

A Syllabus Form

Ben-Gurion University of the Negev
Department of Industrial Engineering and Management

R- a software platform for data analysis

364-2-1201

Dr. Jonathan Rosenblatt

Course Description:

The R environment is becoming one of the most popular software platforms for data analysis, replacing tools such as SPSS, SAS, and Matlab. In this course we will learn data analysis with the R environment and its ecosystem.

By “data analysis” we mean classical statistical methods such as ANOVA, and regression, but also cutting edge machine-learning methods.

Alongside data analysis we will also cover reproducible science, automating reporting, visual analytics, and “big data” RAM and CPU considerations.

Course Objectives:

- The student will be capable of data analysis with R.
- The student will be capable of programming with R.

Course Structure:

Lecture 3

Course requirements:

1. For IE&M students: 364-1-1061 Regression analysis.

2. For other departments: introductory courses in probability and statistics.
 Linear algebra and programming may give an advantage, but is not required.
 Calculus NOT required.

Structure of Final Course-Grade:

	Component	Weight
1.	Exam	50%.
2.	Project	50%.
	_____	_____
Total:		100%

A "Pass" requirement regarding final exam: No

Lecturer Details: Dr. Jonathan Rosenblatt <johnros@bgu.ac.il>

Description of Meetings

Meeti ng	Subject(s)	Details (as necessary)	Chapters (in course textbook)
1	Introduction to R		1-3
2	Exploratory Data Analysis		4
3	Linear Models		5
4	Generalized Linear Models (GLM)		6
5	Linear Mixed Models (LMM)		7
6	Multivariate Analysis		8
7	Supervised Learning		9
8	Unsupervised Learning		10
9	Plotting		11

10	Automating Reports		12
11	BigData	Sparse matrices	13
14	BigData	Out of memory algorithms	14
13	BigData	Parallel processing	15

References:

<http://www.john-ros.com/Rcourse/>