We >eR a community

Development of (local) E-learning platforms in (bio)statistics & data science: The >eR-BioStat platform

Ziv Shkedy
Hasselt University, Belgium

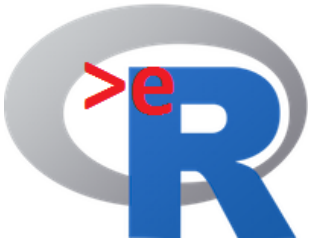
CFAR Symposium on Statistics and Data Science in HIV
June, 5&6, 2023

Visit us on Facebook ER-BioStat

GitHub https://github.com/eR-Biostat

https://erbiostat.wixsite.com/erbiostat

twitter @erbiostat



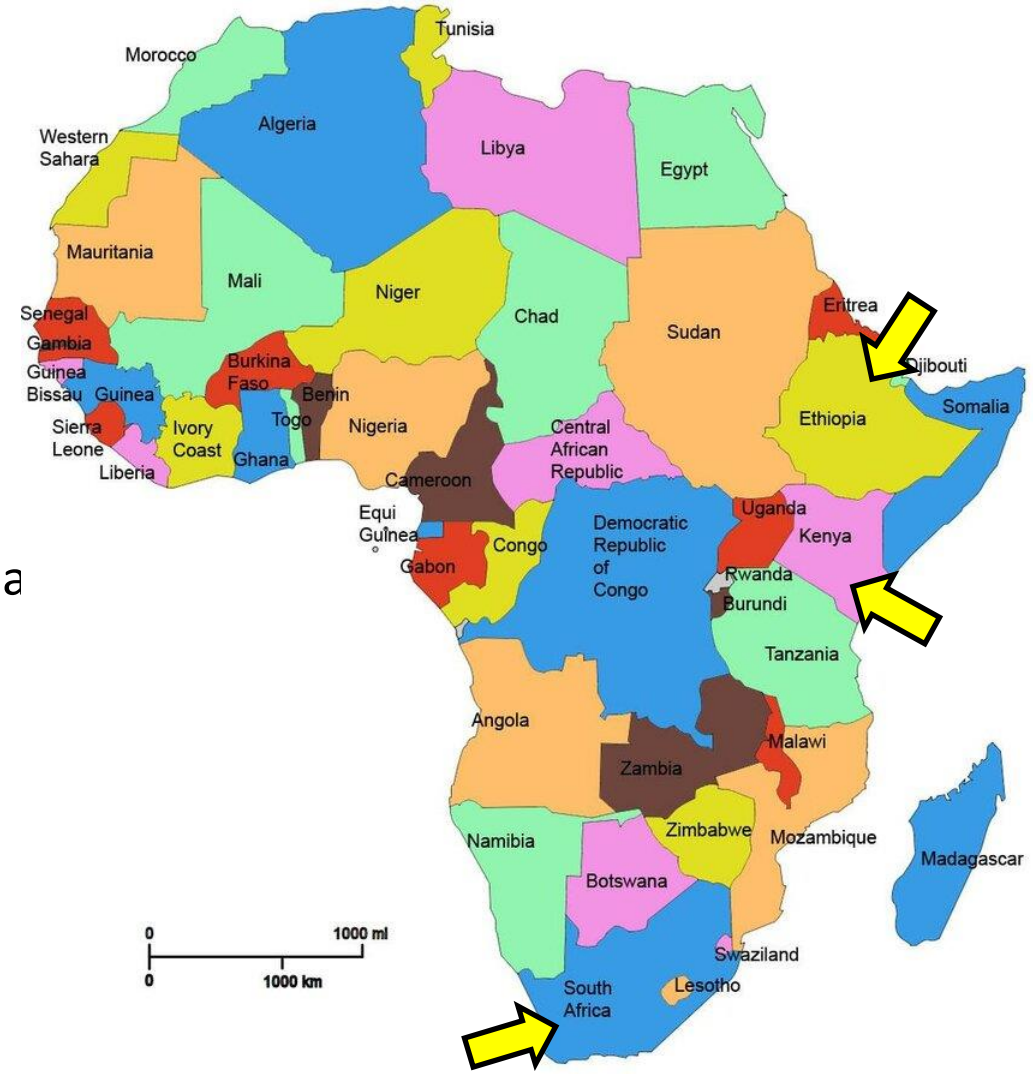
The >eR-BioStat :
short introduction about the ITP project

Our new website

https://erbiostat.wixsite.com/erbiostat

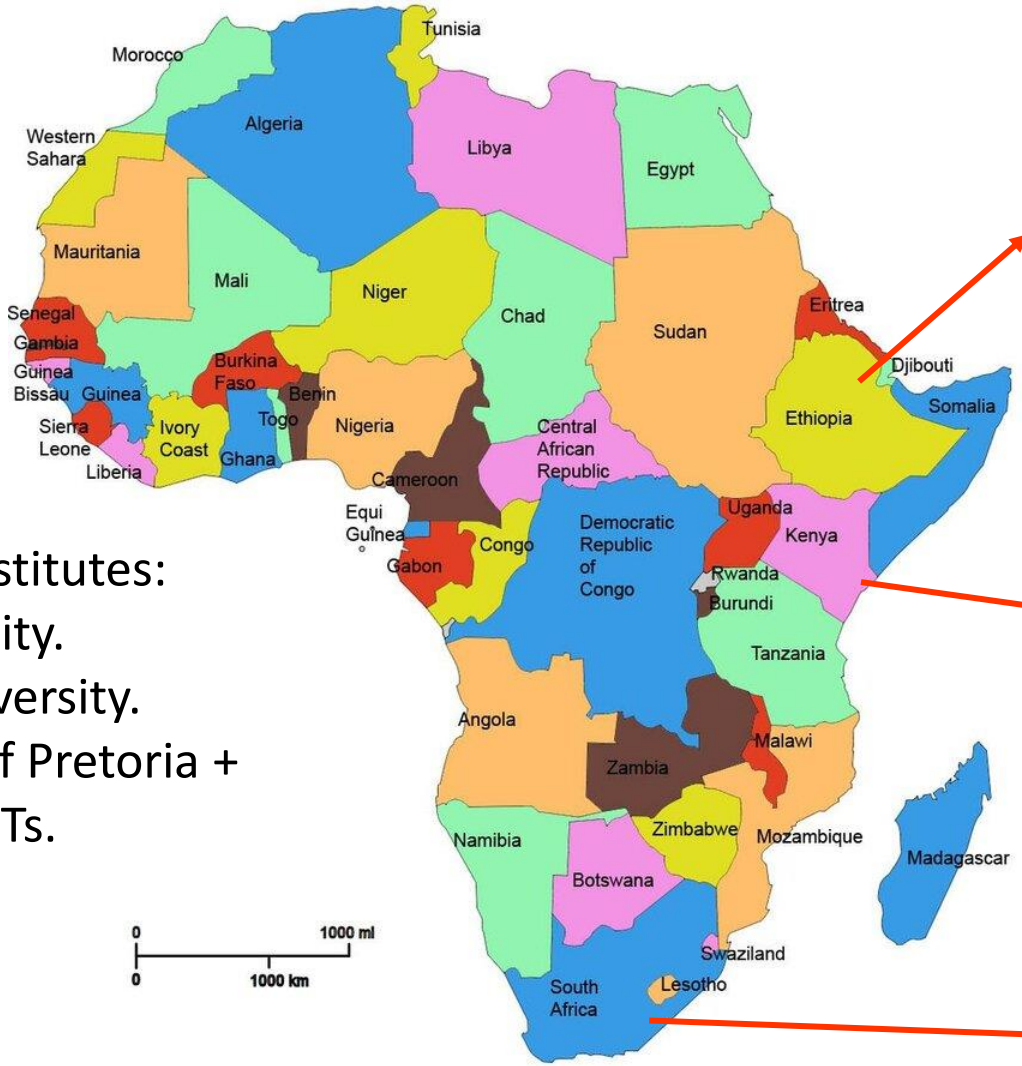
Project's structure

- Three years project.
- Three countries:
  - Kenya.
  - Ethiopia.
  - South Africa.
- In total: 14+ Universities & institutes from the three countries.
- International team: USA, Canada Belgium.



Project's partners

- Three leading institutes:
  - Moi University.
  - Gondar University.
  - University of Pretoria + SA-MRC, WITs.



Gondar University.
Ambo University,
Debre Berhan University
Arba Minch University.
+others….

Moi University.
Karatina University.
JKUAT.
Masinde Muliro Uni.

University of Pretoria.
MRC-SA.
WITS.
Stellenbosch University.
The University of Venda.
The University of Zululand.

Project's focus

- ITP: International Training Program.
- Focus: Development of (local) E-learning and digital education platforms for (bio)statistics & data science.
- All levels of HE:
  - Undergraduate.
  - Graduate.
  - Service courses for non statisticians.

Project's focus

- Development of (local) networks of universities in Kenya, Ethiopia, South Africa that use the digital platform for statistics & DS.
- Establish international network to support local universities.

Deliverable

- A Local version of the >eR-BioStat website.
- Updated global version of https://erbiostat.wixsite.com/erbiostat.
- “Editorial board” for the website.

Activities

- Workshops for training in the local universities:
  - 3 workshops per country (we aim for one per year) = 9 workshops in total.
  - Up to a 5-10 days workshop.
  - ~20 participants (local & international): for the partner institutes + others.
- Workshops will be located in the leading partners.
- Online workshops (~3).

Project duration

- Sponsor: VLIR-UOS.
- Three academic years: 23/24, 24/25 & 25/26.
- First workshop: Gondar (14/08/2023).
- Second workshop: SA-MRC, cape town (end of February 24: workshop + conference).
- Third workshop: Moi.
- …
- …

UHASSELT

KU LEUVEN

Interuniversity Institute for Biostatistics and statistical Bioinformatics

The >eR-BioStat : a very short introduction and main concepts

Our new website

https://erbiostat.wixsite.com/erbiostat

Digital education & E-learning

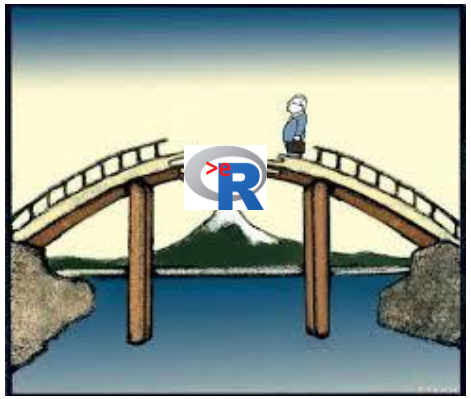
- We DO NOT develop online education program but we develop online materials for on campus education.

- The E-learning Initiative aims to support on campus programs by:
  - Develop accessible course materials in (bio)statistics.
  - Focus on **all education levels**:
    - Undergraduate & master programs, PhD schools.
    - Statisticians & non statisticians.
  - Bring students and teachers costs to minimum by providing **free, high quality and applied** course materials.

We a community

- Build up communities (in south & north).
- Create a bridge between communities:

Academic staff and students in the south.



Development of E- learning capacity

Academic staff in the north.

Concepts

- The E-learning system consists of few components:

1. All course materials are available to the students/teachers online to download.
   - Can be used online or offline.

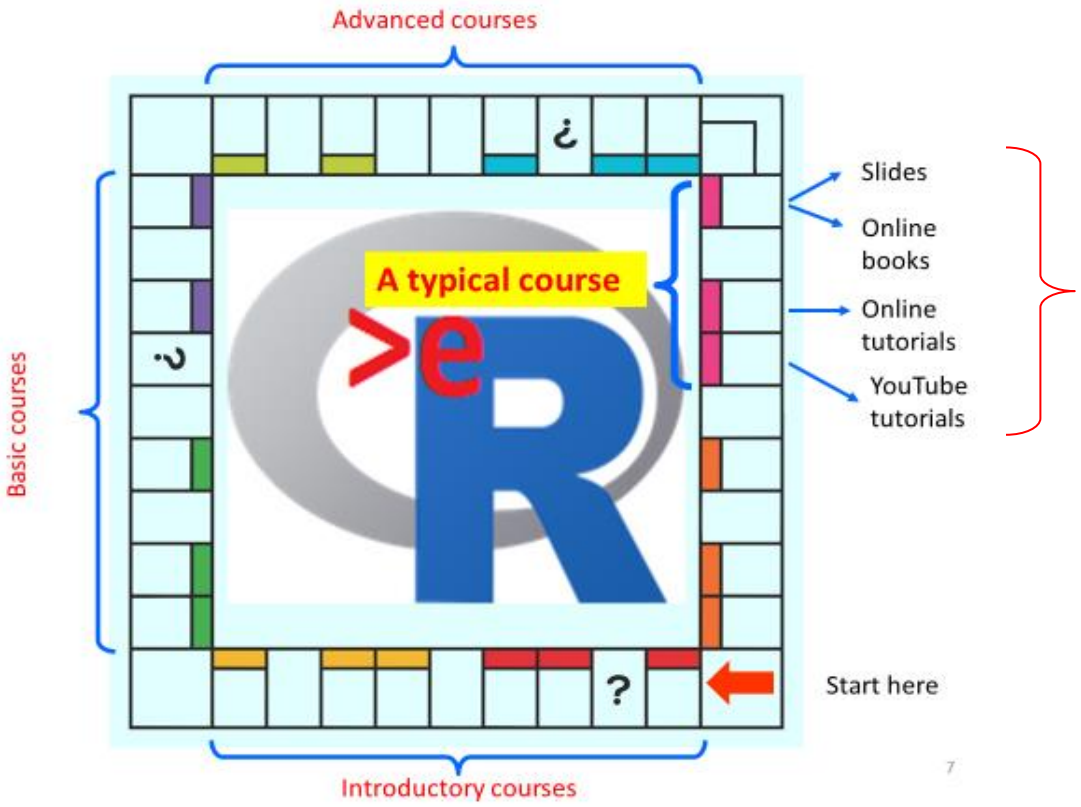
2. Selected courses were/will be developed.

