

### Data Science Course Objectives

- Apply quantitative modeling and data analysis techniques to the solution of real world business problems, communicate findings, and effectively present results using data visualization techniques.
- Apply principles of Data Science to the analysis of business problems.
- Use data mining software to solve real-world problems.
- Employ cutting edge tools and technologies to analyze Big Data.
- Demonstrate use of team work, leadership skills, decision making and organization theory.

### Data Science Course Syllabus (Duration 60 Hours)

#### 1. Introduction to Data Science

1. Introduction To Data Science
2. Real Time Use Cases Of Data Science
3. Who is a Data Scientist??
4. Github Tutorial
5. Skillsets needed for Data Scientist
6. 6 Steps to take in 3 Months for a complete transformation to Data Science from any other domain
7. Machine Learning - Giving Computers The ability to learn from data
8. Supervised vs Unsupervised
9. Deep Learning vs Machine Learning
10. Link to get Free Data to Practice?
11. Some Great self Learning Data Science Resources (Books, Tutorials, Videos, Papers)
12. Software Installation

#### 2. Python Programming

13. Introduction To Python
14. "Hello Python Program" in IDLE
15. Jupyter Notebook Tutorial
16. Spyder Tutorial
17. Introduction to Python
18. Variable, Operators, Data Types
19. If Else, For and While Loops
20. Functions
21. Lambda Expression
22. Filter, Map, Reduce
23. Taking input from keyboard
24. HANDS ON-
25. INTERVIEW QUESTION DISCUSSION

### 3. Python Advanced Topics

#### NumPy

1. Create Arrays
2. Array Item Selection and Indexing
3. Array Mathematics
4. Array Operation
5. HANDS ON

#### Pandas

1. Introduction to Pandas
2. Series
3. Series indexing and Selection
4. Series Operation
5. Introduction to Pandas
6. Data Frames
7. Data Collection from csv,json,html,excel
8. Data Merging,Concatenation,join
9. Group By and Aggregate Function
10. Order By
11. Missing Value Treatment
12. Outlier Detection and Removal
13. Pandas builtin Data Visualisation
14. HANDS ON
15. INTERVIEW QUESTION DISCUSSION

#### 4. Visualisation-matplotlib, seaborn

1. Line Plots
2. Scatter Plots
3. Pair Plots
4. Histograms
5. Heat Maps
6. Bar Plots
7. Count Plots
8. Factor Plots
9. Box Plots
10. Violin Plots
11. Swarm Plots
12. Strip Plots
13. Pandas Builtin Visualisation Library
14. HANDS ON
15. INTERVIEW QUESTION DISCUSSION

#### 6. Statistics

- 7.
1. Descriptive vs Inferential Statistics
2. Mean,Median,Mode,Variance,Std. dev
3. Central Limit Theorem
4. Co-Variance
5. Pearson's Product Moment Correlation
6. R - Square
7. Adjusted R-Square
8. Spearman's. Rank order Coefficient
9. Sample vs Population
10. Standardizing Data(Z-score)
11. Hypothesis Testing
12. Normal Distribution
13. Bias Variance Tradeoff
14. Skewness
15. P Value

16. Z-test vs T-test
17. The F distribution
18. The chi-Square test of Independence
19. Type-1 and Type-2 errors
20. Annova
21. HANDS ON
22. INTERVIEW QUESTION DISCUSSION

## **6 Introduction to Machine Learning**

1. Introduction to Machine Learning
2. Machine Learning Usecases
3. Supervised vs Unsupervised vs Semi-Supervised
4. Machine Learning process Workflow
5. Training a model
6. Validating results
7. Overfitting vs Underfitting
8. Ordinal vs Nominal data
9. Structured vs unstructured vs semi-structured data
10. Intro to scikitLearn
11. HANDS ON

## **7 Supervised**

### **Regression:**

1. Regression Vs Classification
2. Linear regression
3. Multivariate regression
4. Polynomial regression
5. Multi-Collinearity,
6. Auto correlation
7. Heteroscedascity
8. Hands On

### **Classification:**

1. KNN
2. Svm
3. Decision Tree
4. Random Forest
5. Performance tuning of Random Forest
6. Naive Bayse
7. Overfitting Vs Underfitting
8. Hands On

### **8 Validation:**

1. Classification Report
2. Confusion Report
3. ROC
4. RMSE
5. MSE
6. Cross validation
7. Hands On

## **9 Unsupervised (Clustering & PCA:)**

1. Kmeans
2. How to choose number of K in KMeans
3. Hands on
4. PCA
5. Hands on

## 10 Ensemble

1. What is Ensembling
2. Types of Ensembling
3. Bagging
4. Boosting
5. Stacking
6. Random Forest
7. Important Feature Extraction
8. XGBoost
9. HANDS ON

## 11 NLP

1. Tokenizer
2. Stop Word Removal
3. Tf-idf
4. Document similarity
5. Word2vec Model
6. t-SNE visualisation
7. Sentiment Analysis
8. HANDS ON

## 12 Deep Learning

1. Basic of Neural Network
2. Type of NN
3. Cost Function
4. Tensorflow Basics
5. Hands on Simple NN with Tensorflow
6. Image classification using CNN
7. HANDS ON

### Project 1: Sale Prediction

In this project, we will build a predictive model to find out the sales of each product at a particular store.

### Project 2: Predict Taxi Destination

In this project, we will build a predictive framework that is able to infer the final destination of taxi rides based on their (initial) partial trajectories. The output of such a framework will be the final trip's destination employee's attributes change over time.

## Trainer Profile

- 8+ years of Experience in Data Science, currently spearheading the efforts.
- Trained more than 2000+ students on Data Science at Apponix.
- 5-star rating from all Data Science students.
- Well versed in Data Science.
- Excellent training delivery skills with an ability to present information well.
- Demonstrable experience of being student focused and completing projects to hit deadlines and targets.
- Demonstrable proof of enthusiasm, initiative, creativity and problem solving.
- Demonstrable experience in delivering quality training on Data Science.
- Excellent practical experience.

## Data Science Instructor Experience

- Working directly with Google Cloud Platform .
- Support customers related to Google Cloud Platform issues.