

Pre Conference Workshop 1

R and Data Manipulation: Application with Big Data

In conjunction with the joint conferences iCMS2017 & UROS2017, the organizing committee invites participants to our pre conference workshop.

5 – 6 November 2017

Adya Ballroom, Adya Hotel,
Langkawi, Malaysia

HRDF CLAIMABLE

ABOUT

The R platform provides a powerful and comprehensive platform for visualizing data, understanding and evaluating statistical models, and effectively communicating research results to both technical and nontechnical audiences. R is a programming language and software environment for statistical computing and graphics supported by the R Foundation for Statistical Computing. The R language is widely used among statisticians and data miners for developing statistical software and data analysis. R is ready to use for applications in data science and data analytics also for big data sets. In the workshop, some important issues are touched, whereas the focus is the use of efficient tools in R.

Speaker



Matthias Templ, PhD

- Owner of data analysis company OG.
- Lectures at the Institute of Data Analysis & Process Design, Zurich University of Applied Sciences, Switzerland; An associate professor at Vienna University of Technology; Researcher at Statistics Austria.
- Consultant for UNIDO, the World Bank and the OECD.
- Published 45 papers in scientific journals
- Author of R packages – sdcMicro, simPop, robCompositions & VIM
- Author of books: Statistical Disclosure Control for Microdata: Methods and Application in R & Simulation for Data Science with R
- Editor-in-chief of the Austrian Journal of Statistics; associate editor for Springer's Journal Wirtschafts-und Sozialstatistische Archiv; editor for journal The Register.

Who Should Attend?

Lecturers, Researchers, Engineers, Students, Industry Professionals and Scientists of any discipline who wish to explore R and Big Data. Since this is a continuation to our first R pre-conference workshop in 2015, prior experience with R is required since the instructor will not cover the very basics of R but doing an introductory course on efficient methods in R. Thus, please be prepared for knowing some basics on R.

The following URL may help you get started

- either knowledge in basic R (preferred) or doing the examples on <http://tryr.codeschool.com/> or/and package RSwirl beforehand.
- reading <http://adv-r.had.co.nz/> until (and including) the chapter on functions.

Workshop Objectives

It is expected that after this 2-day workshop, participants will be able to:

- use efficient R code
- perform parallel computing using R
- use R packages to perform data analytics
- work with big data sets using R packages like dplyr, Rcpp, ggplot2, and etc

Program Schedule

DAY 1 SUNDAY (5 th Nov 2017)		DAY 2 MONDAY (6 th Nov 2017)
8.15 – 8.45	Registration	Registration
8.45 – 10.45	Brief R Introduction.	How to make your computations faster. A very basic introduction to profiling, Rcpp, and parallel computing.
10.45 – 11.00	COFFEE BREAK AND DISCUSSION	
11.00 – 13.00	Indexing and vectorized computations and data manipulation with base R. The aim is to use efficient R code and functions.	Some comments and issues when working with big data sets. We touch issues on keeping data in a data base and outline some approaches to deal with big data sets.
13.00 – 14.00	LUNCH AND DISCUSSION	
14.00 – 17.00	Efficient and fast data manipulation with <i>dplyr</i> (and optionally <i>data.table</i>) to speed up your data manipulation and use a consistent approach for data manipulation. Seven useful functions when working with <i>dplyr</i> and the use of pipeline syntax. Application of these functions on example data.	Data visualization using <i>ggplot2</i> to produce nice looking graphics with less code. The main aim is to understand the concept of the grammar of graphics implementation of <i>ggplot2</i> (aesthetics, geometric objects, statistical objects, scales, themes). The differences to the graphics package are shown and participants will be trained to use <i>ggplot2</i> on example data sets.
17.00	TEA BREAK	

Workshop Fee Per Participant

CATEGORY	Early Bird	Normal
Local Participant	RM 980	RM 1500
International Participant	USD 250	USD 380
Local Student	RM 890	RM 1200
International Student	USD 230	USD 320

PAYMENT

Bank Name : Bank Islam Malaysia Bhd (Sungai Petani Branch),
 Bank Address : No.71&72 Lorong Lagenda 1, Lagenda Heights, 08000 Sungai Petani, Kedah, MALAYSIA
 Bank Account Number : 02020010021120
 Bank Account Name : Akaun Amanah UiTM Kedah
 Pay to : Bendahari UiTM Kedah
 Swift Code : BIMBMYKL
 GST : 001418756096

For more information, please visit www.icmskdh.net
 To register please go to:
<http://submit.confbay.com/conf/icms2017ws>