

## Data Science using R programming

Day	Topics (2 hours/day)
1	<p><b><u>Introduction to Data Science:</u></b></p> <ul style="list-style-type: none"> <li>• What is data?</li> <li>• Types of data</li> <li>• What is big data?</li> <li>• What is data science and who is data scientist?</li> </ul>
2	<p><b><u>R Programming:</u></b></p> <ul style="list-style-type: none"> <li>• Environmental setup</li> <li>• Operators</li> <li>• Data type</li> <li>• Control flow</li> <li>• Functions</li> </ul>
3	<p><b><u>Exploratory Data Analysis:</u></b></p> <p><b>Measure of central Tendency:</b></p> <ul style="list-style-type: none"> <li>• Mean</li> <li>• Median</li> <li>• Mode</li> </ul> <p><b>Data Visualization:</b></p> <ul style="list-style-type: none"> <li>• Bar Plot</li> <li>• Histogram</li> <li>• Scatter plot</li> <li>• Pie chart</li> </ul>
4	<p><b><u>Probability:</u></b></p> <p><b>Random Variable:</b></p> <ul style="list-style-type: none"> <li>• Discrete</li> <li>• Continuous</li> </ul> <p><b>Sampling funnel:</b></p> <ul style="list-style-type: none"> <li>• Population</li> <li>• Units</li> <li>• Sample size</li> <li>• Sample frame</li> <li>• Sampling bias</li> <li>• Sampling techniques</li> </ul>

**LIVEWIRE**

B2, 21, 3rd Floor, Vivina CHS Super Shopping centre, Opp Andheri Railway Station, Andheri west, Mumbai

02240234833 / +91 9967111302

5	<p><b>Sampling Methods:</b></p> <ul style="list-style-type: none"> <li>• Probability Sampling</li> <li>• Non-Probability Sampling</li> </ul> <p><b>Probability Theory:</b></p> <ul style="list-style-type: none"> <li>• Classical method</li> <li>• Relative frequency method</li> <li>• Subjective method</li> </ul>
6	<p><b>Probability:</b></p> <ul style="list-style-type: none"> <li>• Probability Distribution</li> <li>• Continuous Probability Distribution</li> <li>• Binomial Distribution</li> <li>• Poisson Distribution</li> </ul>
7	<p><b>Understanding data:</b></p> <ul style="list-style-type: none"> <li>• Confident Intervals</li> <li>• Data requirements</li> </ul> <p><b>Test of Significance:</b></p> <ul style="list-style-type: none"> <li>• Hypothesis Test</li> <li>• Null Hypothesis</li> <li>• Alternative</li> <li>• P-value</li> </ul>
8	<p><b>Correlation:</b></p> <ul style="list-style-type: none"> <li>• Coefficient of Correlation</li> </ul> <p><b>T-Test:</b></p> <ul style="list-style-type: none"> <li>• One-Sample T-Test</li> <li>• Independent Sample T-Test</li> <li>• Paired Sample T-Test</li> </ul>
9	<p><b>Analysis of Variance:</b></p> <ul style="list-style-type: none"> <li>• ANOVA</li> </ul> <p><b>Chi-Square Test</b></p>
10&11	<p><b>Regression:</b></p> <ul style="list-style-type: none"> <li>• Regression Analysis</li> <li>• Simple Linear Regression</li> <li>• Multiple Linear Regression</li> </ul>
	<p><b>Regression:</b></p> <ul style="list-style-type: none"> <li>• Non-Linear Regression</li> <li>• Logistic Regression</li> </ul>

12&13	<p><b>Clustering:</b></p> <ul style="list-style-type: none"> <li>• Cluster Analysis</li> <li>• Applications</li> </ul>
	<p><b>Clustering:</b></p> <ul style="list-style-type: none"> <li>• Types of Clustering</li> <li>• Hierarchical Clustering</li> <li>• Agglomerative Clustering</li> <li>• Divisive Hierarchical Clustering</li> </ul>
14	<p><b>Dendrogram:</b></p> <ul style="list-style-type: none"> <li>• Cluster Dendrogram on AGNES vs DIANA</li> </ul> <p><b>K-Means Clustering</b></p>
15	<p><b>Classification:</b></p> <p><b>K Nearest Neighbor:</b></p> <ul style="list-style-type: none"> <li>• Similarity in Distance Measurement</li> <li>• Euclidean Distance</li> </ul> <p><b>Bayesian Classifiers</b></p> <ul style="list-style-type: none"> <li>• Understanding Probability</li> </ul>
16	<p><b>Classification:</b></p> <ul style="list-style-type: none"> <li>• Understanding Joint Probability</li> <li>• Naïve Bayers Classifiers</li> </ul>
17	<p><b>Forecasting: Time series Analysis: Components of Time Series:</b></p> <ul style="list-style-type: none"> <li>• Trend</li> <li>• Cyclical</li> <li>• Seasonal</li> <li>• Irregular</li> </ul> <p><b>Decomposition of Components:</b></p> <ul style="list-style-type: none"> <li>• Additive model</li> <li>• Multiplicative model</li> </ul>
19&20	<p><b>Text mining:</b></p> <ul style="list-style-type: none"> <li>• Importance</li> <li>• DTM &amp; TDM</li> <li>• Word Cloud</li> <li>• Natural Language Processing</li> </ul>

### **What's the course about?**

Data Science lies at the intersection of business and data. It takes business data (such as sales figures or customer relations metrics through statistical and operations analysis and comes out with insights into various business functions.

Data science offers descriptive, predictive, diagnostic and prescriptive action items that make organizations better informed and take better decisions. It helps organizations quickly respond to user needs for availability of data, improves competitiveness and produces a single, unified view of organizational or business information.

### **Why LIVEWIRE is preferred by core Professionals?**

- After completion participant able to perform Data Mining, Text Mining, Forecasting, Time Series Analysis in R programming.
- Practical Training on R Studio on various data sets.
- Anyone with no background of coding or analysis can attend
- ISO 9001:2008 Certified institute.
- Association Partners- NSDC, NIELIT, STED Council, Nasscom, Oracle.
- B2E (Benchmark to Excellence) certified benefited with in depth practical training.
- Industry standard courseware allotted to students.
- Internationally recognized course completion certificate.

### **Who are Learning?**

- Graduates from Management, Engineering, Computer science, Statistics, Economics stream etc to stay updated.
- Working Professionals from Management, Finance, Healthcare, logistic, operation, HR and Sales domain.
- Programming Developers, Analytic Professionals Architects building their product in AI.

#### **LIVEWIRE**

B2, 21, 3rd Floor, Vivina CHS Super Shopping centre, Opp Andheri Railway Station, Andheri west, Mumbai

02240234833 / +91 9967111302