3. Courses can be used either as a complete course or a part of a course.

4. Courses developed up to a class level, i.e. courses are ready to be given in the class.

Typical course structure

- Courses in three levels: introductory, basic & advanced.
- What does it means “fully develop a course…” ?



- Course materials are ready to be used in class.
- Open source policy !!!

Our approach (1): free and publically available

- Reduce costs to zero !!!
- Use publicly available products:
  - Storage course materials: GitHub (https://github.com/).
  - Website development: WIX (https://www.wix.com/).
  - Software: mostly publicly available software.
  - For example:
    - R (https://www.r-project.org/).
    - Python (https://www. python.org/).

All publically available products.

- Free for users(=students/teachers in the south):
  - No password.
  - No registration.

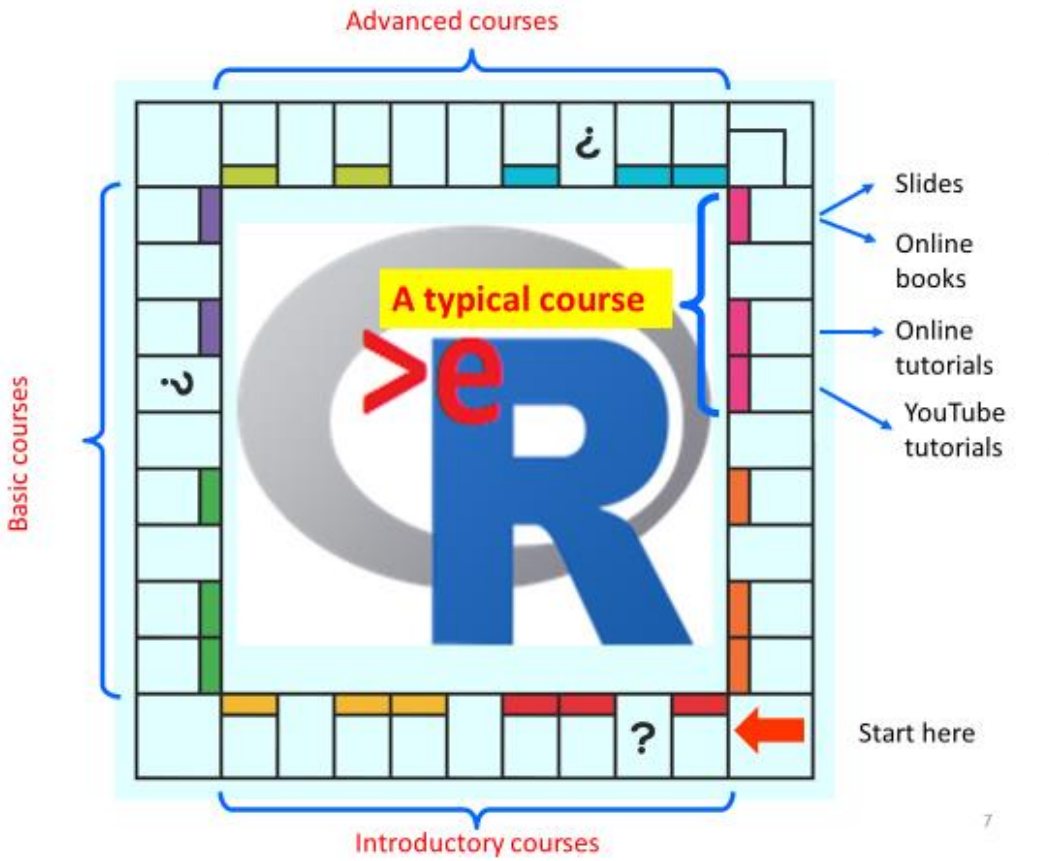
Our approach (2):
open source

- For a given course, everything is available for free:
  - Slides for the class (pdf).
  - Source files to make the slides (PP, Tex, Rmd…) !!!
  - YouTube tutorials.
  - Software programs for the examples in the course.

All publicly available.

- Free for users:
  - No password.
  - No registration.

Our approach (3): Communities



- All source materials are available for FRRE online.
- Everybody can download and use.

- Course materials can be adapted by the users for the local needs.

- Communities of users: students & teachers in the south & ….
- Communities of developers: in both south & north.

The >eR-BioStat : where to find us ?

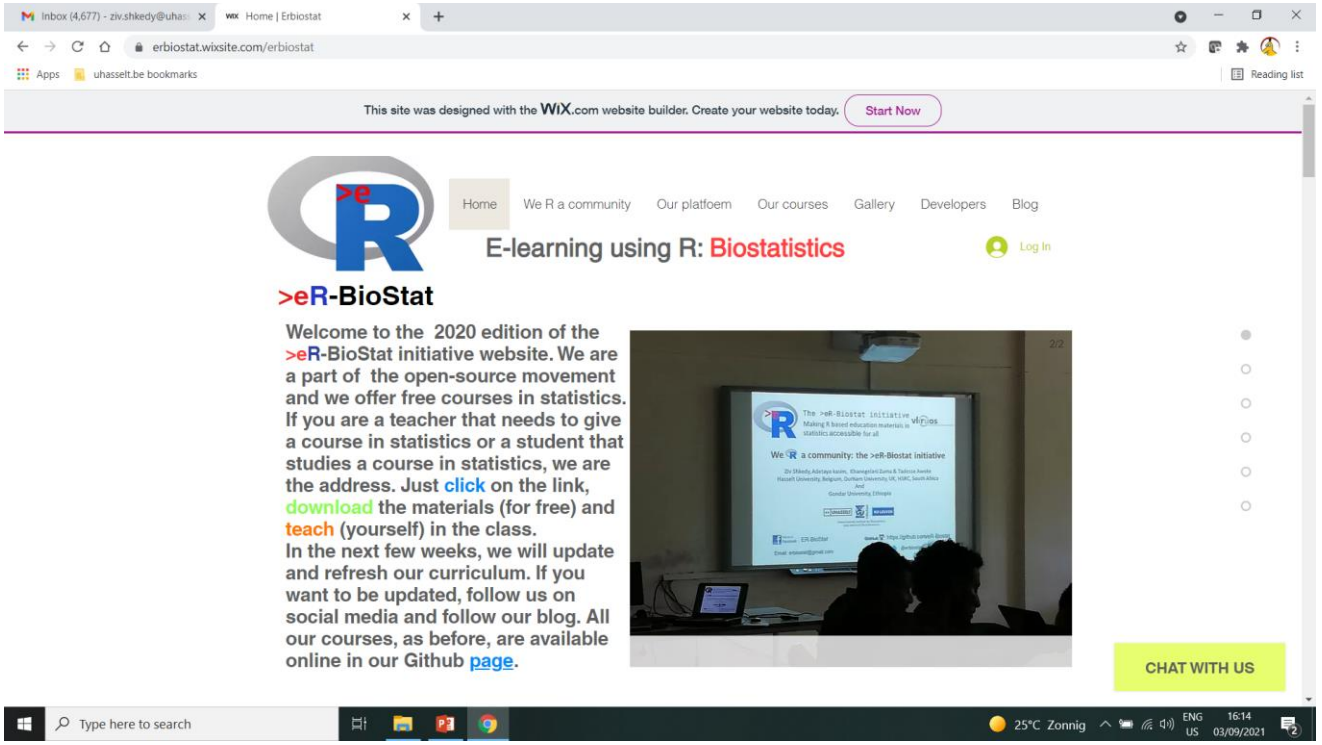
Our new website

https://erbiostat.wixsite.com/erbiostat

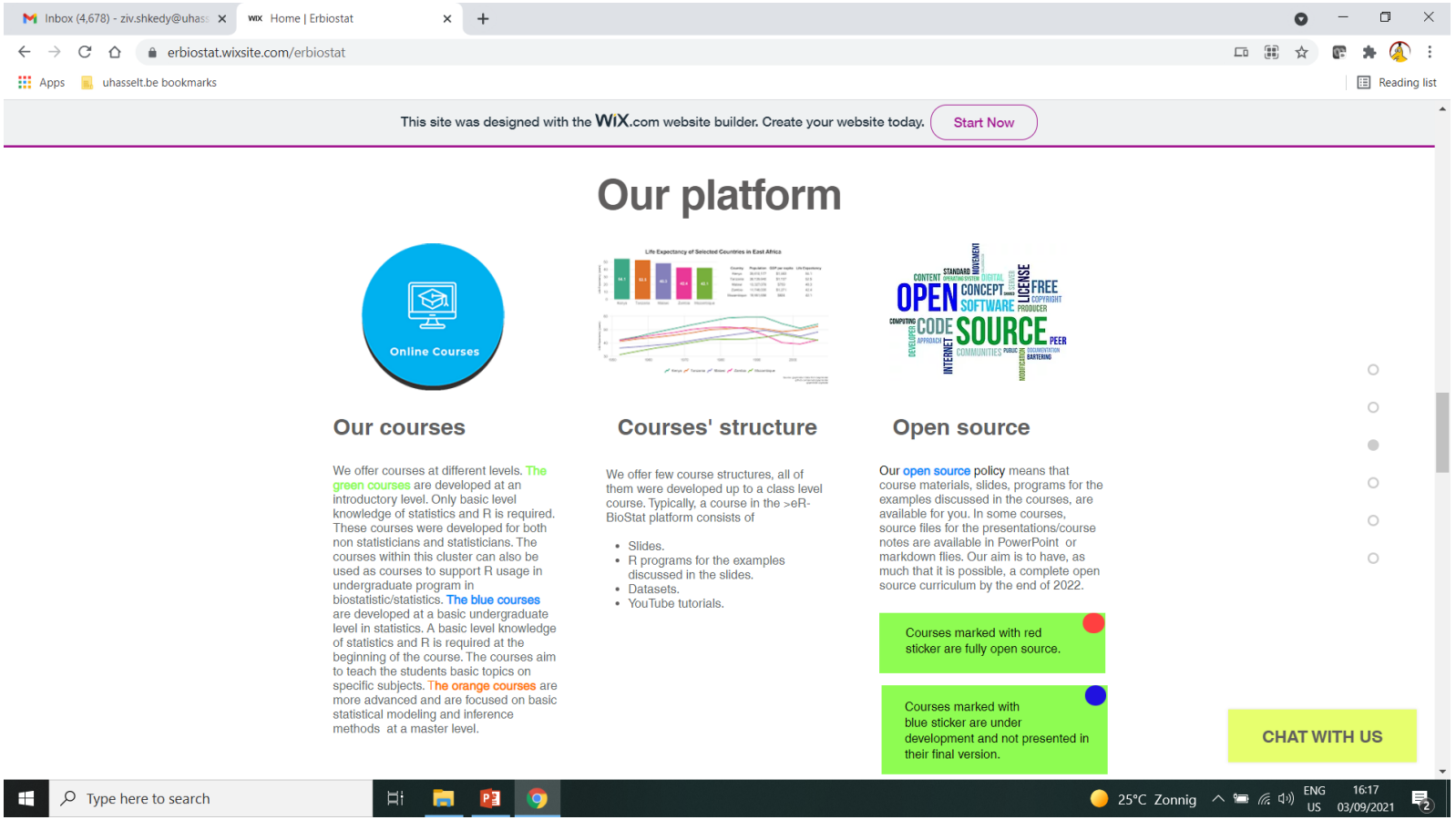
We R online

- We provide an online and free platform:

https://erbiostat.wixsite.com/erbiostat



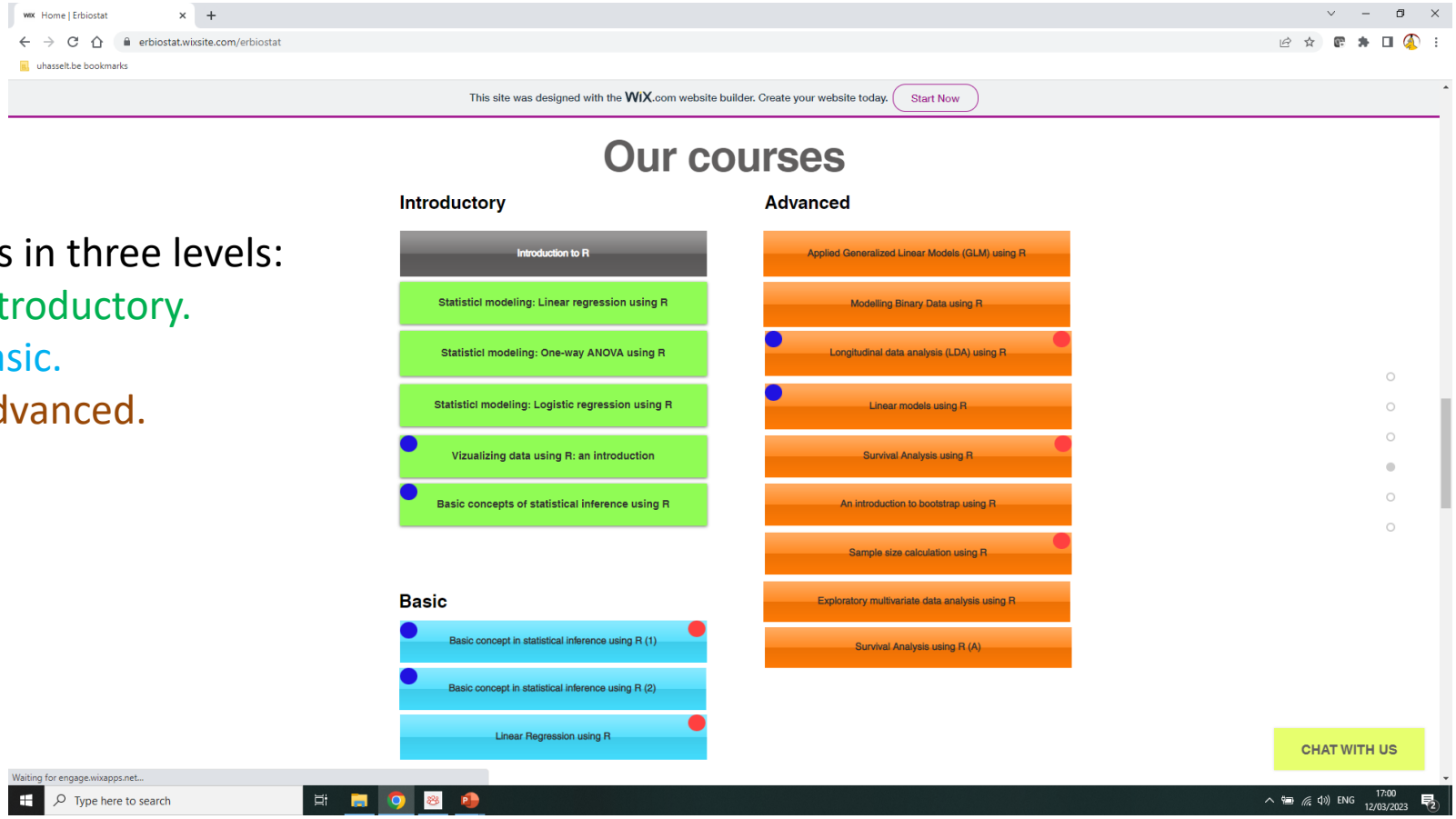
Open source



- We provide the source files for the courses:
  - PPT/Tex/Rmds for slides.
  - Rmds and R programs.

