

ROWAN UNIVERSITY
Department of Mathematics

Master Syllabus

STAT 02.311 Introduction to Statistical Computing (3sh)

CATALOG DESCRIPTION:

This is an introductory course in programming-based statistical software packages, such as SAS, R, Matlab, etc intended for students with statistics background. Students will learn the core of ideas of programming such as objects, data structures, looping, and functions. Students will also learn how to read data from different types of files, format them appropriately and use them to perform basic statistical analyses, such as graphing and computing numerical summaries, or more advanced statistical analyses, such as one and two sample T-tests, Chi-square for comparisons of proportions, regression, non-parametric analyses, bootstrapping, and simulations.

OBJECTIVES:

1. By the end of the course, students should be able to write SAS/R/R programs for data management, presentation, and analysis.
1. Run SAS/R programs on PC environment.
2. Read raw input files from various sources and create SAS/R datasets.
3. Compute new variables from variables read-in.
4. Describe data numerically and graphically.
5. Annotate SAS/R output with informative titles, labels, and formats.
6. Work with SAS/R datasets: sort, subset, merge, and re-format SAS/R datasets.
7. Use SAS/R procedures for descriptive statistics: numerical and graphical summaries.
8. Use SAS/R procedures for basic statistical inference: Chi-square tests, T-Tests, ANOVA, Regression, etc.
9. Use SAS/R procedures fit various forms of regression: linear, logistic, and other.
10. Learn the basic of SAS/R functions, macros, and macro variables.
11. Develop skill in the use of SAS/R graphing methods.
12. Export SAS/R data and output to other software.

PREREQUISITES:

STAT 02.260 Statistics I with grade C- or better, or STAT 02.320 Concepts in Statistical Data Analysis with grade C- or better, or STAT 02.284 Statistics for the Biomedical Sciences with grade C- or better, or STAT 02.280 Biometry with grade C- or better, or STAT 02.290 Probability and Statistical Inference for Computing Systems with grade C- or better.

CONTENT:

1. Introduction:
 - Structure and type of data

- Rules and syntax of commands
 - Creating objects
 - Creating, editing and running a program
2. Advanced programming:
 - User defined functions
 - Vectorization
 - Running programs in batch mode
 - Creating tables for Word, the Web, or Latex
 3. Working with data sets:
 - Reading data from files (Plain text, Excel)
 - Cleaning data
 - Restructuring data
 - Manipulating data (creating additional variables, transformation)
 - Merging, sorting and splitting data sets
 - Exporting datasets to files (excel, plain text)
 - Comparing datasets
 4. Basic statistical analyses:
 - Creating basic graphical displays (Pie charts, Bar charts, Histograms, Boxplots, Scatterplots)
 - Embellishing graphs
 - Saving graphs in files (PDF, EPS, JPG)
 - Computing numerical summaries
 - Compute frequency tables.
 5. Advanced statistical analyses:
 - One and Two sample t-tests
 - Non-parametric tests
 - Cross tabulation and Chi-square tests for categorical data
 - ANOVA (one- and multi-way)
 - Linear regression, residual, and qqplots
 - Logistic regression
 - Survival analysis
 - Bootstrapping
 - Simulations
 6. Special topics:
 - Generalized linear (mixed) models
 - Bayesian analyses
 - Advanced graphing (heatmaps, overlaying graphs)
 - Classification techniques (random forests)

SUGGESTED TEXTS:

1. Learning SAS/R in the Computer Lab, 3rd Edition, RJ Elliott, Duxbury Press (2009).
2. Applied Statistics and the SAS/R Programming Language, 5th Edition, RP Cody, JK Smith (2005) R for Data Science, G Grolemund, H Wickham. Free at <http://r4ds.had.co.nz>

Edited: 10/2024