BROADWAY
INFOSYS
Since 2008

Course Syllabus for
Data Science with Python Training

Presented to:
Nagarjuna College of Information Technology,
Kathmandu, Nepal

Presented by:
Ms. Triveni Jirel
Client Relation Officer

Broadway Infosys
Tinkune, Subidhanagar, Kathmandu, Nepal
info@broadwayinfosys.com
Ph.: 977-1-4117578, 4111849, 9841002000

COURSE OUTLINE: DATA SCIENCE WITH PYTHON TRAINING

Course Outline: Python Programming

Installation

Python version and pip package manager

Introduction to Google Colab, Jupyter Notebook / IDE

- Introduction to markdown

Python Program and statements

Python Arithmetic Operators

- Using Python as calculator

How to define a variable name and Variable Naming convention in Python

Operator, Operands, and Operator Precedence

Changing and updating variable values in Python

Assigning multiple values to multiple variables

Data types in Python

Number data type: int, float, complex

- Number data type
- Taking input from the users
- Type casting and type checking
- Type validation
- Number type with conditionals

Conditions and Recursion

- Modulus operator
- Boolean expressions
- Logical operators
- Conditional execution
- Chained conditionals
- Nested conditionals

- Recursion
- Stack diagrams for recursive functions
- Infinite recursion

Iteration

- Multiple assignments
- Updating variables
- The while statement
- Break
- Debugging
- For loop

Python string

- Introduction
- Single line vs multiline string
- Indexing
- slicing
- len()
- Loop: for loop using range()
- Loop with conditionals
- continue vs break
- characters vs substrings
- Immutable data type
- String methods: .replace(),.lower(), .upper(),.title() .lstrip(), .rstrip(), .strip(), .split(), .join(), .isdigit(), isupper(), islower(), .format()

Python Built-in data types

List

- Introduction to list
- Indexing/Negative indexing
- Slicing
- looping & conditionals, len()
    - Different types of for loop ,while loop, for loop vs while loop

- list of list and nested loop
- Membership operators: in , not in
- Mutable vs Immutable data type with exmaple
- List methods: .insert(), .append(), .remove(), .pop(), .sort(), .extend(), .remove(), .sort()
- List Comprehension
- with if and else

Tuple

- Introduction to tuple
- Indexing , slicing, looping
- list vs tuple
- Typecasting list -> tuple and tuple ->list
- tuple unpacking

Set

- Introduction to set
- .remove() , .add(), .discard() in sets
- Type conversion
- Set operation in Python : union, intersection, difference
- Frozenset vs set

Dictionary

- Introduction
- disctionary methods: .get(), .update(), .keys(), .values(), .pop()
- Loop and dictionary comprehension
- Nested Dictionary

None type

- Identity Operators

Python Functions

- Introduction and syntax: Why function is necessary
- Function definition and function call
- arguments/parameters in function
- return statement in function
- returning multiple value from function
- Handling multiple return values
- Default argument vs non default argument and why it is necessary

- global and local variables
- *args vs **kwargs
- Introduction to Recursion and Recursion tree
- pass keyword

OOP in python

- Class and Objects
- Class attribute and Object attribute
- Initilizing object attribute
- __init__()
- Self keyword and its importance
- Inheritance and its types
- Single Inheritance
- super() method
- Mulitple Inheritance
- Multi-level Inheritance
- Abstraction and Access specifiers
- Polymorphism
- + and len()
- Operator overloading using Dunders/magic method (user defined class)
- function overriding
- Encapsulation

Introduction to Exceptions

- Understanding exceptions in Python
- Types of exceptions and their meaning
- Importance of exception handling
- Handling Exceptions
    - Using try-except blocks to handle exceptions
    - Catching specific exceptions
    - Handling multiple exceptions
    - Using the finally clauses
- Raising Exceptions

File Handling

- open()
- modes:
    - read : 'r'
        - read(), readline(), readlines()
    - write: 'w'
    - append: 'a'
    - create: 'x'
    - Comparison of append and write modes
- File handling on CSV files.
- DictReader and DictWriter
- File handling with exception handling

Others

- Lambda function/Annynomous Function:
    - map() , filter()
- os library
- random library
- math library

Introduction to SQL in python

- Creating database
- Defining table structure with SQL statements and Specifying column names, data types, and constraints
- Inserting new records into tables using SQL INSERT statements
- Retrieving data from tables using SQL SELECT statements
- Modifying existing data in tables using SQL UPDATE statements
- Removing tables data using SQL DELETE statements
- Filtering data using the WHERE clause in SQL SELECT statements

Introduction to git and Github

- Installing Git Bash
    - Overview of Git Bash
    - Installation
- Creating a GitHub Account