Our courses

- Courses in three levels:
  - Introductory.
  - Basic.
  - Advanced.



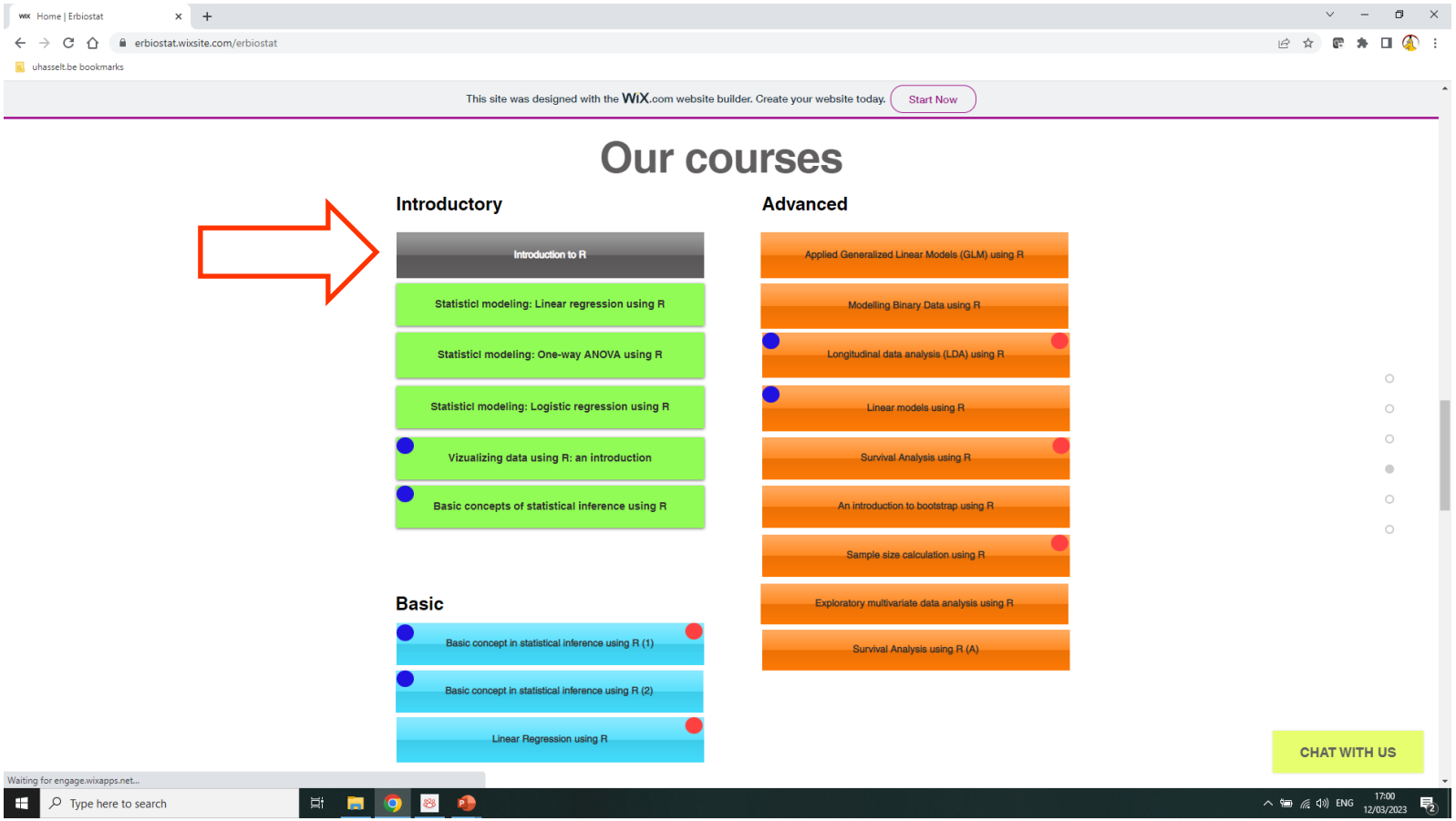
- Target for September 2023: all courses are ready to be given in class.
- To select a course: click on the course name.

The >eR-BioStat : example of one course
Introduction to R

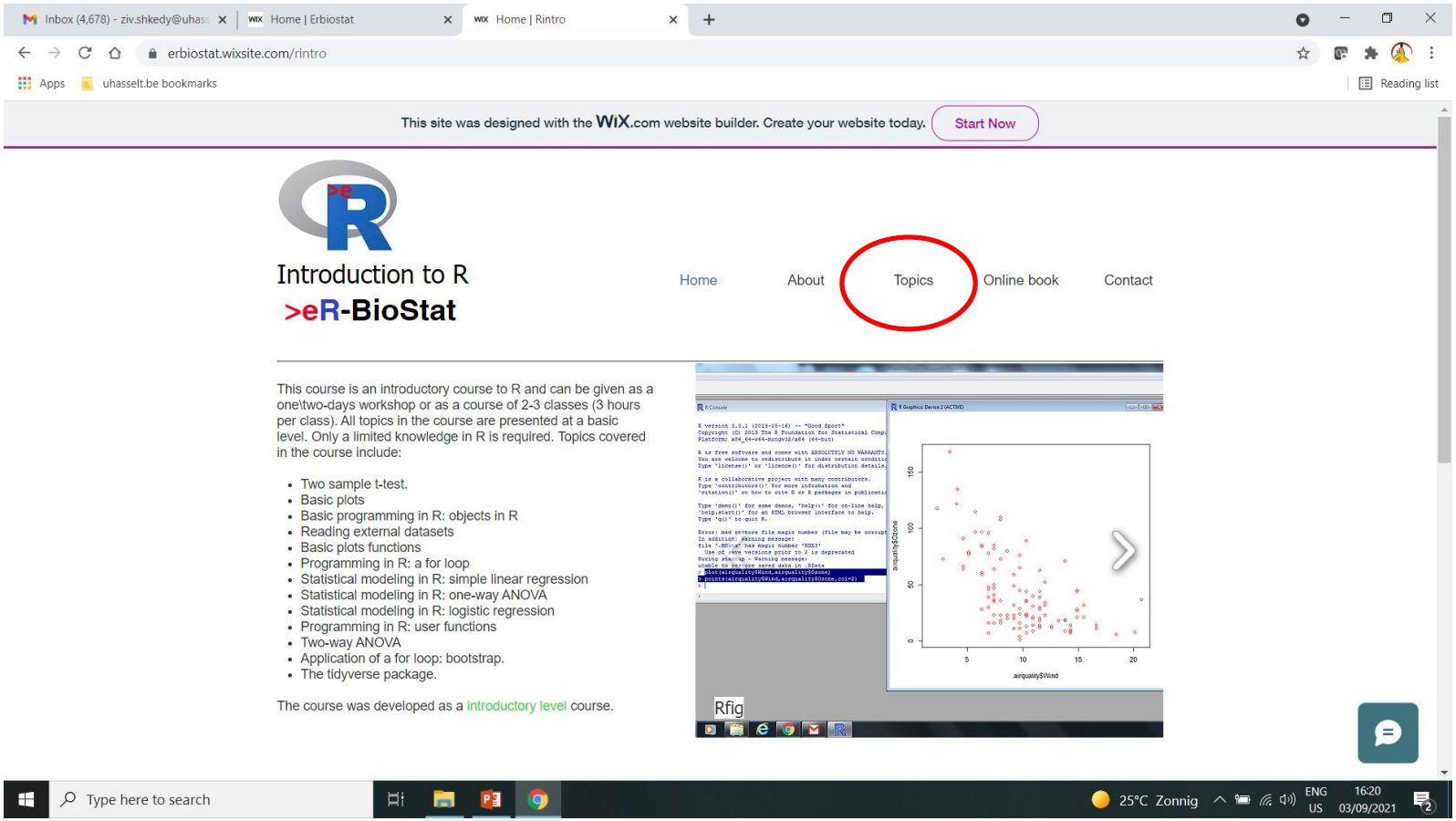
Our new website

https://erbiostat.wixsite.com/erbiostat

Example: introduction to R

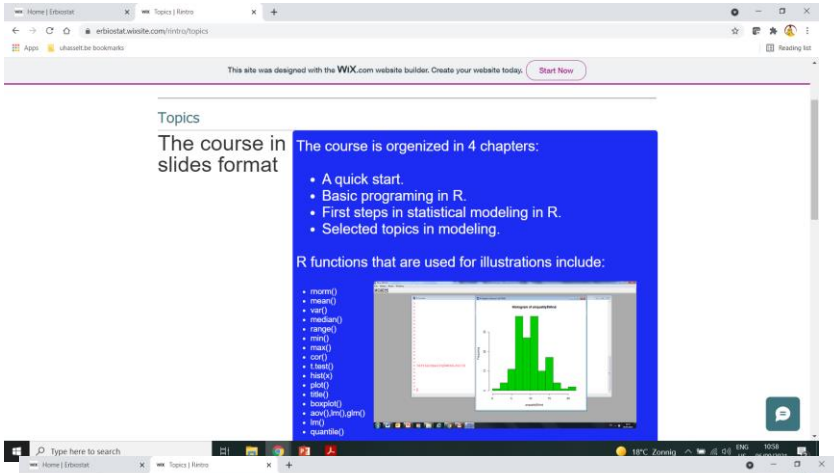


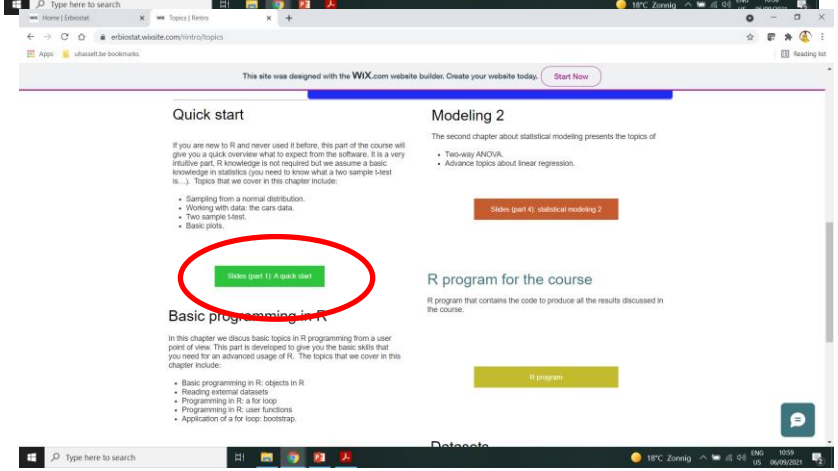
Example: introduction to R



https://erbiostat.wixsite.com/rintro

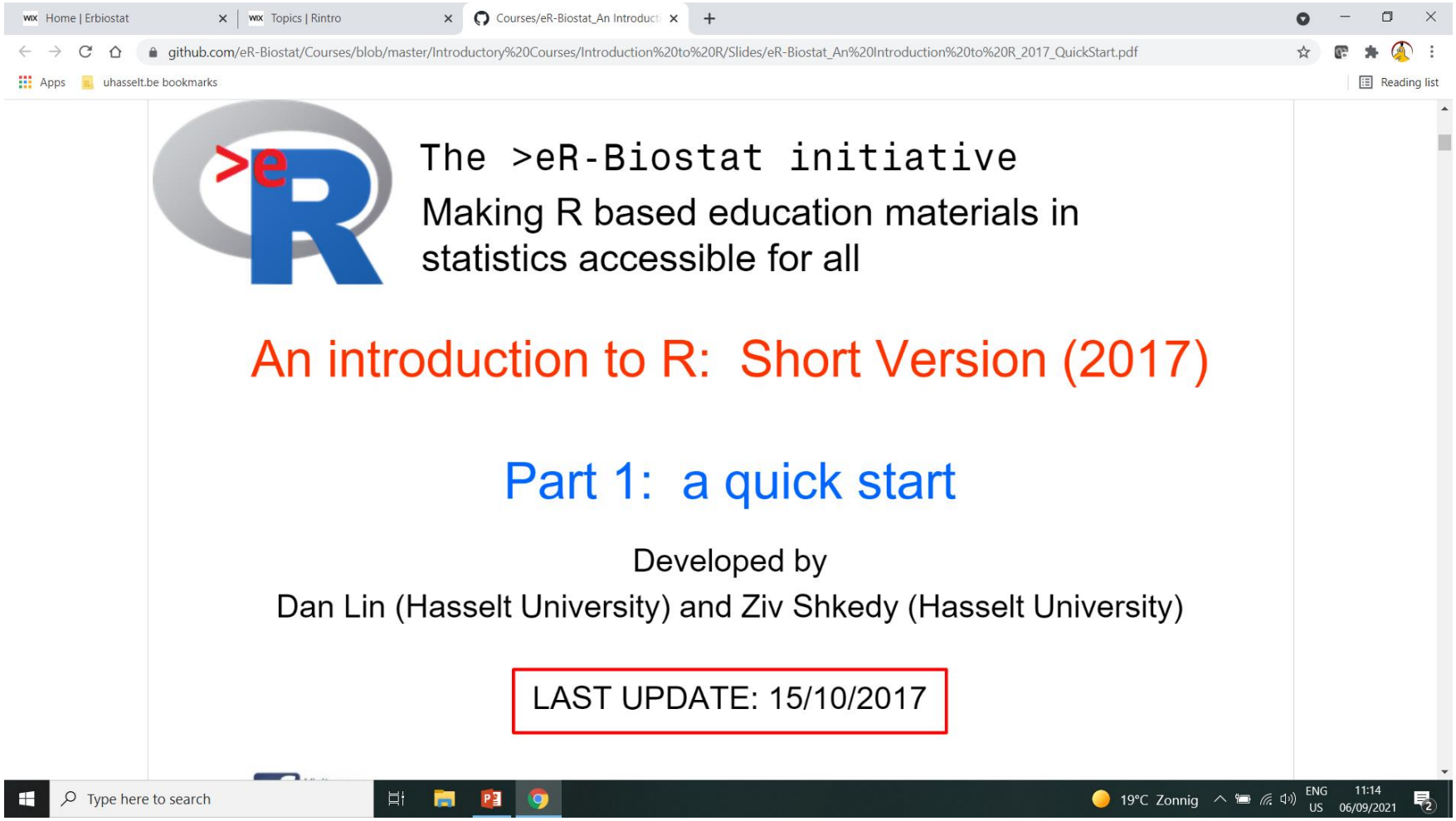
Slide format (in “Topics”)



