08th to 20th March, 2021

V.P.M's B. N. Bandodkar College of Science(Autonomous), Thane.
Department Of Statistics

Report of Online Certificate Course on
"BASICS OF R SOFTWARE"

A Basics of R Software Course is introduced as a certificate course conducted by Department of Statistics, V.P.M's B.N. Bandodkar College of Science, Thane. Due to Pandemic, this year certificate course is conducted using online mode. In choice based credit system the course is of 45 hrs duration. The course was held during 08th to 20th March, 2021 using **Google Meet Online Platform** in between 1pm to 4.30pm.

Any scientific task without the knowledge of software is difficult to imagine and complete in the current scenario. Different software packages are available to analyze the data. It is felt that there is need to well equip the students, teachers and others who are using Statistics and related topics. R is free software that is capable of handling Mathematical and Statistical manipulations. R is an **open-source software environment for Statistical Computing** that is rapidly becoming the tool of choice for data analysis. We intend to give the basics of R software in this course.

Total 10 participants were registered for this online course, one F.Y.B.Sc. student, one S.Y.B.Sc student and four T.Y.B.Sc students of our college and remaining one student was from SYBSc of K.J Somaiya College, one MSc student from NMIMS, one teaching faculty from Pillai HOC College, Rasayani and one administrative staff from St.Xaviers College, Tamil Nadu. Due to online mode of publicity, the responses were also received from states.

The objective of the course were,

1. To familiarize them with basic regular expressions in R.
2. To impart knowledge about creating charts, plots and vectors.
3. To impart knowledge about basic data analysis procedures.
4. To understand the multiple regression models.
5. To understand the procedure to perform all conventional statistical models and analysis tests.
6. To develop hands-on experience on solving problems related to real-life data.

There are 4 units in the course,

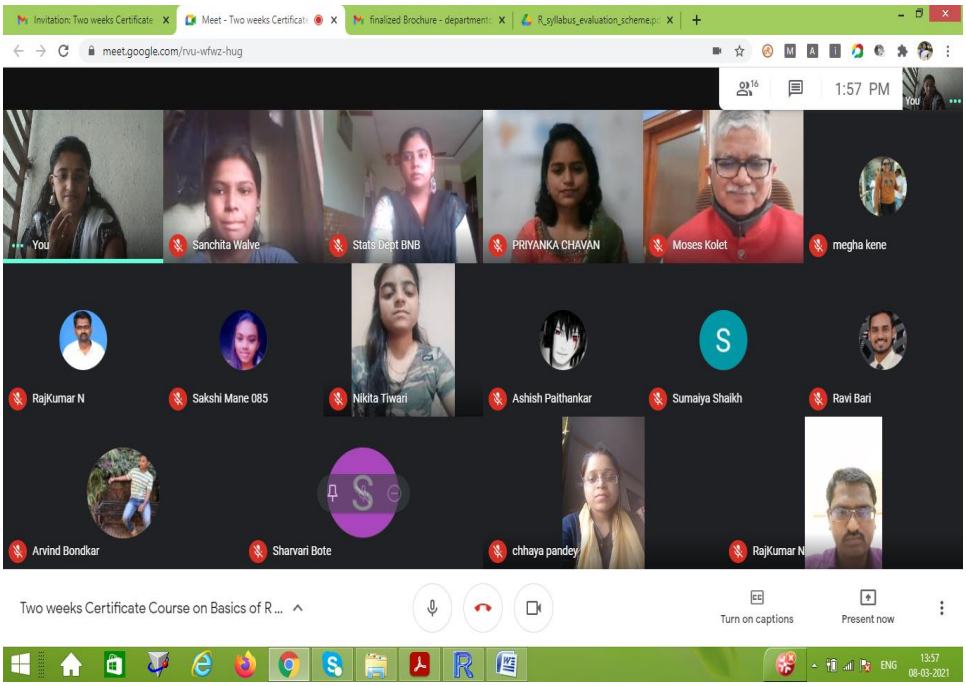
Unit 1: Fundamental of R

Unit 2: Descriptive Statistics

Unit 3: Probability and probability distribution

Unit 4: Correlation analysis, Linear Regression analysis and Curve Fitting.

