

Swiss Institute of  
Bioinformatics

# Data Science in Practice



Isabelle Dupanloup, Rachel Marcone, Frédéric Schütz

April 2022

# About this course-What you will learn:

- Become a real-life bioinformatician
- Work on real-life datasets, with its problems (can be very standard but also less standard problems)
- Get to work in groups to discuss statistical but also computational issues.
- Find where to find answers to your problems.
- Practice data analysis.
- Learn from your neighbors.
- Don't be shy, each one of you will have something to add/offer to any discussion!

# About this course- What you will NOT learn

- Basic introduction to statistics.
- More « intermediate » statistical concepts.
- Standard course with concepts followed by application of the concepts.

# Schedule-tentative

## Day 1:

9h: Welcome - Get to know

9h20: Introduction

10h00: coffee break

10h30-12h00 : qPCR

12h00-13h00 : Lunch

13h00: knitr

13h30: Exercises

14h30-coffee break

14h45- Presentation - GROUP 1

15h30- Dataset 2

17h00- end of day 1

## Day 2:

9h: Welcome - Presentation - GROUP 2

9h45-10h30: Theory

10h30-10h45 -coffee break

10h45-11h45- Dataset 2 (again)

11h45-12h15 Presentation - GROUP 3

12h15-13h15 Lunch

13h15- 14h30 Simulations

14h30-15h00 Presentation-GROUP 4

15h00-16h30 BC

16h30-17h00 Presentation-GROUP 5

# Bioinformatics Core Facility (BCF) soon to be under translational data science TDS



[Home](#)

## Mauro Delorenzi & Frédéric Schütz's group

In the Bioinformatics Core Facility (BCF), we promote trans-disciplinary collaborations between research teams in medicine, molecular biology, genetics, genomics, statistics, and bioinformatics...

We are happy to hear about your projects and problems!  
We work as a facility with transdisciplinary collaborations.  
We are a team of bioinformaticians coming from very different backgrounds.

**CONTACT US!**



# BIOSTATISTICS PLATFORM

Faculty of Biology and Medicine, UNIL

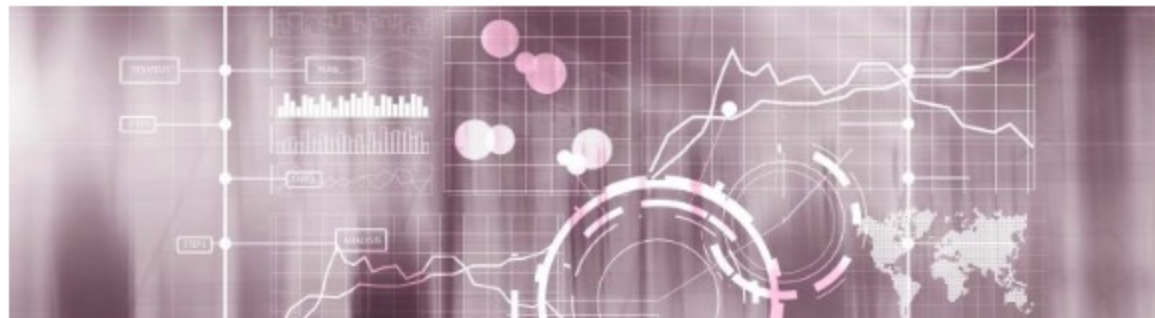
[Home](#)[Presentation](#)[Services](#)[Collaborations](#)[Team and contact](#)[More information ▾](#)

## HOME

<https://wp.unil.ch/biostatistics/>

The Biostatistics platform of the Faculty of Biology and Medicine acts as a main entry point for all your questions related to biostatistics and data analysis. If needed, the platform can redirect you to the people with the best competences in order to answer your questions.

Contact: [Frederic.Schutz@unil.ch](mailto:Frederic.Schutz@unil.ch)



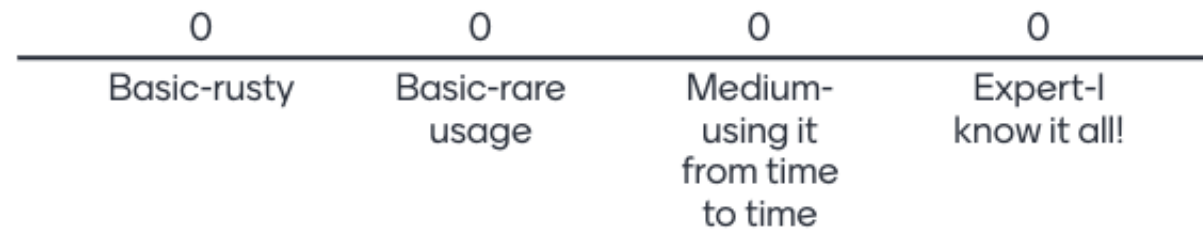
# The teachers- Presentation

- Isabelle Dupanloup
- Rachel Marcone
- Frédéric Schütz

Go to [www.menti.com](https://www.menti.com) and use the code 4032 9138

# My R knowledge

 Mentimeter





Go to [www.menti.com](https://www.menti.com) and use the code 4032 9138

# I am a



# Now a round of table

- Let us know who you are, your background, what you are working on and where, and what type of data you might have to analyse (at some point) or what you expect from this course.

# Small reminder of Introduction to statistics

- \* This is not a full reminder \*
- Start with understanding your data and plotting it *in the right way*.
- Put the question in the form of a biological null hypothesis and alternate hypothesis.
- Put the question in the form of a statistical null hypothesis and alternate hypothesis.
- Choose the appropriate statistical test for your question.
- Test hypotheses of the test you have performed.
- Discuss the results.

# Small reminder of introduction to statistics

- Possible tests you have learned during an Intro to stats course.
- One-sample t-test, two-sample t-test, paired t-test
- What to do with more than two groups comparison (ANOVA, multiple t-tests with corrections etc)
- Parametric and non parametric tests (or what to do if data is not normally distributed, and you think a non parametric test would suit your question), like Wilcoxon rank-sum test.
- The importance of multiple testing corrections
- Correlation/regression as well as clustering and PCA.

# What to do if you already feel overwhelmed:

- Ask your neighbors (Interactive course)
- Do a quick search for refreshment on google on some of the stats concepts
- Ask us!