

Workshop on reproducible science: Motivations

Gbadamassi G.O. Dossa

last updated: 2022-11-04

Acknowledgements

The content of this module are based on materials from:



olivier gimenez's materials

The problems

a reproducible workflow



Credit: Ignasi Bartomeus, PhD

Why reproducible science?

SPECIAL

[▶ See all specials](#)

CHALLENGES IN IRREPRODUCIBLE RESEARCH

Science moves forward by corroboration – when researchers verify others' results. Science advances faster when people waste less time pursuing false leads. No research paper can ever be considered to be the final word, but there are too many that do not stand up to further study.

There is growing alarm about results that cannot be reproduced. Explanations include increased levels of scrutiny, complexity of experiments and statistics, and pressures on researchers. Journals, scientists, institutions and funders all have a part in tackling reproducibility. *Nature* has taken substantive steps to improve the transparency and robustness in what we publish, and to promote awareness within the scientific community. We hope that the articles contained in this collection will help.



Pour une recherche reproductible, publiez vos codes et données

LE MONDE SCIENCE ET TECHNO | 15.07.2013 à 14h42 • Mis à jour le 10.10.2013 à 14h35 |

Antoine Blanchard et Elifsu Sabuncu (Fondateurs de Deuxième Labo) et Yvan Stroppa (Ingénieur CNRS, cofondateur et "chief computing officer" de RunMyCode)

Abonnez vous à partir de 1 €

Réagir

Ajouter

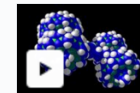


Partager (266)

Tweeter

En avril, un pavé a été jeté dans la mare de la recherche en économie : l'un des articles économiques les plus cités ces dernières années par les partisans de l'austérité, selon lequel le taux de croissance d'un pays devient négatif quand sa dette dépasse 90 % du produit intérieur brut (PIB), serait entaché d'erreurs de calcul et de manipulations de données. Ce sont trois économistes de l'université du Massachusetts qui ont découvert la fraude après avoir obtenu l'accès à la feuille de calcul Excel que les chercheurs de Harvard Carmen Reinhart et Kenneth Rogoff avaient utilisée pour bâtir leur démonstration. En cause : des erreurs de formules Excel, une exclusion arbitraire de certaines données et une pondération non conventionnelle - sans lesquelles le résultat ne serait plus du tout probant et ne plaiderait pas en faveur d'un maintien de la dette en deçà de 90 % du PIB.

Vidéo



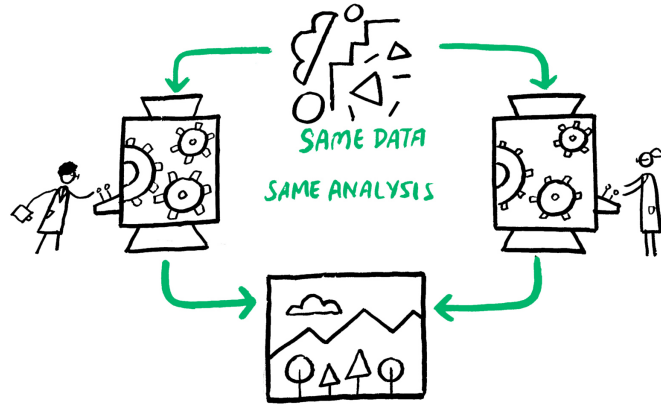
La course des plus petits bolides du monde

Les plus partagés

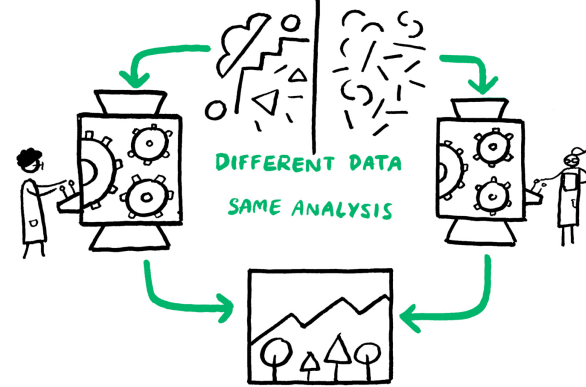
- 1 « A quoi sert François Hollande ? », surprenant sujet d'un parti de droit 5938
- 2 La France doit se préparer à l'émigration de ses cerveaux 3327
- 3 Justin Shults, 30 ans, #EnMémoireBruxelles 2540
- 4 Natation : Camille Lacourt sacré champion d'Europe du 100 m dos 1608
- 5 Issad Rebrab : en Algérie, « on veut tout museler, tout 1308

What?

