



Elective Courses for PhD students 2018

Data Analysis and Visualization with R	
Offered by:	International Max Planck Research School for Cell, Developmental and Systems Biology (IMPRS-CellDevoSys)
Lecturer:	Ian Henry, Holger Brandl & Lena Hersemann, MPI-CBG Scientific Computing Facility
Date:	23 – 25 April 2018
Time:	9:00 am – 5:00 pm
Location:	CSBD, SR1 (ground floor)
Target audience:	DIPP PhD students
No of participants:	12
Registration deadline:	15 March 2018
Pre-course work:	
Course requirements:	Prepare your computer as described in the INDICO system

COURSE AIM:

Research in the field of biology now often involves the acquisition of large datasets. In order to get the most out of such large and complex datasets it is useful to know how to efficiently structure such data and the methods available to manipulate and visualize large datasets. This course is aimed at those with limited or no programming experience and aims to teach participants the basics of data analysis and visualization using the R statistical programming language.

LEARNING OUTCOMES:

This course will enable students to get started with data analysis using R. Topics covered include, importing data into R, formatting data for easy further analysis, manipulating tabular data and performing basic calculations, creating publication quality figures, and creating complete data analysis reports.

COURSE CONTENT/ METHODS ENCOUNTERED DURING THE COURSE:

- Importing data into R (e.g. from Excel)
- Formatting data for easy further analysis
- Manipulating tabular data and performing basic calculations
- Creating publication quality plots
- Creating complete analysis reports



Elective Courses for PhD students 2018

COURSE STRUCTURE:

The course will include presentations to introduce important topics, live coding sessions where the instructor and students write code together, and exercises for students to work on individually or in small groups. The course is intended to be very hands on with plenty of opportunity for students to ask questions.

TIMETABLE:

Monday

9:00 – 12:00: Introduction
10:00 -12:00: RStudio Fundamentals
13:00 – 15:00: R basics
15:00 – 17:00: Importing Data into R

Tuesday

9:00 – 12:00: Formatting Data for Easy Analysis (tidyr)
13:00 – 17:00: Data Manipulation (dplyr)

Wednesday

9:00 – 12:00: Plotting data (ggplot2)
13:00 – 17:00: Generating Data Analysis reports (RMarkdown)

BACKGROUND READING:

<http://r4ds.had.co.nz/>