



**Porsesh Research and Studies Organization**  
**Capacity Building Unit**  
**Course: Introduction to data analysis with R**  
**3-21 JUNE 2017**

---

**Facilitators:** Ramazan Bahrami  
**Time:** 4-June to 21 June (Saturday, Monday and Thursday) 4-5:30 pm – Possibly course will extend  
**Venue:** Porsesh office  
**Contacts:** Ramazan Bahrami | 0793127497 | [bahram.ramazanali@gmail.com](mailto:bahram.ramazanali@gmail.com)

## Introduction

This is 7 module course on “Data Analysis with R”. R is one of the most popular programming languages for data science and statistics, which is extensively used by data analysts and professionals to map marketing trends, develop models, extract useful information and do different kinds of statistical analysis with the data. It is open sourced and supported by a large community of professional users around the world. The program is intended for those curious about data science and R.

Knowledge of some basic mathematics is assumed.

## Expectations

Since we do not assume prior knowledge of R, we will start by introducing R and RStudio. Next, we will cover basic R syntax, data types and structures. Then, we will introduce basic operations and data analysis techniques. We will also demonstrate data visualization capabilities of R. Later, we will cover probability and statistics briefly in theory and implementation in R. Finally, you will work on projects of your choosing to get experience and the needed exposure to R.

Each module comprises of several lessons, readings and labs.

## Logistic Requirements

- You are required to bring your own laptop in which R and RStudio can be installed. Installation files will be provided.
- Fee: 5000 Afs

## Course Topics

### Module 1: Introduction to R basics

- Installation of R and RStudio, introduction to RStudio interface
- Data Structures: variables, vectors, matrices, lists, data frames, data tables, etc.
- Data types: numeric, integer, string, factor, logical
- Importing and exporting data from external sources
- R library and packages
- Perform basic data manipulation: create, name and select elements from vectors
- Learn how to work with matrices and perform basic computations with them
- Learn how to create, subset and compare categorical data.
- Learn how to work with lists
- learn how to work with Data Frames

### Module 2: Descriptive Statistics

- frequency and proportion tables
- Descriptive statistics with R: mean, median, std. dev, covariance, correlation, etc.

### **Module 3: Graphics**

- Graphics in base R: scatter plot, bar plot, line plot, etc.
- Professional graphics using ggplot2: scatter plot, bar plot, line plot, etc.

### **Module 4: Probability**

- Introduction to discrete distributions, continuous distributions, multivariate distributions, sampling distributions, estimation, and hypothesis testing.
- Learn how to calculate each of these probability distributions in R

### **Module 5: Statistics**

- Simple Linear Regression, Multiple Linear Regression, Resampling Methods.
- Learn Statistical methods in R.

### **Module 6: Time-series analysis**

- Introduction to time-series analysis in R

### **Module 7: Final Projects**

- Hands-on experience working with real data