Data Analysis for Cognitive Neuroscience

Universitat Pompeu Fabra Barcelona

Klaus Wimmer
kwimmer@crm.cat

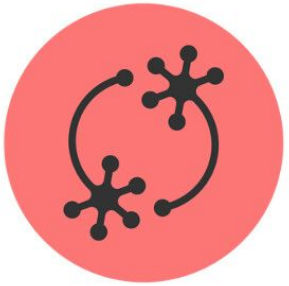
Alexandre Hyafil
ahyafil@crm.cat

CENTRE DE RECERCA MATEMÀTICA

Inspirations

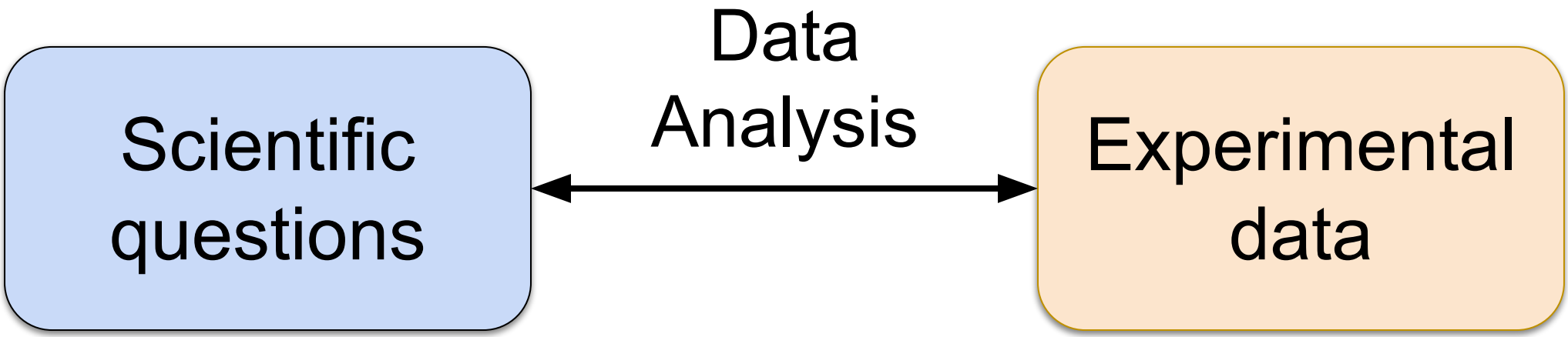
Albert Compte

www.bambschool.org



neuromatch
academy

https://academy.neuromatch.io/


Scientific questions
Data Analysis
Experimental data

What is the aim of this course?

- Data analysis is fun!
- Data analysis is creative!
- Data analysis has constraints:
  - rigor
  - some structure
    - data checking
    - preprocessing
    - exploratory data analysis
    - statistical inference
    - model-based analyses
    - presentation of results

- Learn specific general-purpose techniques

How will we achieve that??

- Working on **real data** as examples
- Using **Python** as a programming language

- We will learn: statistical inference, linear models, logistic regression, dimensionality reduction, cluster analysis (exploratory), spike train analysis

- A specific homework/class dynamics:
  - assignments to prepare at home and discuss in class
  - online quizzes

Classroom / Assignments dynamics

- Before each class: we'll send the theoretical material (mostly videos) to be watched *prior* to the class
- During the class:
  - we'll make a brief recap and address your questions
  - Then we'll start doing the tutorial in class
- After the class: You will have to complete it and send it before the next class

Why learn coding?

- Because it opens a world of opportunities for data analysis, programming experiments, etc., that are simply not possible with button-clicking solutions (SPSS, Presentation,...)
- Because you reach a deeper understanding of the analysis
- Because this has become a very valuable asset in the job market (specifically Python for data scientists)

Schedule (preliminary)

Date	Topic	Discuss in class/Submit	Get
January 12	Introduction to Python		Assignment 0
January 19	Descriptive statistics	Assignment 0	Assignment 1
January 26	Linear models		
February 2	Models of choices / Generalized linear models	Assignment 1	Assignment 2
February 9 (3h)	Statistical inference	Assignment 2	Assignment 3
~~February 16~~	no class	-	-
February 23 (3h)	Spike train analysis	Assignment 3	Assignment 4
March 2	Spike train analysis / Decoding analyses		
March 9	Dimensionality reduction	Assignment 4	Assignment 5
March 16	Journal club (paper discussion)	Assignment 5	
April ?	Exam		