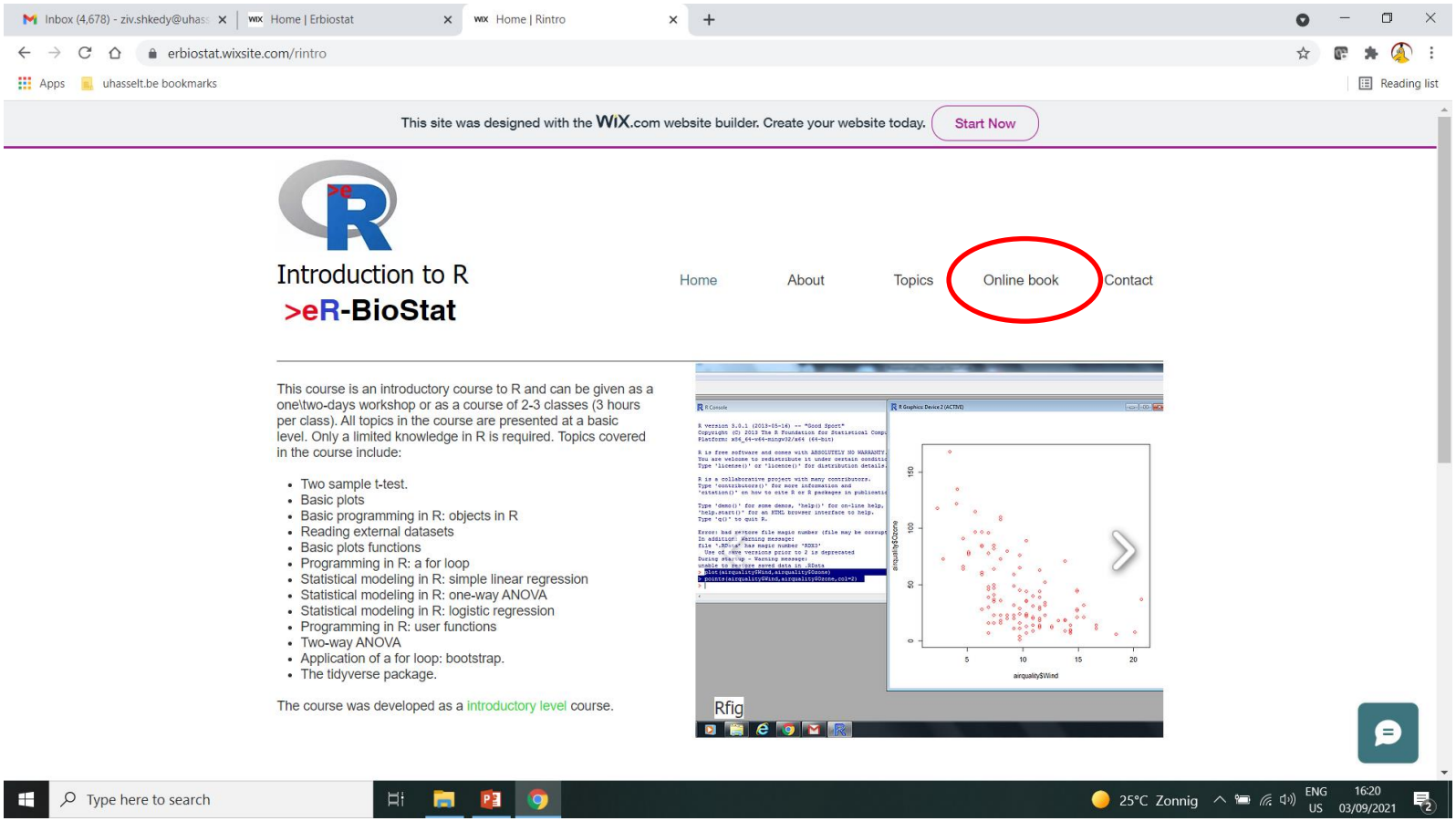


- The course in a usual slides format:
  - Slides (PDF).
  - Slides (Power point).
  - R program to produce the results presented in the slides.

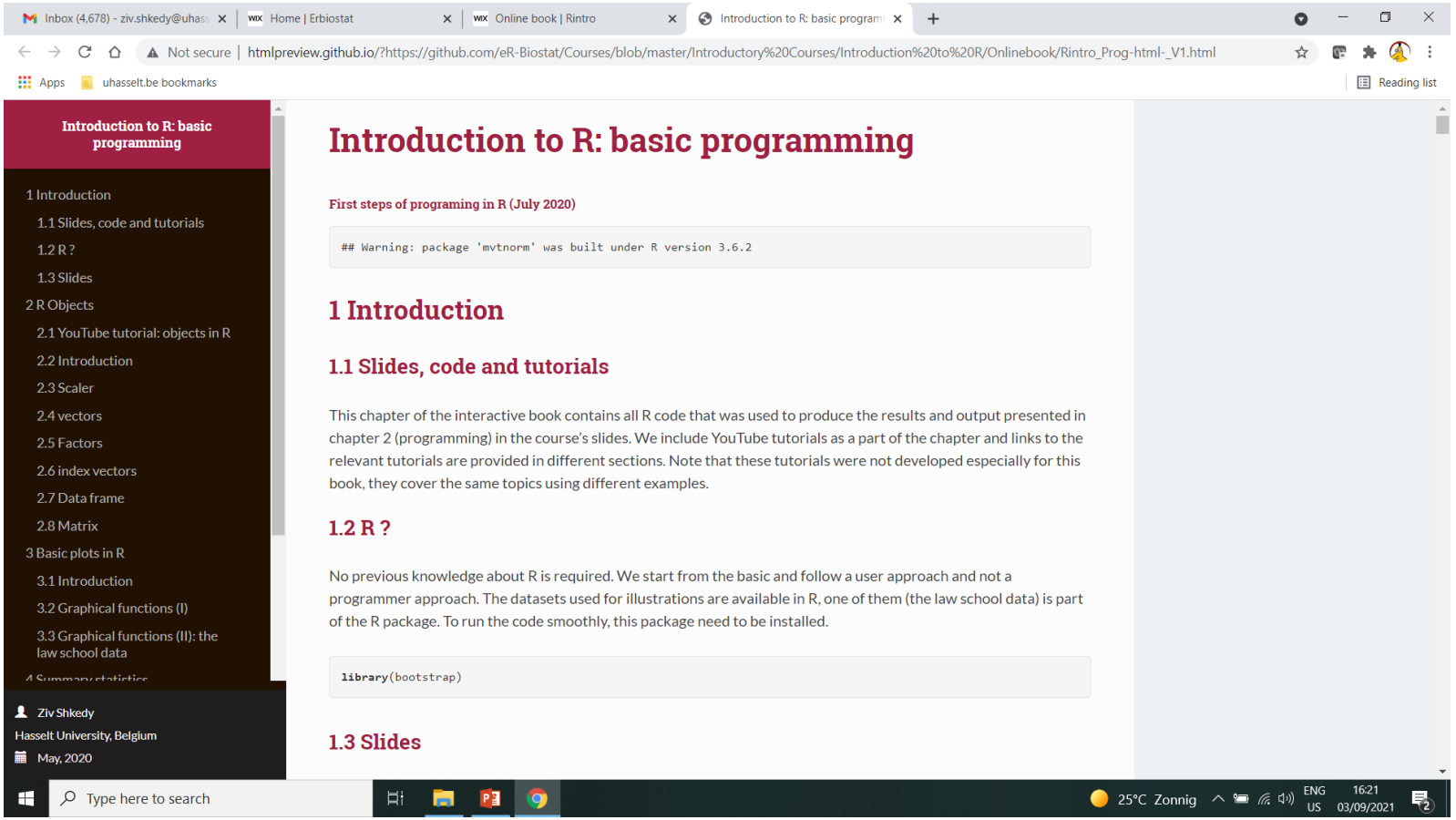
Slide format (PDF)



Online book

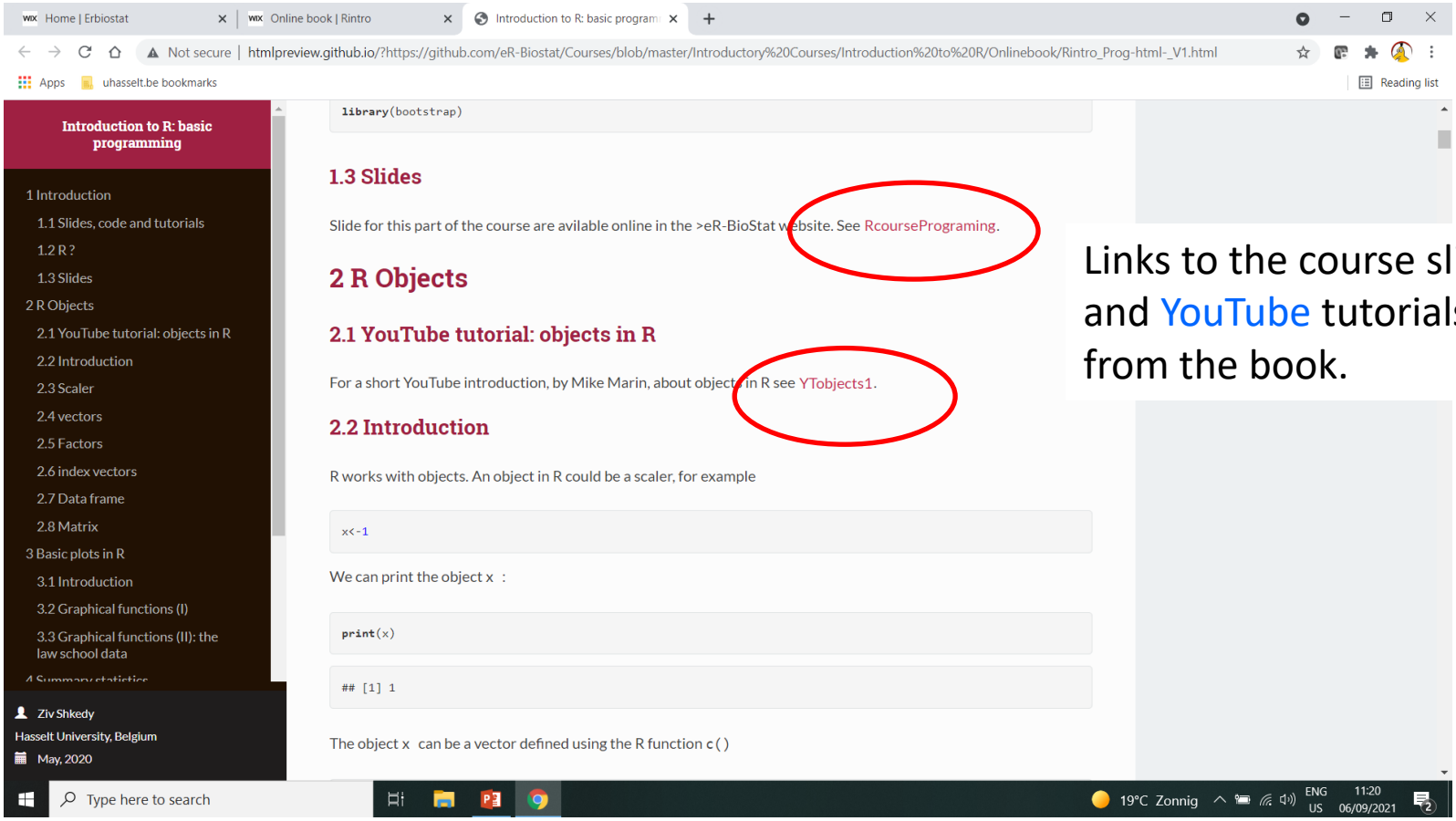

Inbox (4,678) - ziv.shkedy@uhas
Home | Erbiostat
Home | Rintro
erbiostat.wixsite.com/rintro
Apps
uhasselt.be bookmarks
Reading list
This site was designed with the WiX.com website builder. Create your website today.
Start Now
Introduction to R
>eR-BioStat
Home
About
Topics
Online book
Contact
This course is an introductory course to R and can be given as a one\two-days workshop or as a course of 2-3 classes (3 hours per class). All topics in the course are presented at a basic level. Only a limited knowledge in R is required. Topics covered in the course include:
Two sample t-test.
Basic plots
Basic programming in R: objects in R
Reading external datasets
Basic plots functions
Programming in R: a for loop
Statistical modeling in R: simple linear regression
Statistical modeling in R: one-way ANOVA
Statistical modeling in R: logistic regression
Programming in R: user functions
Two-way ANOVA
Application of a for loop: bootstrap.
The tidyverse package.
The course was developed as a introductory level course.
Rfig
airquality$Ozone
airquality$Wind
Type here to search
25°C Zonnig
ENG
US
16:20
03/09/2021

Online book



- Available in
  - HTML.
  - PDF.
  - Rmd to reproduce the book on your laptop.

