ML tools

Bank of Uganda

University of Cape Town

Short Course

November 2019

Aims:

Introduce:

     o R (a language)

     o RStudio (the Integrated Development Environment)

     o Version control (a process), and

     o Git (the tool for version control)

UNIVERSITY OF CAPE TOWN
IYUNIVESITHI YASEKAPA · UNIVERSITEIT VAN KAAPSTAD

R

What is it?

- Language and environment for statistical computing and graphics

- Open source (free)

R is an integrated suite of software facilities for data manipulation, calculation and graphical display. It includes

- an effective data handling and storage facility,

- a suite of operators for calculations on arrays, in particular matrices,

- a large, coherent, integrated collection of intermediate tools for data analysis,

- graphical facilities for data analysis and display either on-screen or on hardcopy, and

- a well-developed, simple and effective programming language which includes conditionals, loops, user-defined recursive functions and input and output facilities.

UNIVERSITY OF CAPE TOWN
IYUNIVESITHI YASEKAPA · UNIVERSITEIT VAN KAAPSTAD

Where can you get it?

http://r.adu.org.za/

What is RStudio?

"An integrated development environment for R, with a console, syntax-highlighting editor that supports direct code execution, and tools for plotting, history, debugging and workspace management."

Basically a nice environment to run R in

UNIVERSITY OF CAPE TOWN
IYUNIVESITHI YASEKAPA · UNIVERSITEIT VAN KAAPSTAD

Where can you get it?
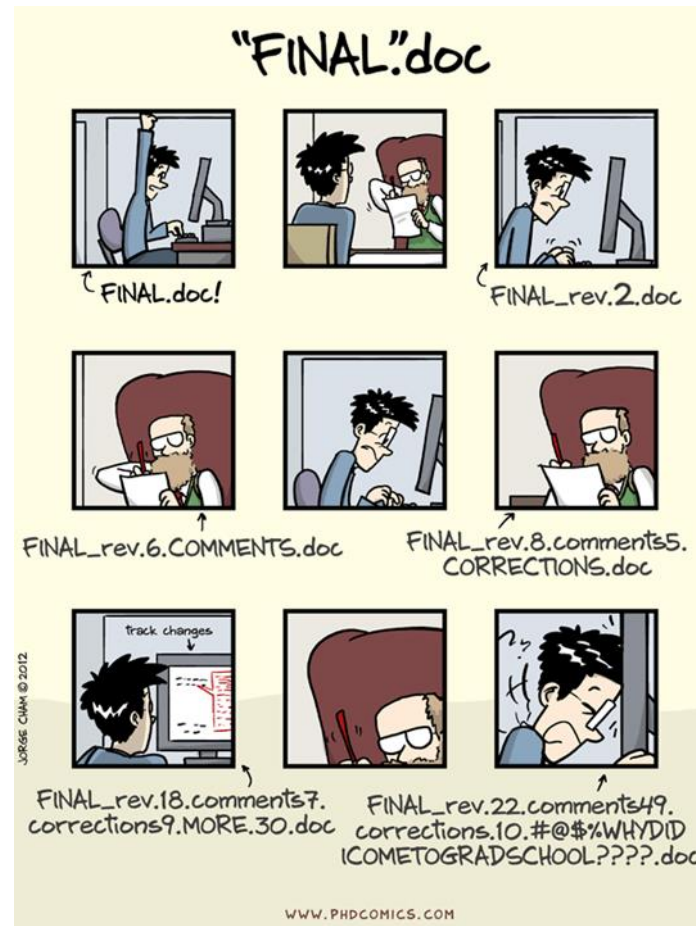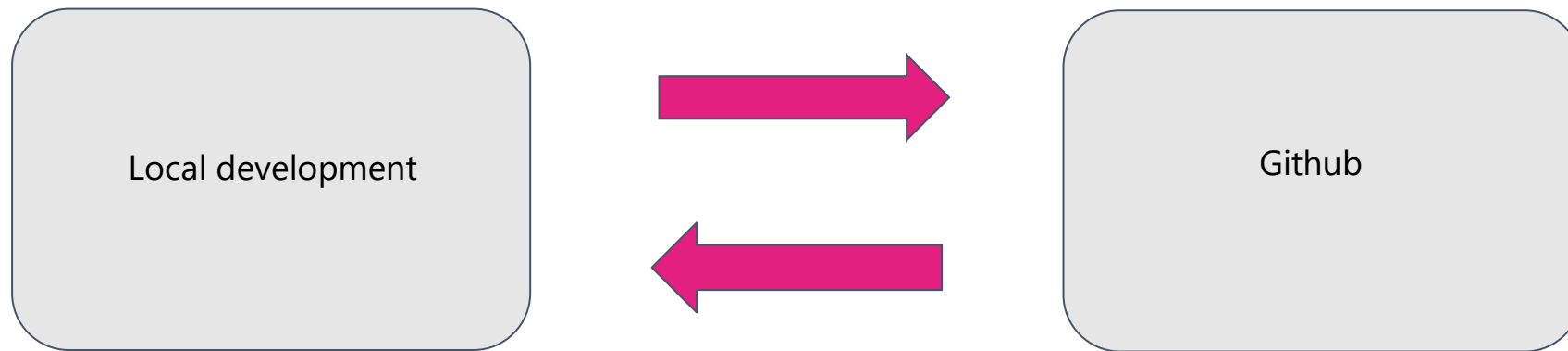
https://rstudio.com/products/rstudio/

UNIVERSITY OF CAPE TOWN
IYUNIVESITHI YASEKAPA · UNIVERSITEIT VAN KAAPSTAD

A quick tour

Why we need version control

The 'version control' process

Local development → Github

In a self-contained 'project'

No 'hard coded' paths

UNIVERSITY OF CAPE TOWN
IYUNIVESITHI YASEKAPA · UNIVERSITEIT VAN KAAPSTAD

What do I need to do?

Read:

- Git with R (https://happygitwithr.com/)

Then do Chapters 4 - 15