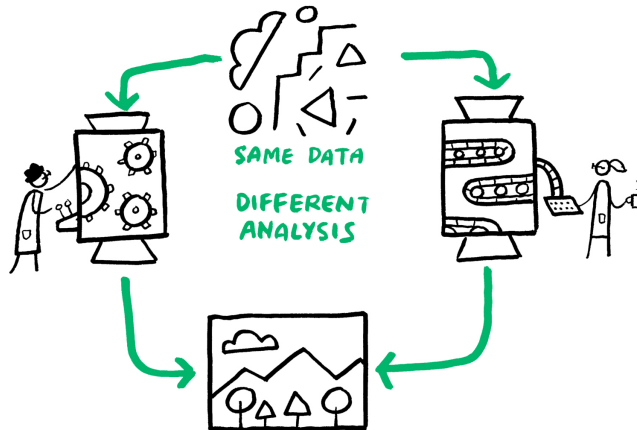
REPRODUCIBLE



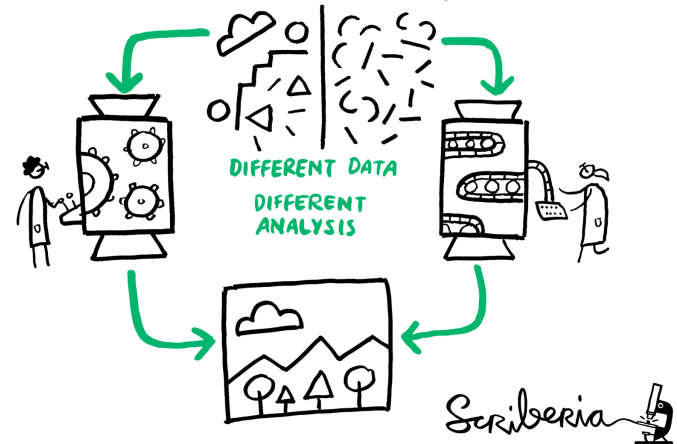
REPLICABLE



ROBUST



GENERALISABLE



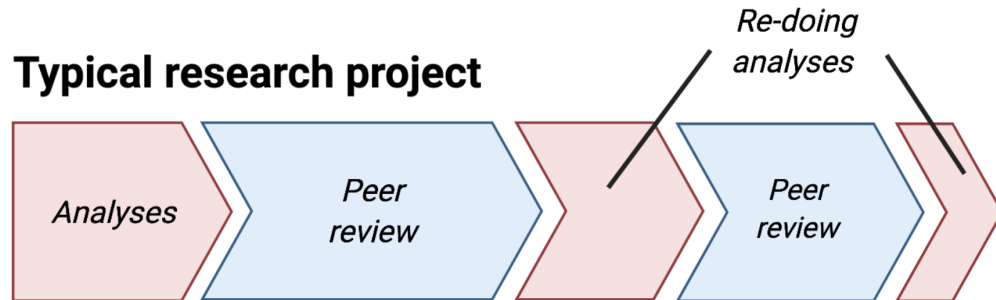
Scriberia 

Benefits?

