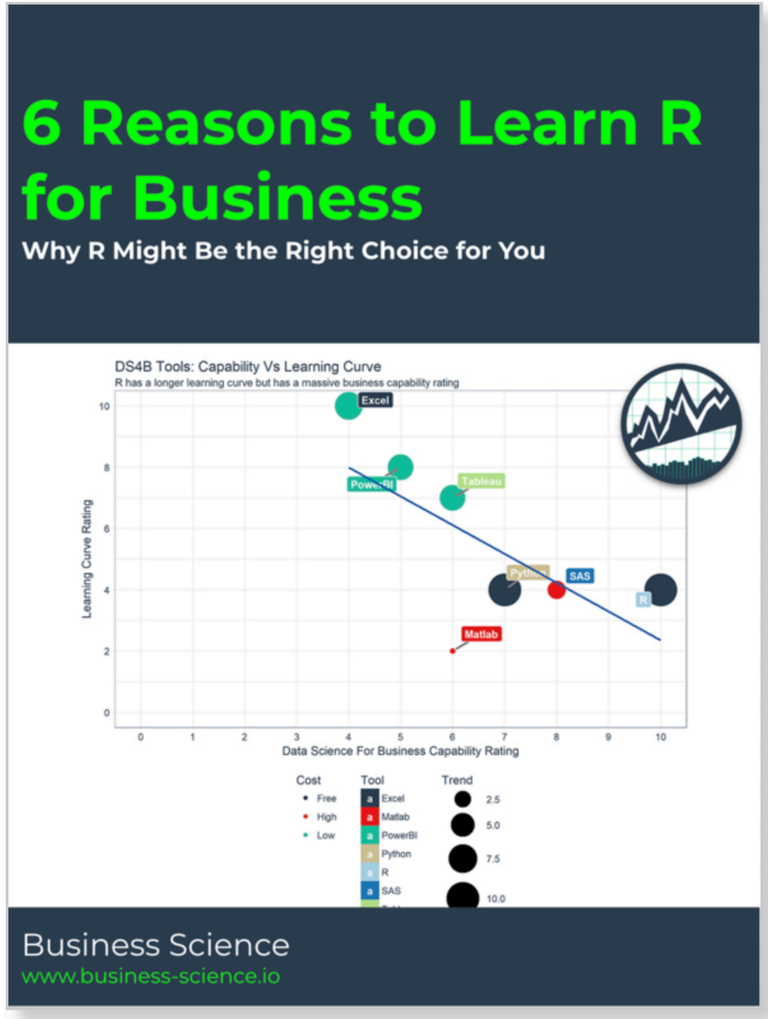
Data Analysis in Biology

BIO144
Prof. Owen Petchey
and others

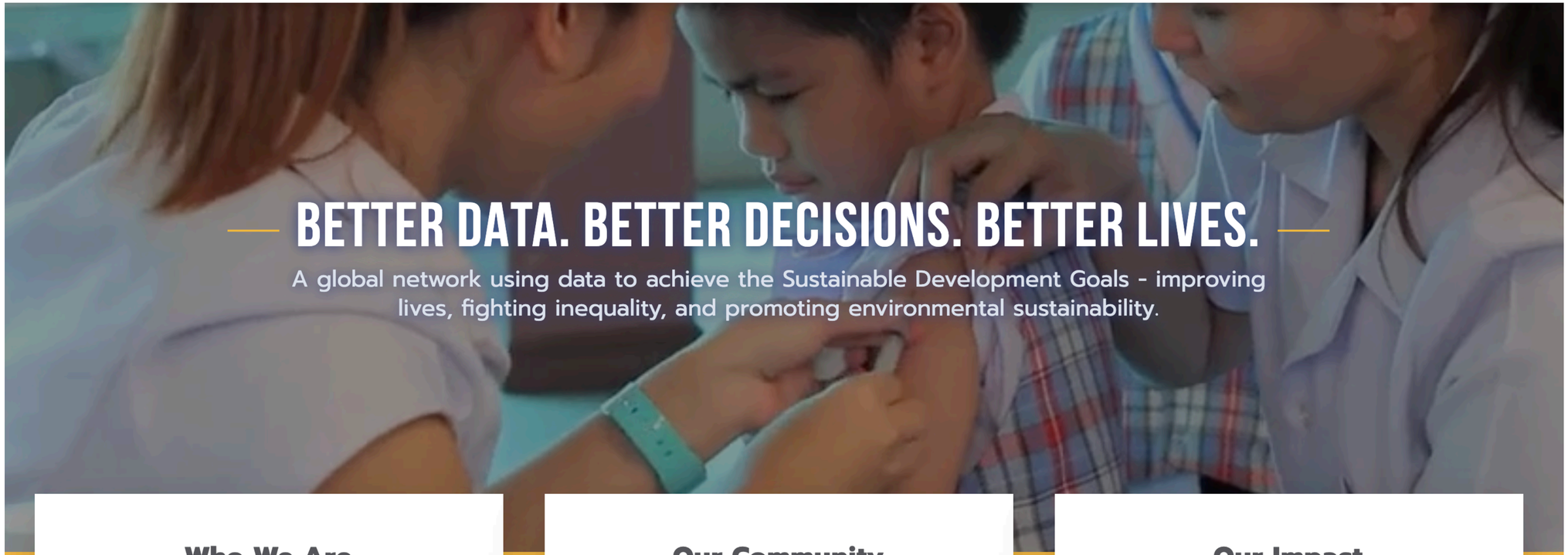

The "hottest skill"
that got people
hired in 2014?
Statistical Analysis
Source: LinkedIn
0:06 / 3:00

6 Reasons To Learn R For Business [2021]

Written by Matt Dancho on December 17, 2020



Global Partnership for Sustainable Development Data



BETTER DATA. BETTER DECISIONS. BETTER LIVES.

A global network using data to achieve the Sustainable Development Goals - improving lives, fighting inequality, and promoting environmental sustainability.

Who We Are

The Global Partnership for Sustainable Development Data is a global network working together to ensure the new opportunities of the data revolution are used to achieve the Sustainable Development Goals.

ABOUT US

Our Community

Our hundreds of partners from governments, the private sector, and civil society organizations are joining forces to take action, galvanize political commitment, build trust, and spur innovation in the booming data ecosystems of the 21st century.