General Course Information

Tauler d'avisos

Schedule, program and evaluation

Course Homepage on Github

All information about the course will be on github. By clicking on the URL you will get to a readme file that contains the information about how to get started in Python and several Python tutorials. In the future, we will post the assignments there as well.

before January 12

Basic Python programming quiz

Marca com a feta

Before the course starts on January 12, do the quick quiz about your Python programming skills, so you know if you need to do an extra push before the class starts (**see the link to our github page above for links to tutorials etc.**). We will not consider this quiz for the final evaluation of the course. It is just for you to see where you stand and what concepts you should be familiar with in order to benefit from the class. If you

https://github.com/wimmerlab/MBC-DataAnalysis/

MBC-DataAnalysis

This is the repository for the materials of the Data Analysis class at the UPF Masters for Brain & Cognition taught by Klaus Wimmer & Alex Hyafil, largely using material from Albert Compte.

Coding in Python

Why should a cognitive psychologist / neuroscientist learn how to code? See the arguments in these different sources:

- "The next generation of neuroscientists needs to learn how to code" article from Neuron journal
- "Why it's important to learn to code in contemporary biology?" article from Wire
- "Why every (psychology) student should learn to code" blog post

You can brush up your Python skills by following one (or some) of these suggested Python tutorials:

- Introduction to Python from the Department of Cognitive Science of UC San Diego
- Python for Everybody Specialization (free Coursera course starts Dec 20)
- LearnPython.org
- Datacamp - Intro to Python for Data Science
- Python training at the Neuromatch Academy: Day 1 and Day 2, tutorials with videos
- Introductory material from the "Advanced Scientific Programming in Python" summer school
- If your master's project relies heavily on programming or you continue along with a PhD, we strongly recommend reading quietly through Patrick Mineault's wonderful Good Research Code Handbook

Evaluation

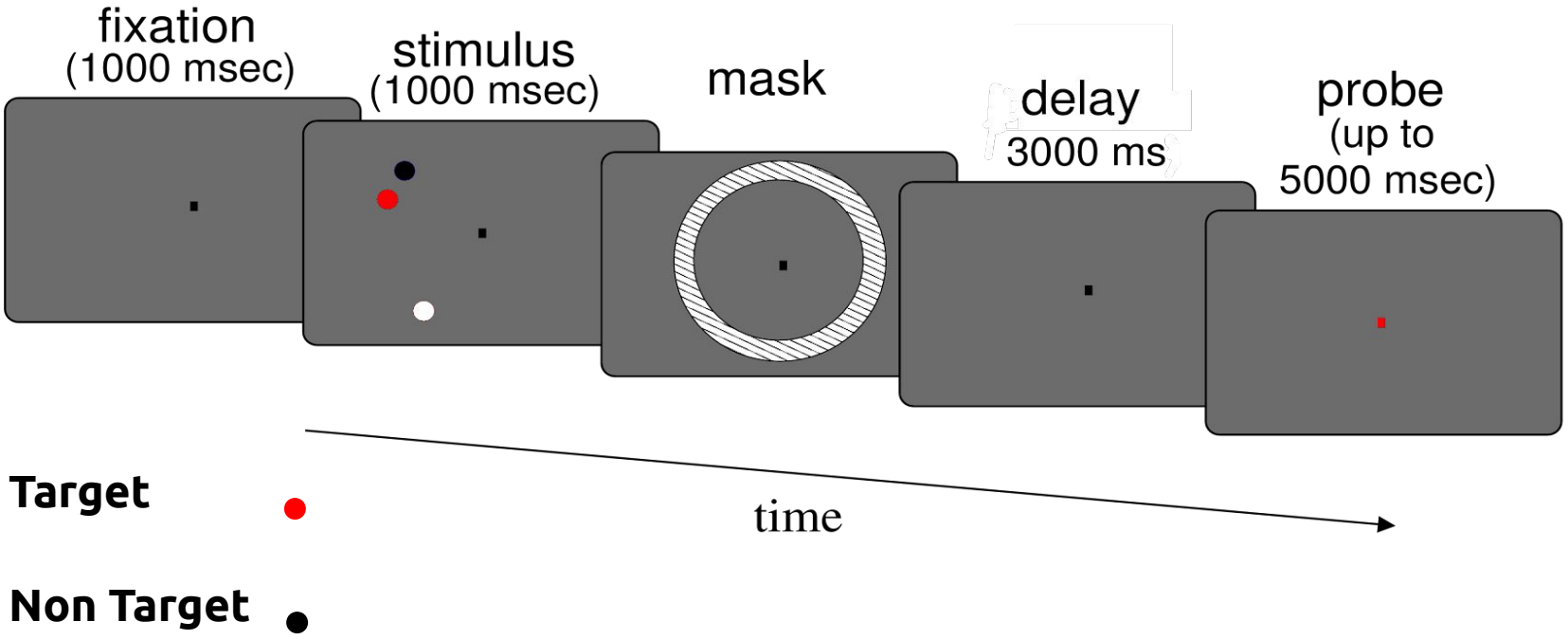
The grade of the course will be computed based on:

