CT5102

Programming for Data Analytics @NUI Galway

This github site is a resource for Progamming for Data Analytics (CT5102), a module that is part of the M.Sc. Computer Science (Data Analytics). This is an advanced MSc programme that provides graduates of Computer Science and closely related areas with in-depth knowledge and skills in the emerging growth area of Data Analytics. There are twelve topics, and these will be updated as the course progresses. The course structure has main three elements, as shown below.

- Base R (Vectors, Functions, Lists, Matrices and Data Frame)
- Data Science, with the tidyverse packages in R (ggplot2, dplyr, readr, tidyr, lubridate and stringr)
- Advanced R, including closures, object systems (S3, S4 and RC), and building packages

For tips on good programming style, see Hadley Wickham's web resource

For an excellent resource to discover courses and learning materials for learning and teaching R, see rstudio-education/rstats-rd

Datasets used as part of the course can be viewed here

Details on the individual lectures are:

(1) Base R

- *Lecture 01* - **Atomic Vectors** (Materials & Summary) (Code) (Problem Sheet)
- *Lecture 02* - **Lists and Functions** (Materials & Summary) (Code)
- *Lecture 03* - **Matrices and Functionals** (Materials & Summary) (Code)
- *Lecture 04* - **Data Frames** (Materials & Summary) (Code)

(2) Data Science with R

- *Lecture 05* - **ggplot2** (Materials & Summary) (Code)
- *Lecture 06* - **dplyr** (Materials & Summary) (Code)
- *Lecture 07* - **Relational Data** (Materials & Summary) (Code)
- *Lecture 08* - **stringr** (Materials & Summary) (Code)

(3) Advanced R

- *Lecture 09* - **S3 Classes** (Materials & Summary) (Code)
- *Lecture 10* - **Packages** (Materials & Summary) (Code)
- *Lecture 11* - **Environments, Functions and Closures** (Materials & Summary) (Code)
- *Lecture 12* - **Markdown, RShiny and Course Summary** (Materials & Summary) (Code)