SEE OUR PARTNERS

Our Impact

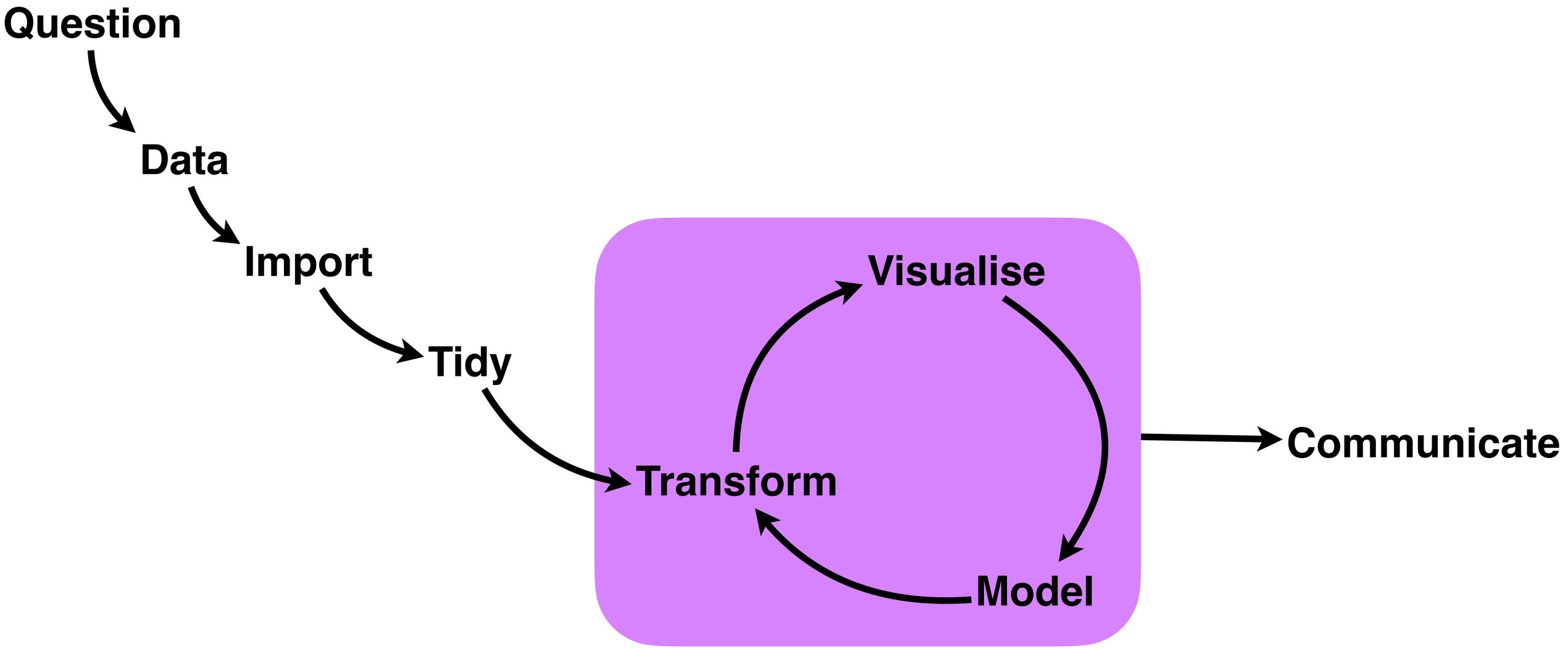
Since our founding in 2015, our network has improved data to monitor and achieve the Sustainable Development Goals, created incentives for new commitments to fund and share data, and enabled knowledge-sharing, bringing partners together to make change.

LEARN MORE

Question
Puzzle
Problem

Data
+
Analysis

Answer
Solution

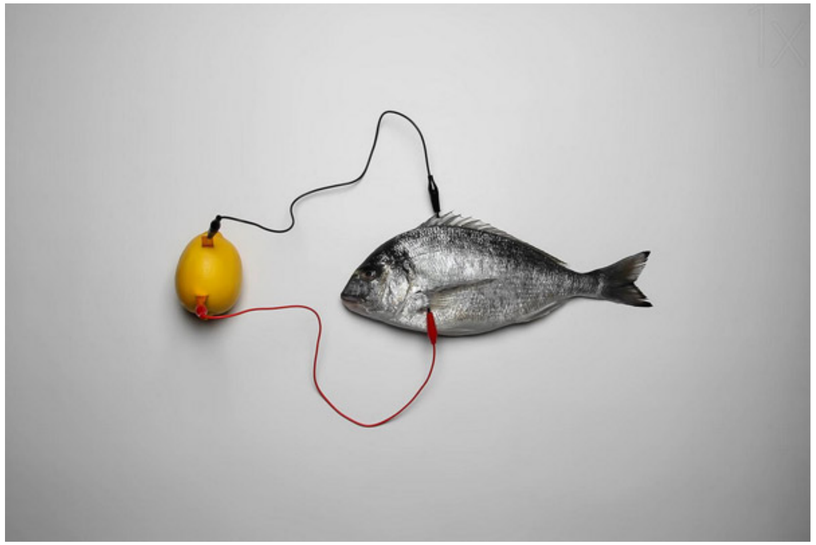

Question
Data
Import
Tidy
Visualise
Transform
Model
Communicate

Efficient
Consistent
Repeatable
Reliable
Readable
Robust
Persistent
Sharable
Scalable

=

Equip you with the knowledge of how to learn more.
And the confidence that you can.