- Sign up for a GitHub account
- Set up profile

- Creating an Empty Repository
  - Create a new empty repository on GitHub
- Initializing a Git Repository Locally
  - Initialize a Git repository on your local machine using Git Bash
- Tracking Files
  - Add files to the staging area
  - Commit changes to the repository
- Configuration of Global User Information
  - Configure your global user.name and user.email for Git
- Branching (main)
  - Understand the concept of branches
  - Work with the default main branch
- Adding a Remote
  - Connect your local repository to the remote repository on GitHub
  - Configure the remote repository URL
- Pushing Changes
  - Push your local changes to the remote repository on GitHub
- Cloning
  - Clone a repository from GitHub to your local machine using Git Bash
- Creating a New Branch
  - Create a new branch for making changes
  - Switch between branches
- Pushing Changes to a Branch
  - Push your changes to the remote repository on a specific branch

Pandas

- Introduction to Pandas
- DataFrame Data Structure
- Reading and writing csv files using DataFrame

- Manipulating DataFrame

Basic Data Visualization

- Introduction to Matplotlib and Seaborn and plotly
- Basic plotting using any of these library

Project Work (one of the following):

- Web Scraping project + Database
  OR
- Any desktop application: eg. Data Entry application

Data Science Course

Introduction

- Prelude
- The problem landscape
- Defining Data Science
- Demystifying Data Science, Decision Science, AI, ML and DL
- Overview of Data Scientist's Toolbox

Data Science Toolbox

- Python - Quick recap ? Python 2.7.x or 3.x ?
- Installation and setup
- Data types, functions and important packages
- Data manipulation & Data Engineering
- Data Visualization

Probability and Statistics

- Statistics (90% Theoretical Concept + 10% Practical)
- Introduction
- Data Description
- Population and Sample
- Variables and Variable Measurements Scales
- Data Distribution
- Central measure of Tendency (mean, median, mode)

- Measure of dispersion (Variance, standard deviation)
- Gaussian Normal Distribution
- P values
- Type 1 and Type 2 error
- 1-tailed and 2-tailed Test
- Statistical Test (z-test,t-test, chi-square test)
- Pearson Correlation Coefficient
- Spearman's rank correlation
- Addition Rule and Multiplication rule
- Permutation and Combination
- Function of random variables
- Log-Normal Distribution
- Bernoulli Distribution
- Binomial Distribution
- Pareto Distribution
- Poisson distribution

Numpy

- Introduction to Numpy
- Random Data Generation
- Numpy Array, Indexing & Operations
- Array Data Structures in Numpy
- Array operations and methods
- Course Assignment

Pandas

- Importing Datasets
- Data Wrangling
- Exploratory Data Analysis and Model Development

Scipy and Seaborn

- Introduction to Scipy

- Numerical Computaions
- Exploratory Data Analysis
- Model Generation

Plotting, Charting & Data Visualization

- Principles of Information Visualization
- Basic Charting
- Charting Fundamentals
- Applied Visualizations

Tableau Basics

- Introduction to Tableau
- Download and Install Tableau Public
- Load Data from Excel
- Creating Charts and Graphs
- Basic Visual Analysis

Exploratory Data Analysis (EDA) and Hypothesis Testing

- Overview of the Machine Learning methodology
- Exploratory Data Analysis (EDA)
- Introduction to Feature Engineering
- Statistical Inference, Probability Distributions
- Hypothesis Testing

Text Mining in Python

- Basic Natural Language Processing
- Working with NLTK
- Text Preprocessing
- Text Cleaning and regular expression
- Regex Introduction
- Regex codes
- Text extraction with Python Regex

- Stop Word Removal
- Stemming
- Lemmatization
- POS Tagging

MACHINE LEARNING INTRODUCTION

- ML core concepts
- Unsupervised and Supervised Learning
- Clustering, Classification, and Regression
- Supervised Vs Unsupervised

Supervised Learning

- Introduction to Linear Regression
- Regression and Best Fit Line
- Modeling and Evaluation in Python
- Introduction to Logistic Regression
- Classification & Sigmoid Curve Modeling and Evaluation
- Introduction to SVM
- Modeling and Evaluation of SVM in Python

Unsupervised Machine Learning

- Understanding Clustering (Unsupervised)
- K Means Algorithm
- K Means theory
- Modeling in Python

ML Web App development Streamlit

- Introduction to Flask
- URL and App routing
- Streamlit application – ML Model Deployment

Projects

- Exploratory Data Analysis (EDA) and Hypothesis Testing
- Regression: Predict Employee Salary using regression
- Text classification
- Topic Modeling or Customer Segmentation