

Data Science in Practice

L. Torgo

ltorgo@fc.up.pt

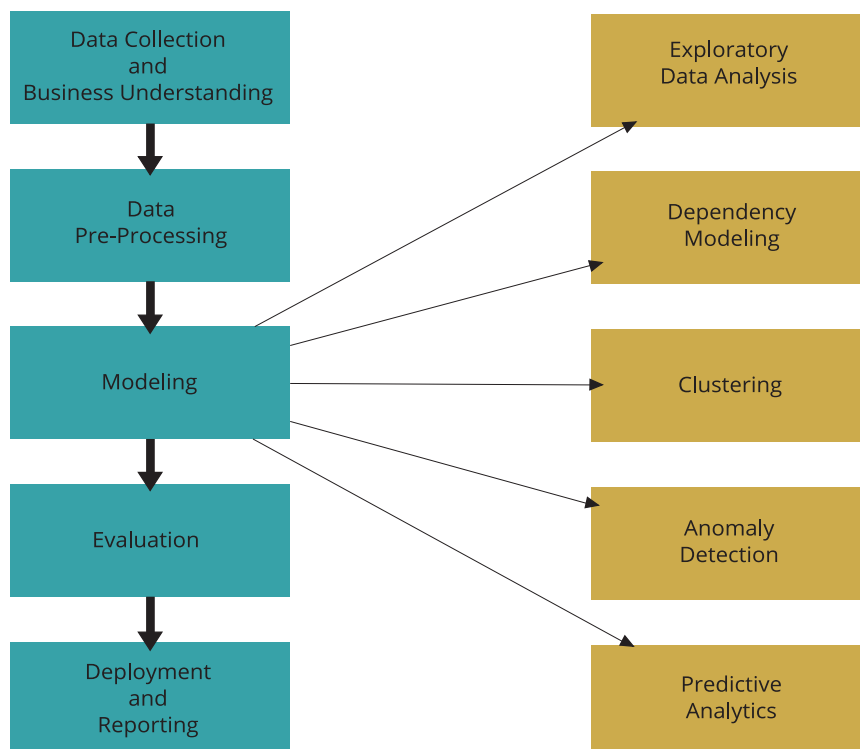
Faculty of Sciences / LIAAD-INESC TEC, LA
University of Porto, Portugal

Aug, 2017



Data Science

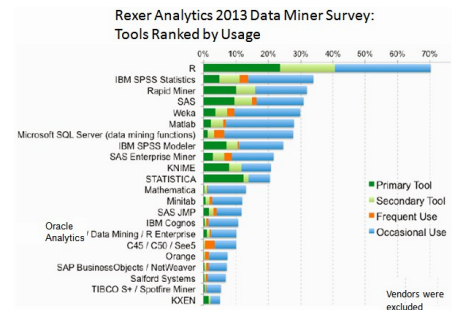
The Typical Data Science Workflow



The R Project (<http://www.R-project.org>)

What is R?

- R is a programming language and environment for data analysis.
- It is a free and open source software tool.
 - Trustable software; reproducible and adaptable analysis
- One of the most used data analysis tools.
- Strong impact on both academia and industry



RStudio

RStudio

An excellent way of using R is through the integrated development environment provided by RStudio which is also freely available for Windows, OS X e Linux. This is probably the most user-friendly way of using R. RStudio can be obtained at <http://www.rstudio.org>

