



GeeksforGeeks

# COMPLETE **DATA SCIENCE** PROGRAM



Detailed  
Course Syllabus

# CONTENTS

## PYTHON FOR DATA SCIENCE

### INTRODUCING PYTHON

#### Python Basics

- Python Introduction
- print() in python
- Variables in python
- input() in python
- Arithmetic Operators in python
- Type () in python
- Type Conversion in python
- Comments in python
- if, else and elif in python

#### Operators

- Arithmetic Operators
- Logical Operators
- Identity Comparison Operators
- Membership Test Operators
- Bitwise Operator Part 1
- Bitwise Operator Part 2

#### Loops in Python

- Loops In Python
- While Loops in python
- range() in python
- For Loop In python
- Break In Python
- Continue In Python
- Nested Loop in Python

#### Functions in Python

- Functions in Python

# CONTENTS

## Strings in Python

- Strings in Python
- Escape sequence & Raw Strings
- String Operations Part 1
- String Operations Part 2
- String Comparison
- Pattern Searching

## Data Structure

- List Introduction
- Tuples in Python
- Set in Python
- Dictionary in Python
- Slicing (List, Tuple And String)
- Comprehensions in python

## Object Oriented Programming

- Introduction to OOPs
- Classes and Objects
- Encapsulation
- Decorators
- Class method and static members
- Inheritance
- Types of Inheritance
- Multiple Inheritance
- Polymorphism
- Abstraction
- Operator Overloading
- Abstract Class

## DATA TOOLKIT

### Getting Started with Files

- Reading Data from Text-File
- Reading Data from text-file Corpus
- Text Preprocessing
- Writing Data On a text-file
- Writing Data on a text-file with New Line

# CONTENTS

## Inventory Management System with Files

- Inventory Management with Files - Overview
- Inventory Management with Files - Product Details
- Inventory Management with Files - Updating Inventory
- Inventory Management with Files - Add Functionalities
- Inventory Management with Files - Generating Sales
- Inventory Management with - Conclusion

## Inventory Management System with JSON

- Inventory Management System - Overview
- Inventory Management System - Generating Bill
- Inventory Management System - Updating Inventory
- Inventory Management System - Saving Record on JSON
- Inventory Management System - Loading Record from JSON
- Inventory Management System JSON - Adding Functionalities
- Inventory Management System JSON - Generating Sales Structure
- Inventory Management System JSON- Generating Sales file
- Inventory Management System JSON- Conclusion
- Inventory Management System JSON - Discount

## Mastering Numpy Arrays

- Getting Started with Numpy
- Reshape and Random Number Generator
- Arithmetic Operations on Array
- Arithmetic Operations on Multiple Arrays
- Array Sorting
- Array Merging
- Array Slicing - DAP
- Automating using Numpy

## Getting Started with OS

- Introduction to OS, CLI and GUI
- OS Commands on Mac - Directories
- OS Commands on Mac - Files
- OS commands on Windows

# CONTENTS

## Jupyter Notebook Setup

- Jupyter Notebook Setup
- Jupyter Notebook Walkthrough

## OS with Python

- OS Library - Directories
- OS Library - List Directories
- OS Library - Bulk Directories Creation
- OS Library - Hierarchical Bulk Directories Creation
- Bulk Text-file Reading
- Bulk Text-file Data Combining

## OS with Python

- OS Library - Directories
- OS Library - List Directories
- OS Library - Bulk Directories Creation
- OS Library - Hierarchical Bulk Directories Creation
- Bulk Text-file Reading
- Bulk Text-file Data Combining

## DATA ANALYSIS WITH PYTHON

### Getting Started with Pandas

- Getting Started with Pandas
- Dataset Walkthrough

### Statistics

- Mean, median mode
- Standard Deviation and Variance
- Normal Distribution

### Data Preprocessing

- Data Preprocessing - Removing Null Value Rows
- Data Analysis - Numeric
- Data Analysis - Categorical
- Data Analysis - Automatic Categorical
- Null Values Handling - Numeric
- Null Values Handling - Categorical
- Null Values Handling on GooglePlaystore Dataset

# CONTENTS

## Data Analysis

- Data Analysis with Multiple Columns
- Data Analysis using Conditions
- Group By in Pandas

## Data Visualization on Heart Disease Dataset

- Heart Disease EDA - Introduction to Kaggle
- Heart Disease EDA - Age (Distort)
- Heart Disease EDA - Categorical Columns (Pie Charts)
- Heart Disease EDA - Violin Plot
- Heart Disease EDA - Correlation (Heatmap)
- Heart Disease EDA - Correlation (Pair Plot)
- Heart Disease EDA - Correlation - (Joint Plot)

