Open Science using the Statistical Package JASP

Herbert Hoijtink
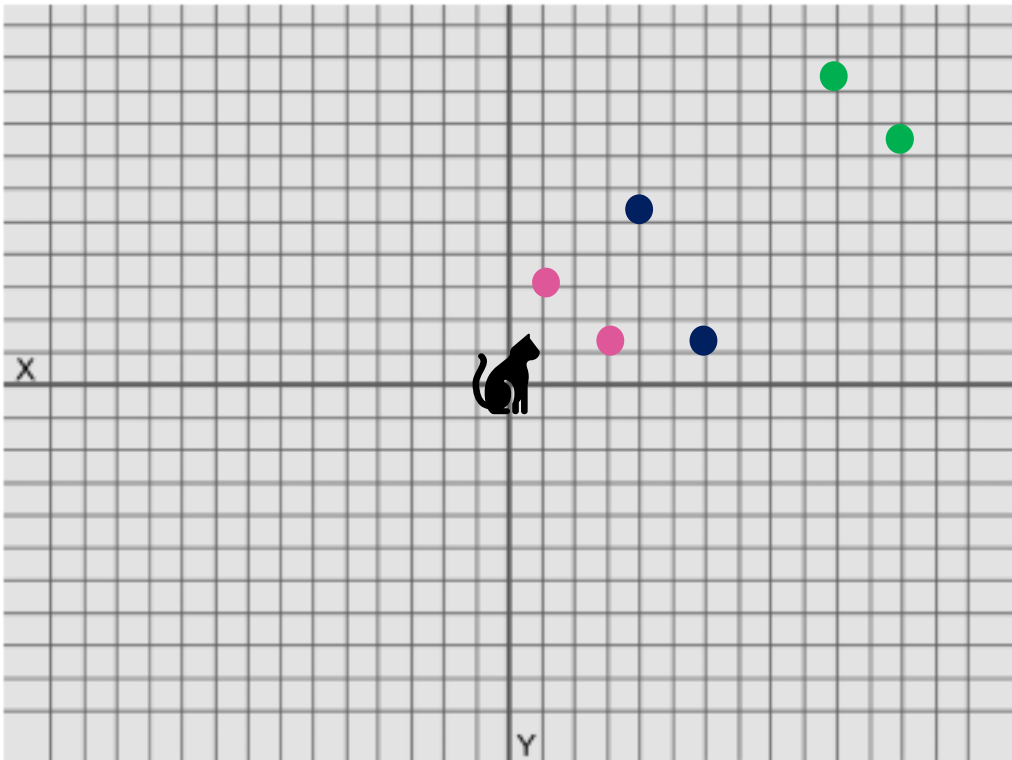
H.Hoijtink@uu.nl

An Experiment and its Replication

An experiment with three conditions:

- 🔴 The "close" condition
- 🔵 The "intermediate" condition
- 🟢 The "distant" condition

Participants Rated:

Attachment to:
- Siblings
- Parents
- Home-town

on a

1 (not at all strong) – 7 (extremely strong)
Likert scale

which are
averaged to obtain the dependent variable

The description given here is a modification of and inspired by the actual experiment executed by Williams, L.E. and Bargh, J.A. (2008). Keeping One's Distance. The Influence of Spatial Distance Cues on Affect and Evaluation. *Psychological Science, 19,* 302-308.

$H_0: \mu_{close} = \mu_{intermediate} = \mu_{distant}$,

that is, the three means are equal

rendering

p-value = .01, that is, smaller than .05, that is, the means are significantly different

with

m_{close} = 5.61 , $m_{intermediate}$ = 🐱 , $m_{distant}$ = 4.86

and

η^2 = .11 , that is, the three conditions explain 11% of the variation in attachment, which is a medium to strong effect of condition

p-value = .79

η^2 = .00

Joy-Gaba, J., Clay, R., and Cleary, H. (2016). Replication of keeping one's distance: The influence of spatial distance cues on affect and evaluation by Williams L.E. and Bargh J.A. (2008) *Psychological Science, 19,* 302-308). Retrieved from https://osf.io/a78bm/

The Replication Crisis

This is only one of 100 psychological experiments of which only about 33% were successfully replicated (OSC, 2015).

This resulted in a reduced trust in science by scientists and society: The replication crisis was born.

