

Data Analysis in Biology

BIO144

Prof. Owen Petchey

Dr Stefanie Muff

Teaching Assistants

You All



TED

LOG IN



Arthur Benjamin:

Teach statistics before calculus!

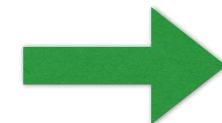
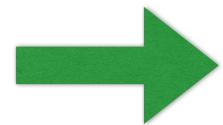


The “**hottest skill**”
that got people
hired in 2014?

Statistical Analysis

A video player interface with a dark grey background. At the top, there is a light grey header containing the text "The “hottest skill” that got people hired in 2014?". Below this, the title "Statistical Analysis" is displayed in a large, bold, teal font. In the bottom left corner, there is a progress bar with a red scrubber, followed by three white control icons: a play button, a forward button, and a volume button. To the right of the progress bar, the text "Source: LinkedIn" is written in a small, teal font. In the bottom right corner, there is a set of five white control icons: a document, a gear, a square, a screen, and a double arrow.

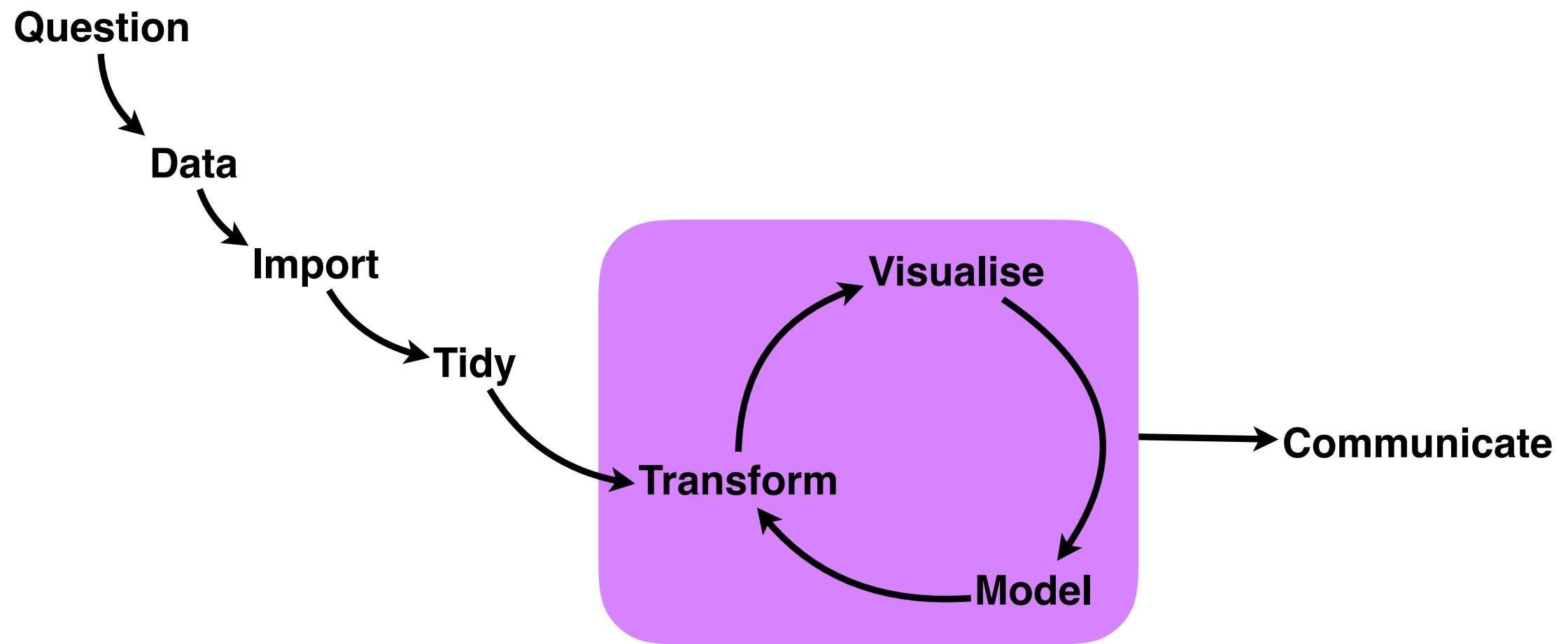
Source: LinkedIn



Question
Puzzle
Problem

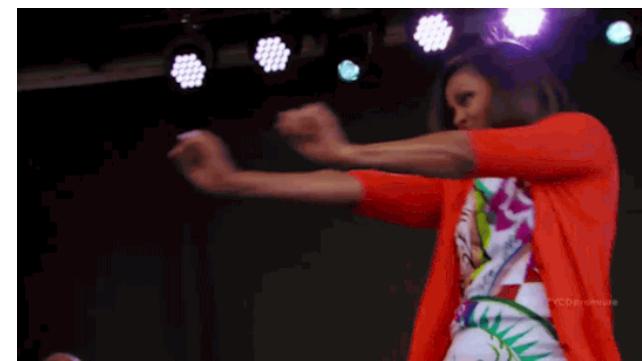
Data
+
Analysis

Answer
Solution

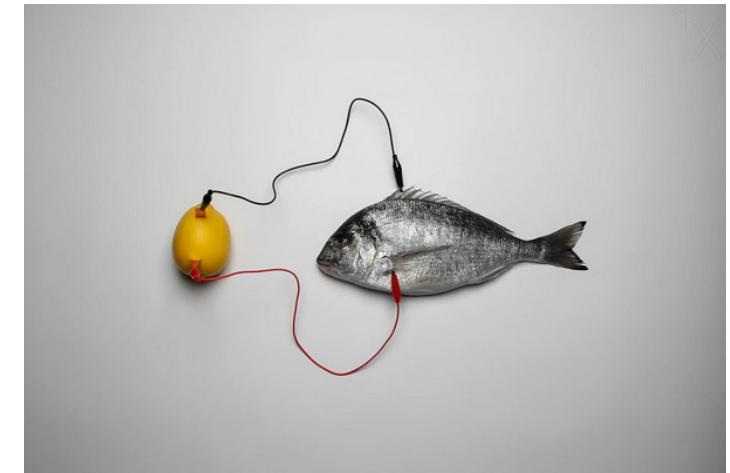


Efficient
Consistent
Repeatable
 Reliable
Readable
 Robust
Persistent
Sharable
 Scalable

==



The conceptual side



3 - 5pm

The practical side



1 - 3pm

