COURSE CONTENT

TECHINFOPLACE
A radical new idea

1. INTRODUCTION OF DATA SCIENCE

- Welcome/General Discussion about the expectation from course
- Definition of Data
- Difference between data management and data analytics
- Data Science components

2. PROGRAMMING USING PYTHON

Q&A/PROJECT DISCUSSION BASED ON PREVIOUS DAYS

- Python Overview
- Python Data Types
- Python operations using Numbers, String, Logical, Arithmetic and so on
- Python Strings
- Python Lists
- Python Tuple
- Python Dictionary
- FOR and WHILE loops
- IF/THEN/ELSE in Python
- Data Manipulation using Numpy and Pandas

+91-9714462409
http://www.techinfoplace.com/datascience/

COURSE CONTENT

TECHINFOPLACE
A radical new idea

3. STATISTICS USING PYTHON

Q&A/PROJECT DISCUSSION BASED ON PREVIOUS DAYS

- Levels of Measurement and Variable types
- Descriptive Statistics and Picturing Distributions
- Confidence Interval for the Mean

4. HYPOTHESIS TESTING AND ANOVA USING PYTHON

Q&A/PROJECT DISCUSSION BASED ON PREVIOUS DAYS

- One-Sample T-Test of Comparing Means
- Two-Sample T-Test of Comparing Means
- One Way ANOVA
- Assumptions of ANOVA Modeling
- N-Way ANOVA
- ANOVA Post Hoc Studies

5. EXPLORATORY DATA ANALYSIS USING PYTHON

Q&A/PROJECT DISCUSSION BASED ON PREVIOUS DAYS

- Data Exploration by using Scatter Plots
- Pearson and Spearmen Correlations

COURSE CONTENT

TECHINFOPLACE
A radical new idea

6. LINEAR REGRESSION USING PYTHON

Q&A/PROJECT DISCUSSION BASED ON PREVIOUS DAYS

- Fit Simple Linear Regression Model
- Assumptions of Linear Regression Model
- Analyze the output of the Linear Regression
- Producing Predicted Values
- Difference between Simple Linear Regression and Multiple Linear Regression Models
- Fit Multiple Linear Regression Model
- Stepwise Regression/Model Selection Techniques

7. REGRESSION DIAGNOSTICS USING PYTHON

Q&A/PROJECT DISCUSSION BASED ON PREVIOUS DAYS

- Residual Analysis
- Influential Observation
- Difference between Influential Observation and Outliers
- Collinearity Diagnostics

MODEL BUILDING PROCESS USING PYTHON

8. CATEGORICAL DATA ANALYSIS USING PYTHON

Q&A/PROJECT DISCUSSION BASED ON PREVIOUS DAYS

- Examining Distributions
- Test of Associations by using the chi-square test
- Fisher's Exact p-values for Pearson Chi-square test

COURSE CONTENT

TECHINFOPLACE
A radical new idea

9. LOGISTIC REGRESSION USING PYTHON

Q&A/PROJECT DISCUSSION BASED ON PREVIOUS DAYS

- Odds and Odds Ratio
- Simple Logistic Regression
- Multiple Logistic Regression with categorical predictors
- Analyze the output of Logistic Regression

10. MEASURE MODEL PERFORMANCE USING PYTHON

Q&A/PROJECT DISCUSSION BASED ON PREVIOUS DAYS

- Apply the principles of honest assessment to model performance measurement
- Rare event adjustments
- Assess classifier performance using the confusion matrix
- Model selection and validation using training and validation data
- Create and interpret graphs (ROC, lift, and gains charts) for model comparison and selection
- Establish effective decision cut-off values for scoring

11. DECISION TREE MODELING USING PYTHON

Q&A/PROJECT DISCUSSION BASED ON PREVIOUS DAYS

- Introduction to Decision Tree Modeling
- Model essential for Decision Tree Models
- Decision Tree Model Development by using CHAID, Entropy/Information Gain, and Gini
- Decision Tree Model Tuning

COURSE CONTENT

TECHINFOPLACE
A radical new idea

12. GRADIENT BOOSTING (XGBOOST) USING PYTHON

Q&A/PROJECT DISCUSSION BASED ON PREVIOUS DAYS

- Introduction to Boosting
- Example of Boosting
- Regression Decision Tree
- Gradient Boosted Trees Regression.

13. TIME SERIES FORECASTING MODELS USING PYTHON

Q&A/PROJECT DISCUSSION BASED ON PREVIOUS DAYS

- Introduction to Time Series Forecasting
- Component Factors affecting Time Series
- Moving Average (MA)
- Exponential Smoothing
- Trend Fitting Models (Linear trend, Quadratic trend, and Exponential trend)
- Autoregressive Integrated Moving Average (ARIMA) Model
- Vector Autoregression (VAR) Model
- Autoregressive Conditional Heteroskedasticity (ARCH) Model
- Generalized Autoregressive Conditional Heteroskedasticity (GARCH) Model
- Long Short-Term Memory (LSTM) Model