## Black Friday Sales Data Analysis

- Walkthrough
- Analyzing Columns
- Analyzing Gender
- Analysing Age & Marital Status
- Multi Column Analysis
- Occupation and Products Analysis
- Combining Gender & Marital Status

## GDP Analysis Assignment & Solution

- GDP Analysis - Assignment
- GDP Analysis - Dataset Walkthrough
- GDP Analysis - GDP Growth of a Country
- GDP Analysis - GDP Growth on whole Dataset
- GDP Analysis - Plotting Graphs Using Polly
- GDP Analysis - Plotting Graphs in Bulk
- GDP Analysis - Compare GDP across Countries
- GDP Analysis - Compare GDP across Countries Advanced
- GDP Analysis - Compare GDP Growth Comparison

## EXCEL

### Introduction to Excel

- Interface of Excel
- Entering Data
- Changing the Structure of Worksheet
- Data Styling

# CONTENTS

## Data Entry in Excel

- Entering Data
- Tables in Excel
- Insert and Create Objects in Excel
- Managing Large Worksheets
- Find, Replace
- Autofill, Custom List

## Data Formatting & Validation

- Data Validation In Excel
- Conditional Formatting
- Data Consolidation vs 3D Sum
- Printing Options In Excel
- What-If-Analysis

## Functions in Excel

- Text Functions In Excel
- Date Functions In Excel
- Logical Functions In Excel
- BODMAS Rule and Count Functions
- Index Function Vs Match Function
- Financial Functions In Excel
- Statistical Functions in Excel
- Maths and Trigonometric Functions
- VLOOKUP In Excel
- HLOOKUP

## Hyperlinks & Illustration in Excel

- How to add hyperlinks in excel
- Excel illustration
- Excel Illustrations Part-1
- Excel Illustrations Part-2
- Errors in Excel

## Pivot Table & Charts in Excel

- Pivot Table Part-1
- Pivot Table Part-2
- Charts in Excel- Part 1
- Charts in Excel- Part 2
- Charts in Excel- Part 3
- Spark lines Chart Vs Pivot Chart

# CONTENTS

## Shortcuts in Excel

- Shortcut Keys (Ctrl A to Ctrl Z)
- Shortcut Keys for Accessing the Tabs
- Some more Alt Short Cut Keys
- F 1 to F 12 Shortcut Keys
- Shift F1 to Shift F12 Shortcut keys
- Ctrl F 1 to Ctrl F 12 Shortcut keys
- Ctrl 1 to Ctrl 0 Shortcut keys
- Ctrl Single key Shortcut keys
- Ctrl Shift Single key Shortcut keys

## Visual Basic Analysis

- Introduction to Macros and VBA
- How to record a Macro
- How to save a Macro
- Code Your First Macro ,Add/Edit Cell Content
- Copy/Paste Sheet Content and WITH Block using Macro
- Border and Alignment Using Macro
- How to Change Font Color and Cell color using Macro
- Change Orientation, Wrap Text and Merger/Unmerge Cells using Macro
- Clear and Delete Cells Using Macro
- Change Column Width and Row Height Using Macro Activate Cells
- Add New Sheets
- Copy, Move Sheets using Macro
- Change Colors and Hide/Unhide Tabs using Macro
- Activate and Password Protect Sheets using Macro
- Create , Save and Close Workbooks using Macro
- Open & Delete Workbooks using Macros
- Comments & Variables in Macro
- If & For in Macro
- Name and Rename a Sheet using Macro

## SQL

### What is DBMS?

- Databases
- Databases vs FS
- Relational and Non Relational Databases
- Database widely used(Examples)



# CONTENTS

## Installing MYSQL(MacOS/Windows)

- Installing MYSQL(MacOS/Windows)

## Introduction to SQL

- What is SQL
- Basic Terminologies
- ER Diagrams

## Types of Commands

- Types of Commands

## Creating Tables and Databases

- Creating a Database
- Creating a Table
- What is Schema?
- Modifying Database

## Inserting Data

- Inserting Data
- Keys

## Retrieving Data

- SELECT Statement
- Queries based on conditions (Simple)