<https://openedx.mnf.uzh.ch>

The screenshot shows the homepage of the OpenEdX platform at the University of Zurich (UZH). The top navigation bar includes links for various academic databases and services, such as ERC app, ScribblePost, Hyp, PUR, rOpenSci, GO, feedly, Iris, MI, Github, I2R, DE, MM, Scopus, WoS, GS, EMEH, TTL, RSeek, UZH WM, Altmet, OLAT+, phd funding, Work, and pers. A search bar is also present.

The main header features the University of Zurich logo and a "REGISTER NOW" button. On the right, there is a "Sign in" button. The background is dark blue with white text and icons. At the top, there are category labels: "Geographie", "Chemie", "Mathematik", and "Biologie". Below these categories are several hexagonal icons representing different subjects like Chemistry, Biology, and Mathematics.

A central banner displays the "University of Zurich" logo and the text "WELCOME TO OPENEDX @ UZH!". Below the banner, there are four course cards:

- UZH BIO144 DATA ANALYSIS IN BIOLOGY**
Starts: Jul 27, 2016
- UZH BIO134 PROGRAMMING IN BIOLOGY**
Starts: Sep 01, 2016
- UZH MULTI_MODULE INTRODUCTION TO R**
Starts: Sep 11, 2016
- UZH AST241 INTRODUCTION TO ASTROPHYSICS**
Starts: Sep 12, 2016

At the bottom of the page, there are two additional course cards:

- UZH** (with a small image of a resistor grid)
- UZH** (with a small image of a cell)