1) **completed assignments** after revision in class. Students will submit the code generated for each assignment, after it has been revised in class. There will be an online test in the following week. Within-deadline submission and quality of the assignments and tests will count towards **60% of the grade**.

2) **final exam**. At the end of the 2nd trimester there will be an exam, it will include both a Python Notebook assignment and an online test similar to what we will have done through the course. It will count towards **40% of the grade**.

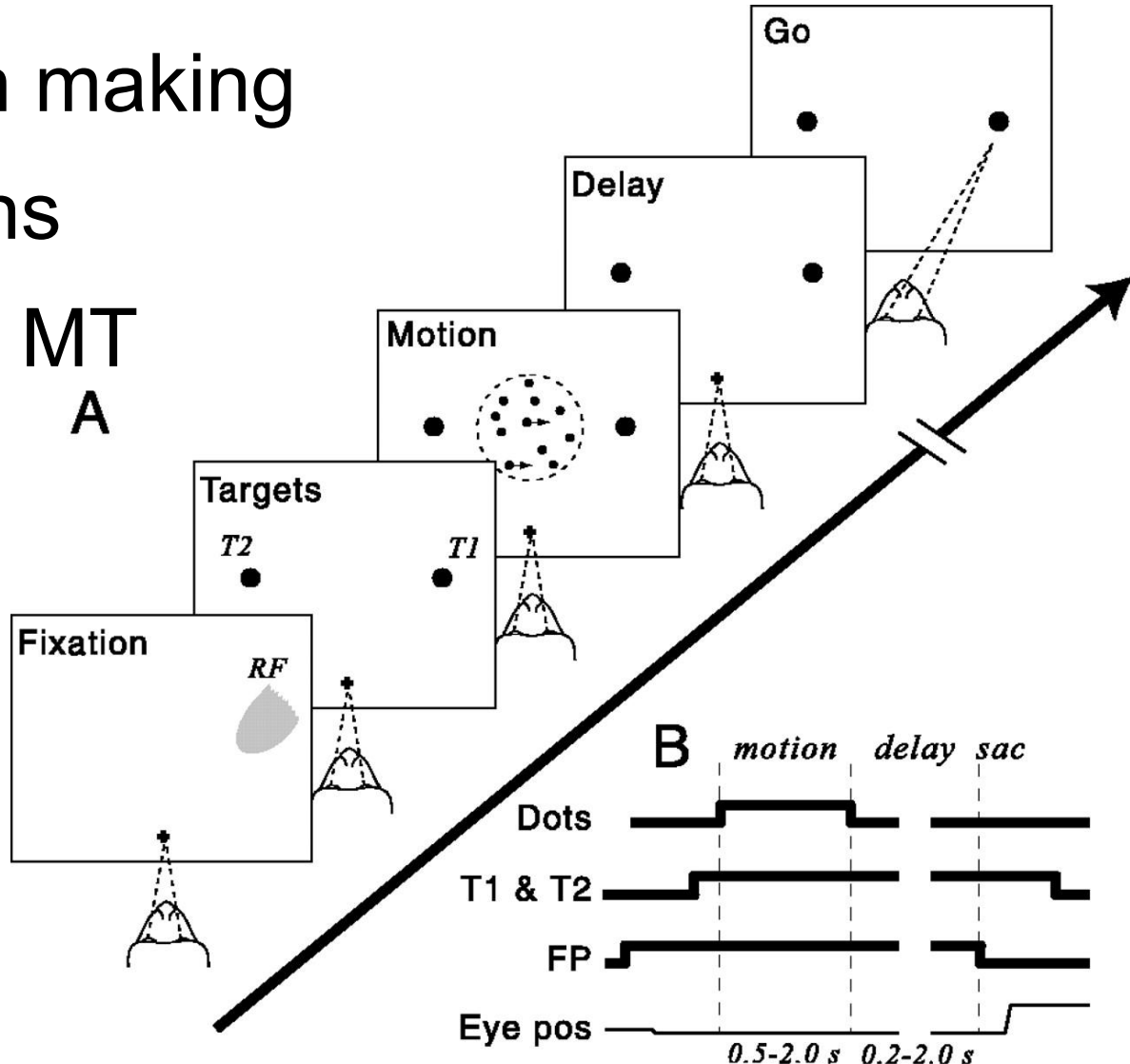
Dataset 1: psychophysics data

- Almeida et al. *J Neurophysiol* 2015
- 2 Working Memory (WM) experiments: Yes/No and continuous report
- 9 subjects



Dataset 2: spiking data

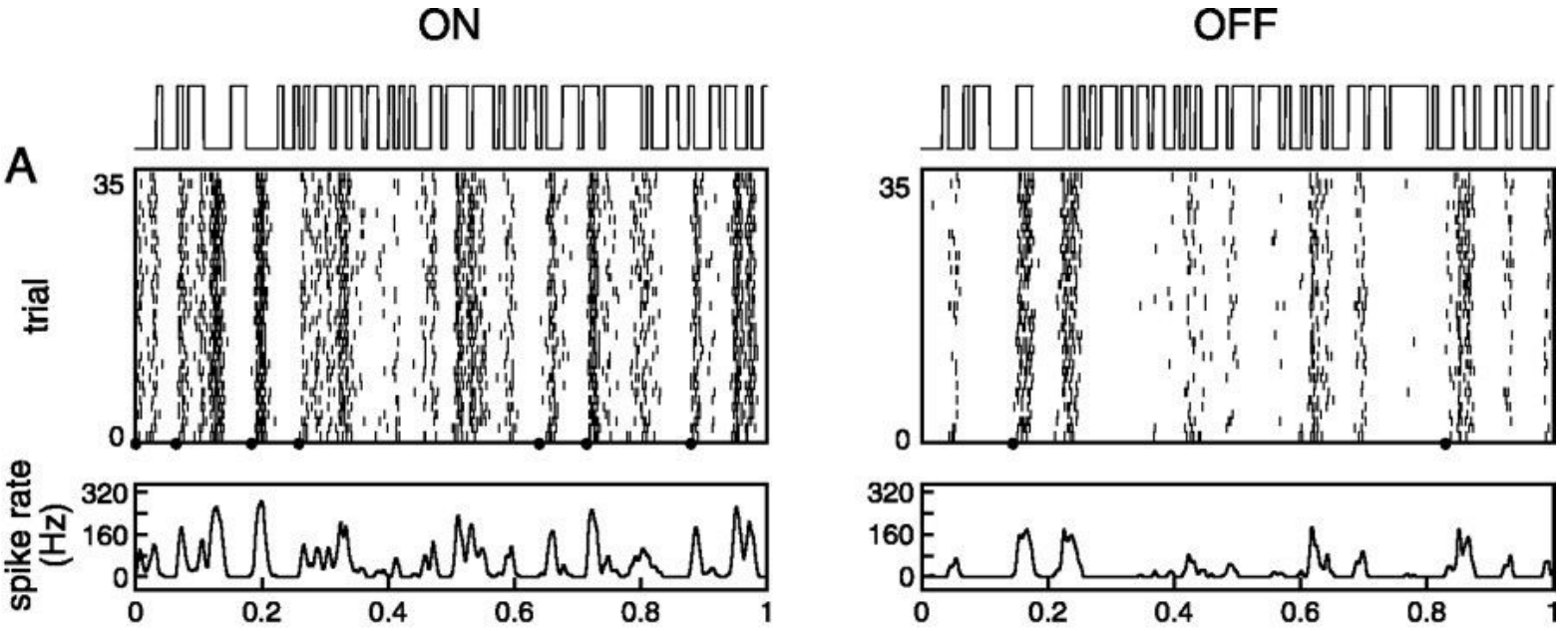
- Zohary et al. Nature 1994
- Perceptual decision making
- Random dot patterns
- Single-units in area MT
- 76 neurons,
- 41 pairs of neurons