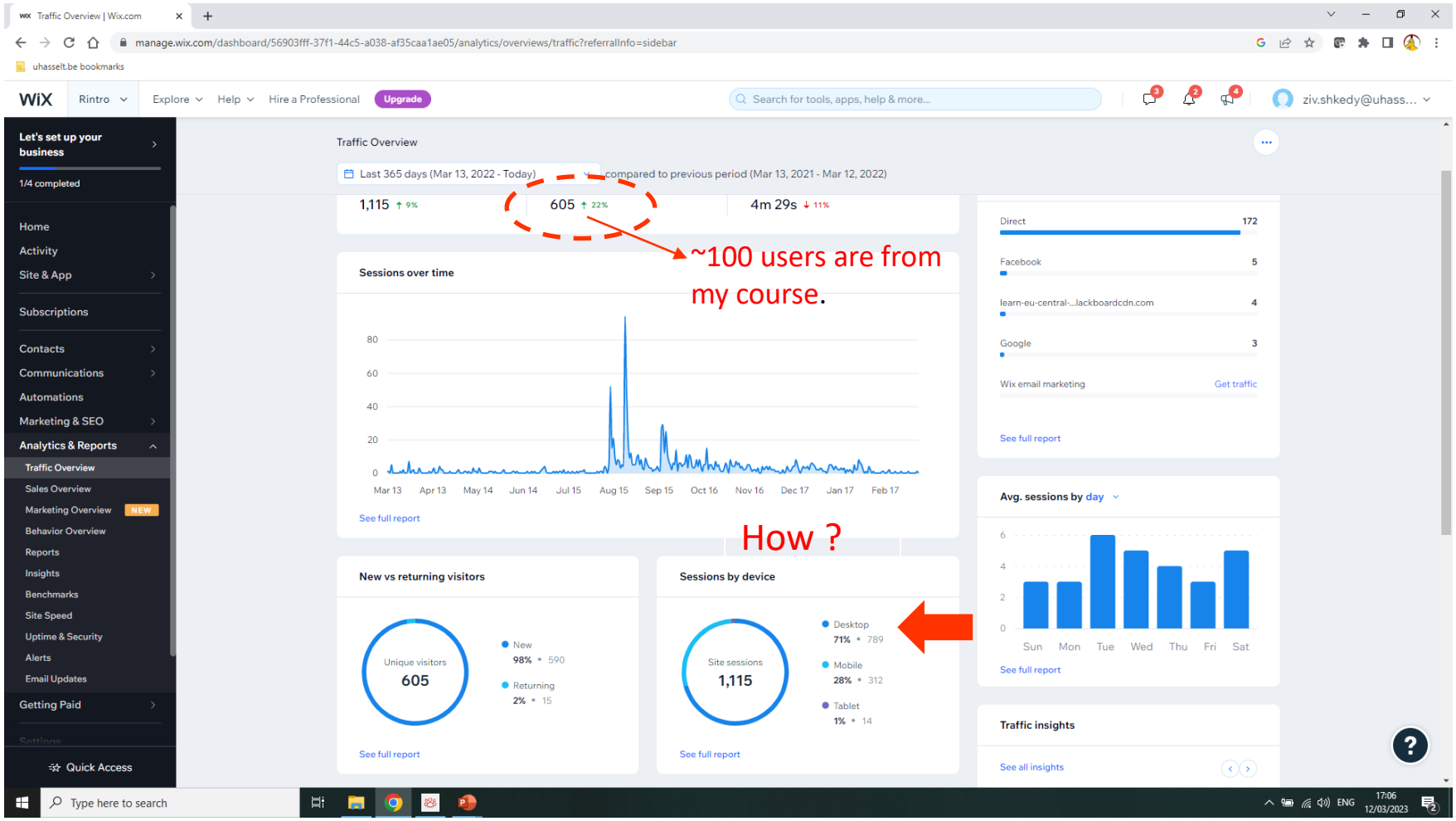
Online book



Links to the course slides and YouTube tutorials from the book.

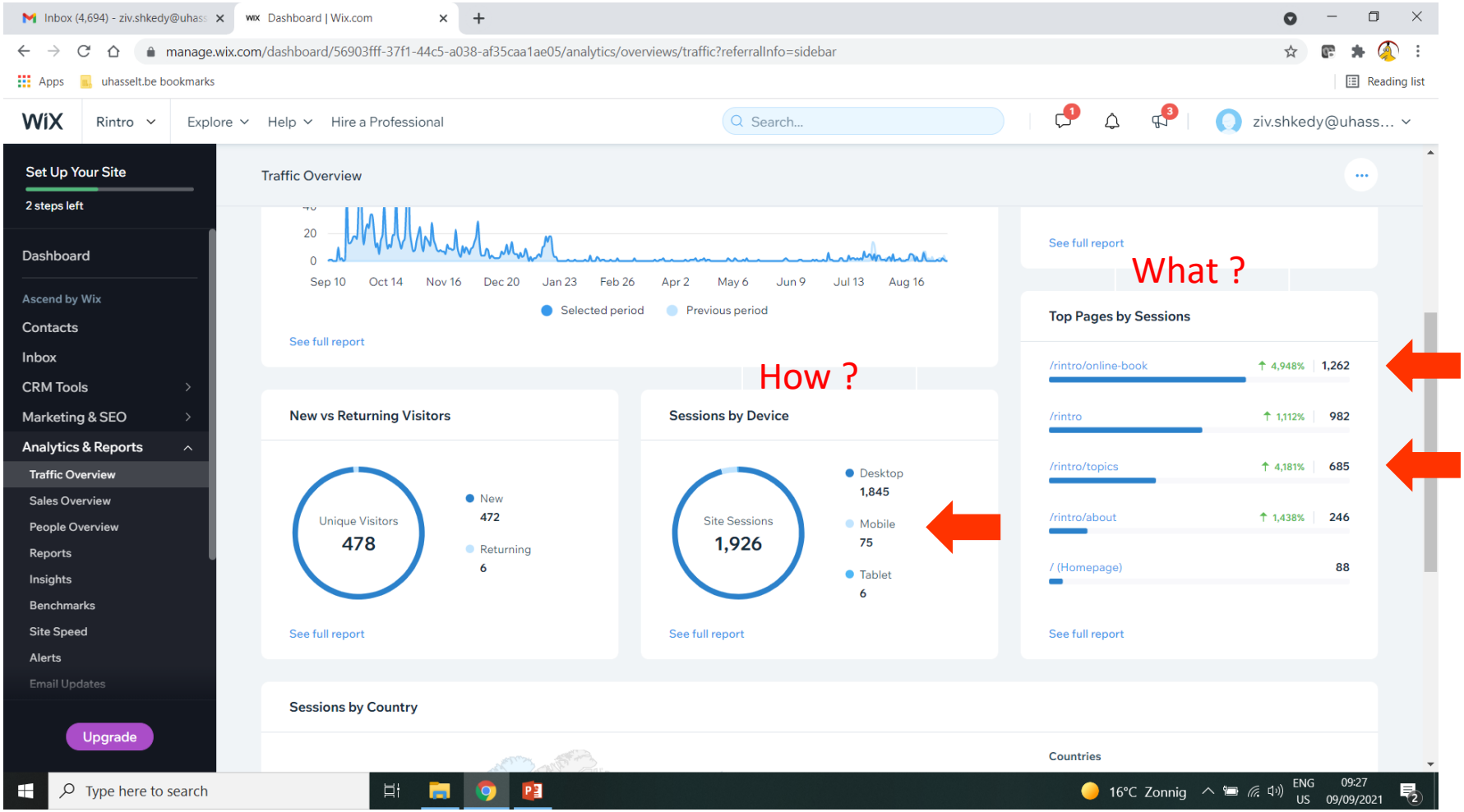
R course: users in the last 365 days

13/03/2022-12/03/2023:605 users.



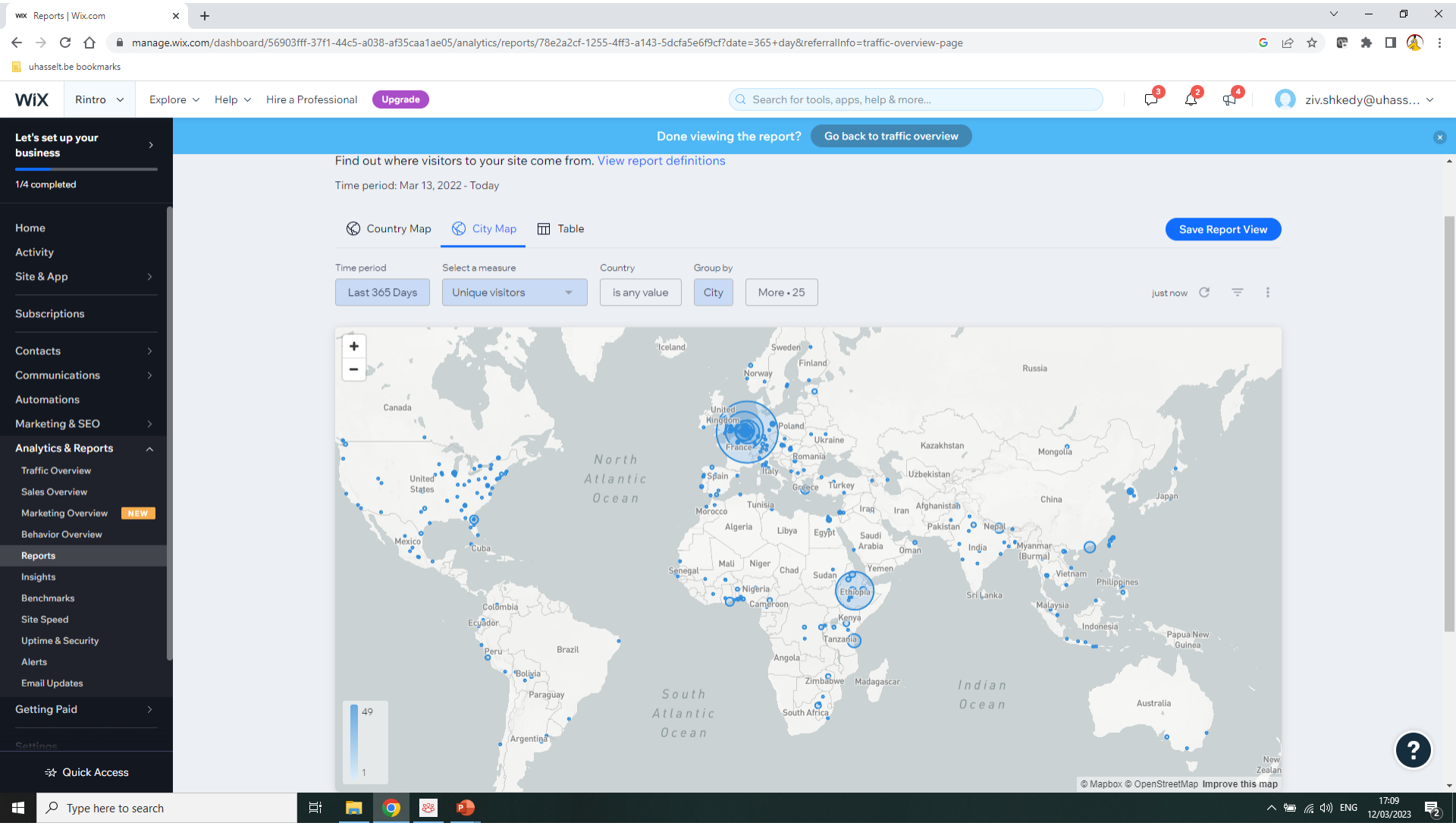
How our users use the website ? (Introduction to R)

For 2021



Who are our users? (Introduction to R)

13/03/2022-12/03/2023



Stay connected

Interuniversity Institute for Biostatistics and statistical Bioinformatics

- Look for more examples & courses in our website:

  https://erbiostat.wixsite.com/erbiostat

- Users:
  - Want to take part in our initiative ?
  - Want to include your university in our initiative ?
  - Want to use our courses for your classes ?
- Developers:
  - Want to contribute a course for this initiative ?
- Send us an email !!!

  ziv.shkedy@uhasselt.be

- Follow us on social media:

Visit us on Facebook ER-BioStat & twitter @erbiostat

UHASSELT

I-BioStat

KU LEUVEN

Interuniversity Institute for Biostatistics and statistical Bioinformatics

vlir uos

SHARING MINDS, CHANGING LIVES

Thank you ve y much !!