The details



Are all on openedX@MNE

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



switch

Live data analysis demonstration

BIO144
Week 1

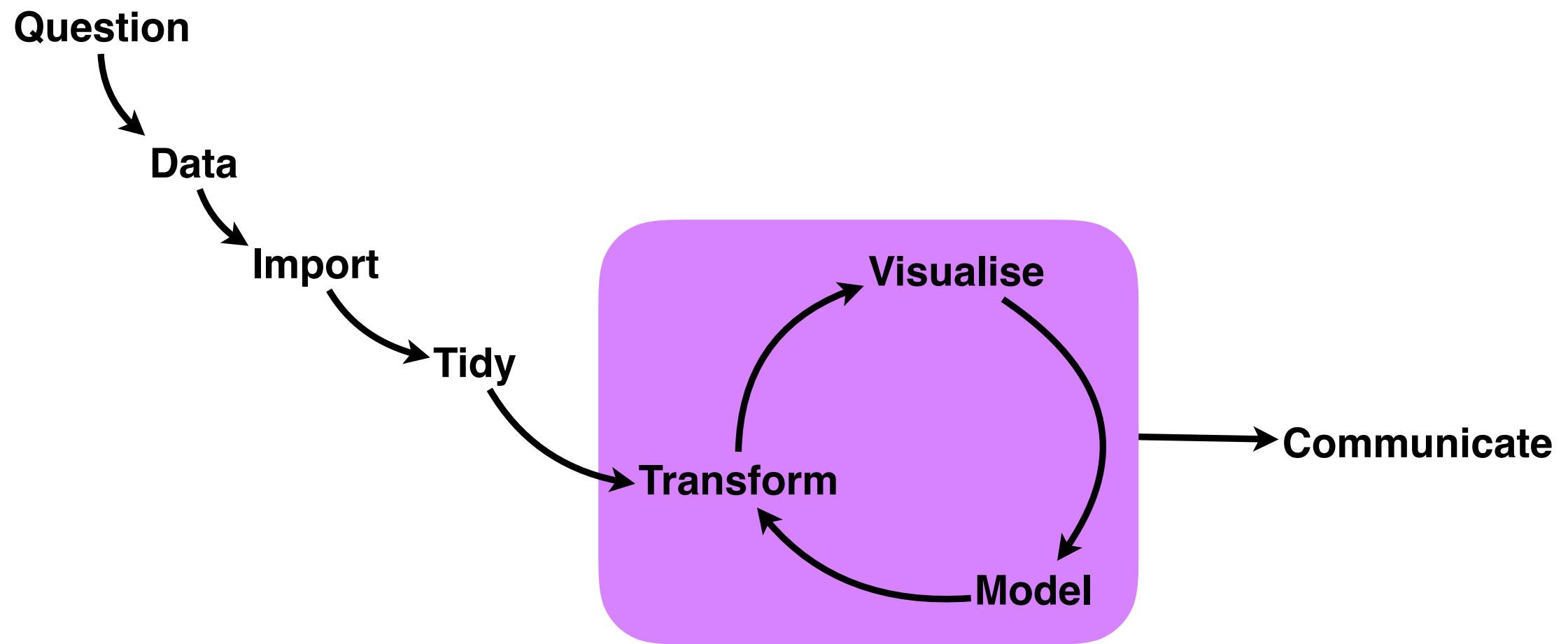
Its a demonstration...

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.



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 know 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

