**Summer Training 2020**

**Daily Log**

Submitted by:

**Bhartendu Dubey (9917103102)**

Under the supervision of:

**Ms. Ambalika Sarkar**



**Department of CSE/IT**

**Jaypee Institute of Information Technology University, Noida**

**May 2020**

**DATE: 01/05/2020**

* What is Data Science & Data Analytics?
* Exploring the need of Data Analytics.
* Understanding scope of Big Data Analytics.

**DATE: 02/05/2020**

* Understanding the R Environment
* Installation of R
* Installation of R-Studio.

**DATE: 03/05/2020**

* Basic Data Types of R.
* Variables
* Practicing Some simple examples on iris dataset.

**DATE: 04/05/2020**

* Vectors
* Matrices & Arrays
* 3D-Arrays
* List

**DATE: 05/05/2020**

* Data Frames in R.
* Data Sets in R.
* Importing the Data in R-Studio.

**DATE: 06/05/2020**

* Exploring R Packages
* Understanding concept of Session Management & Realization

**DATE: 07/05/2020**

* Dealing with Expressions & Objects.
* Understanding Factors & their role in handling categorical data.

**DATE: 08/05/2020**

* Working of Functions in R
* Logical Comparisons

**DATE: 09/05/2020**

* Conditional Selection (Sub-setting)
* Modifying Objects

**DATE: 10/05/2020**

* Indexing by a vector of positive Integer.
* Indexing by a vector of negative Integer.
* Indexing by a logical vector.
* Indexing by a vector of names.

**DATE: 11/05/2020**

* Loops (if, ifelse, switch, for, while)
* Loop Control (Break & Next)

**DATE: 12/05/2020**

* Functions & Data Frames
* Combining Data Frames (cbind)

**DATE: 13/05/2020**

* Understanding Data Manipulation
* Implementation of Data Manipulation with dplyr

**DATE: 14/05/2020**

|  |
| --- |
|  |

* Understanding Difference of Qualitative and Quantitative Data
* Understanding the need of Qualitative and Quantitative Data

**DATE: 15/05/2020**

* Data Visualization - Continuous One Variable
* Data Visualization - Discrete One Variable

**DATE: 16/05/2020**

* Data Visualisation - Continuous X and Y

**DATE: 16/05/2020**

* Plotting of Histograms on Iris dataset.
* Plotting of Density plots on Iris dataset.
* Plotting of Pie chart on Iris dataset.
* Plotting of Bar chart on Iris dataset.

**DATE: 17/05/2020**

* Plotting of Box Plot on Iris dataset.
* Plotting of Scatter Plot on Iris dataset.
* Plotting of Scatter Plot with jitter.
* Plotting a matrix of Scatter Plots.

**DATE: 18/05/2020**

* Plotting of Heat Maps on Iris dataset.
* Plotting of Level Plot on Iris dataset.
* Plotting of Contours.

**DATE: 19/05/2020**

* Plotting 3D-surface diagrams.
* Plotting parallel coordinates.
* Simple Descriptive Statistics (mean, median, Inter quartile range)
* Understanding Covariance & Correlation.

**DATE: 20/05/2020**

* Understanding Data Mining, it’s need & data warehousing
* R Packages and Functions for Data Mining

**DATE: 21/05/2020**

* Implementing Clustering (K-Means Clustering) in R.