Call Us: +91-9885561042
Email: info@skillvidya.com

Course Content

Data Science with Python

Introduction to Data Science

- What is Analytics and Data Science?
- Overview of Data Science and Analytics
- Why Analytics is is becoming popular now?
- Application of Analytics in business
- Analytics Vs Data warehousing and MIS Reporting
- Various Terminology in Analytics
- Various Analytics Methodology
- How businesses are using the power of Analytics?
- Various Analytics tools and their usage

Introduction to Data Analytics with Python

- Installing Python Anaconda Distribution
- Python native data types
- Basic programming concepts
- Python data science packages overview

Python Basics: Basic Syntax, Data Structures

- Python Objects
- Math and Comparison Operators
- Conditional Statement
- Loops
- Functions
- Exception handling

Python Concepts (Core)

- Overview and history of python
- Python Installation
- Introduction to python Editors and IDE'S (Canopy, Pycharm, Jupyter, Rodeo etc.)
- Understand Jupyter notebook and customize settings
- Concept of Packages/Libraries – Important packages (NumPy, SciPy, Scikit-learn, Pandas, Matplotlib etc.)
- Installing and loading Packages and Name Spaces
- Data Types and Data objects/structures (strings, tuples, lists, dictionaries)
- List and Dictionary Comprehensions
- Variables and Value Labels - Date and Time Values
- Basic Operations – Mathematical – String – Date
- Reading and writing data
- Simple plotting
- Control flow and conditional statements
- Debugging and Code profiling
- How to create class and modules and how to call them?

Numpy Package

- What is Numpy?
- Importing Numpy
- Numpy overview
- Numpy Array creation and basic operation
- Numpy universal function
- Selecting and retrieving data
- Data slicing
- Iterating Numpy Data
- Shape Manipulation
- Stacking and Splitting Arrays
- Copies and Views: no copy, shallow copy, deep copy
- Indexing : Arrays of indices, Boolean Arrays

Introduction to Pandas

- Selecting data from Pandas DataFrame
- Slicing and dicing using Pandas
- GroupBy/Aggregate
- Strings with Pandas
- Cleaning up messy data with Pandas
- Dropping Entries
- Selecting Entries

Data Manipulation using Pandas

- Data Alignment
- Sorting and Ranking
- Summary Statistics
- Missing Values
- Merging data
- Concatenation
- Combining DataFrames
- Pivot
- Duplicates
- Binning

Pandas Package

- Importing Pandas
- Pandas overview
- Object creation : Series Object , DataFrame Object
- View Data
- Selecting data by Label and Position
- Data Slicing
- Boolean Indexing
- Setting Data

Python Advance: Data Mugging with Pandas

- Applying functions to data
- Histogramming
- String methods
- Merge Data: Concat, Join and Append
- Grouping and Aggregation
- Reshaping
- Analysing Data for missing values
- Filling missing values: fill with constant, forward filling, mean
- Removing Duplicates
- Transforming Data

Python Advance: Visualization with Mat Plot Lib

- Anatomy of a MatplotLib Plot
- Matplotlib basic plots and it's containers
- A Matplotlib figure, it's components and properties
- Axes and other graphical objects

Python Advance: Visualization with Mat Plot Lib

- Pylab and Pyplot
- Data for Matplotlib Plots
- What is a Subplot?
- Modifying size of figures
- Plotting routines with pyplot
- Customizing your pyplot
- Deleting an Axes
- Setting up Plot Title, Axes Labels, Legend, Layout
- Showing, Saving and Closing your Plot
- Save a Plot to an image file and pdf file
- Use cla(), clf() or close.