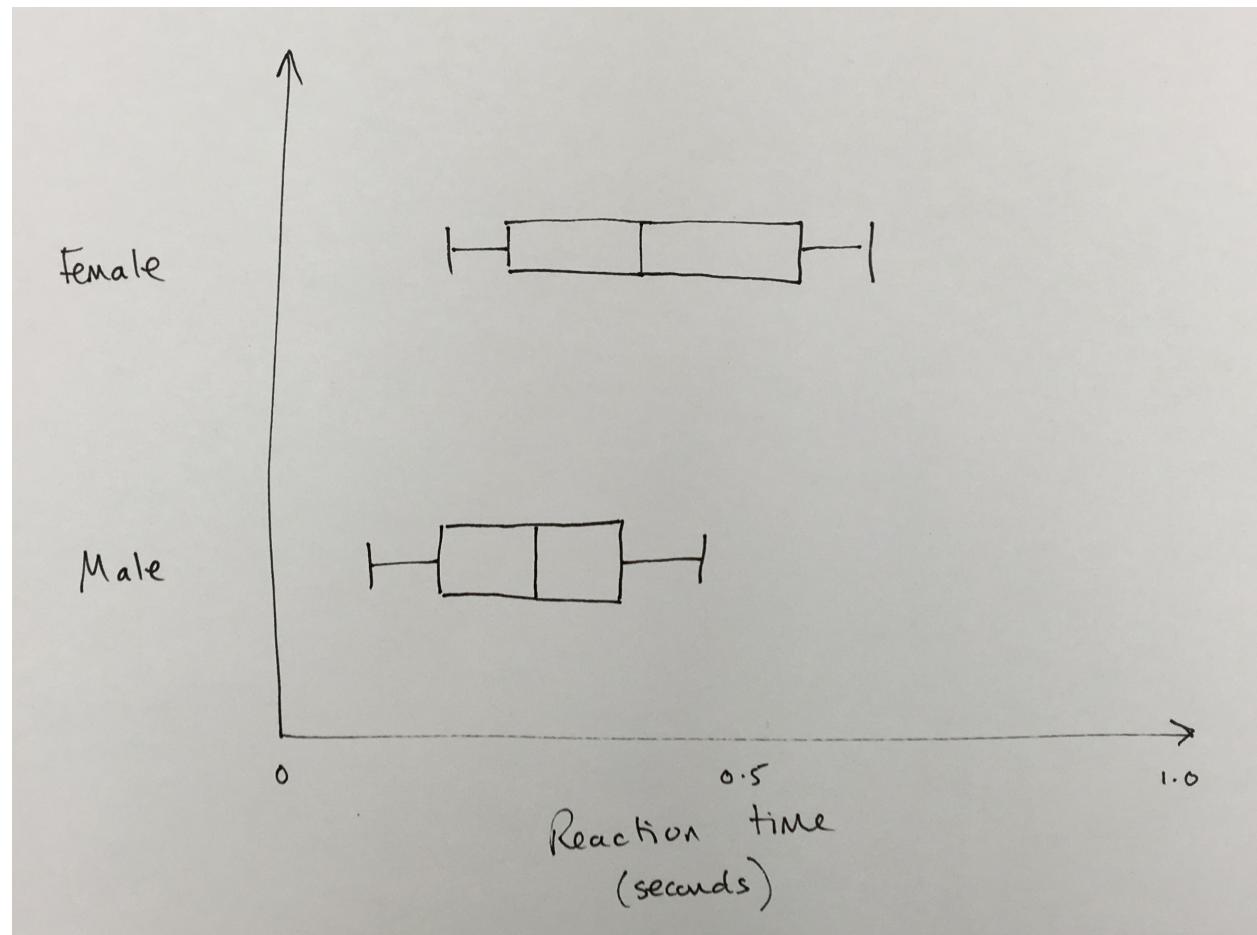
Data wrangling

Visualise

Statistical test

Critical thinking

Report / communicate



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

The screenshot shows a web browser window with the URL humanbenchmark.com. The page features a large green background image of a brain. At the top center is a yellow circle with a white lightning bolt icon. Below it, the text "How powerful is your brain?" is displayed in a large, bold, white font. Underneath, a smaller text reads "Use our targeted tests to find out for free." A "GET STARTED" button is located below this text. At the bottom of the page, there is a white callout box containing four circular icons with text descriptions:

- Number Memory**: Represented by a red circle with the number "123". Description: "Remember the longest number you can."
- Reaction Time**: Represented by a yellow circle with a lightning bolt icon. Description: "Test your visual reflexes."
- Verbal Memory**: Represented by a green circle with a book icon. Description: "Keep as many words in short term memory as possible."
- Visual Memory**: Represented by a blue circle with a grid of squares icon. Description: "Remember an increasingly large board of squares."

At the very bottom of the page, the copyright information is visible: "Copyright 2007-2016 Human Benchmark" followed by links to "contact@humanbenchmark.com" and "Licensing".

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

<http://bit.ly/2gnXxfo>

The screenshot shows a Google Forms survey titled "My Human Benchmark results". The survey is designed for a live data analysis demonstration. It includes fields for entering a unique ID code, gender (Female, Male, Other), average reaction time in seconds, Verbal Memory test score, and Number Memory test score. The survey is part of a larger Google Sheets document titled "Human Benchmark" and is located in the "BIO144 - Google Drive" folder.