Dataset 2: spiking data

Uzzell & Chichilnisky *J Neurophysiology* 2004

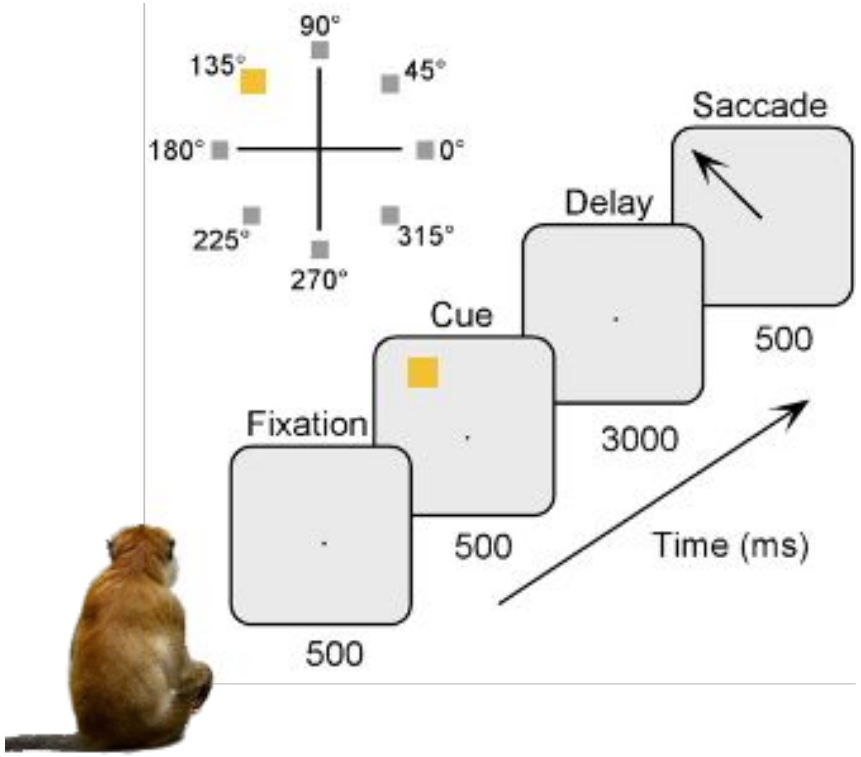
- Spatially uniform display
- Spiking activity of macaques' retinal ganglion cells
- 19 On-cells and 22 Off-cells

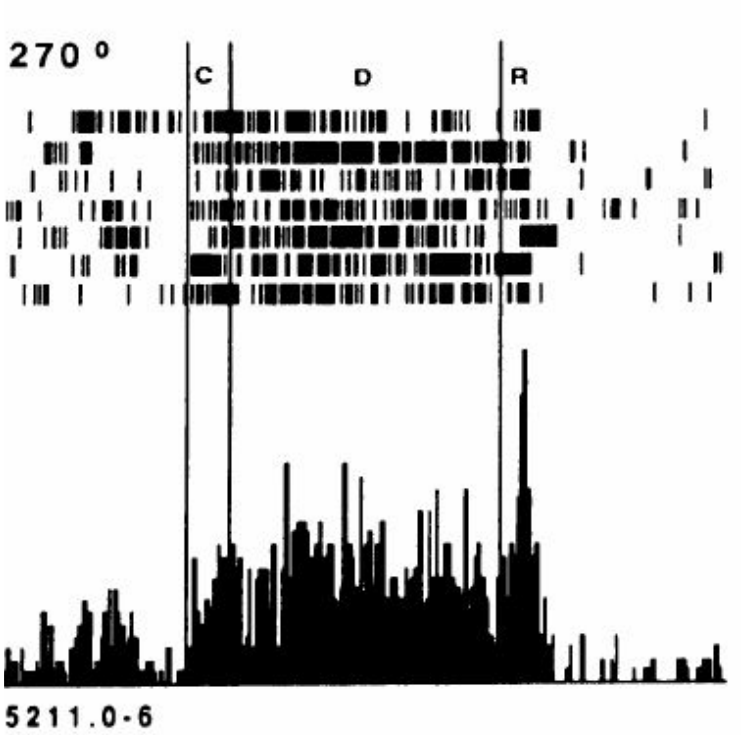


Dataset 3: working memory

Constantinidis et al. *J. Neuroscience* 2001

>200 neurons recorded from prefrontal cortex





Assignment 0
Things to learn about Python

- Basic math
- Matrices, vectors
- Indexing
- Graphics
- Functions and scripts

Assignment 1

- Descriptive statistics
  - Measures of centrality
  - Measures of dispersion
- Statistical inference
  - T-tests
  - Permutation tests
- Linear regression