Table 1: Training Plan for R

Module	Section	Sub Section	Type
1	R Introduction	Downloading and Installing R	Instructor Lead training
		What is R	
		R environment	
		Getting Help on a function	
		Viewing Documentation	
		General issues in R	
		Packages Management	
		R VS SAS	
2	Data Inputting in R –PART A	Changing Directories	Instructor Lead training
		Basic Types of Data in R	
		- Vectors, List, Factors, Data Frames, Environment	
		Initializing a data frame	
	Data Inputting in R –PART B	Selecting Data frame cols by position and name	Instructor Lead training
		Creating a vector and vector operations	
		Writing various types of data	
		Matrices & Arrays	
		Reading data files (e.g. csv, .sas7bdat, etc.)	
3	Data Manipulation in R- PART A	Logical Operations	Instructor Lead training
		Relational Operators	
		Sub-setting of data	
	PART B	Appending data to a vector	Instructor Lead training
		Combining multiple vectors	
		Accessing variables	
		Concatenation of data	
		Merging the datasets	
		Sorting the data	
		Handling NAs and Missing values	
	PART C	Data Transformation (Long to wide; wide to long)	Instructor Lead training
		Character Manipulation	
		String and Dates	
		Value Transformation (Numeric to Character vice versa)	
4	Functions and Programming in R- PART A	Flow Control: For Loop	Instructor Lead training
		If condition	
		While and repeat loop	
		Debugging tool in R	
	Functions and Programming in R- PART B	Sapply, apply, tapply functions	Instructor Lead training
		Control Structures	
		Computing basic statistics	

		Summarizing data	
		Cross Tabulation	
5	Basic Statistical Analysis in R	T-Test, ANOVA	Self-Learning
		Introduction to Linear Model	
		Classical testing of hypotheses	
		Re-directing R output	
6	R and Databases	Basics of SQL	Self-Learning
		RODBC and DBI package	
		Performing Queries	
		Advanced Data Handling	
7	Data Visualization in R	Creating basic charts and plots	Self-Learning
		Plotting with basic graphics	
		Plotting with Lattice	
		Basic Statistical Plots(Kaplan –Meier, Time series ACF, PACF)	
8	Statistical Methods	Linear and Multiple Regression	Self-Learning
		Mixed effect models	
		Logistic Regression	
		Hierarchical Regression	
		K-fold Validation	
		Simulation	
		Random number generation	
		Random Sampling	
		Bayesian Modelling	
		Time Series Analysis	
		EM algorithm	
		PCA, FA, K-nn, LDA, QDA, KDA, Clustering	

1. *Relevant packages needed column will be updated while creating the training material
2. There will be a total of 8 training decks for each section
3. Dependency of packages to be checked and documented