My Human Benchmark results

For live data analysis demonstration, BIO144, Data Analysis in Biology

*Required

Please enter the unique ID code you gave yourself. *

Your answer _____

What is your gender? *

Female

Male

Other: _____

Please enter your average reaction time in seconds (e.g., 0.326). *

Your answer _____

Please enter your score on the Verbal Memory test. *

Your answer _____

Please enter your score on the Number Memory test

Your answer _____

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

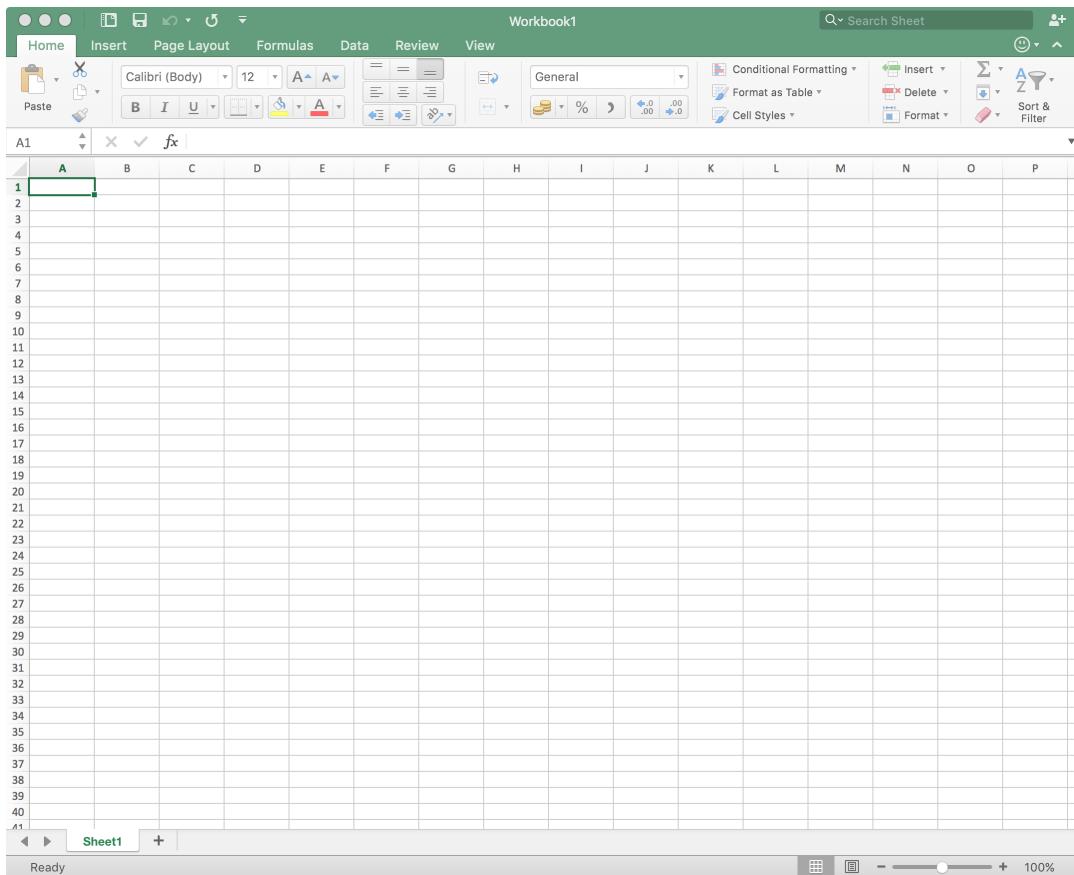
Data wrangling
Visualise
Statistical test
Critical thinking
Report / communicate

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**



R version 3.2.2 (2015-08-14) -- "Fire Safety"
Copyright (C) 2015 The R Foundation for Statistical Computing
Platform: x86_64-apple-darwin13.4.0 (64-bit)

R is free software and comes with ABSOLUTELY NO WARRANTY.
You are welcome to redistribute it under certain conditions.
Type 'license()' or 'licence()' for distribution details.

Natural language support but running in an English locale

R is a collaborative project with many contributors.
Type 'contributors()' for more information and
'citation()' on how to cite R or R packages in publications.

Type 'demo()' for some demos, 'help()' for on-line help, or
'help.start()' for an HTML browser interface to help.
Type 'q()' to quit R.

[R.app GUI 1.66 (6996) x86_64-apple-darwin13.4.0]

[Workspace restored from /Users/owenpetchey/.RData]
[History restored from /Users/owenpetchey/.Rapp.history]

>

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