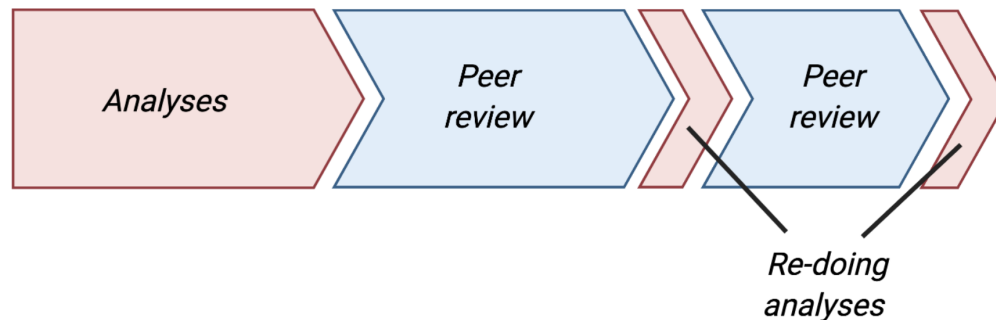
Among other advantages

- Forces you to double-check your stuff (you share data and code)
- Your future-self will thank you (repeat analyses, write papers)
- By being reproducible, you build up your credibility and reputation
- Reproducibility favors trust in the scientific endeavour
- Faster scientific progress

You don't *lose* time doing reproducible science, you just *relocate* where you spend it



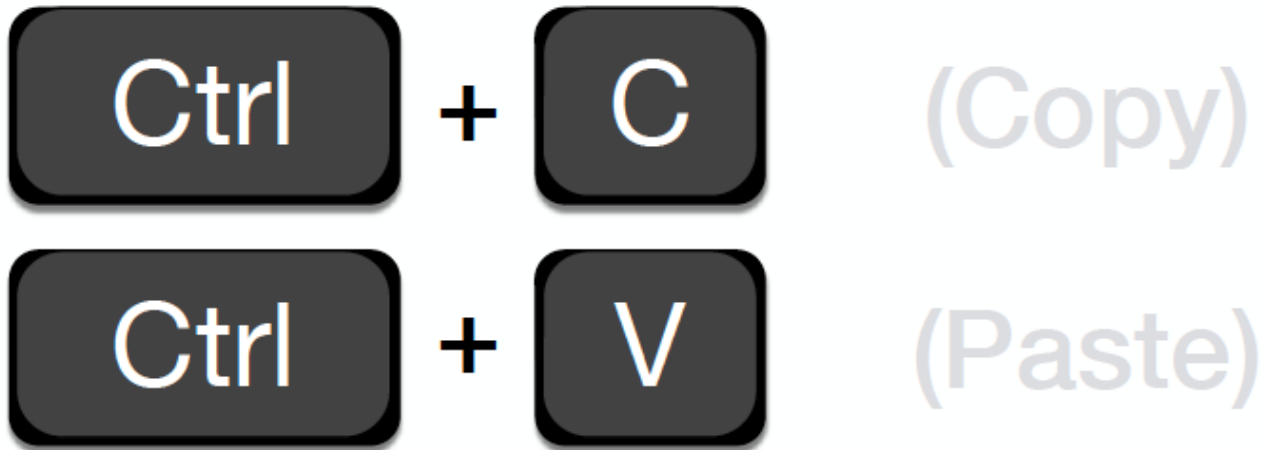
Research project using reproducible practices



@dsquintana

How?

How can you reproduce results?



What do we need to make research reproducible?

- Data in some coherent format
- Programming language (R, Python)
- Text, figures and code in same environment (literate programming)
- Continuous and transparent editions and updates (version control)

This workshop

- Motivations (these slides)
- Manipulating data in the tidyverse
- Visualising data in the tidyverse
- Writing dynamic and reproducible documents with R Markdown
- Versioning with Git and GitHub in RStudio
- Take-home messages

Further reading

- Alston & Rick (2021). [A Beginner's Guide to Conducting Reproducible Research](#). Bull Ecol Soc Am.
- Baker (2016). [1,500 scientists lift the lid on reproducibility](#). Nature. 533:452-4.
- Lowndes et al. (2017). [Our path to better science in less time using open data science tools](#). Nat Ecol Evol 1:0160.
- Wilson et al. (2017). [Good enough practices in scientific computing](#). PLoS Comput Biol 13(6): e1005510.
- Markowetz (2015). [Five selfish reasons to work reproducibly](#). Genome Biol 16: 274.
- BES Guides to Better Science: [Reproducible Code](#).
- Casajus, Dray, Gimenez, Guéry, Guilhaumon & Schiettekatte (2020). [Formation FRB CESAB / GDR ECOSTAT : Bonnes pratiques pour une recherche reproductible en écologie numérique](#)