Course No.	Course Name	L-T-P Credits	Year of Introduction
RLMCA133	**Applied Statistics Lab**	**0-0-4-1**	**2016**

Course Objectives

- To introduce students to modern statistical tools
- Prepare students for big data analysis course

Syllabus

Companion course of RMCA103, Practical aspects of RMCA103 to be covered in the laboratory Environment.

Expected Outcome

- Students will be able to apply statistical methods to real life problems

References

1. Jared P Lander, "R for everyone", Pearson education, 1st Edition (2014).
2. Dr. Mark Gardener, "Beginning R: The Statistical Programming Language", Wiley (2013)
3. Gnuu PSPP Team, "GNU PSPP Reference Manual", Samurai Media Limited (2015)

Web Resources

1. PSPP
   www.gnu.org/s/**pspp**/manual/**pspp**.pdf
2. Simple R
   http://www.math.csi.cuny.edu/Statistics/R/simpleR/

Suggested MOOCs

1. https://www.edx.org/course/analyzing-visualizing-data-excel-microsoft-dat206x-1
2. https://www.coursera.org/learn/analytics-excel
3. Instructors can also use the simulations material at
   http://wiki.stat.ucla.edu/socr/index.php/SOCR_EduMaterials

Course plan		
Experiment No.	**Description**	**Hours Allotted**
1	Visualizing Data Tables, charts and plots. Visualising Measures of Central Tendency, Variation, and Shape. Box plots, Pareto diagrams. How to find the mean median standard deviation and quantiles of a set of observations. Students may experiment with real as well as artificial data sets.	50
2	Probability Distributions. Set operations, simulation of various properties. Bays' rule. Generate and Visualize Discrete and continuous distributions using the statistical environment. Demonstration of CDF and PDF uniform and normal, binomial Poisson distributions. Students are expected to generate artificial data using the chosen statistical environment and explore various distribution and its properties. Various parameter changes may be studied.	
3	Random samples. How to generate random numbers. Study how to select a random sample with replacement from normal and uniform distribution. Students can use the built in functions to explore random sample selection.	
4	Study of binomial distribution. Plots of density and distribution functions. Normal approximation to the Binomial distribution. Central limit theorem.	
5	Study of confidence intervals. How to compute confidence intervals for the mean when the standard deviation is known.	
6	How to perform tests of hypotheses about the mean when the variance is known. How to compute the p-value. Explore the connection between the critical region, the test statistic, and the p-value.	
7	How to find quantiles of the t-distribution. How to perform a significance test for testing the mean of a population with unknown standard deviation.	
8	Compare populations means from two Normal distributions with unknown variance Tests of Hypotheses for One Proportion Tests of Hypotheses for Comparing Two Proportions	
9	How to calculate the correlation between two variables. How to make scatter plots. Use the scatterplot to investigate the relationship between two variables	
10	Find the least-squares regression line. How to calculate and plot the residuals.	

	Note : This laboratory is to be conducted with a suitable statistical software. The colleges can choose the statistical software. Some of the suggested environments are R, SciPy, SPSS Excel, or any other statistical analysis software depending on availability. ***The students are expected to write code for statistical applications using the chosen environment. The instructor may choose a standard data set and ask the students to work with it.***	