Introduction to Python

Overview of Python- Starting with Python

Introduction to installation of Python

Introduction to Python Editors & IDE's(Canopy, pycharm, Jupyter, Rodeo, Ipython etc…)

Understand Jupyter notebook & Customize Settings

Concept of Modules/Libraries - Important packages(NumPy, SciPy, scikit-learn, Pandas,

Matplotlib, etc)

Installing & loading Packages & Name Spaces

Data Types & Data objects/structures (strings, Tuples, Lists, Dictionaries)

List and Dictionary Comprehensions

Variable & Value Labels – Date & Time Values

Basic Operations - Mathematical - string - date

Reading and writing data

Control flow & conditional statements

Errors and exception handling

Accessing/Importing and Exporting Data using python modules

Importing Data from various sources (Csv, txt, excel, access etc)

Database Input (Connecting to database)

Viewing Data objects - subsetting, methods

Manipulating data

Combining data

Exporting Data to various formats

Important python modules: Pandas

Data Manipulation – cleansing – Munging using Python modules

Cleansing Data with Python

Data Manipulation steps(Sorting, filtering, duplicates, merging, appending, subsetting, derived variables, sampling, Data type conversions, renaming, formatting etc)

Data manipulation tools(Operators, Functions, Packages, control structures, Loops, arrays etc)

Python Built-in Functions (Text, numeric, date, utility functions)

Python User Defined Functions

Normalizing data

Formatting data

Important Python modules for data manipulation (Pandas, Numpy, math, string, datetime etc)

Implementation of stats methods and Visualization using Python

Basic Statistics - Measures of Central Tendencies and Variance

Inferential Statistics -Sampling - Concept of Hypothesis Testing

Exploratory data analysis

Descriptive statistics, Frequency Tables and summarization

Creating Graphs- Simple plotting/Bar/pie/line chart/histogram/ boxplot/ scatter etc)

Important Packages for Exploratory Analysis(NumPy Arrays, Matplotlib, Pandas and scipy.stats etc)

Important modules for statistical methods: Numpy, Scipy, Pandas

Machine Learning

Introduction to Machine Learning

Origin and the history of machine learning

Differences between AI and machine learning

Differences between data science, statistics, data mining and machine learning

Applications of machine learning

Limitations of machine learning

Machine learning is the future

Machine Learning in Python

Machine Learning Process

Collecting data

Pre-processing and preparing data

Exploring data

Choosing a model

Training the model

Evaluating the model

Improving the performance of model

Machine Learning Theories and Algorithms

Meaning of algorithm

Importance of algorithms in machine learning

Types of Machine Learning Algorithms

Supervised learning

Unsupervised learning

Semi-supervised learning

Reinforcement learning

Supervised Learning Tasks and Algorithms

Classification

Nearest neighbor (non-parametric /instance-based)

Decision trees (non-metric /symbolic)

Naive bayes theorem **(**parametric /probabilistic)

Numeric Prediction

Linear Regression

Multiple Regression

Logistic Regression

Time Series Forecasting

Unsupervised Learning Tasks and Algorithms

Pattern detection

Association rules (rule based learning)

Apriori Algorithm

Clustering

K Means Clustering

Hierarchical clustering

Black Box Method

Support Vector Machines(SVM)

Ensemble Methods

Random Forest

Bagging

Boosting

Data Preprocessing in Python

Standardization and Normalization

Missing value replacement

Resampling

Discretization

Feature Selection

Dimensionality Reduction