Join our community and look for us online:

Visit us on Facebook ER-BioStat

GitHub https://github.com/eR-Biostat

https://erbiostat.wixsite.com/erbiostat

twitter @erbiostat

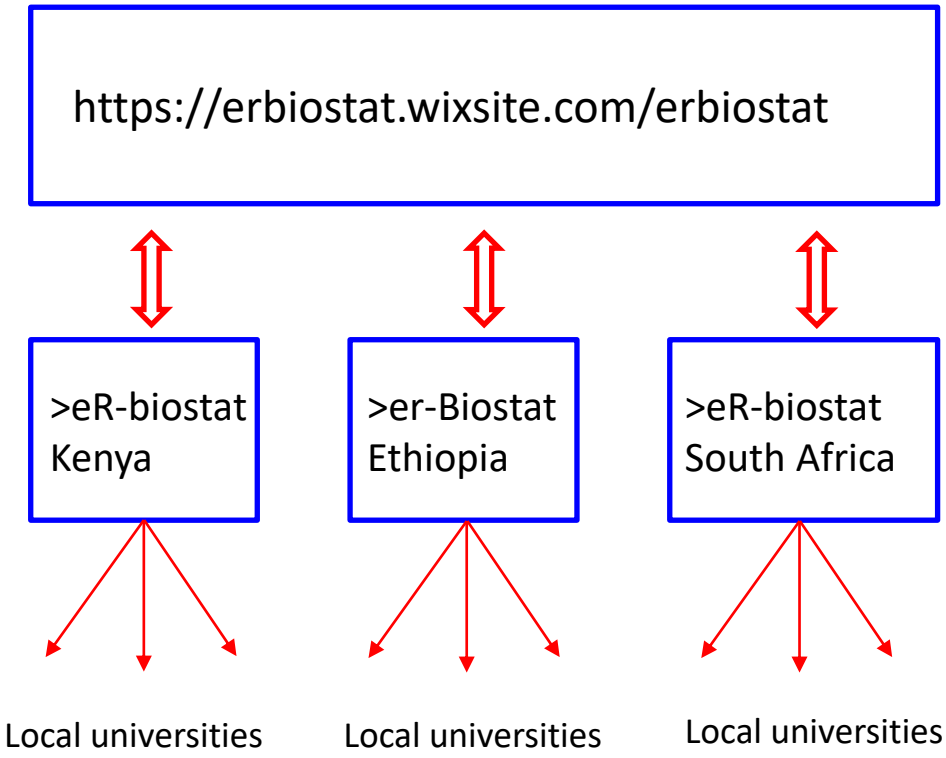
Some extra slides about the ITP project

The >eR-BioStat platform: global & local websites & colaborators

Our new website

https://erbiostat.wixsite.com/erbiostat

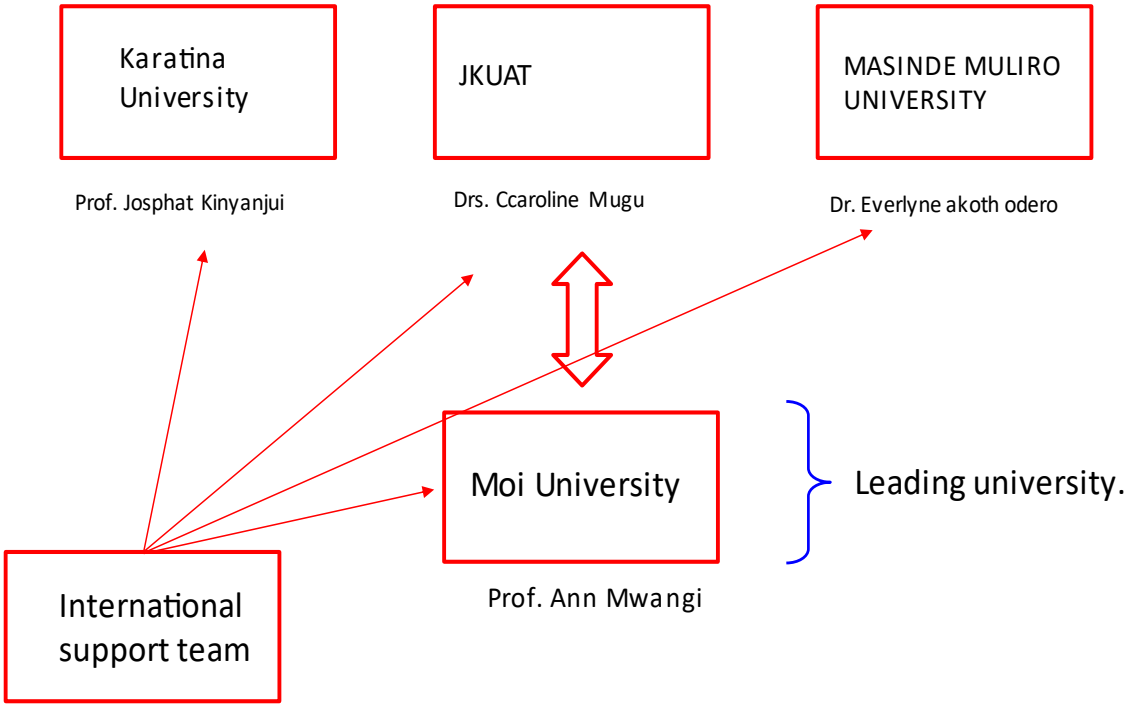
Online structure



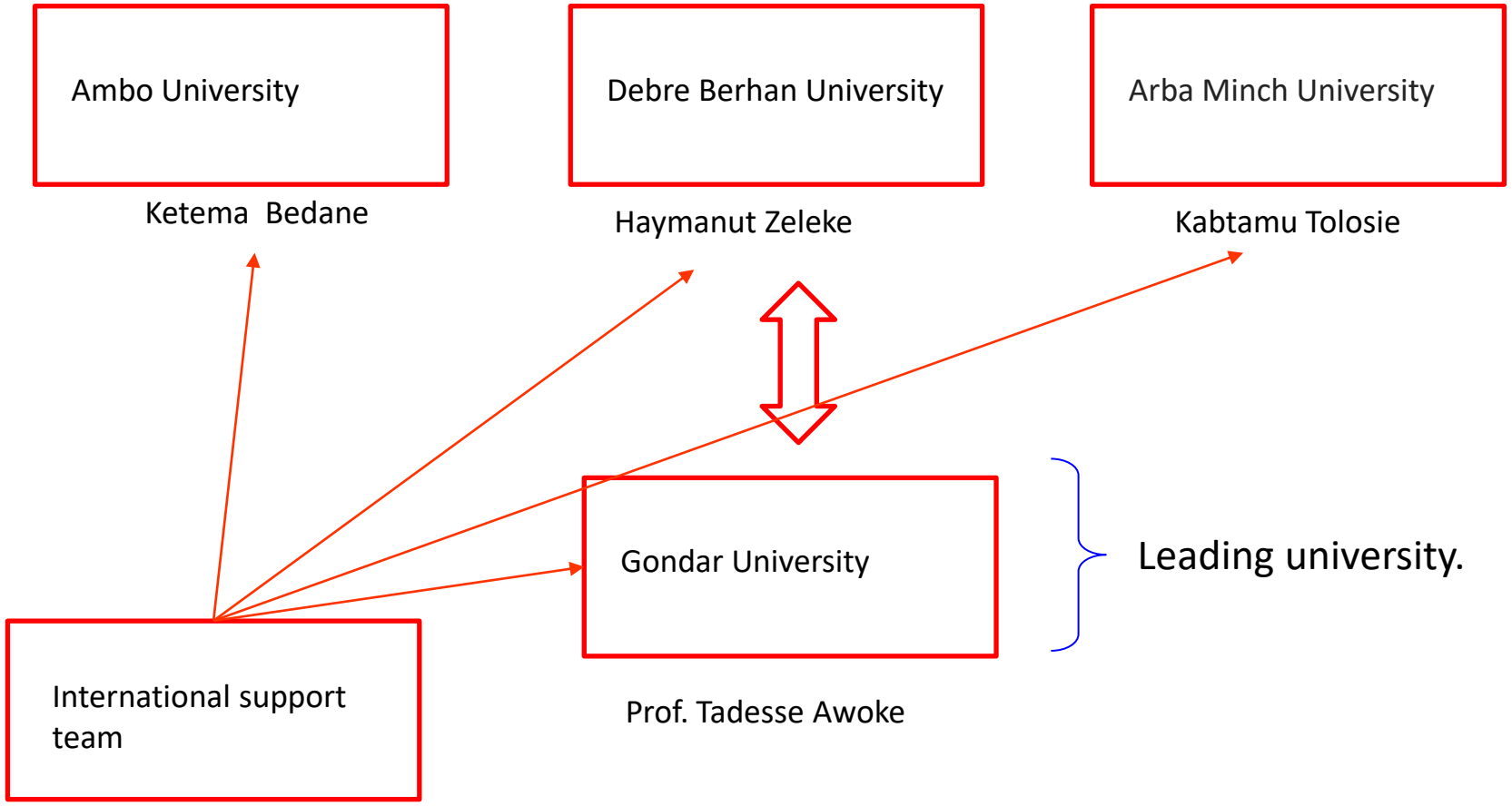
Collaborators network (South partners)

- Kenya:
  - Moi University.
    - Karatina University, The Department of Mathematics, Statistics and Actuarial Science.
    - Jomo Kenyatta University of Agriculture and Technology (JKUAT) ,The department of Statistics and actuarial science.
    - Masinde Muliro University of Science and Technology, department of statistics.
- Ethiopia:
  - Gondar University.
    - Ambo University, The department of statistics.
    - Arba Minch University, The Department of Statistics.
    - Debre Berhan University (DBU), Statistics Department.
- South Africa:
  - University of Pretoria.
  - MRC-SA.
  - WITS.
    - Stellenbosch University (SU), the division of epidemiology and biostatistics.
    - The University of Venda, The Department of Mathematical and Computational Sciences.
    - The University of Zululand (UNIZULU), The Department of Mathematical Science.

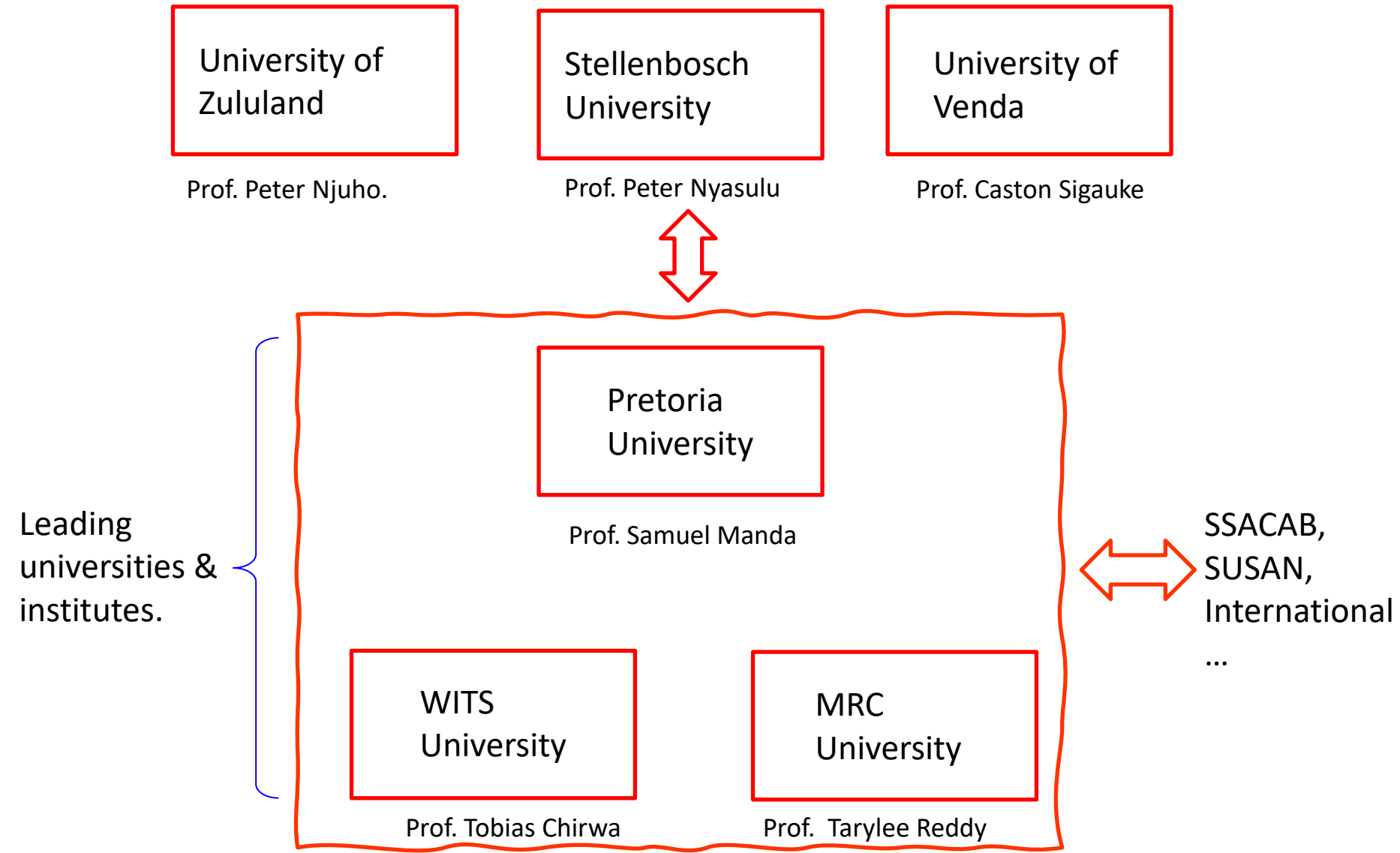
Collaborators network (Kenya)



Collaborators network (Ethiopia)



Collaborators network (South Africa)



Collaborators network 1 (International)

- Existing collaboration links among partners:

Leading partner and South promoter

- Brown University, USA

Kenya

Belgium

- UHasselt ⟷ UGent, KULeuven

- SAACAB PARTNERS
- NIH, Maryland, USA
- Georgetown University, Washington DC, USA

South Africa

Ethiopia

- McMaster University, Hamilton, Canada
- University of Kentucky, USA
- University, Truro, NS, Canada

Collaborators (SSACAB, Gambia workshop, DS paper etc.)

- USA:
  - Misrak Gezmu, Statistician, NIAID, NIH, Maryland, USA.
  - Joseph Hogan, Brown University, Rhode Island, USA.
  - Mahlet Tadesse, Georgetown University, Washington DC, USA.
  - Solomon W. Harar, Department of Statistics, University of Kentucky, Lexington, Kentucky, USA.
  - Mekbib Altaye and Tesfaye B. Mersha, Cincinnati Children's Hospital Medical Center, University of Cincinnati, Cincinnati, OH, USA.
- Canada:
  - Joseph Beyne, Department of Health Research Methods, Evidence, and Impact, McMaster University, Hamilton, ON, Canada.
  - Tessema Astatkie, Faculty of Agriculture, Dalhousie University, Truro, NS, Canada.
- Others:
  - .....

