLIGENCE AND DASHBOARDING

Module 17

This module will introduce you to Business Intelligence Reporting and Dashboarding using Power BI Desktop. You will know how to extract data from various sources and establish connections with Power BI Desktop, perform transformation operations on data and the Role of Query Editor in Power BI. Finally, you will learn to work with complex databases, report elements and build end-to-end industry styled Dashboards.

APPLIED BIG DATA ANALYTICS

BIG DATA PRIMER

Module 18

What is Big Data, Big data in marketing, analytics, retail, hospitality, consumer good, defense etc., Technologies for Handling Big Data, Introduction to Hadoop, Functioning of Hadoop, Cloud computing (features, advantages, applications).

BIG DATA STORAGE AND PROCESSING

Module 19

Big Data storage systems, Relational Databases, NoSQL Databases: HBase, Graph DB, Distributed File Systems/HDFS, Cloud storage. Introduction to Big Data processing platforms - Data Volume: Hadoop, Spark, Data Velocity: Storm, Complex Event Processing, Cloud platforms.

DATA ENGINEERING

Module 20

Hadoop And MapReduce Programming, Data Management And Relational Database Modelling, NOSQL Databases And Apache HBase, Data Warehousing, Data Ingestion With Apache Sqoop And Apache Flume, Building And Querying Data Warehouse With Apache Hive.

BUSINESS USE-CASES IN BIG DATA

Module 21

Applied Case Studies to different domains and areas such as Social Media, Marketing Analytics, Financial Analytics, Supply Chain, HR Analytics, Google Analytics