Enter as Trainees....
Exit as Professionals

Data Science with Python(10-12 Weeks)

1. Python for Data Science
2. Introduction to Statistics
    - Types of Statistics
    - Analytics Methodology and Problem Solving Framework
    - Populations and samples
    - Parameter and Statistics
    - Uses of variable: Dependent and Independent variable
    - Types of Variable: Continuous and categorical variable
3. Descriptive Statistics
4. Picturing your Data
    - Histogram

    Normal Distribution
    - Skewness, Kurtosis
    - Outlier detection
5. Inferential Statistics
6. Hypothesis Testing
7. Analysis of variance (ANOVA)
    - Two sample t-Test
    - F-test
    - One-way ANOVA
    - ANOVA hypothesis
    - ANOVA Model
    - Two way ANOVA
8. Regression
    - Exploratory data analysis
    - Hypothesis testing for correlation
    - Outliers, Types of Relationship, Scatter plot
    - Missing Value Imputation
    - Simple Linear Regression Model
    - Multiple Regression
    - Model Building and Evaluation
9. Model post fitting for Inference
    - Examining Residuals
    - Regression Assumptions
    - Identifying Influential Observations
    - Detecting Collinearity
10. Categorical Data Analysis
    - Describing categorical Data
    - One way frequency tables
    - Association
    - Cross Tabulation Tables
    - Test of Association
    - Logistic Regression
    - Model Building
    - Multiple Logistic Regression and Interpretation
11. Model Building and scoring for Prediction
    - Introduction to predictive modeling
    - Building predictive model

- Scoring Predictive Model
- Introduction to Machine Learning and Analytics

12. Introduction to Machine Learning
    - What is Machine Learning?
    - Fundamental of Machine Learning
    - Key Concepts and an example of ML
    - Supervised Learning
    - Unsupervised Learning
13. Linear Regression with one variable
    - Model Representation
    - Cost Function
    - Parameter Learning
    - Gradient Descent
14. Linear Regression with Multiple Variable
    - Computing parameter analytically
    - Ridge, Lasso, Polynomial Regression
15. Logistic Regression
    - Classification
    - Hypothesis Testing
    - Decision Boundary
    - Cost Function and Optimization
16. Multiclass Classification
17. Regularization
    - Overfitting, Under fitting
18. K-Nearest Neighbor – Classification and Regression
19. Support Vector Machine
20. Introduction to Naïve Bayes, Random Forest
21. Model Evaluation and Selection
    - Confusion Matrix
    - Precision-recall and ROC curve
    - Regression Evaluation
22. Unsupervised Learning
    - Clustering
    - K-mean Algorithm
23. Dimensionality Reduction
    - Principal Component Analysis and applications
24. Introduction to text analytics
25. Introduction to Neural Network
27. Problems with Data Analytics Project