## Data Types in SQL

- Data Types in SQL

## Constraints in SQL

- NOT NULL
- PRIMARY KEY
- UNIQUE
- FOREIGN KEY
- CHECK
- DEFAULT

# CONTENTS

## Updating-Data

- SELECT CLAUSE
- WHERE CLAUSE
- AND & OR CLAUSE
- LIKE CLAUSE
- TOP CLAUSE
- Updating a single row
- Updating Multiple rows
- DELETE QUERY

## Nested Queries

- Scalar
- Column
- Row
- Exists
- Correlated

## Operators in SQL

- Arithmetic operators
- Comparison operators
- Logical Operators
- Wildcard Operators

## Aggregation

- What is aggregation
- Min, Max, SUM, Avg.
- COUNT, DISTINCT
- ORDER BY
- GROUP BY
- HAVING
- CASE When
- Sorting Results

## Joins

- What are joins?
- Types of joins
- Left join
- Right join
- Inner join
- Outer join
- Natural join

# CONTENTS

## Unions

- Unions

## Alias

- Alias

## Indexes

- Indexes

## Alter Command

- Alter Command

## Truncate and Drop

- Truncate and Drop

## Transaction

- Transactions

## Clone Tables

- Clone Tables

## Handling Duplicates

- Handling Duplicates

## Injection

- Injection

## DLC Commands

- GRANT and REVOKE

## TLD Commands

- COMMIT
- SAVE POINT
- ROLLBACK

# CONTENTS

## Function in SQL

- DATE Functions
- SUBSTRING
- LCASE, UCASE. CONCAT etc

## Views

- What are views?
- Advantages of Views
- CRUD In Views

## Miscellaneous

- Comment
- Using regex
- Stored Procedures
- Triggers
- CTE

## Normalization in DBMS

- Normalization in DBMS

## Banking Project using MySQL and Python

- Banking Project using MySQL and Python

## Banking PostgreSQL and Python

- Banking PostgreSQL and Python

## NoSQL

- NoSQL

## TABLEAU

### Introduction to Tableau

- Tableau Installation
- Connecting Tableau with Data Navigating Tableau
- Exporting the Worksheet
- Dashboards
- Workbook

# CONTENTS

## Understanding the Parameters

- Introduction to Parameters
- Checking the data parameter and format
- Measure vs Dimension
- Continuous data vs Discrete data

## Basic Plots in Tableau

- Creating Bar Graph
- Creating Line Plot
- Creating Scatter Plot.

## Fundamentals of Tableau

- Marks Cards
- Encoding to Marks
- Labelling and Tool Tips addition
- Applying Filters to the plot
- Data Hierarchies
- Need of Calculated Fields
- Add Calculated Fields
- Table Calculations
- Highlighting in Tables
- Sets in Tableau
- Way to implement Sets
- Detailed Expressions in Tableau
- Conditional Formatting
- Groups
- Sparklines

## Designing the plots

- Heatmap
- Histogram
- Box and Whisker Plot
- Dual Axis Combo Chart
- Tree Map
- Bullet Graph
- Stacked Area Chart
- Pie Chart
- Donut Chart

# CONTENTS

- Funnel Chart
- Gantt Chart
- Waterfall Chart
- Graph Pareto Chart
- Control Chart
- Dumbbell Charts
- Jitter Points
- Clustering in Tableau
- Word Cloud
- Bubble Chart
- Bump Chart
- Choropleth Map
- Symbol Map
- Dual-Axis Map

## Project 1: Superstore Sales Analysis Dashboard

- Introduction to Superstore Sales Analysis Dashboard
- Understanding the data
- Designing - Overall Sales Choropleth Map
- Sales vs Profit Line Chart
- Discount Histogram Distribution
- Profit Histogram Distribution
- Sales Donut Chart
- Profit Donut Chart
- Discount Donut Chart
- Preparing the Final Dashboard

## Project 2: COVID-19 Worlds Dashboard

- Introduction to Covid-19 World Dashboard
- Understanding the Data - Project Covid
- Designing - Symbol Map showing Covid Cases in the worlds
- Covid Case Trends with Stack Area Chart
- Top-10 Covid Affected Countries and their trends
- Confirmed Cases and Death Cases Pie Chart
- Preparing the Final dashboard - Project Covid

# CONTENTS

## WEB SCRAPING

### To Scrape

- How to scrap a Web-Page
- Scraping Quotes
- Scraping Quotes with Author Details
- Scraping Author Info
- Scraping Quotes from Multiple Pages
- Book Scraper I Scraping Books data from Home-Page
- Book Scraper I Scraping Books data from Multiple Pages
- Book Scraper I Individual Page Scraper
- Books Scraper I Data Combining