On 08th March 2021 around 1.45 pm, the inauguration of Online Certificate Course was done by our Incharge Principal Capt Dr. Moses Kolet by giving valuable suggestions and motivational speech.



Unit 1 session was scheduled on 08th March 2021 from 2pm onwards and the Resource Person was **Sumaiya A. Shaikh**. In Unit 1 the following techniques on Fundamentals of R were learned:

- What is R? How to install R and to input R syntax?
- To import and export data from excel files, text files to R.
- Function to be applied in R.
- Data frames which displays the data in matrix form.
- The command to access and analyse with the existing data sets.

Unit 3 session was scheduled on 10th March,2021 and 12th March 2021 and the Resource Person was **Chhaya P. Panday** and **Divya V Nair** respectively. In Unit 3 the following techniques on Probability and Probability Distributions and Testing Of Hypothesis were learned:

- Application of Probability using R syntax.
- The syntax **"choose"** which gives the value of number of combination of n objects taken r at a time(order is unimportant) and the syntax **"factorial"** which gives the value of number of n objects taken r at a time (order is important)
- Commonly used distributions with their R name.
- Functions in connections with different distributions like p , d , q and r.
- Plotting and its interpretation of Discrete and Continuous Distributions.
- In order to take decisions a statistical technique known as hypothesis testing which is useful in research work.

- The concept of **"P value"** to test the hypothesis i.e. either to reject or not to reject the hypothesis.
- Testing hypothesis concerning means, proportion and variance.
- Testing hypothesis on association between two variables.
- To test the significant difference on means of two or more population groups using ANOVA technique.

Unit 2 session was scheduled on 15th March, 2021 and the resource was **Priyanka M Chavan**. In Unit 2 the following techniques on Descriptive Statistics were learned:

- Scale of Measurement.
- Types of variables.
- Absolute and Relative frequencies which creates the sample frequency of the data of the data set.
- Codes to denote categorical variables.
- The use of syntax "**table**" which creates a bivariate distribution table.
- Representation of data by using suitable Graphs and Diagrams.
- The basic idea about averages , median and modal values using R syntax.
- Measures of Central Tendency using R language.
- Measures of Dispersion using R language.
- Measures of Skewness and Kurtosis using R language.
- Dealing with missing observation.
- A box plot which gives graphical technique showing summary of Univariate distribution.

Unit 4 session was scheduled 17th March, 2021 and the resource person was **Priyanka M. Chavan**. In Unit 4 the following techniques of Correlation Analysis and Linear Regression were learned:

- Correlation is concerned with the relationship between the two variables.
- It describes the strength of relationship between two variables.
- Scatter plot to explain how much one variable is affected by another.
- Regression analysis is a statistical process which is used to estimate the relationship among variables.
- Fitting of Regression Analysis in order to get the desired output.
- Application of Multiple Regression model which predicts dependent variable using independent variable.

On 18th March, 2021 assignments on all units of 20 marks were submitted by sending those pdfs to our department Email ID which also cleared their difficulties regarding R syntax.
On 19th March, 2021 a practical examination was conducted of 80 marks from 1.30 pm to 4 pm as part of the course which consists of 40 multiple choice questions of all units and each question with 2 marks using Google Forms Quiz. All students were present and attended and qualified the practical examination.

On 20th March, 2021 at 1.00pm certificates where distributed for successfully completing the course. All students actively participated and provided valuable feedback for the entire course.

After completion of this course participant would be able to

1. Edit and save changes in existing data.
2. Summarise and graphically display data.
3. Perform all conventional statistical analysis tests.
4. Read existing datasets into R or create new ones.
5. Handle any form of data.