Statistics Tutorial

Day 1

Prabesh Dhakal

08 April 2019

WHO AM I?

PRABESH DHAKAL

DATA SCIENCE MASTER

WHAT I DO

- Study in the data science master program at Leuphana
- Work in Henrik's lab and support everyone with their data related challenges: data collection, preparation, analysis, and communication

WHAT I USE

R

Python

WHAT ARE WE DOING TODAY?

INTRODUCE THE COURSE

What topics are we going to cover during the future weeks in this semester?

SETUP TOOLS

Setup the tools required for the module on the devices. Familiarize ourselves with the basic setup of the tools.

ANSWER SOME QUESTIONS

Discuss a few basic things to get a gentle start to the tutorials.

WHAT WILL WE COVER IN THE FUTURE?

Setup Tools

1. Setup the tools required for the module on the devices.
2. Familiarize ourselves with the basic setup of the tools.

INSTALL SOFTWARES

You start using R-Studio by doing the following things:

Step 1: Install R

Step 2: Install R-Studio

Windows Users

Download Links:

tinyurl.com/R-Leuphana-W

tinyurl.com/RStudio-Windows

Mac Users

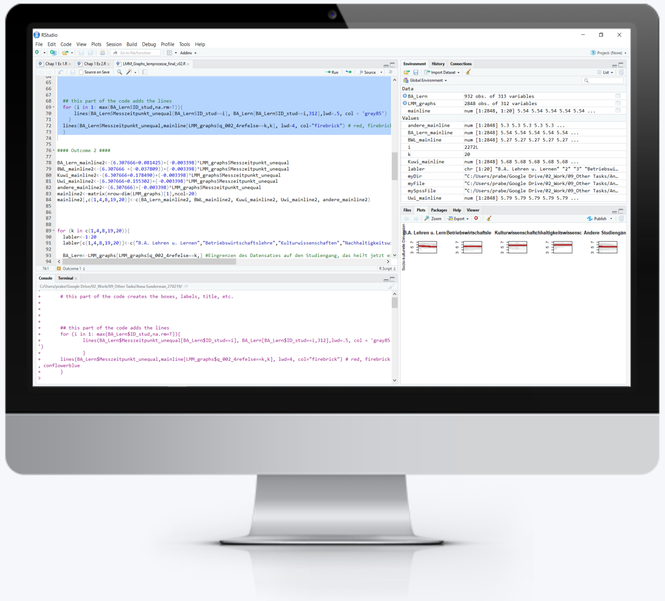
Download Links:

tinyurl.com/R-Leuphana-M

tinyurl.com/RStudio-Mac

Linux Users

You know what to do.



PROGRAMMING PARADIGM: R

WHAT IS R?

Short version: a calculator on steroids

A slightly longer explanation:

R is a programming language **purpose built for statistics.**

It is **easy to learn**, and it works in all desktop environments.

Additionally, it has a great community of users and supporters.

MAIN TOOL WE WILL USE: RStudio®

Prepackaged Tools

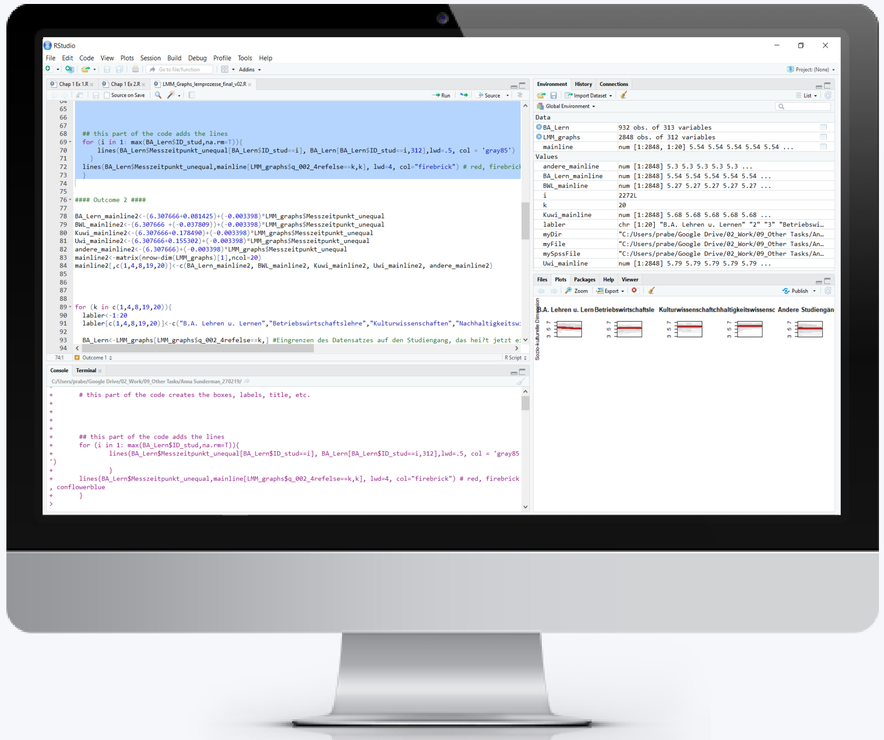
Its graphical interface makes working with R even easier.