### Wikipedia Scraper

- Wikipedia Article Scraping
- Google search Link Generator
- Wikipedia Scraping by Title

### Selenium

- Getting Started with Selenium

### YouTube Scraper

- YouTube Web Scraping I Understanding the Tags
- YouTube Web Scraping I Data from Channel Page
- YouTube Web Scraping I Video Data Scraping
- YouTube Web Scraping I Saving Dataset

### Stock Image Scraper

- Stock Image Scraper I Link Scraper
- Stock Image Scraper I Image Scraper

### Stock Image Scraper Infinite Scroll

- Stock Images Infinite Scroll - Website Walkthrough
- Stock Images Infinite Scroll - Auto Infinite Scroll
- Stock Images Infinite Scroll - Finding the Bottom

# CONTENTS

- Stock Images Infinite Scroll - Scraping the Data
- Stock Images Infinite Scroll - Saving the Dataset
- Stock Images Infinite Scroll - Dynamic Name Allocation
- Stock Images Infinite Scroll - Downloading All Images
- Stock Images Infinite Scroll - Saving the CSV

## Image Dataset Creation

- Image Dataset Creation - Finding all the Tags
- Image Dataset Creation - Creating Folder for all tags
- Image Dataset Creation - finding source and destination path for each Image
- Image Dataset Creation - Creating Final Dataset
- Image Dataset Creation - Dataset with Threshold Images

## MACHINE LEARNING & AI

### Introduction to AI

- What is AI
- Subsets of AI

### How Data Science Comes into Play

- What is Data Science
- AI vs ML vs DL

### Linear Regression

- Linear Regression Intuition
- Forward Propagation and Cost Function in Linear Regression
- Gradient Descent in Linear Regression
- Updating the Parameters in Linear Regression
- Detailed Mathematics behind Linear Regression
- Linear Regression Model from Scratch
- Linear Regression Model Training
- Linear Regression Model Prediction
- Linear Regression Model using ScikitLearn library

### Multiple Linear Regression

- Multiple Linear Regression Intuition
- Multiple Linear Regression using Hands On
- Linear Regression Model Assumption
- Linear Regression Assumptions Hands On
- Ordinary Least Square (OLS) Method
- Multiple Linear Regression using OLS



# CONTENTS

## Polynomial Linear Regression

- Polynomial Linear Regression Intuition
- Polynomial Linear Regression Hands On

## Support Vector Machine

- Support Vector Regression Intuition
- Support Vector Regression Hands On

## Decision Tree

- Decision Tree Regression Intuition
- Decision Tree Regression Hands On

## Random Forest

- Random Forest Regression Intuition
- Random Forest Regression Hands On

## Classification Algorithm

- Logistic Regression Intuition
- KNN Algorithm Intuition
- Naive Bayes Intuition
- Project Titanic - Classification

## Clustering Algorithm

- K means intro
- K means Initialise Centres
- E step in K-Means
- Plotting Clusters
- M Step in K-Means
- Random Init improvement in K-Means

## Feature Engineering

- Feature Selection - with Correlation Matrix
- Feature Selection - with Extra Tree Classifier
- Feature Selection - with SelectKBest Method
- Principal Component Analysis (PCA) Intuition
- PCA Implementation
- TSNE Intuition
- TSNE Implementation
- K-Fold Cross Validation Intuition
- K-Fold Cross Validation Implementation

# CONTENTS

## MNIST Handwritten Digit Recogniser

- Intro - MNIST Dataset
- Dataset Introduction
- Introduction to images in python
- Feature engineering in Images
- Evaluating the model

## Titanic Survival I EDA

- Dataset Intro
- Data Visualisation - 1
- Data Visualisation - 2
- Feature Engineering -1
- Feature Engineering -2
- Feature Engineering -3
- ML Modelling and submission

## PUBG Game Prediction

- Introduction
- Libraries and data understanding
- Data Wrangling
- Feature Engineering
- Cat Boost Model prediction and evaluation

## Human Activity Recognition using Smartphone Data

- Project Introduction
- Libraries and Data Understanding - Data Wrangling -
- EDA - Analysing how acceleration strongly relates to the body activity
- EDA - Understanding how body x-axis and gravity is linked to the body activity
- EDA - Understanding how body y-axis and gravity is linked to the body activity
- Analysing data using PCA
- Analysing data using tSNE
- Data preparation for ML models
- Logistic regression model with Hyperparameter tuning and cross validation
- SVM model with Hyperparameter tuning and cross validation
- Decision Tree and Random Forest with Hyperparameter tuning and cross validation