The >eR-BioStat : example of one course
GLM

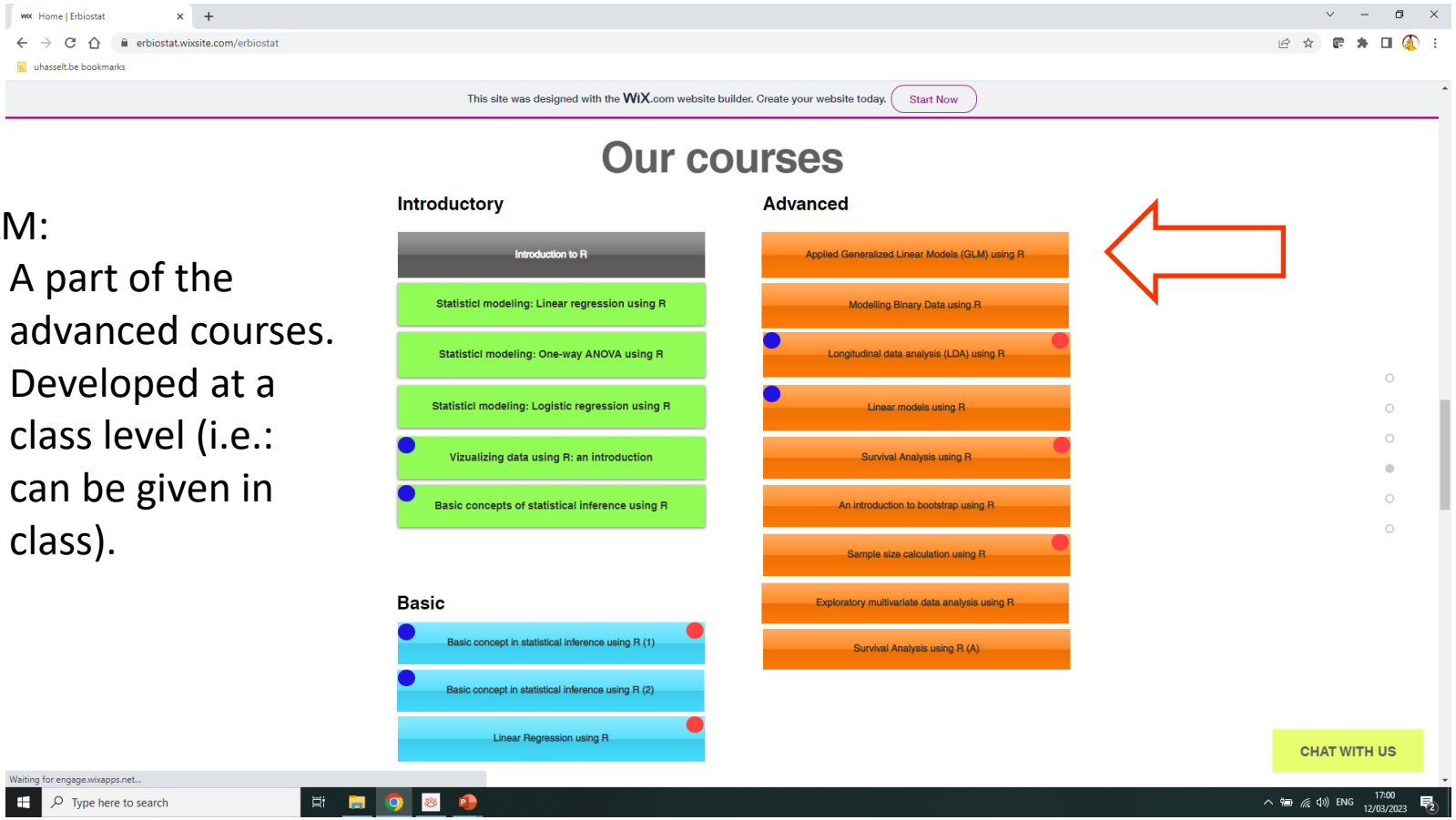
Our new website

https://erbiostat.wixsite.com/erbiostat

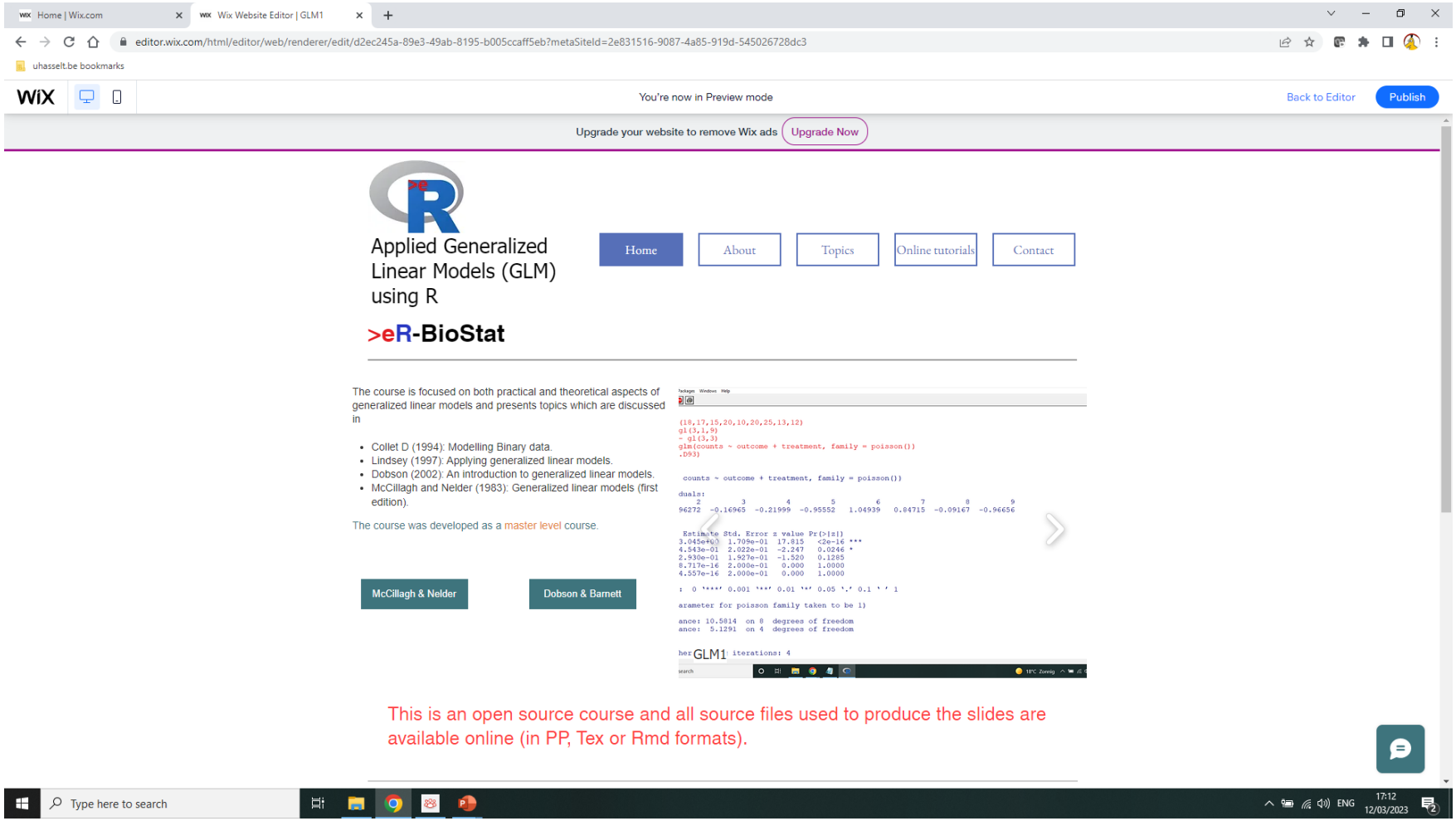
Example of the course: GLM

GLM:

- A part of the advanced courses.
- Developed at a class level (i.e.: can be given in class).

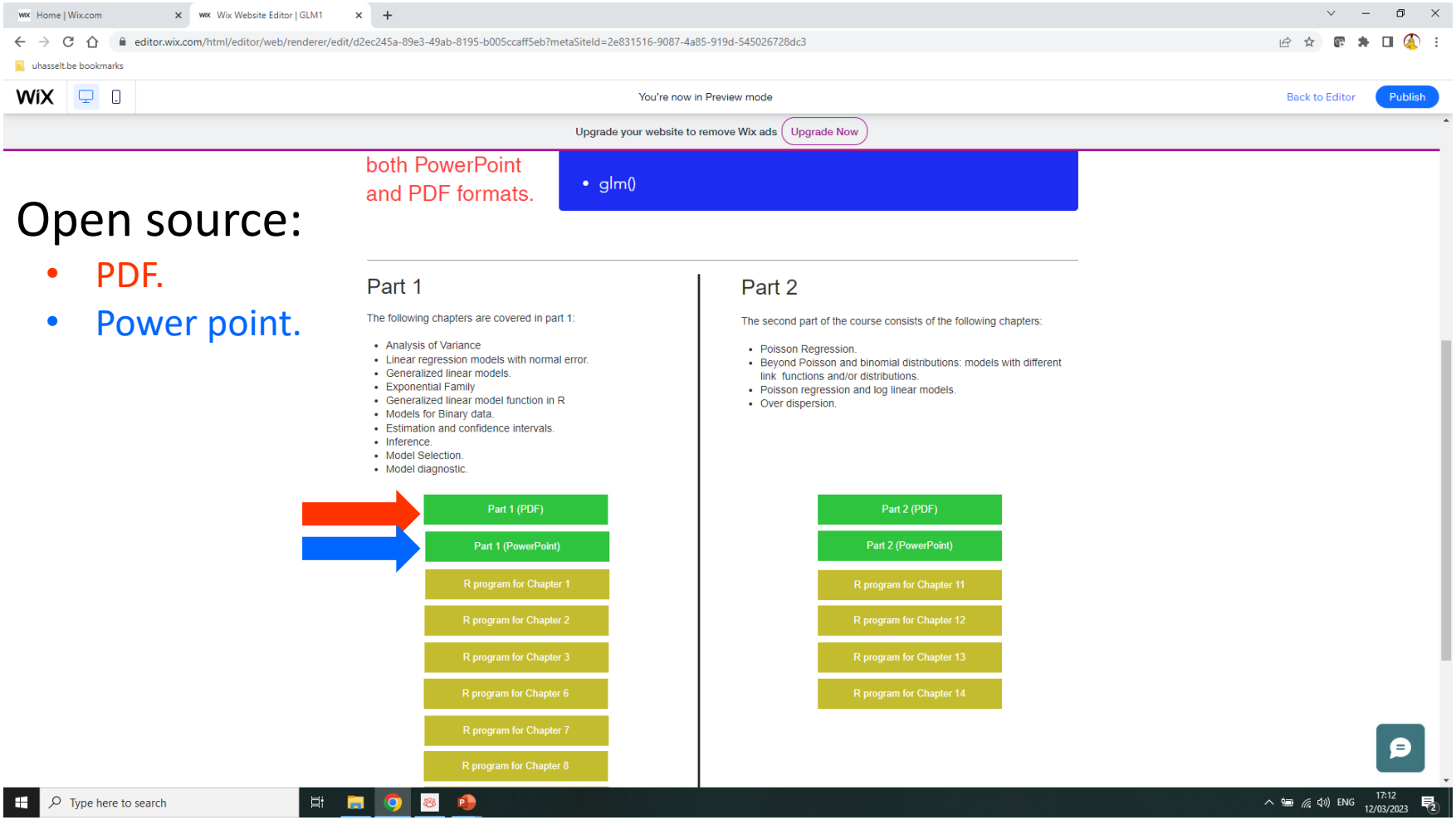


The GLM page


Home | Wix.com
Wix Website Editor | GLM1
editor.wix.com/html/editor/web/renderer/edit/d2ec245a-89e3-49ab-8195-b005ccaff5eb?metaSiteId=2e831516-9087-4a85-919d-545026728dc3
uhasselt.be bookmarks
WiX
You're now in Preview mode
Back to Editor
Publish
Upgrade your website to remove Wix ads
Upgrade Now
Applied Generalized Linear Models (GLM) using R
>eR-BioStat
Home
About
Topics
Online tutorials
Contact
The course is focused on both practical and theoretical aspects of generalized linear models and presents topics which are discussed in
Collet D (1994): Modelling Binary data.
Lindsey (1997): Applying generalized linear models.
Dobson (2002): An introduction to generalized linear models.
McCillagh and Nelder (1983): Generalized linear models (first edition).
The course was developed as a master level course.
McCillagh & Nelder
Dobson & Barnett
counts ~ outcome + treatment, family = poisson())
Estimate Std. Error z value Pr(>|z|)
GLM1
This is an open source course and all source files used to produce the slides are available online (in PP, Tex or Rmd formats).
Type here to search
ENG
17:12
12/03/2023