Scientists are alerted:

- Estimating the reproducibility of psychological science (OSC, 2015)
- An open investigation of the reproducibility of cancer biology research (Errington et al., 2014)

"Society" is alerted:

- Is psychology a real science? (Is psychologie wel een echte wetenschap, Volkskrant, 12-8-2016)
- Public Trust in Science (Rathenau Instituut, August 28, 2018)

Open Science Collaboration. (2015). Estimating the reproducibility of psychological science. *Science, 349*, 6251. https://osf.io/ezcuj/

Errington, T.M., Iorns, E., Gunn, W., Tan, F.E., Lomax, J., and Nosek, B.A. (2014). An open investigation of the reproducibilty of cancer biology research. *eLIFE, 3*, e04333. https://elifesciences.org/collections/9b1e83d1/reproducibility-project-cancer-biology

Volkskrant (2016). https://www.volkskrant.nl/columns-opinie/is-psychologie-wel-een-echte-wetenschap~b9978e6c

Rathenau Instituut (2018). Public Trust in Science. https://www.rathenau.nl/en/science-figures/impact/trust-science/public-trust-science

Addressing the Replication Crisis: The Open Empirical Cycle

Step 1: Preparation (Observation)

Step 2: Formulate Research Hypotheses (Induction)

Step 3: Planning of Data Collection and Analysis

Step 4: Get Approval from the (Medical) Ethical Testing Committee and Data Management Plan

Step 5: Preregistration and Preregistered Reports

Step 6: Execution of Data Collection and Analyses (Testing)

Step 7: Publish Data and Analyses

Step 8: Write an Open Access Report (Evaluation)

Step 9: Having Your Report Reviewed

Step 1: Preparation

- Review literature
- Gaps in the literature
- Quality of the literature
- Variables that are not covered in the literature

- Writing it up sharpens thoughts

- Write down you research questions:
1. New questions
2. Replication studies
3. Exploratory studies

Step 2: Formulate Research Hypotheses

A research hypothesis is a verbal representation of the expected relations between the variables resulting from Step 1.

Step 3: Planning of Data Collection and Analysis

- Describe population
- Describe research design
- Describe variables
- Derived variables
- Statistical model
- Formal hypotheses
- Power Analyses or Updating
- Missing data
- Data exclusion
- Exploration
- Replication

Step 4: (M)ETC and Data Management Plan

- Much of the previous three steps
- Informed consent forms
- Data management plan

Step 5: Preregistration and Preregistered Reports

This is the pre-data-collection or pre-data-access account of all that has been covered in the previous four steps.

Step 6: Execution of Data Collection and Analysis

Proof that data collection takes place after preregistration.

Proof that data access takes place after preregistration.

Step 7: Publish Data and Analyses

Data and analyses should be FAIR, that is, accessible to all interested parties. Most of the remainder of this presentation discusses how that can be achieved.

Step 8: Write an Open Access Report

Include a link to your pre-registration.

Include a link to your data-analyses repository.

Publishing open access is an important feature of open science.

Being able to unobstructedly obtain everything (also the report) related to a research project will enable anyone to:
- benefit from your research
- reuse (parts) of your research
- engage in a fully informed discussion about your research.

This should both increase the impact of your research and increase the trust in your research and the trust in science in general.

Step 9: Having your Report Reviewed

May lead to changes in your report and possibly deviations from your pre-registration.

These deviations can be highlighted in your report using footnotes.

Continued: Step 7: Publish Data and Analyses

Opening, that is, publishing, your data, analyses input, analyses results, and interpretation of the results

Others can:
- Inspect your data
- Reproduce your analyses (and get all the information, cf., the missing $m_{intermediate}$ in Williams and Bargh, 2008)
- Read your interpretation of the analyses
- Data are available for meta analyses and "null-findings" become accessible

Being open will add to the trust in your research, that of the group to which you belong, and science in general

Open the Data Analyses of your Thesis

... can easily be done using:

JASP
A Fresh Way to Do Statistics

... applies to:

All research based on quantitative data

Open Data Analyses should be FAIR

- **Findable:** Place your data and analyses in a public repository
- **Accessible:** Make certain your data come with a codebook, and your analyses with annotations
- **Interoperable:** Ensure that data and analyses can be opened on different types of computers
- **Reusable:** include a license, that is, make clear what others are (not) allowed to do with your data

For further elaboration and information see "How to make your data FAIR" (Research Data Management Support, Utrecht University).

Interoperable

https://jasp-stats.org/download/

JASP

DOWNLOAD | SUPPORT | TEACHING | BLOG | DONATE
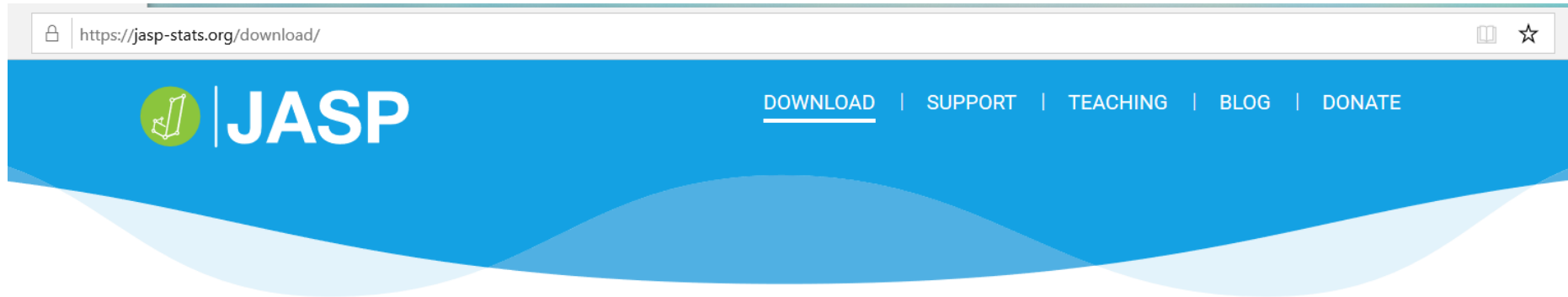
JASP 0.13.1

Released July 16th, 2020.