The conceptual side

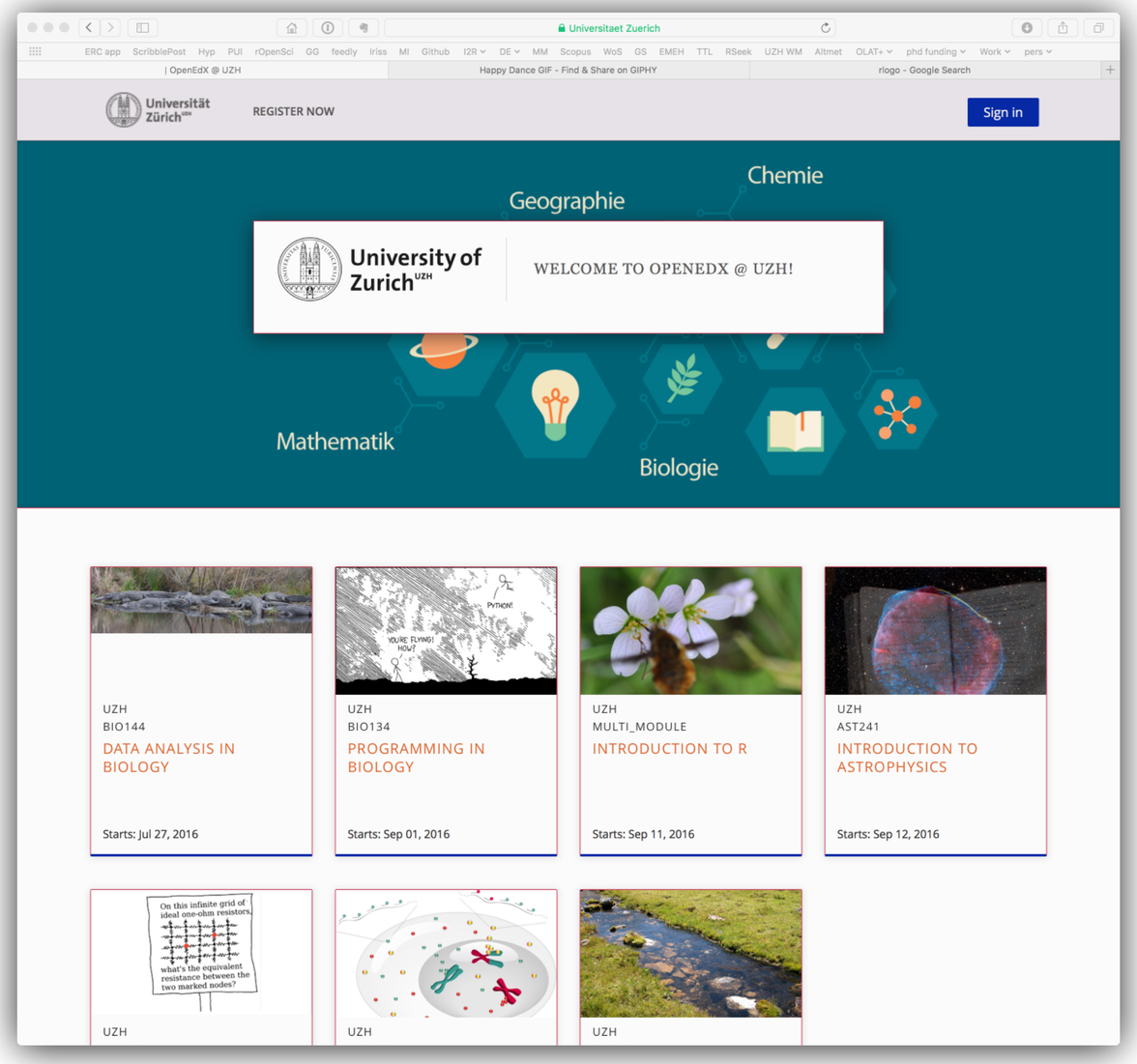


1 - 2:45pm Mondays

The practical side

1 - 3pm Thurs. or Fri.

Details on openedX



Learning objectives
Schedule
Weekly structure / activities
Assessment
Getting help
Giving feedback
Attendance
Email and matriculation number

https://edu-exchange.uzh.ch/

Owen sent a welcome email

Graded assessment questions

Hyperlink

Some “play” questions

Online learning until further notice

Lectures:

- by ZOOM, recorded, available to watch on Switchtube
  (be patient with me getting them posted there).
- Mics off, video on or off.
  Speak up if you have a question
  (I likely won't see a raised hand or ZOOM chat comment).
- Google doc is another way for comments, questions, jokes.
  (show link and doc)

This can be very monotonous:

- reading
- watching videos
- doing online exercises

Online learning until further notice

Practicals:

- You call in to us to ask questions.
- I and a team of TAs will be waiting in online rooms.
- Online rooms will be listed in a google doc (show link and doc)
- You will very likely need to share your screen, perhaps even give control.

> “I really appreciate the effort of making online video calls available, but in this case I didn't really know what exactly to ask.”
> - A student, FS 2020

What worked well last year:

- Some students called in together, pairs, threes, fours (they organised this).
- Some students compiled a list of questions to ask.

Remember:

- We love to talk with you!
- For us, you are the main thing that makes the course interesting!!!
- You can call in to talk about anything… you may have a specific question, but “I just don’t get it” and “my head feels like its gonna explode” are also great ways to start a conversation.
- ***TAs have been trained in how to help.***

Online learning until further notice

This can be rather monotonous

reading
watching videos
doing online exercises

Low diversity in your environment.

We will return to on-site, in-person, when we can.
I do not just assume we must be on-line only.

Lots of good feedback, but here is the more critical/constructive:

Feedback	Mitigation
What is in the course and exam was unclear.	Learning Objectives. Mock exam.
Some variation in amount and type of BC material.	Refer to Learning Objectives for what is core. Please ask if you are uncertain.
A summary of R commands and corresponding library would be good.	You should make one as you go along.
Diverse resources: various books, documents, videos, seem a bit thrown together.	Refer to Learning Objectives for what is core. Please ask if you are uncertain.
Solution scripts: Please give us solution scripts (promptly). Having to wait a week for help is demotivating (please give solution scripts).	There are some. I will try to make more. Not all though… see post in discussion forum Work on the practical parts when the help is available.
Lack of feedback with IC exercises: some do not have instant-feedback, no way to check if we made the correct analysis or graph (give us a solution script). “Felt a bit alone.”	Work on the practical parts at the same time that help is available, i.e. during the scheduled practical sessions.
Amount of work: The IC parts were too short, add 30 mins. All the material is a bit too much. BC reading sometimes rather long. Some variation in amount and type of BC material.	Work steadily, attend live lectures, attend practicals, keep up, ask for help.
Video/podcast quality: in 2020, podcasts from 2019 had to be used. They were of poor quality. In 2021 all lectures will be live and recorded, with good quality.	All lectures will be live (synchronous), will be recorded, and will be made available.
Critical analysis: would be good to work through a paper together, to assess the data analytic methods used.	I have not had time to implement this. It would require removing other content, so is not a simple change to make.
Prof. Petchey talks for hours on end.	Prof. Petchey tries to be more concise.