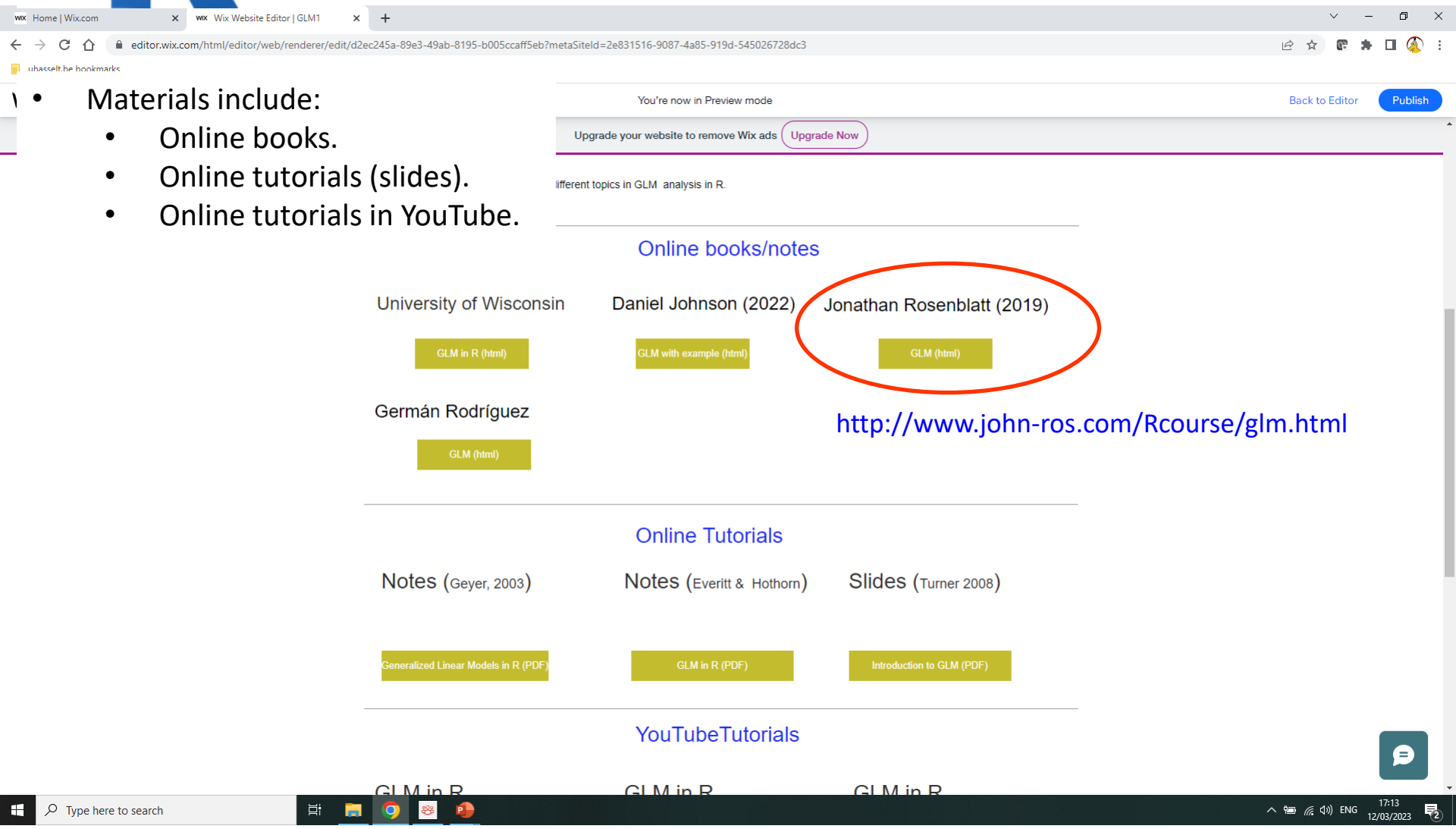
Course content + slides and R programs

- Open source:
  - PDF.
  - Power point.



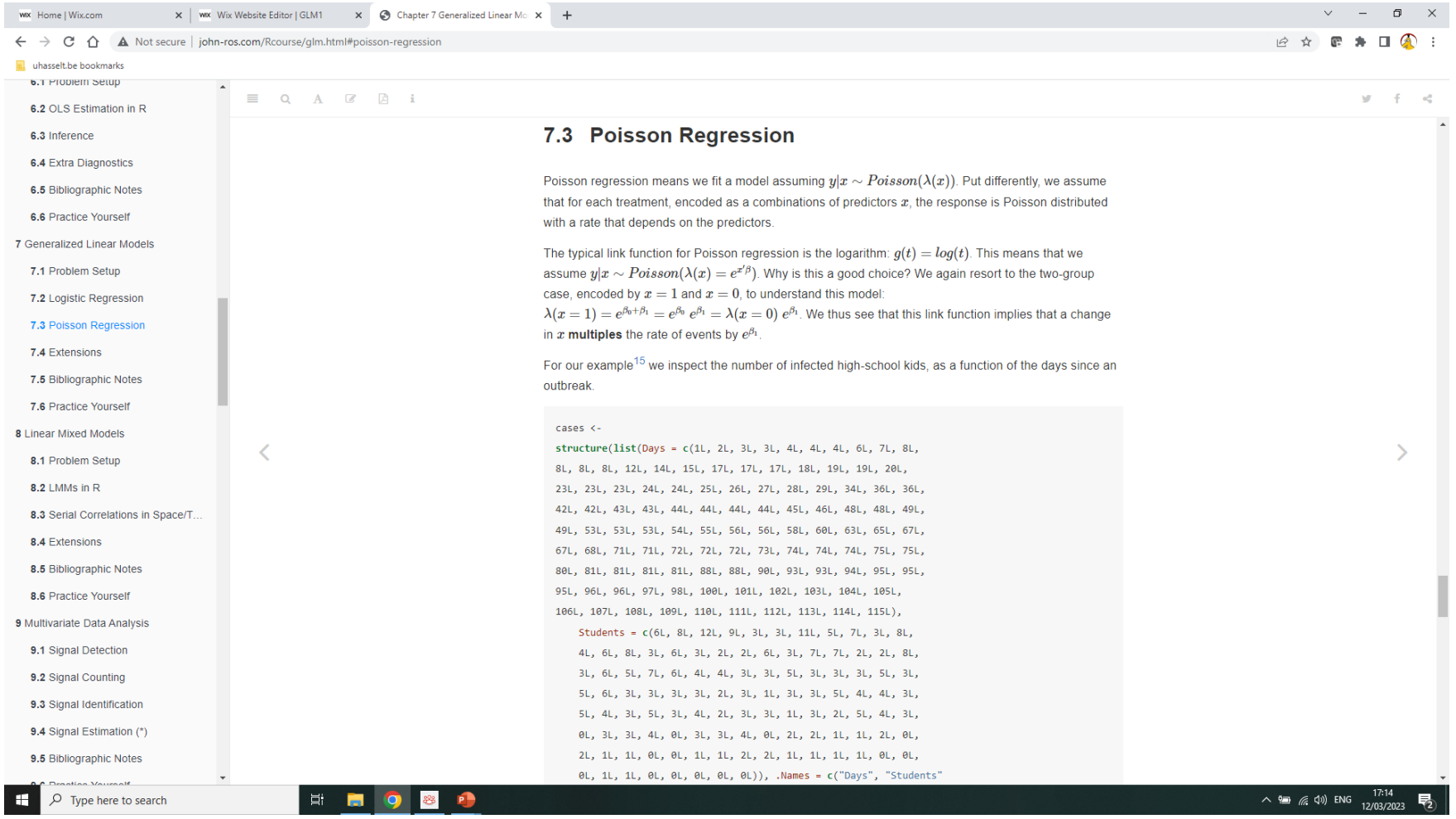
Online material

- Materials include:
  - Online books.
  - Online tutorials (slides).
  - Online tutorials in YouTube.



Example of a chapter

(book of Jonathan D. Rosenblatt)



Home | Wix.com | Wix Website Editor | GLM1 | Chapter 7 Generalized Linear Mo

Not secure | john-ros.com/Rcourse/glm.html#poisson-regression

uhasselt.be bookmarks

6.1 Problem Setup
6.2 OLS Estimation in R
6.3 Inference
6.4 Extra Diagnostics
6.5 Bibliographic Notes
6.6 Practice Yourself
7 Generalized Linear Models
7.1 Problem Setup
7.2 Logistic Regression
7.3 Poisson Regression
7.4 Extensions
7.5 Bibliographic Notes
7.6 Practice Yourself
8 Linear Mixed Models
8.1 Problem Setup
8.2 LMMs in R
8.3 Serial Correlations in Space/T...
8.4 Extensions
8.5 Bibliographic Notes
8.6 Practice Yourself
9 Multivariate Data Analysis
9.1 Signal Detection
9.2 Signal Counting
9.3 Signal Identification
9.4 Signal Estimation (*)
9.5 Bibliographic Notes

7.3 Poisson Regression

Poisson regression means we fit a model assuming $y|x \sim Poisson(\lambda(x))$. Put differently, we assume that for each treatment, encoded as a combinations of predictors x, the response is Poisson distributed with a rate that depends on the predictors.

The typical link function for Poisson regression is the logarithm: $g(t) = log(t)$. This means that we assume $y|x \sim Poisson(\lambda(x) = e^{x'\beta})$. Why is this a good choice? We again resort to the two-group case, encoded by $x = 1$ and $x = 0$, to understand this model:
$\lambda(x = 1) = e^{\beta_0+\beta_1} = e^{\beta_0} e^{\beta_1} = \lambda(x = 0)\, e^{\beta_1}$. We thus see that this link function implies that a change in x **multiples** the rate of events by e^{β_1}.

For our example[15] we inspect the number of infected high-school kids, as a function of the days since an outbreak.

```
cases <-
structure(list(Days = c(1L, 2L, 3L, 3L, 4L, 4L, 4L, 6L, 7L, 8L,
8L, 8L, 8L, 12L, 14L, 15L, 17L, 17L, 17L, 18L, 19L, 19L, 20L,
23L, 23L, 23L, 24L, 24L, 25L, 26L, 27L, 28L, 29L, 34L, 36L, 36L,
42L, 42L, 43L, 43L, 44L, 44L, 44L, 44L, 45L, 46L, 48L, 48L, 49L,
49L, 53L, 53L, 53L, 54L, 55L, 56L, 56L, 58L, 60L, 63L, 65L, 67L,
67L, 68L, 71L, 71L, 72L, 72L, 72L, 73L, 74L, 74L, 74L, 75L, 75L,
80L, 81L, 81L, 81L, 81L, 88L, 88L, 90L, 93L, 93L, 94L, 95L, 95L,
95L, 96L, 96L, 97L, 98L, 100L, 101L, 102L, 103L, 104L, 105L,
106L, 107L, 108L, 109L, 110L, 111L, 112L, 113L, 114L, 115L),
    Students = c(6L, 8L, 12L, 9L, 3L, 3L, 11L, 5L, 7L, 3L, 8L,
    4L, 6L, 8L, 3L, 6L, 3L, 2L, 2L, 6L, 3L, 7L, 7L, 2L, 2L, 8L,
    3L, 6L, 5L, 7L, 6L, 4L, 4L, 3L, 3L, 5L, 3L, 3L, 3L, 5L, 3L,
    5L, 6L, 3L, 3L, 3L, 3L, 2L, 3L, 1L, 3L, 3L, 5L, 4L, 4L, 3L,
    5L, 4L, 3L, 5L, 3L, 4L, 2L, 3L, 3L, 1L, 3L, 2L, 5L, 4L, 3L,
    0L, 3L, 3L, 4L, 0L, 3L, 3L, 4L, 0L, 2L, 2L, 1L, 1L, 2L, 0L,
    2L, 1L, 1L, 0L, 0L, 1L, 1L, 2L, 2L, 1L, 1L, 1L, 1L, 0L, 0L,
    0L, 1L, 1L, 0L, 0L, 0L, 0L, 0L)), .Names = c("Days", "Students"
```

Type here to search

ENG 17:14 12/03/2023

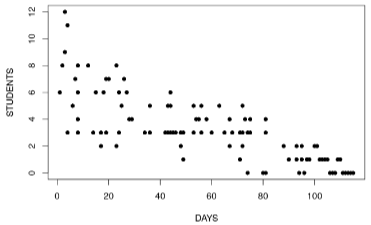
R code for modeling

http://www.john-ros.com/Rcourse/glm.html#poisson-regression

6.1 Problem Setup
6.2 OLS Estimation in R
6.3 Inference
6.4 Extra Diagnostics
6.5 Bibliographic Notes
6.6 Practice Yourself
7 Generalized Linear Models
7.1 Problem Setup
7.2 Logistic Regression
7.3 Poisson Regression
7.4 Extensions
7.5 Bibliographic Notes
7.6 Practice Yourself
8 Linear Mixed Models
8.1 Problem Setup
8.2 LMMs in R
8.3 Serial Correlations in Space/T...
8.4 Extensions
8.5 Bibliographic Notes
8.6 Practice Yourself
9 Multivariate Data Analysis
9.1 Signal Detection
9.2 Signal Counting
9.3 Signal Identification
9.4 Signal Estimation (*)
9.5 Bibliographic Notes

- What is the sign of the effect of time on the number of sick students?
- Can we assume a linear effect of time?

```
plot(Days, Students, xlab = "DAYS", ylab = "STUDENTS", pch = 16)
```



We now fit a model to check for the change in the rate of events as a function of the days since the outbreak.

```
glm.3 <- glm(Students ~ Days, family = poisson)
summary(glm.3)
```

```
##
## Call:
## glm(formula = Students ~ Days, family = poisson)
##
## Deviance Residuals:
##      Min        1Q    Median        3Q       Max
## -2.00482  -0.85719  -0.09331   0.63969   1.73696
##
## Coefficients:
```

Stay connected

- Look for examples in our website:

https://erbiostat.wixsite.com/erbiostat

- Want to take part in our initiative ?
- Want to include your university in our initiative ?
- Want to use our courses for your classes ?
- Send us an email !!!

ziv.shkedy@uhasselt.be

- Follow us on social media:

Visit us on Facebook ER-BioStat & twitter @erbiostat

UHASSELT

I-BioStat

KU LEUVEN

Interuniversity Institute for Biostatistics and statistical Bioinformatics

vlir uos

SHARING MINDS, CHANGING LIVES

Thank you ve y much !!

Join our community and look for us online:

Visit us on Facebook ER-BioStat

GitHub https://github.com/eR-Biostat

https://erbiostat.wixsite.com/erbiostat

twitter @erbiostat