This version adds mixed models, the reliability module, and the R console. For a complete list of all improvements and bug fixes per release, see the release notes.

Having trouble installing JASP under Mac OS X?

Take a look at our installation guide.

Want to go back?

You can download many of JASP's previous versions.

Download JASP

Entirely for free, no strings attached.

Windows

⬇ Windows 64bit

⬇ Windows 32bit

The pre-installed 64-bit or 32-bit version can be used if the msi fails.)

MacOS

⬇ Catalina

⬇ Mojave & High Sierra

For older versions of MacOS (Sierra and before), download JASP 0.9.2. We recommend upgrading your system though.

Linux

Flatpak Installation Guide

A Crash Course into JASP
How to use JASP
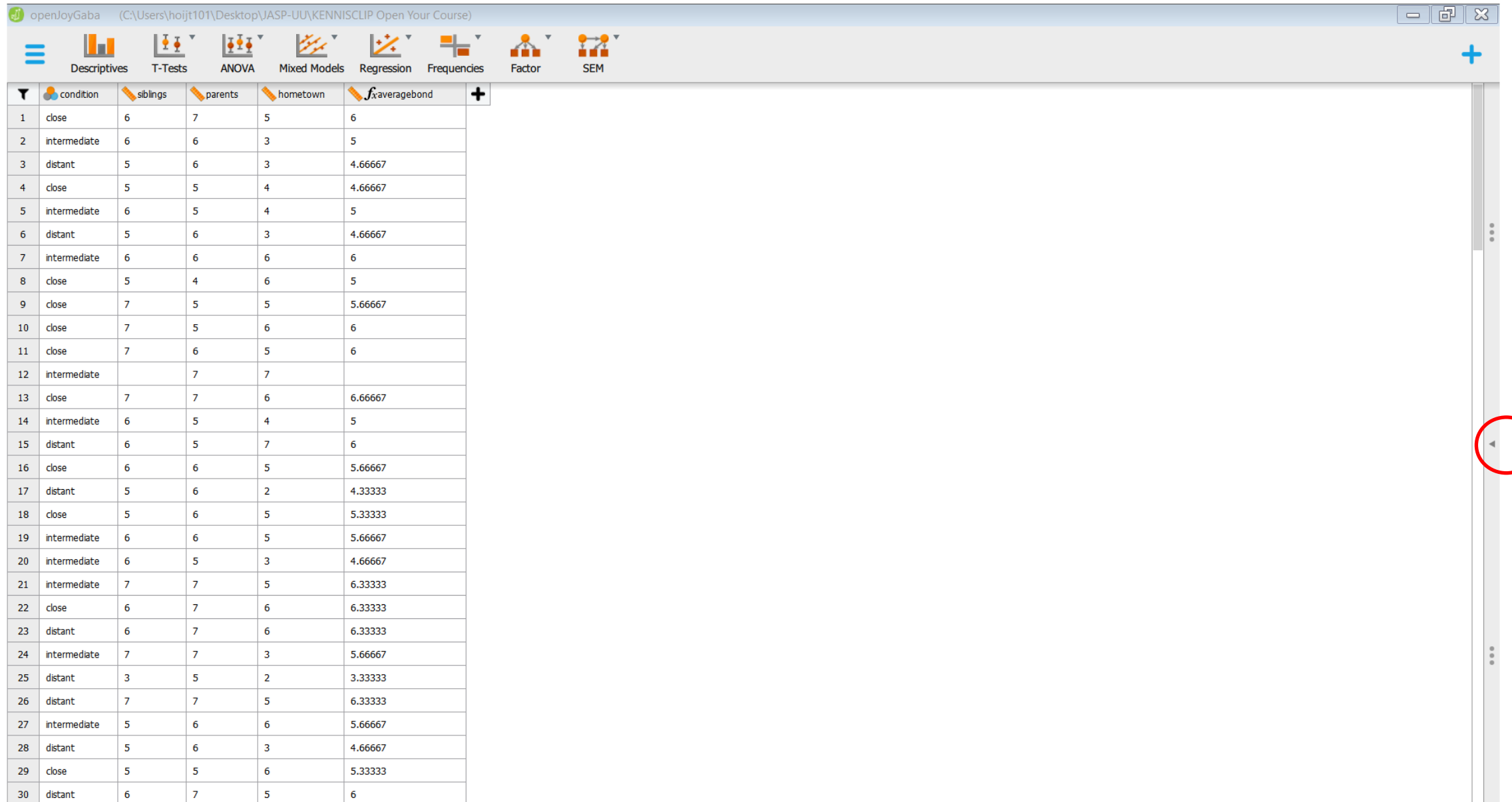
Accessible