The team…

You

Director
Instructor

Communications

Statistics consultant

Computing consultant

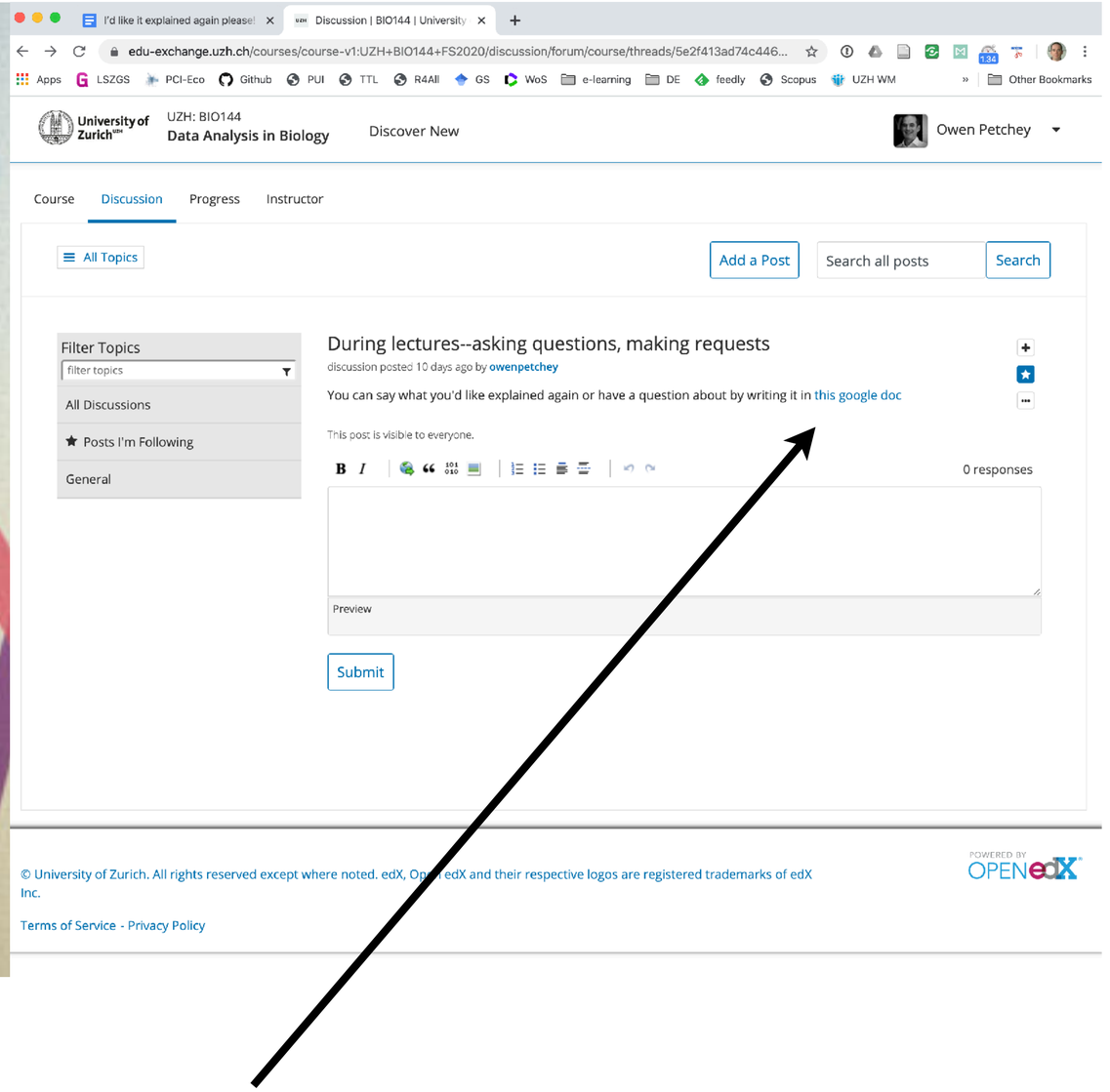
OpenEdx

Assessment
Examinations

Practicals
Teaching Assistants

Comms channels:

- OpenEdx Forum
- bio144uzh@gmail.com
- Lectures & Practicals



Ask questions / make requests here.

switch to Steffi’s lecture

Live data analysis demonstration

BIO144
Week 1

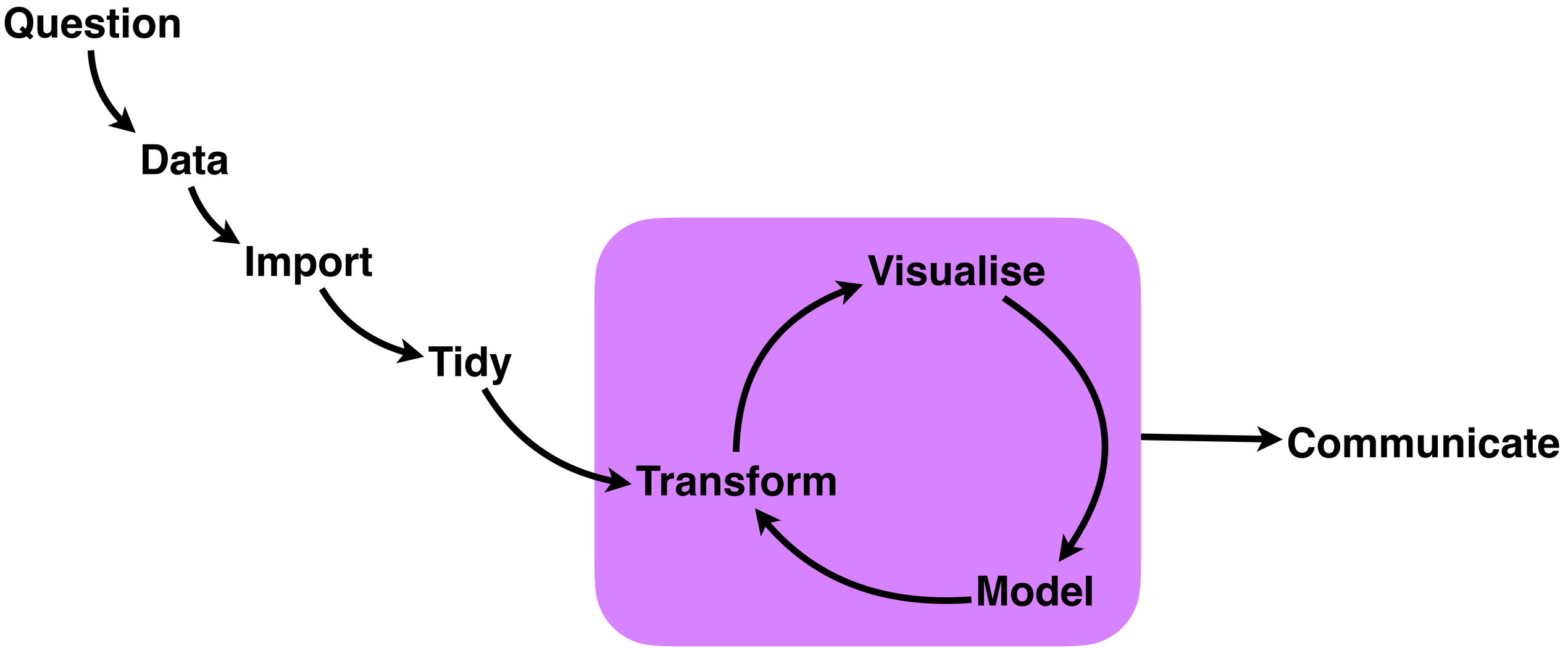
Its a demonstration…

The idea is to give you a feel of what is involved in data analysis.