Built Specifically for R

It is purpose built to run with R and all the supported libraries.

Free and Open Source

Its graphical interface also makes working with R even easier.



A walkthrough of R
in R-Studio

Tutorial & Discussion

1. Start of the tutorial.
2. General Q&A

STATISTICS: WHAT IS IT?

What is statistics?

Statistics is the science of collecting, preparing, analyzing, and interpreting data.
2 key parts:

Collecting and Preparing

People who work in marketing try to get the attention.

Analyzing and Interpreting

People who work in marketing try to get the attention.

STATISTICS: WHY?

Why should we care about statistics?

Statistics is really one of the essential tools that we can use to make sense of the world around us. It allows us to make discoveries, make decisions, and predict events.

[2 Exercises:]

Factfulness Test

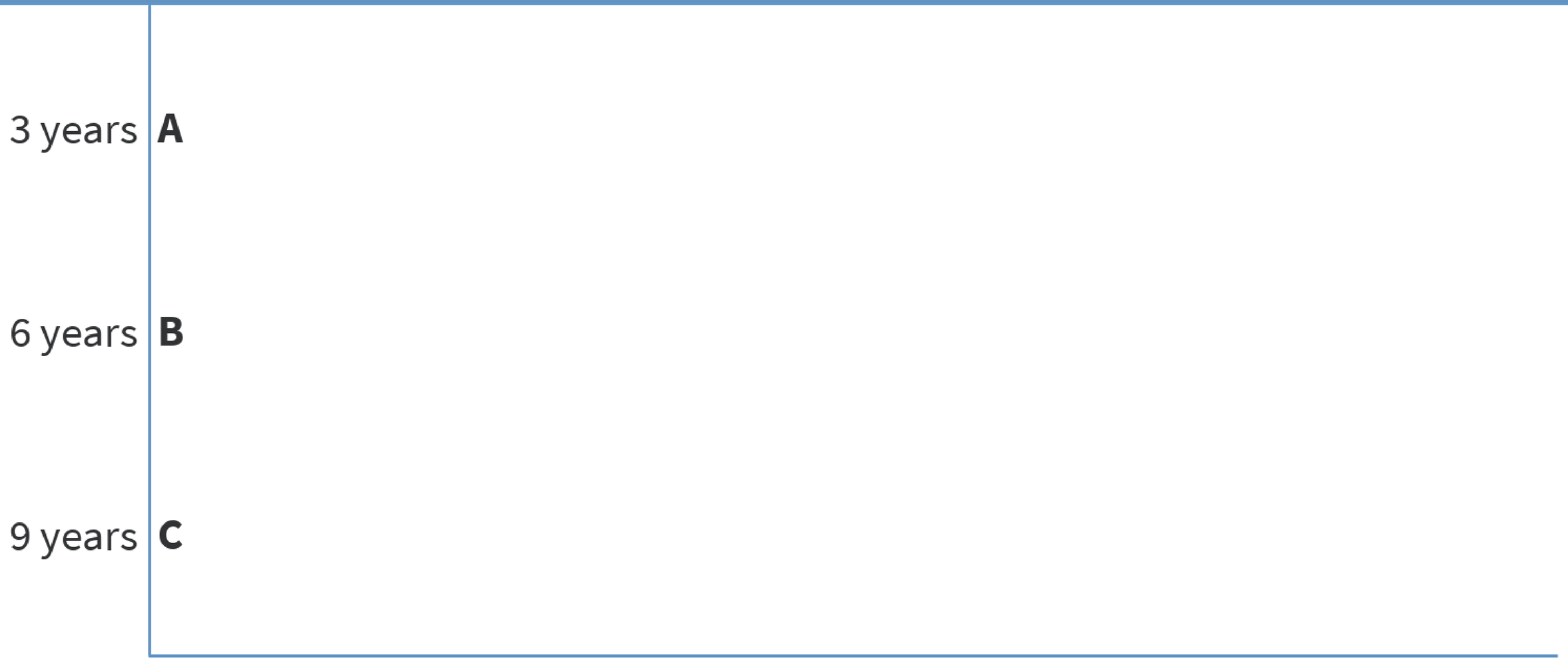
Link to the poll: **tinyurl.com/Pollev-April-08**

[Full test: tinyurl.com/Factfulness-Leuphana]

Video

3 Ways to Spot a Bad Statistic by Mona Chalabi

Worldwide, 30 year old men have spent 10 years in school on average, how many years have women of the same age spent in school?



SUMMING UP

1. This course will complement the lectures and we will apply the concepts discussed there with real data.
2. The tools we will use to work with the data and perform calculations is going to be R-Studio.
3. Statistics is the science of collecting, preparing, analyzing, and interpreting data.
4. Statistics allows us to: (i) make sense of the world around us, (ii) make decisions, and (iii) make predictions.

PLAN FOR NEXT WEEK

Next week, we are going to discuss:

1. Types of data
2. What analysis techniques apply to a given data type

That’s it for today! :-)

If you want to reach me, mail me at:

`prabesh.dhakal@stud.Leuphana.de`