Joy-Gaba, Clay, and Cleary (2016) replicated Willams and Bargh (2008)

The replication data and analyses are contained in openJoyGaba.jasp

The Data Collected by Joy-Gaba, Clay, and Cleary (2016)

	condition	siblings	parents	hometown	averagebond
1	close	6	7	5	6
2	intermediate	6	6	3	5
3	distant	5	6	3	4.66667
4	close	5	5	4	4.66667
5	intermediate	6	5	4	5
6	distant	5	6	3	4.66667
7	intermediate	6	6	6	6
8	close	5	4	6	5
9	close	7	5	5	5.66667
10	close	7	5	6	6
11	close	7	6	5	6
12	intermediate		7	7	
13	close	7	7	6	6.66667
14	intermediate	6	5	4	5
15	distant	6	5	7	6
16	close	6	6	5	5.66667
17	distant	5	6	2	4.33333
18	close	5	6	5	5.33333
19	intermediate	6	6	5	5.66667
20	intermediate	6	5	3	4.66667
21	intermediate	7	7	5	6.33333
22	close	6	7	6	6.33333
23	distant	6	7	6	6.33333
24	intermediate	7	7	3	5.66667
25	distant	3	5	2	3.33333
26	distant	7	7	5	6.33333
27	intermediate	5	6	6	5.66667
28	distant	5	6	3	4.66667
29	close	5	5	6	5.33333
30	distant	6	7	5	6

Data collected by, License, Data analyzed by, Code book, and Descriptive Statistics

Include Annotations of your Analyses

Explain using annotations which analyses were executed and what your interpretation of the outcomes was

Annotated Analyses

By doing this you Create an Annotated Logbook of your Analyses

As shown on the previous two slides JASP tracks the analyses that you execute

- There is a Descriptive Statistics button on the left side of the screen
                                        rendering the data overview on the right side of the screen
- There is an ANOVA button on the left side of the screen
                                        rendering analyses with annotations on the right side of the screen

If you continue executing analyses and adding annotations to your analyses, you create a logbook that will help you to track your traces, therefor it will be completely clear to others:

- Which analyses you did
- How you did them (this should shortly be explained in the annotations)
- How you interpret the results (this should shortly be explained in the annotations)

In this manner you make the data and analyses from your bachelor thesis accessible

Findable Using the Open Science Foundation

Reusable

Fully Open Data Analyses: Include the CC0 1.0 Universal License With Your Data

The license can be included in your mydata.jasp file.

The CC0 1.0 Universal Public Domain Dedication:
- Is truly open, that is, anybody can use your data for whatever purpose
- If you include a reference to a paper or your contact information in the mydata.jasp file, anybody using your data can refer to you

The CC0 1.0 Universal Public Domain Dedication: https://creativecommons.org/publicdomain/zero/1.0/

The Conflict between Fully Open Data Analyses and Privacy Regulations

Personal data can only be published if privacy is guaranteed. One way to achieve this is, if data are truly anonymized (General Data Protection Regulation: https://gdpr-info.eu/).

- Truly anonymized data is no longer personal data and thus no longer subject to the GDPR
- It involves the complete and utter removal of all personal identifiers in a database
- Anonymized data can no longer be attributed to any particular individual by any means

If anonymization cannot be achieved, you can still "Open your Analyses" by publishing only your code book, results, references, license and annotated analyses (these are often called meta data). This can be done using a myresults.html file that can be created based on your mydata.jasp file.

This page was constructed using quotes and input from
Handling Personal Data in Research (Utrecht University)

The Conflict between Fully Open Data Analyses and your Future Research Plans

- Therefore, you may want to impose restrictions on the use of your data, e.g., "can only be used to reproduce the analyses you executed"
- However, this cannot be arranged via the application of a license, data are often considered to be facts and facts cannot be copyrighted (OpenAIRE)
- In such cases you can consider publishing only a part of your data, e.g., only the data and analyses that are used in a specific research report (but do provide complete meta data, most importantly, the code book of your complete data set)
- Then the unused data are only available for you and thereby you avoid being scooped out of your next paper

OpenAIRE: https://www.openaire.eu/how-do-i-license-my-research-data

Links to …

Open your Course/Bachelor Thesis

JASP
A Crash Course into JASP
How to use JASP

REPOSITORIES
Open Science Foundation
Yoda (Utrecht University)

LICENCES
Creative Commons
OpenAIRE

ANONIMIZED AND FAIR DATA
How to Make you Data FAIR
Handling Personal Data in Research (Utrecht University)
General Data Protection Regulation

OPEN SCIENCE
Open Science Community Utrecht
Special Interest Group Open Statistical Software: JASP and R

Utrecht University

How To …

Use one of the example data sets that downloaded with this presentation: data.sav, data.txt, or, data.csv.