You will understand some of the demonstration.

You will not understand some of it.

Keep notes about what you don't understand.


Question
Data
Import
Tidy
Visualise
Transform
Model
Communicate

Live data analysis demonstration

The whole data analysis workflow in one hour!!!

- **Question**
- Expectation
- Planned presentation & analysis
- Selection of subjects
- How will data be collected?
- Ethics / permissions
- Data collection
- Data wrangling
- Visualise
- Statistical test
- Critical thinking
- Report / communicate

Live data analysis demonstration

The whole data analysis workflow in one hour!!!

- **Question**
- Expectation
- Planned presentation & analysis
- Selection of subjects
- How will data be collected?
- Ethics / permissions
- Data collection
- Data wrangling
- Visualise
- Statistical test
- Critical thinking
- Report / communicate

The question

- What should be our question?
- As always, there are some influences and some constraints.
- We should ask a question of interest to us, and of some importance.
- And we should be able to collect the data, within our current constraints, necessary to answer the question.
- The question we will address is "***do male and female reaction times of students at the University of Zurich differ?***".
- Why this question? Reaction times are important, safety, sport…

Live data analysis demonstration

The whole data analysis workflow in one hour!!!

- Question
- **Expectation**
- Planned presentation & analysis
- Selection of subjects
- How will data be collected?
- Ethics / permissions
- Data collection
- Data wrangling
- Visualise
- Statistical test
- Critical thinking
- Report / communicate

Expectation

- Quite a lot of work on this already.
- Generally, males tend to have faster reaction times than females. So we expect that to be the same for students at the University of Zurich.
- Given that you know this pattern, and you are the subjects, its interesting to see if you women can buck the trend, perhaps by trying especially hard. Though now the men know you might do this, it probably won't work!

Live data analysis demonstration

The whole data analysis workflow in one hour!!!

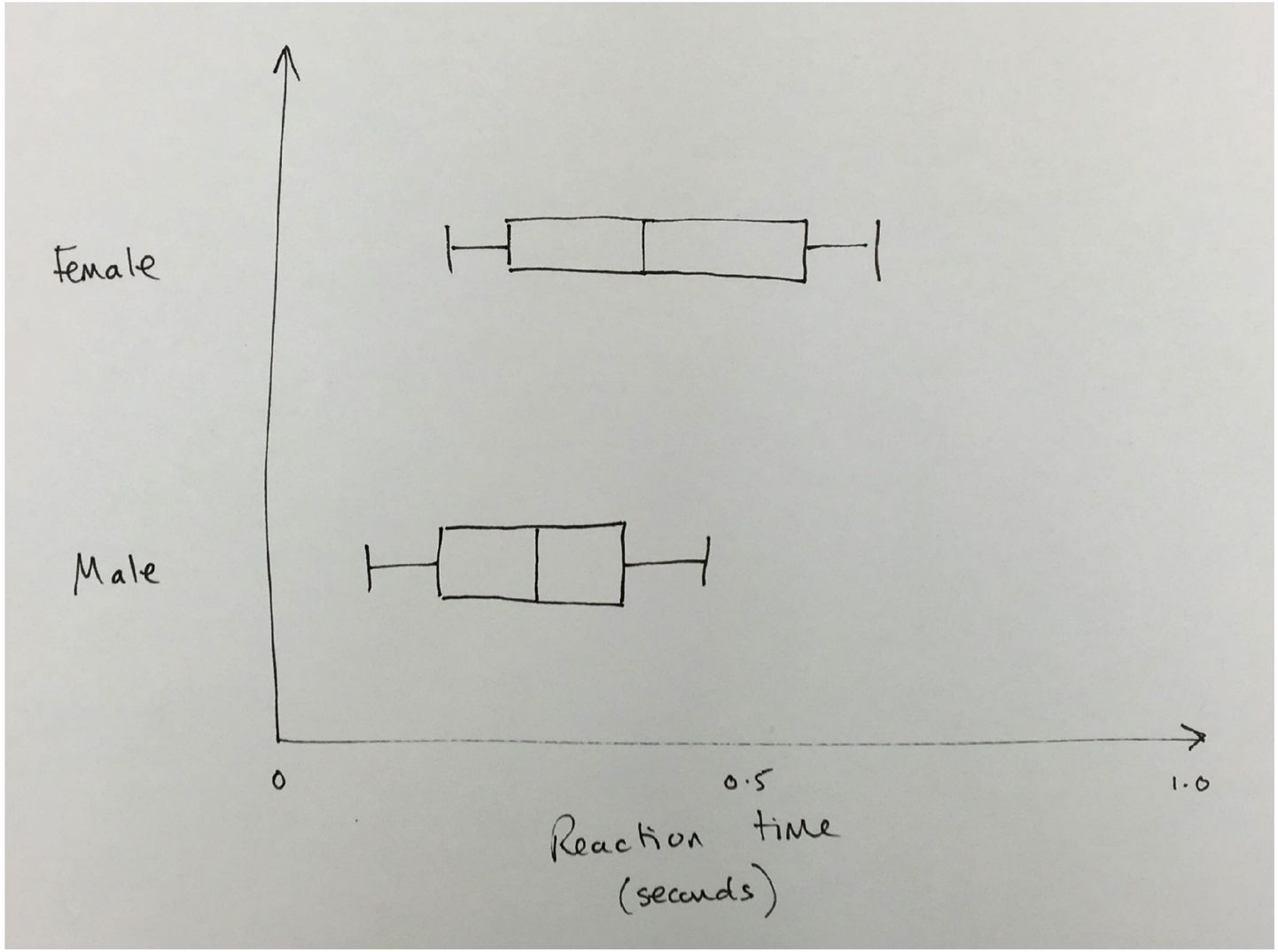
- Question
- Expectation
- **Planned presentation & analysis**
- Selection of subjects
- How will data be collected?
- Ethics / permissions
- Data collection
- Data wrangling
- Visualise
- Statistical test
- Critical thinking
- Report / communicate

What graph?

