

Basic Statistics for Biomedical Research

Santiago Perez-Hoyos, Miriam Mota, Marcos Esteve

Statistics and Bioinformatics Unit. Vall d'Hebron Institut de
Recerca

Readme

- License: Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License <http://creativecommons.org/licenses/by-nc-sa/4.0/>
- You are free to:
 - **Share** : copy and redistribute the material
 - **Adapt** : rebuild and transform the material
- Under the following conditions:
 - **Attribution** : You must give appropriate credit, provide a link to the license, and indicate if changes were made.
 - **NonCommercial** : You may not use this work for commercial purposes.
 - **Share Alike** : If you remix, transform, or build upon this work, you must distribute your contributions under the same license to this one.

Section 1

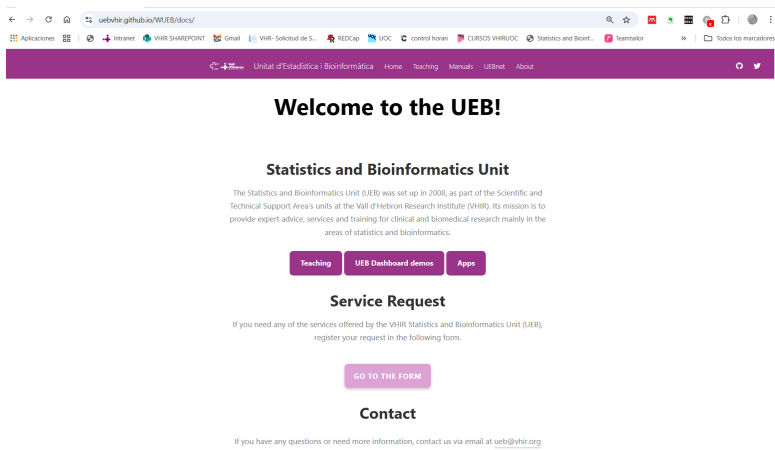
Introduction

Outline

- Introduction
 - Who are we (“we”=teachers & students)
 - Why are we here (Why learn R?)
- How will we proceed: Methodology
- HW Data Science approach to using R
- References & Resources

Who are we: The Statistics and Bioinformatics Unit

<https://uebvhir.github.io/WUEB/docs/>

A screenshot of a web browser displaying the homepage of the UEB (Statistics and Bioinformatics Unit). The browser's address bar shows the URL 'uebvhir.github.io/WUEB/docs/'. The page has a purple header with navigation links: 'Unitat d'Estadística i Bioinformàtica', 'Home', 'Teaching', 'Manuals', 'UEBnet', and 'About'. The main content area features a large heading 'Welcome to the UEB!' followed by 'Statistics and Bioinformatics Unit'. A paragraph describes the unit's establishment in 2008 and its mission. Below this are three purple buttons: 'Teaching', 'UEB Dashboard demos', and 'Apps'. A 'Service Request' section follows, with a paragraph explaining the unit's services and a 'GO TO THE FORM' button. The 'Contact' section at the bottom provides an email address: 'uebvhir@vhir.org'.

uebvhir.github.io/WUEB/docs/

Aplicaciones Intranet VHIR SHAREPOINT Gmail VHIR- Solicitud de S... REDCap UOC control horari CURSOS VHIRUOC Statistics and Bioin... Teambitor Todos los marcadores

Unitat d'Estadística i Bioinformàtica Home Teaching Manuals UEBnet About

Welcome to the UEB!

Statistics and Bioinformatics Unit

The Statistics and Bioinformatics Unit (UEB) was set up in 2008, as part of the Scientific and Technical Support Area's units at the Vall d'Hebron Research Institute (VHIR). Its mission is to provide expert advice, services and training for clinical and biomedical research mainly in the areas of statistics and bioinformatics.

Teaching UEB Dashboard demos Apps

Service Request

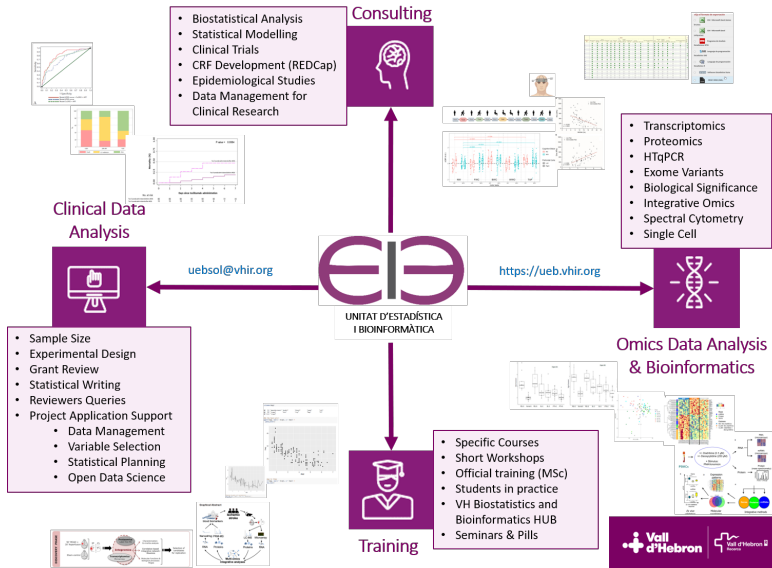
If you need any of the services offered by the VHIR Statistics and Bioinformatics Unit (UEB), register your request in the following form.