Execute each of the steps in the following slides.

How to Create a mydata.jasp File

Make certain that JASP is installed on your computer. If not, install it from https://jasp-stats.org/download/

If JASP is installed, start the program

Open a .csv, .sav, or .txt file containing your data

After the data spreadsheet opens, click
on Descriptives to open the analysis screen
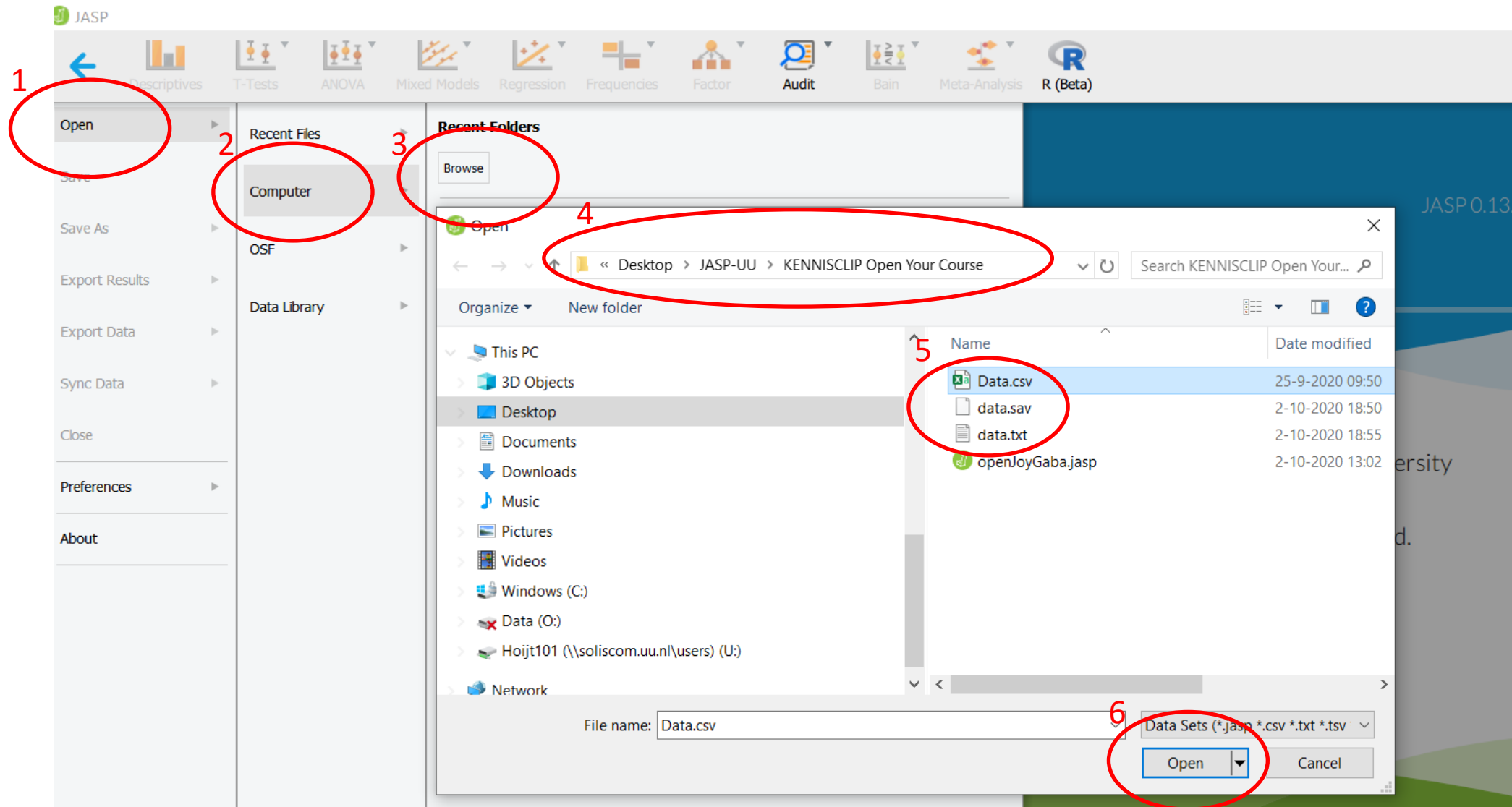
openJoyGaba   (C:\Users\hoijt101\Desktop\JASP-UU\KENNISCLIP Open Your Course)