Live data analysis demonstration

The whole data analysis workflow in one hour!!!

- Question
- Expectation
- **Planned presentation & analysis**
- Selection of subjects
- How will data be collected?
- Ethics / permissions
- Data collection
- Data wrangling
- Visualise
- Statistical test
- Critical thinking
- Report / communicate



Live data analysis demonstration

The whole data analysis workflow in one hour!!!

- Question
- Expectation
- **Planned presentation & analysis**
- Selection of subjects
- How will data be collected?
- Ethics / permissions
- Data collection
- Data wrangling
- Visualise
- Statistical test
- Critical thinking
- Report / communicate

What statistical test?

Live data analysis demonstration

The whole data analysis workflow in one hour!!!

- Question
- Expectation
- Planned presentation & analysis
- **Selection of subjects**
- How will data be collected?
- Ethics / permissions
- Data collection
- Data wrangling
- Visualise
- Statistical test
- Critical thinking
- Report / communicate

Live data analysis demonstration

The whole data analysis workflow in one hour!!!

- Question
- Expectation
- Planned presentation & analysis
- Selection of subjects
- **How will data be collected?**
- Ethics / permissions
- Data collection
- Data wrangling
- Visualise
- Statistical test
- Critical thinking
- Report / communicate

Live data analysis demonstration

The whole data analysis workflow in one hour!!!

- Question
- Expectation
- Planned presentation & analysis
- Selection of subjects
- How will data be collected?
- **Ethics / permissions**
- Data collection
- Data wrangling
- Visualise
- Statistical test
- Critical thinking
- Report / communicate

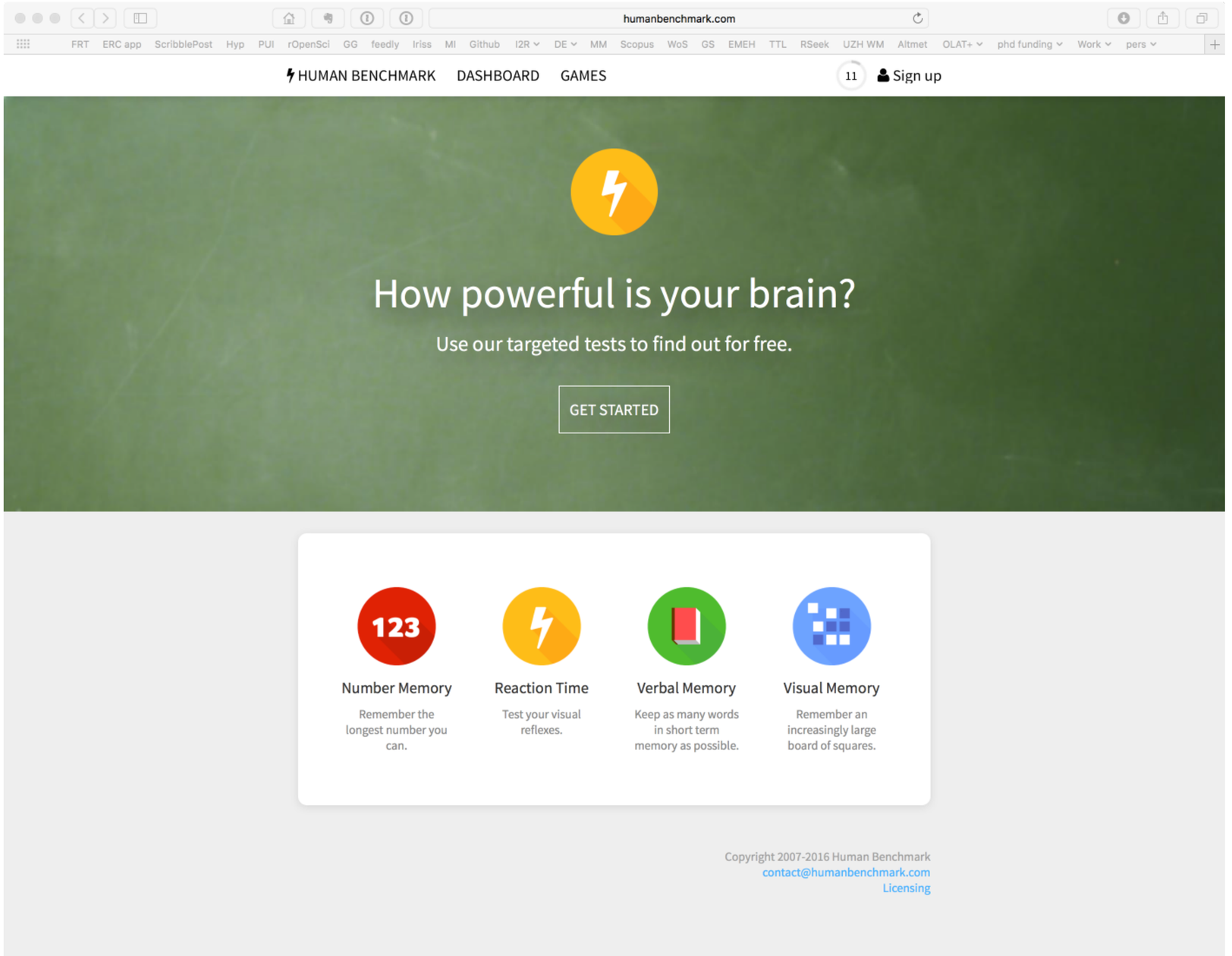
Make up a unique ID code for yourself.
It should not be anything that could identify you.
Keep it safe.

Live data analysis demonstration

The whole data analysis workflow in one hour!!!