Exploratory Analysis

- Measures of Central Tendencies
- Mean, Median and Mode
- Data Variability: Range, Quartiles, Standard Deviation
- Calculating Standard Deviation
- Z-Score/Standard Score
- Empirical Rule
- Calculating Percentiles
- Outliers

Distributions

- Distribution Introduction
- Normal Distribution
- Cental Limit Theorem

- Histogram - Normalization
- Other Distribution: Poisson, Binominal etc.
- Normality Testing
- Skewness
- Kurtosis
- Measure of Distance

Hypothesis and Computational Techniques

- Hypothesis Testing
- Null Hypothesis, P-Value
- Need for Hypothesis Testing in Business
- Two tailed, Left tailed and Right tailed test
- Hypothesis Testing Outcomes: Type 1 & 2 errors
- Parametric vs Non-Parametric Testing
- Parametric Tests, T-Tests : One sample, two sample , paired
- One way ANOVA
- Importance of Parametric Tests
- Non Parametric Tests: Chi-Square, Mann-Whitney, Kruskal-Wallis etc
- Which Test to Choose?
- Ascerting accuracy of data

Correlation and Regression

- Introduction to Regression
- Types of Regression
- Hands on of Regression with Python
- Correlation
- Weak and Strong Correlation
- Finding Correlation with Python

Predictive Modeling Concepts – Introduction

- How to use Predictive Modeling in Python
- Linear Regression
- Logistic Regression
- Model Selection
- Scoring

Predictive Modeling Concepts – Introduction

- Predictive Modeling Techniques
- Different phases of Predictive Modeling
- Business Case Study (Predictive Modeling)

Introduction to Machine Learning

- What is Machine Learning?
- Data Science Vs Machine Learning
- Fundamentals of Machine Learning
- Converting business problems to data problems
- Understand supervised and unsupervised learning with examples
- Understanding biases associated with any Machine Learning algorithm
- Drivers of Machine Learning algorithms
- Cost Functions
- Brief introduction to gradient descent
- Importance of model validation
- Overview of cross validation
- Model performance metrics

Unsupervised Learning

- K-Means Clustering: Theory, Euclidean method
- K-Means hands on with python
- K-Means Advantages and Disadvantages

Supervised Learning

- Simple Linear Regression : Implementing in Python , Working on use case
- Multiple Linear Regression : Implementing in python, Working on use case
- K-Nearest Neighbour : Implementing in Python, KNN advantages, Working on use case.
- Decision Trees: Implementing in python, Decision Tree Pros and Cons,
- Working on use case

Supervised Learning

- Tuning with Hyper parameters
- Popular ML algorithms
- Clustering, Classification and regression

Supervised Learning

- Supervised vs unsupervised
- Choice of ML algorithm
- Grid Search vs Random search cross validation

Principle Component Analysis (PCA)

- Key concepts of dimensionality reduction
- PCA theory
- Hands on coding
- Case study on PCA

Random Forest – Ensemble

- Key concepts of Random Forest
- Hands on coding
- Pros and Cons
- Case study on Random Forest

Support Vector Machine (SVN)

- Key concepts of Support Vector Machine
- Hands on coding
- Pros and Cons
- Case study on SVM

Deep Learning Concepts

- Neural Networks - understand basic of Deep Learning
- Back Propagation in Deep learning -understanding fundamentals and implementation details.
- Implement Neural Networks to classify images -Deep Learning approach
- Introduction to Tensor Flow- Deep learning for image classification
- Optimizing Algorithms
- Foundation of Convolutional Neural Network

- Natural Language Processing

SQL Concepts

Getting Started and Selecting & Retrieving Data with SQL.

- Filtering, Sorting, and Calculating Data with SQL
- Subqueries and Joins in SQL
- Modifying and Analyzing Data with SQL

POWER BI

- Introduction to Power BI -Introducing Power BI
- -Getting Acquainted with Power BI
- -Power BI installation -quick insights to explore a Dataset
- -viewing reports -Exploring a dashboard
- 2) Getting Data
- 3)Transforming Data
- -Transform Data
- -Rename Queries
- -Combining Queries
- -Fixing Metadata
- -Filtering Rows
- -Eliminating Rows
- -merging queries
- 4)Modelling Data for Analysis
- -Modelling Data
- -Creating Relationships
- -Creating Relationships using Power BI excel.
- -Defining new columns -concatenating column using Power BI

TABLEAU

- Understanding the fundamental concepts
- Learn how to create charts in Tableau
- Learn dimensions and measures
- Learn the different type of joins
- Creating aliases working with the data source.
- Learn about the different filters


Thank You!