Descriptives · T-Tests · ANOVA · Mixed Models · Regression · Frequencies · Factor · SEM

	condition	siblings	parents	hometown	fx:averagebond
1	close	6	7	5	6
2	intermediate	6	6	3	5
3	distant	5	6	3	4.66667
4	close	5	5	4	4.66667
5	intermediate	6	5	4	5
6	distant	5	6	3	4.66667
7	intermediate	6	6	6	6
8	close	5	4	6	5
9	close	7	5	5	5.66667
10	close	7	5	6	6
11	close	7	6	5	6
12	intermediate		7	7	
13	close	7	7	6	6.66667
14	intermediate	6	5	4	5
15	distant	6	5	7	6
16	close	6	6	5	5.66667
17	distant	5	6	2	4.33333
18	close	5	6	5	5.33333
19	intermediate	6	6	5	5.66667
20	intermediate	6	5	3	4.66667
21	intermediate	7	7	5	6.33333
22	close	6	7	6	6.33333
23	distant	6	7	6	6.33333
24	intermediate	7	7	3	5.66667
25	distant	3	5	2	3.33333
26	distant	7	7	5	6.33333
27	intermediate	5	6	6	5.66667
28	distant	5	6	3	4.66667
29	close	5	5	6	5.33333
30	distant	6	7	5	6

Execute your Analyses – Start with Simple Descriptives – and Add Notes to your Results

Data*   (C:\Users\hoijt101\Desktop\JASP-UU\KENNISCLIP Open Your Course)

Descriptives   T-Tests   ANOVA   Mixed Models   Regression   Frequencies   Factor   SEM

1

▶ **Descriptive Statistics**

2   ▼ **ANOVA**

siblings
parents
hometown

Dependent Variable
✎ averagebond

Fixed Factors
condition

WLS Weights

Display
☑ Descriptive statistics
☑ Estimates of effect size
   ☑ η²   ☐ partial η²   ☐ ω²
☐ Vovk-Sellke maximum p-ratio

▶ Model

▶ Assumption Checks

▶ Contrasts

▶ Post Hoc Tests

Results ▼

Data collected by: Joy-Gaba ...

Descriptive Statistics

Descriptive Statistics

	condition	siblings	parents	hometown	averagebond
Valid	133	125	133	133	133
Missing	0	8	0	0	0
Minimum	1.000	1.000	1.000	1.000	2.333
Maximum	3.000	7.000	7.000	7.000	7.000

4: click the black triangle and choose annotate

ANOVA ▼

5

B I U 🔗   Normal ☷ ☰   A 🖊   x₂ x²   ❞ ⫶⊟ ⫶⊟   Large ⇕ Tₓ

The null- hypothesis ...

ANOVA - averagebond

Cases	Sum of Squares	df	Mean Square	F	p	η²
condition	0.199	2	0.099	0.092	0.912	0.001
Residuals	140.232	130	1.079			

Note. Type III Sum of Squares

3

Descriptives ▼

Descriptives - averagebond ▼

condition	Mean	SD	N
1	5.420	0.877	46
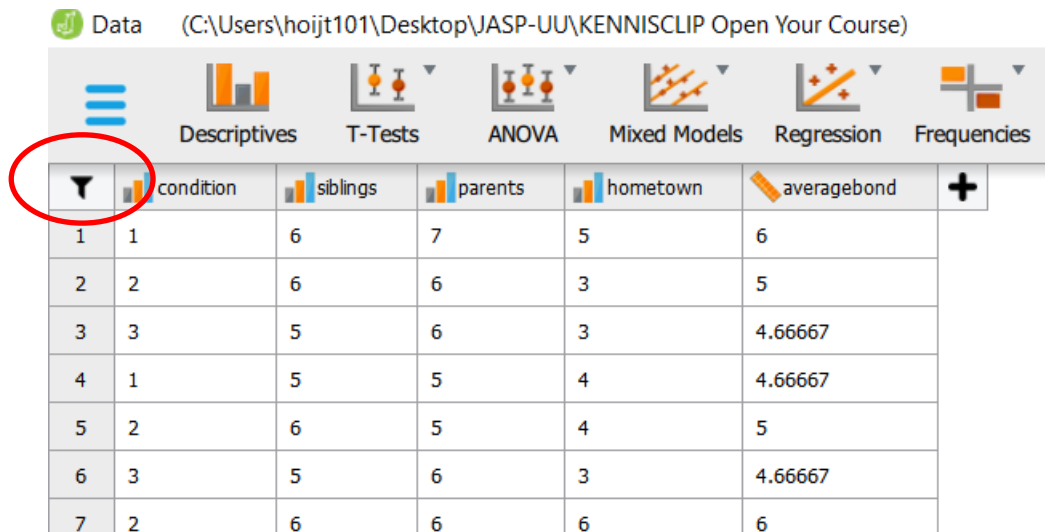
Currently you have to make a different mydata.jasp file for each selection of cases you make, e.g., before filtering (marked by the red ellipse), and before deleting or adding cases. Otherwise, all your analyses will be based on the latest selection of cases created.