- Question
- Expectation
- Planned presentation & analysis
- Selection of subjects
- How will data be collected?
- Ethics / permissions
- **Data collection**
- Data wrangling
- Visualise
- Statistical test
- Critical thinking
- Report / communicate

humanbenchmark.com

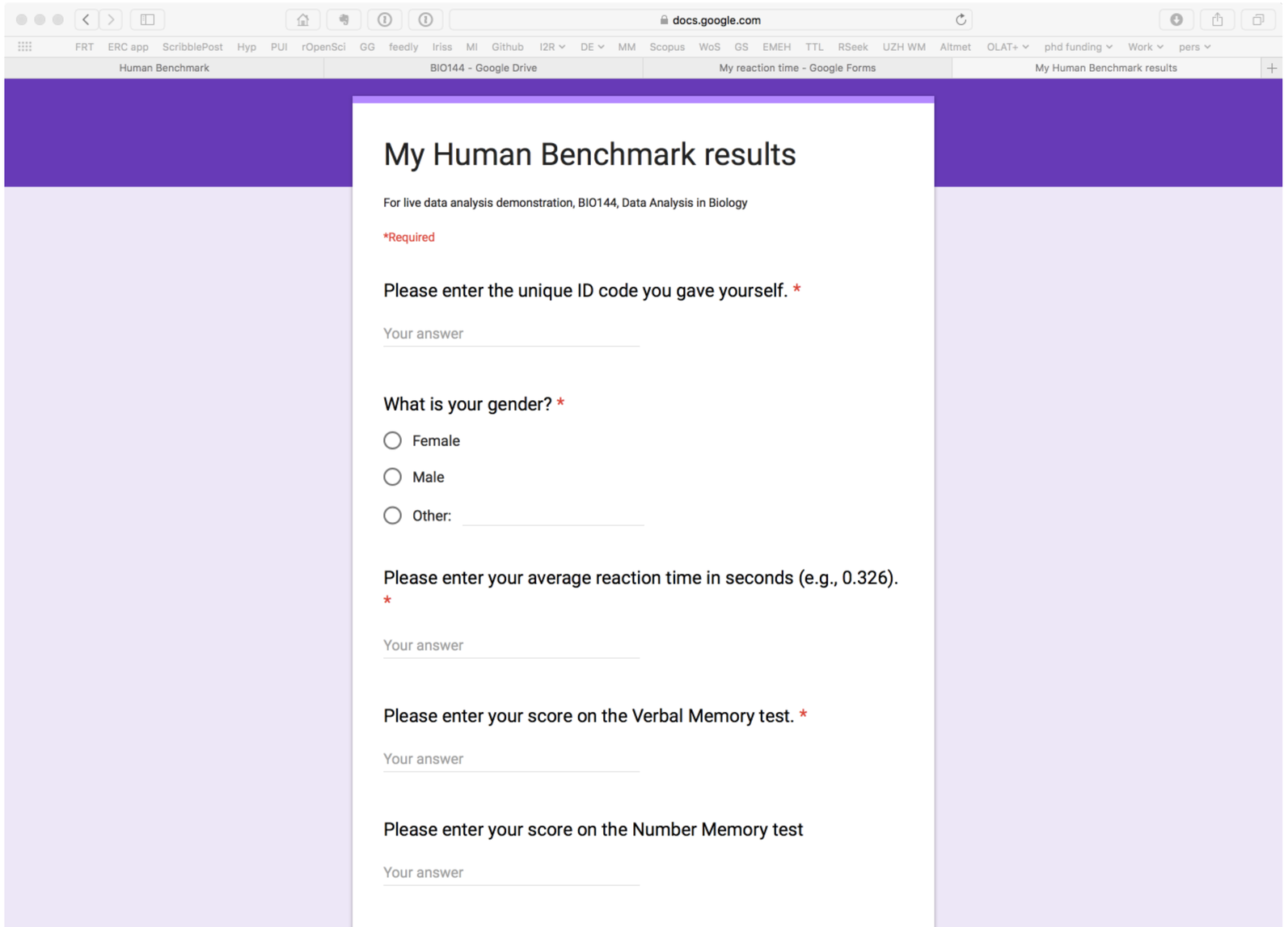


Live data analysis demonstration

The whole data analysis workflow in one hour!!!

- Question
- Expectation
- Planned presentation & analysis
- Selection of subjects
- How will data be collected?
- Ethics / permissions
- **Data collection**
- Data wrangling
- Visualise
- Statistical test
- Critical thinking
- Report / communicate

https://bit.ly/2Educmy



Live data analysis demonstration

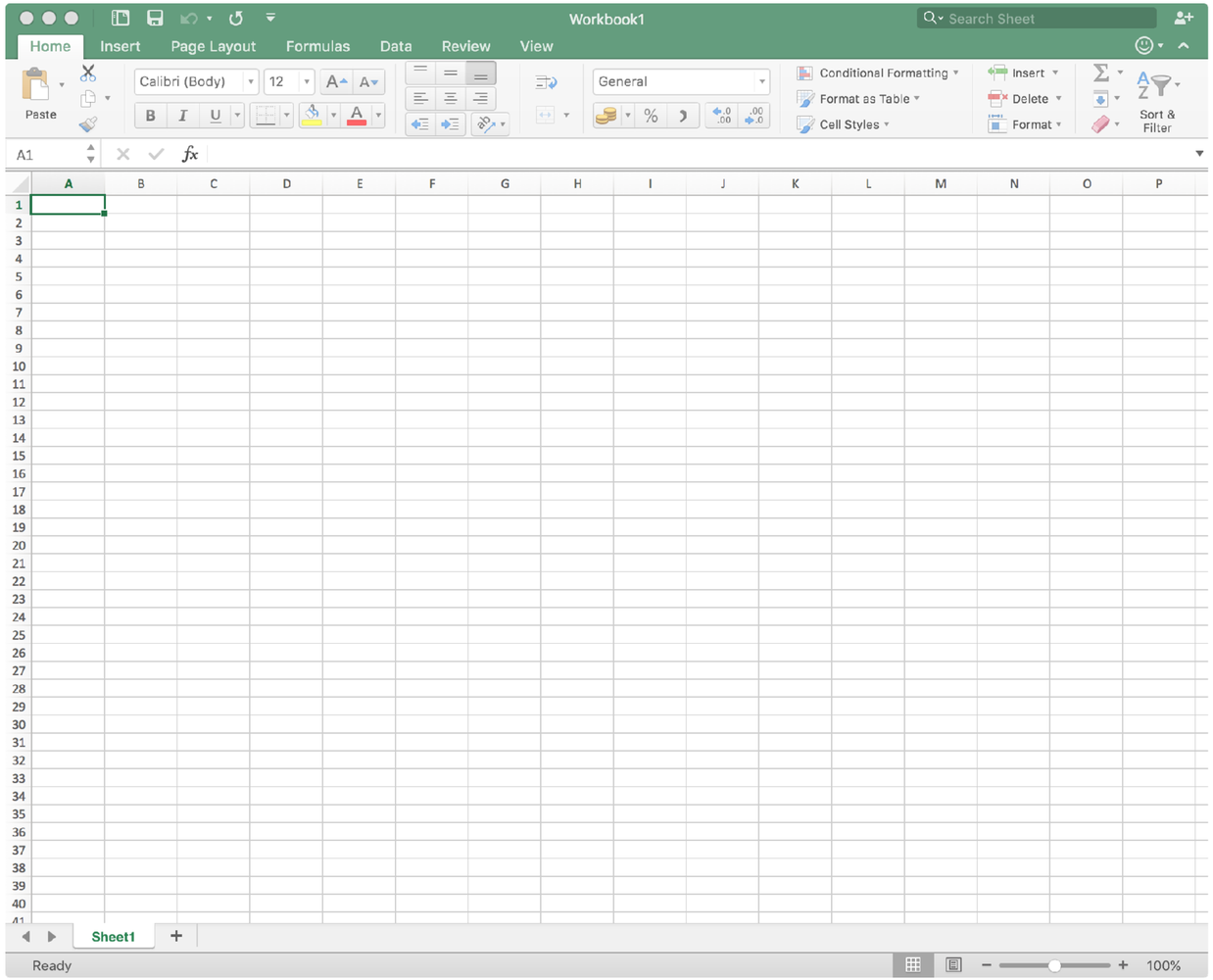
The whole data analysis workflow in one hour!!!

