

Website: [www.dw-learnwell.com](http://www.dw-learnwell.com)

Contact: +91 8411002339/+91 7709292162

Email: [info@dw-learnwell.com](mailto:info@dw-learnwell.com)

Classroom

| Corporate

| Online

## Data Analytics/Data Scientist Syllabus – Python

---

### What is Python?

- Birth and Rise of Python
- Links for the necessary software
- GUI of Python: IDLE and Statistical Analysis Interfaces
- Python Notebooks
- Anaconda Python Distribution

### Basic Operations in Python

- Expressions: Basic Idea
- Constant Values: Numeric & Strings
- Arithmetic: Operations and BODMAS
- Common Mathematical functions
- Conditions: Equality, Greater Than, Less Than, etc.
- Function Calls: Introduction to Python Functions
- Symbols & Assignment, Declaring Python Variables
- Reserved Keywords
- Naming a Variable: Generally accepted conventions

### Data Types & Data Structures in Python

- Basic data types bool (Boolean), int (Integer/Long), float, complex
- Type conversions: into to float, float to int, etc.
- String definition and manipulation commands
- Lists
- Tuples
- Dictionary
- Sets

Website: [www.dw-learnwell.com](http://www.dw-learnwell.com)

Contact: +91 8411002339/+91 7709292162

Email: [info@dw-learnwell.com](mailto:info@dw-learnwell.com)

Classroom

| Corporate

| Online

## Data Analytics/Data Scientist Syllabus – Python

---

### Control structures and user defined Functions

- For Loop
- While Loop
- Range
- If- else construct
- User defined functions

### Numpy

- Introduction to Numpy: What is numpy
- ndarray
- arange
- list vs numpy
- zeros
- linspace
- random\_sample
- Statistical operations on numpy ( min, max, sum etc.)
- Reshaping
- Indexing
- Slice and dice numpy arrays

### Pandas

- Introduction to Pandas: what is Pandas
- Series
- Slice and Dice pandas Series
- any() and all()
- apply()

Website: [www.dw-learnwell.com](http://www.dw-learnwell.com)

Contact: +91 8411002339/+91 7709292162

Email: [info@dw-learnwell.com](mailto:info@dw-learnwell.com)

Classroom

| Corporate

| Online

## Data Analytics/Data Scientist Syllabus – Python

---

- id() – referencing of python objects
- Deep copy vs Shallow copy
- timeit usage
- Pandas Data Frame
- Slice and Dice Data Frame.
- Creating calculated columns
- Map()
- Drop()
- Reading data from CSV/TSV files

### Data Visualizations

- Matplotlib
- Seaborn
- Scatter plot
- Box Plot
- Stripplot
- Violinplot
- PairPlot

### Data Pre-Processing

- Treating Missing data
- Isnull()
- dropna()
- Imputer object
- Treating categorical variables
- Label Encoder
- Ordinal vs Nominal treatment

Website: [www.dw-learnwell.com](http://www.dw-learnwell.com)

Contact: +91 8411002339/+91 7709292162

Email: [info@dw-learnwell.com](mailto:info@dw-learnwell.com)

Classroom

| Corporate

| Online

## Data Analytics/Data Scientist Syllabus – Python

---

- One Hot Encoding
- Get\_dummies
- Standardization
- Normalization

### Machine Learning Using Python

- Machine Learning: Supervised Vs Unsupervised
- Correlation
- SciKit learn library (sklearn)
- Data Pre processing before Machine Learning
- Sampling
- Training and Testing split
- Linear Regression
- Multiple Linear Regression
- Polynomial Regression
- Measuring goodness of fit
- SSE (Sum of squared Errors)
- Classification models vs Regression models
- Logistic Regression
- Decision Trees
- Support Vector Machines
- KNN
- Naïve Bayes
- Measuring the performance of classifier
- Precision vs Recall
- Confusion Matrix
- True Positives, True Negatives, False Positives, False Negatives.

Website: [www.dw-learnwell.com](http://www.dw-learnwell.com)

Contact: +91 8411002339/+91 7709292162

Email: [info@dw-learnwell.com](mailto:info@dw-learnwell.com)

Classroom

| Corporate

| Online

## Data Analytics/Data Scientist Syllabus – Python

---

- Unsupervised Learning: Dimensionality reduction technique-PCA
- Unsupervised Learning: K-Means Clustering

### Text Mining with Python

- Introduction to Text Mining
- Unstructured data cleaning to remove clutter
- Extracting features from unstructured text
- Understanding the Document Term Matrix
- Keyword search
- Word cloud
- Sentiment Analysis
- Trend analysis using Twitter
- Finding global trends
- Finding Trends based on geographic location
- Finding Trends for a given topic
- Finding product sentiment in market using Text analysis
- Twitter Data Analysis – Case Study.

### Natural Language Processing with Python

- Introduction to nltk
- nltk.download()
- Tokens: sent\_tokenize, word\_tokenize
- Stopwords removal
- Stemming
- Lemmatization
- POS tagging
- PunktSentenceTokenizer

Website: [www.dw-learnwell.com](http://www.dw-learnwell.com)

Contact: +91 8411002339/+91 7709292162

Email: [info@dw-learnwell.com](mailto:info@dw-learnwell.com)

Classroom

|

Corporate

|

Online

## Data Analytics/Data Scientist Syllabus – Python

---

- Chunking
- Named Entity Recognition
- N-Grams