This will be remedied in one of the future JASP releases.

A knowledge clip explaining the main features of JASP (selecting cases, computing variables, recoding, opening a data file, saving a mydata.jasp file, executing analyses, getting help, etc.) can be found at https://osf.io/z7tbg/

A virtually complete elaboration of all the features of JASP can be found at https://jasp-stats.org/how-to-use-jasp/

Save as a mydata.jasp File on your Computer – contains data, analyses input and annotated analyses results
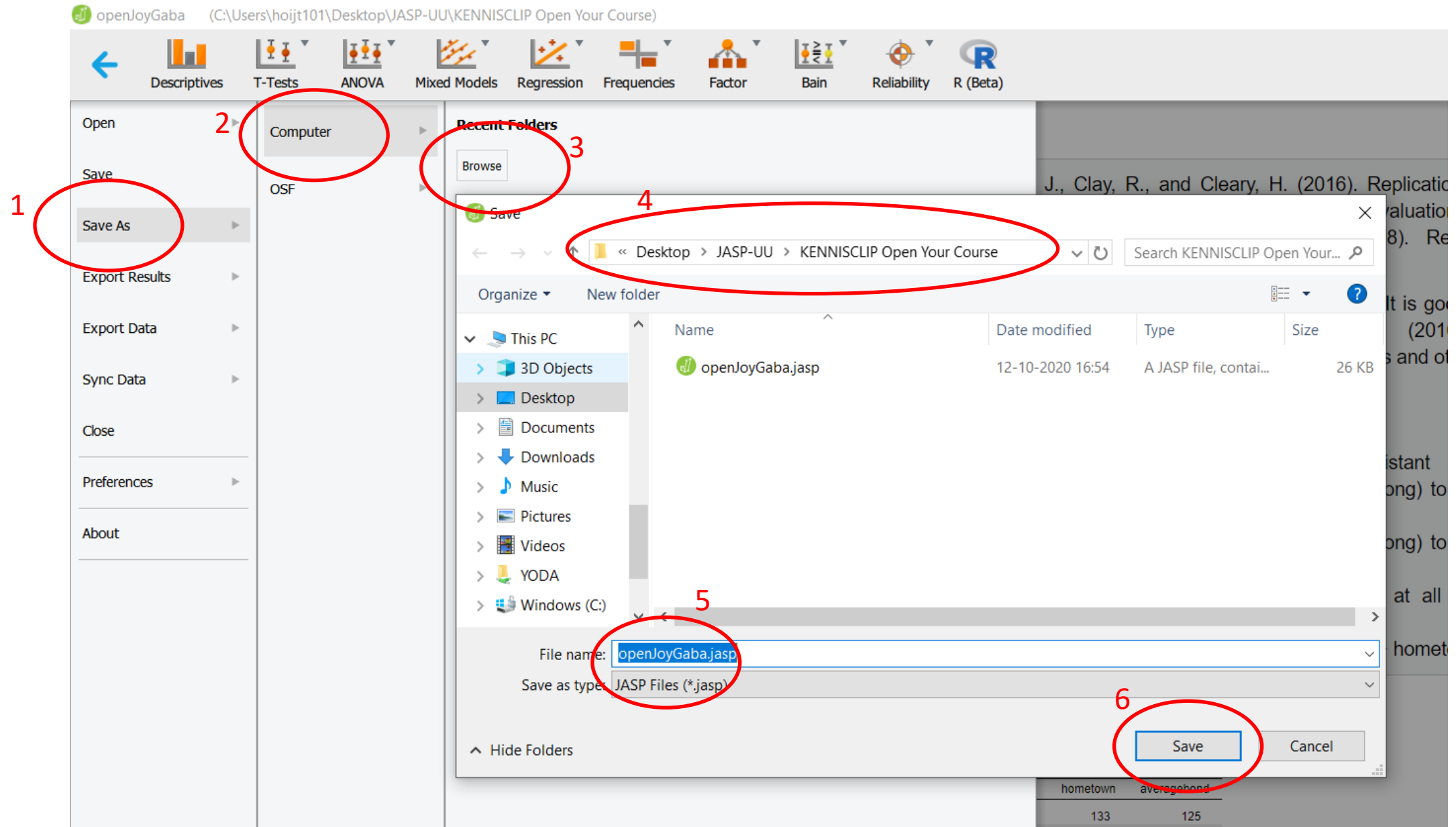
Save a mydata.jasp file as a myresults.html File on your Computer – only contains the annotated results