- Question
- Expectation
- Planned presentation & analysis
- Selection of subjects
- How will data be collected?
- Ethics / permissions
- **Data collection**
- Data wrangling
- Visualise
- Statistical test
- Critical thinking
- Report / communicate

Check the data in the spreadsheet

Efficient
Consistent
Repeatable
Reliable
Readable
Robust
Persistent
Sharable
Scalable

=



Efficient

Consistent

Repeatable

Reliable

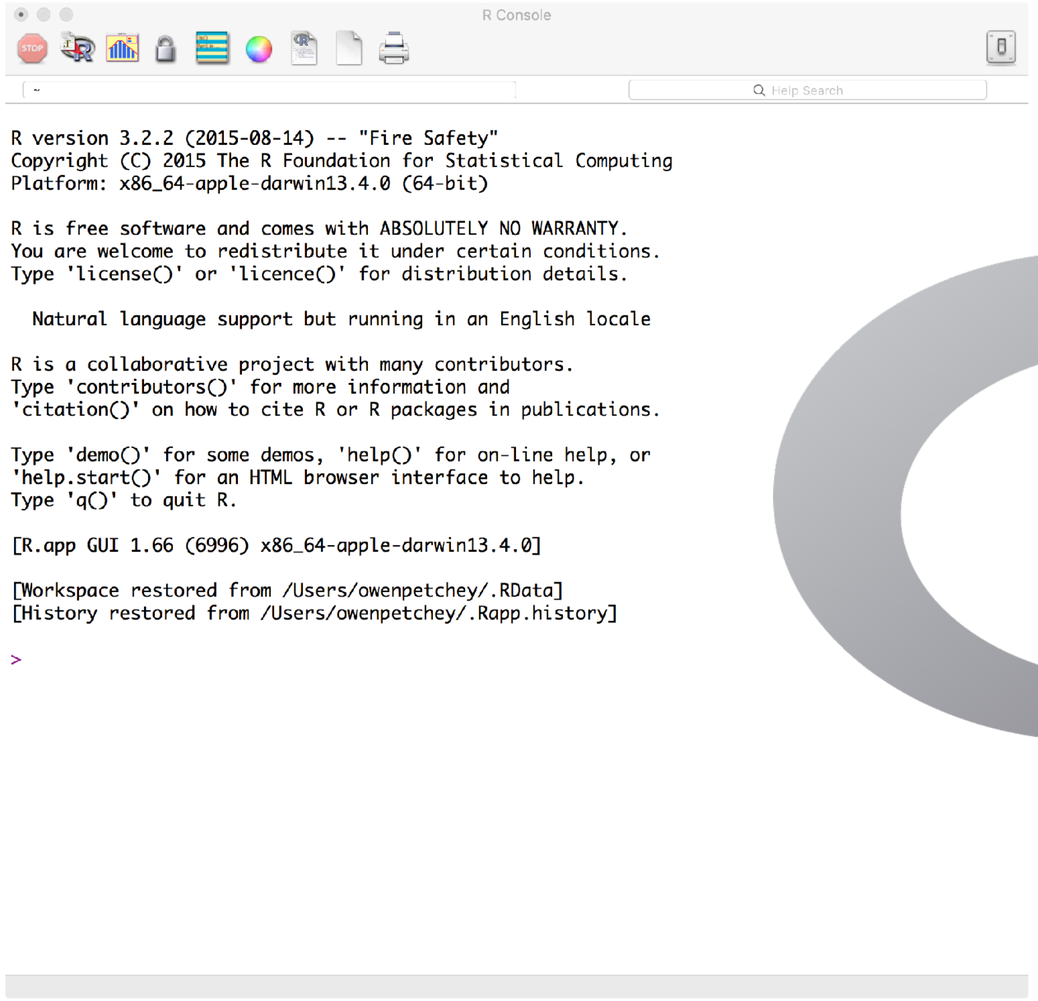
Readable

Robust

Persistent

Sharable

Scalable



Efficient

Consistent

Repeatable

Reliable

Readable

Robust

Persistent

Sharable

Scalable

Live data analysis demonstration

The whole data analysis workflow in one hour!!!

- Question
- Expectation
- Planned presentation & analysis
- Selection of subjects
- How will data be collected?
- Ethics / permissions
- Data collection
- **Data wrangling**
- **Visualise**
- **Statistical test**
- Critical thinking
- Report / communicate

Live in RStudio

Live data analysis demonstration

The whole data analysis workflow in one hour!!!

- Question
- Expectation
- Planned presentation & analysis
- Selection of subjects
- How will data be collected?
- Ethics / permissions
- Data collection
- Data wrangling
- Visualise
- Statistical test
- **Critical thinking**
- Report / communicate

Live data analysis demonstration

The whole data analysis workflow in one hour!!!

- Question
- Expectation
- Planned presentation & analysis
- Selection of subjects
- How will data be collected?
- Ethics / permissions
- Data collection
- Data wrangling
- Visualise
- Statistical test
- Critical thinking
- **Report / communicate**

Live in RStudio

Live data analysis demonstration

The whole data analysis workflow in one hour!!!

- Question
- Expectation
- Planned presentation & analysis
- Selection of subjects
- How will data be collected?
- Ethics / permissions
- Data collection
- Data wrangling
- Visualise
- Statistical test
- Critical thinking
- Report / communicate