# CONTENTS

## Predicting Solar Irradiance

- Understanding the project
- Libraries and data understanding -
- Data Wrangling -
- Feature Selection using Correlation Matrix
- Feature Selection using SelectKBest Method
- Feature Selection using Extra Tree Classifier
- Feature Engineering with BoxCox, Log, Min-Max and Standard transformation
- Preparing data - Standardisation and Splitting
- Prediction with XGBoost
- Using Multilayer Perceptron for prediction

## IMAGE PROCESSING

### Fundamentals of Image Processing

- Matrix Vs Image
- DIY High Res Grayscale
- RGB colour scale
- Create Colors- RGB
- Adding Transitions to RGB
- Create Custom Colors

### Image Processing Techniques

- BGR vs RGB
- Frame Extraction
- Display image in OpenCV

### Image Processing on Live Web Cam

- Working with Webcam
- Webcam - Flip and Crop
- Webcam - Frame Extraction

### Taking a selfie with OpenCV

- Clicking a selfie using OpenCV
- Clicking multiple selfies using OpenCV

# CONTENTS

## Image Manipulation

- Draw Shapes with OpenCV
- Edge Detection
- Image Blur
- Edge Detection with Blur
- Image Scaling

## DIY Instagram Filters

- Brightness Control
- Warm and Cool(Video)
- Warm and Cool(Image)
- Merging Images

## Masking

- Thresholding on Greyscale
- Colour Masking - Images
- Colour Masking - Videos

## Adding Logo on a Live Video

- Adding Logo on Live Video
- Analysing Aspect Ratio
- Auto Fit
- All Direction Fit
- Dynamic Fit
- Final Fit
- Adding Transparency

## Face Detection and Manipulation

- Face Detection-HaarCascade
- Face Crop-HaarCascade
- Face Blur-HaarCascade
- Face Black-Round and Square
- Extract Face from an Image

# CONTENTS

## DEEP LEARNING

### Perceptrons

- Introduction- Neurons vs Artificial Neural Networks
- Learning of ANN
- Gradient Descent of ANN
- Implementation and Visualization of perceptron

### Multi Layer Perceptron Architecture

- Architecture and Introduction
- Layer Architecture
- Why we need Multilayer Perceptron
- Architecture of Multilayer Perceptron
- Forward Propagation in Multilayer Perceptron
- Backward Propagation in Multilayer Perceptron
- Final Equation for Multilayer Perceptron
- Activation Function and Derivatives
- Titanic Survival Prediction using ANN

### Convolutional Neural Networks

- Intro Video
- Introduction to CNN
- Why we need CNN
- Convolutional Layer, Filters, Stride-Part 1
- Convolutional Layer, Filters, Stride-Part 2
- Pooling
- Overall Model
- Malarial Cell Detection using CNN

## NLP

### Getting Started with NLP

- Getting Started with NLTK and Tokenization
- Stemming & Lemmatisation
- StopWords Removal from Scratch
- Corpus & Vocabulary
- Vocabulary with Keras

# CONTENTS

## Mastering Strings and ASCII codes

- Getting started with ASCII Codes
- ASCII Code to String Conversion and Vice Versa A-Z with ASCII Codes
- DIY Functions - capitalize()
- DIY Functions - upper() and lower()
- DIY Function - Checking the Data
- DIY Function - Title
- DIY Library for String Operations

## Regular Expression from SCRATCH

- Getting started with Regular Expressions
- Pattern Matching with Alphanumeric
- Text Preprocessing with RE
- Email Pattern Matching with re
- DIY Pattern Matching - Continuous Numbers
- DIY Pattern Matching - Words Finding
- DIY Pattern Matching - Words starting from specific characters
- DIY Pattern Matching - Email Extraction

## Getting Started with Spacy

- Getting started with Spacy library
- Stop-Words Removal with Spacy
- Synonyms and Antonyms

## Text Sequencing using Word Cloud

- Word Cloud
- Text Encoding - Decoding
- Text Encoding - Decoding I Without Stop Words

## Guessing the Title of a Corpus Project

- Guessing Title I Most Frequent Word

## Spell Checker Project

- Finding Probability Distribution
- Spell Checking Architecture
- Splitting and Deletion Operation
- Swap, Replace and Insert Operation I Spell Checker from Scratch
- Predicting the Correct word I Level - 1 Edit
- Predicting the Correct word I Level - 2 Edit