GO TO THE FORM

Contact

If you have any questions or need more information, contact us via email at uebvhir@vhir.org

What we do: Teachers



Why this course (1)



“We are drowning in information but starving for knowledge”

Why this course (2)

- **(Biomedical) research**, as well as many other human activities (social networks, sports, COVID ...) generate **huge quantities** of -often complex- **data**.
 - Although sometimes we will also have small datasets
- We believe that data leads to information that leads to knowledge, but we need to be able to extract one from the other.
- This can be attempted in many ways, artificial intelligence, machine learning, data science or something which is common to all of them: **Plain Statistics!!!**

What are our goals (1)

- The main goal of this course is to introduce a variety of statistical methods and tools, which is good enough to:
 - Help you analyze your own data when it makes sense
 - Suggests you when the analysis is complex enough to contact an expert statistician (such as those in the UEB)
 - Help you to distinguish one from the other
- A secondary, but not least important objective: Show how to do it using R

Readme

- License: Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License <http://creativecommons.org/licenses/by-nc-sa/4.0/>
- You are free to:
 - **Share** : copy and redistribute the material
 - **Adapt** : rebuild and transform the material
- Under the following conditions:
 - **Attribution** : You must give appropriate credit, provide a link to the license, and indicate if changes were made.
 - **NonCommercial** : You may not use this work for commercial purposes.
 - **Share Alike** : If you remix, transform, or build upon this work, you must distribute your contributions under the same license to this one.

Section 2

Introduction

Outline

Sessió	Data	Modalitat	Professor	Títol
1	20/05	En línia	Santi	Introducció a R i RStudio i Tidyverse
2	21/05	En línia	Miriam	Estadística descriptiva I: Resums numèrics, taules i gràfics.
3	22/05	En línia	Miriam	Estadística descriptiva II: Gràfics i taules bivariants. ggplot2
4	27/05	En línia	Miriam	Més manegament de dades amb R. Automatització amb scripts
5	29/05	Presencial	Santi	Introducció a la inferència estadística. Intervals de confiança
6	3/06	En línia	Santi	Proves d'hipòtesis I: Conceptes bàsics
7	4/06	En línia	Santi	Proves d'hipòtesis II: Variables quantitatives
8	5/06	En línia	Miriam	Proves d'hipòtesis III: Variables qualitatives
9	10/06	En línia	Miriam	Tests diagnòstics: Sensibilitat, especificitat i corbes ROC
10	12/06	Presencial	Santi	Exercici resum d'anàlisis de dades reals

Horari: 13:00h a 15:00h

Sessions presencials: *Sala de Junes*

Sessions en línia: *via plataforma virtual*

Course contents

- Sessió 1. Introducció a R i RStudio i Tidyverse.
- Sessió 2. Estadística descriptiva I: Resums numèrics, taules i gràfics. Reproduïbilitat amb R
- Sessió 3. Estadística descriptiva II: Gràfics i taules bivariants. Visualització de dades amb ggplot2.
- Sessió 4. Més manegament de dades amb R. Automatització de tasques amb scripts..
- **Sessió 5. Sessió presencial de pràctiques en R i resolució de dubtes**
- Sessió 6. Introducció a la inferència. Intervals de Confiança.
- Sessió 7. Proves d'hipòtesis I: Conceptes bàsics
- Sessió 8. Proves d'hipòtesis II: Variables quantitatives.
- Sessió 9. Proves d'hipòtesis III: Taules de contingència, χ^2 , Tests diagnòstics: Sensibilitat, especificitat i corbes ROC.
- **Sessió 10. Sessió presencial de resolució d'exercici d'anàlisi de dades reals.**
- If there is anything else you would like to learn, let us know and we'll try not to let it out.

Why doing statistics with R

- R has become a “de facto” standard for statistical analysis
- R is free to use!!
- Practically all existing statistical methods available
- Powerful graphics that can be used interactively to explore data
- Possibility of scripting analysis → **Reproducibility**
- Enormous and collaborative Community

How we will work

- ① Our session will have the following structure (all but the first)
 - 1st part: We introduce a few new theoretical ideas
 - 2nd part: We wil show how to do it in R
 - 3rd part: Practice exercises and start working on the case study suggested or if you want in your data.
- ② Mastering R requires as many other disciplines
 - Time
 - Study, and
 - Practice.

So take it easy. You can go back to R once the course finish

Class attendance and participation (class attendance should be at least 80%)

The materials

You can find the presentations, scripts , exercises in the teaching section of our web ueb.vhir.org

Teaching Activities at the UEB

Welcome to the Statistics and Bioinformatics Unit Teaching Activities web site. This page links with the web pages of different courses we are teaching right now or have taught in the past. In these pages you will find all the materials we use for the courses -unless of course they are copyrighted or under some type of confidentiality.

Current Activity

[2025-05: Curs d'Estadística Bàsica per a la Recerca Biomèdica](#)

Past Courses

- Resources and references

There is a huge variety of resources to learn R, books, tutorials, free online courses, etc.

- This course is based on the book Data Science for R.
- Other interesting books + Using R and RStudio for Data Management, Statistical Analysis, and Graphics, 2nd edition
- Online courses + Coursera's Data Science Specialization
- A list of R tutorials and courses

Materials del curs

Sessió 0 : Presentació del curs

- [Presentació](#)

Sessió 1. Introducció a R i RStudio i Tidyverse

- [Presentació](#)
- Datasets
 - [Osteoporosis \(csv\)](#)
 - [Diabetes \(Excel\)](#)

Sessió 2. Estadística descriptiva I: Resums numèrics, taules i gràfics. Reproduïbilitat amb R

- [Presentació](#)
- [Notebook](#)