How to Make your Data Analyses Findable Using the Open Science Foundation

Create a Project to Store your Data at the Open Science Foundation in Four Simple Steps
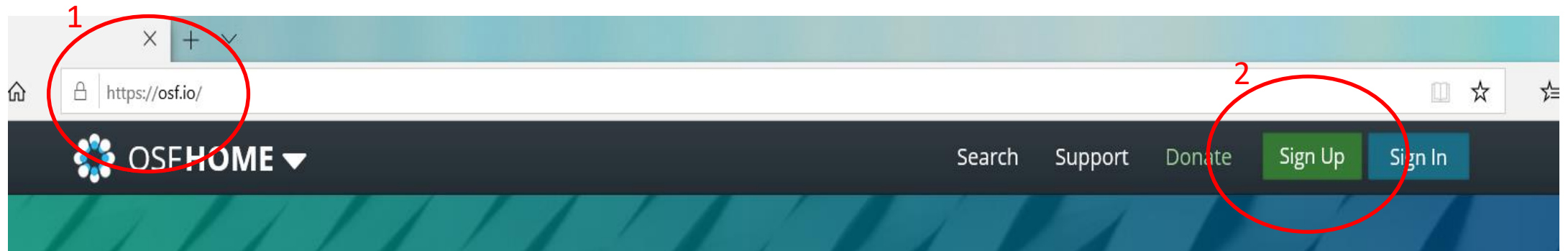
Step 1: Create a Free Account

Create a free account

Sign up using:

| ORCID | Institution |

OR

1

Full name

Contact email

Confirm email

Password
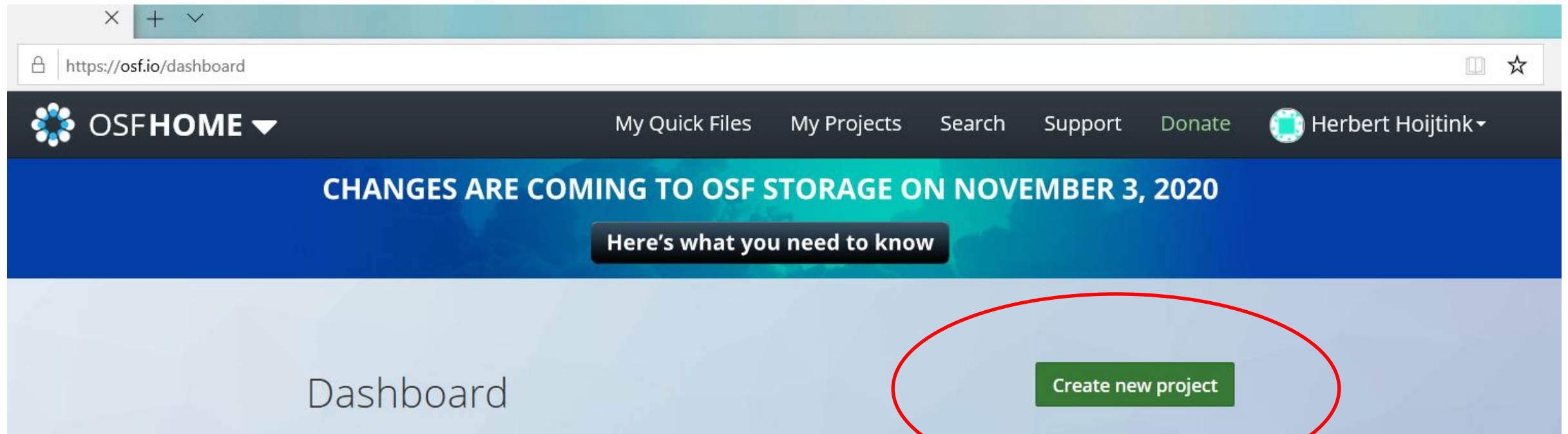
2 ☐ I have read and agree to the **Terms of Use** and **Privacy Policy**.

3 ☐ Ik ben geen robot

reCAPTCHA
Privacy - Voorwaarden

4 Sign up

Step 2: Create a New Project

Step 2: Create a New Project (continued)

Step 3: Make your Project Publicly Available

https://osf.io/z7tbg/ **1: anybody can surf here to access your data analyses if ...**

OSFHOME ▼   My Quick Files   My Projects   Search   Support   Donate   Herbert Hoijtink ▼

Open Your Course   Files   Wiki   Analytics   Registrations   Contributors   Add-ons   Settings

Open Your Course

112.5KB   Make Private   Public   ⅃ 0   ...

2: ... you make your project publicly available

Contributors: **Herbert Hoijtink**

Date created: 2020-10-05 02:49 PM | Last Updated: 2020-10-21 09:44 AM

Create DOI

Category: 📦 Project

Description: Add a brief description to your project

License: CC0 1.0 Universal   **3**

Wiki	↗
Add important information, links, or images here to describe your project.	

Files	↗

Click on a storage provider or drag and drop to upload

🔍 Filter   ℹ

Name ⌃ ⌄	Modified ⌃ ⌄
📦 Open Your Course	

Citation	⌄

Components	Add Component Link Projects

Add components to organize your project.

4

Tags

Open Science ✕   Teaching ✕   Add a tag

Drag your Data and other Relevant Files into the Project you Created

Drag files onto the blue bar
to add them to your project

The End