Barriers to Open Research (and how to overcome them)

@kirstie_j
doi: https://dx.doi.org/10.6084/m9.figshare.5914444

- Research fellow at the Alan Turing Institute for Data Science
- Senior research associate in the Brain Mapping Unit, Department of Psychiatry, University of Cambridge
- 2016/17 Mozilla Fellow for Science

@kirstie_j
doi: https://dx.doi.org/10.6084/m9.figshare.5914444

@kirstie_j
doi: https://dx.doi.org/10.6084/m9.figshare.5914444

Reproducible vs Replicable

@kirstie_j
doi: https://dx.doi.org/10.6084/m9.figshare.5914444

		Data	
		Same	Different
Code	Same		
	Different		

@kirstie_j
doi: https://dx.doi.org/10.6084/m9.figshare.5914444

		Data	
		Same	Different
Code	Same	Reproducible	
	Different		

@kirstie_j
doi: https://dx.doi.org/10.6084/m9.figshare.5914444

		Data	
		Same	Different
Code	Same	Reproducible	Replicable
	Different		

@kirstie_j
doi: https://dx.doi.org/10.6084/m9.figshare.5914444

		Data	
		Same	Different
Code	Same	Reproducible	Replicable
	Different	Robust	

@kirstie_j
doi: https://dx.doi.org/10.6084/m9.figshare.5914444

		Data	
		Same	Different
Code	Same	Reproducible	Replicable
	Different	Robust	Generalisable

@kirstie_j
doi: https://dx.doi.org/10.6084/m9.figshare.5914444

Barriers to reproducible research

@kirstie_j
doi: https://dx.doi.org/10.6084/m9.figshare.5914444

Is not considered for promotion

Requires additional skills

Plead the 5th

Support additional users

Barriers to reproducible research

Takes time

Held to higher standards than others

Publication bias towards novel findings

@kirstie_j

doi: https://dx.doi.org/10.6084/m9.figshare.5914444

Barriers to reproducible research

Is not considered for promotion

Requires additional skills

Plead the 5th

Support additional users

Takes time

Held to higher standards than others

Publication bias towards novel findings

@kirstie_j

doi: https://dx.doi.org/10.6084/m9.figshare.5914444

Is not considered for promotion

Requires additional skills

Plead the 5th

Support additional users

Barriers to reproducible research

Takes time

Held to higher standards than others

Publication bias towards novel findings

@kirstie_j

doi: https://dx.doi.org/10.6084/m9.figshare.5914444

Is not considered for promotion

Requires additional skills

Plead the 5th

Support additional users

Barriers to reproducible research

Takes time

Held to higher standards than others

Publication bias towards novel findings

@kirstie_j

doi: https://dx.doi.org/10.6084/m9.figshare.5914444

Is not considered for promotion

Requires additional skills

Plead the 5th

Support additional users

Barriers to reproducible research

Takes time

Held to higher standards than others

Publication bias towards novel findings

@kirstie_j

doi: https://dx.doi.org/10.6084/m9.figshare.5914444

Is not considered for promotion

Requires additional skills

Plead the 5th

Support additional users

Barriers to reproducible research

Takes time

Held to higher standards than others

Publication bias towards novel findings

@kirstie_j

doi: https://dx.doi.org/10.6084/m9.figshare.5914444

Is not considered for promotion

Requires additional skills

Plead the 5th

Support additional users

Barriers to reproducible research

Takes time

Held to higher standards than others

Publication bias towards novel findings

@kirstie_j

doi: https://dx.doi.org/10.6084/m9.figshare.5914444

Is not considered for promotion

Requires additional skills

Plead the 5th

Support additional users

Barriers to reproducible research

Takes time

Held to higher standards than others

Publication bias towards novel findings

@kirstie_j

doi: https://dx.doi.org/10.6084/m9.figshare.5914444

Aim for 50% comments in your code

@kirstie_j
doi: https://dx.doi.org/10.6084/m9.figshare.5914444

Share your comments with the original author

@kirstie_j
doi: https://dx.doi.org/10.6084/m9.figshare.5914444

(which is almost always going to be YOU!)

@kirstie_j
doi: https://dx.doi.org/10.6084/m9.figshare.5914444

** It is embarrassing to me how often I forget not just details of experiments, but entire experiments. For example, for the manuscript I am working on now, I forgot that we had done an experiment to test for vertical transmission of the parasite. Fortunately, the undergrad who has been working on the project remembered and had it in his writeup!

https://dynamicecology.wordpress.com/2015/02/18/the-biggest-benefit-of-my-shift-to-r-reproducibility

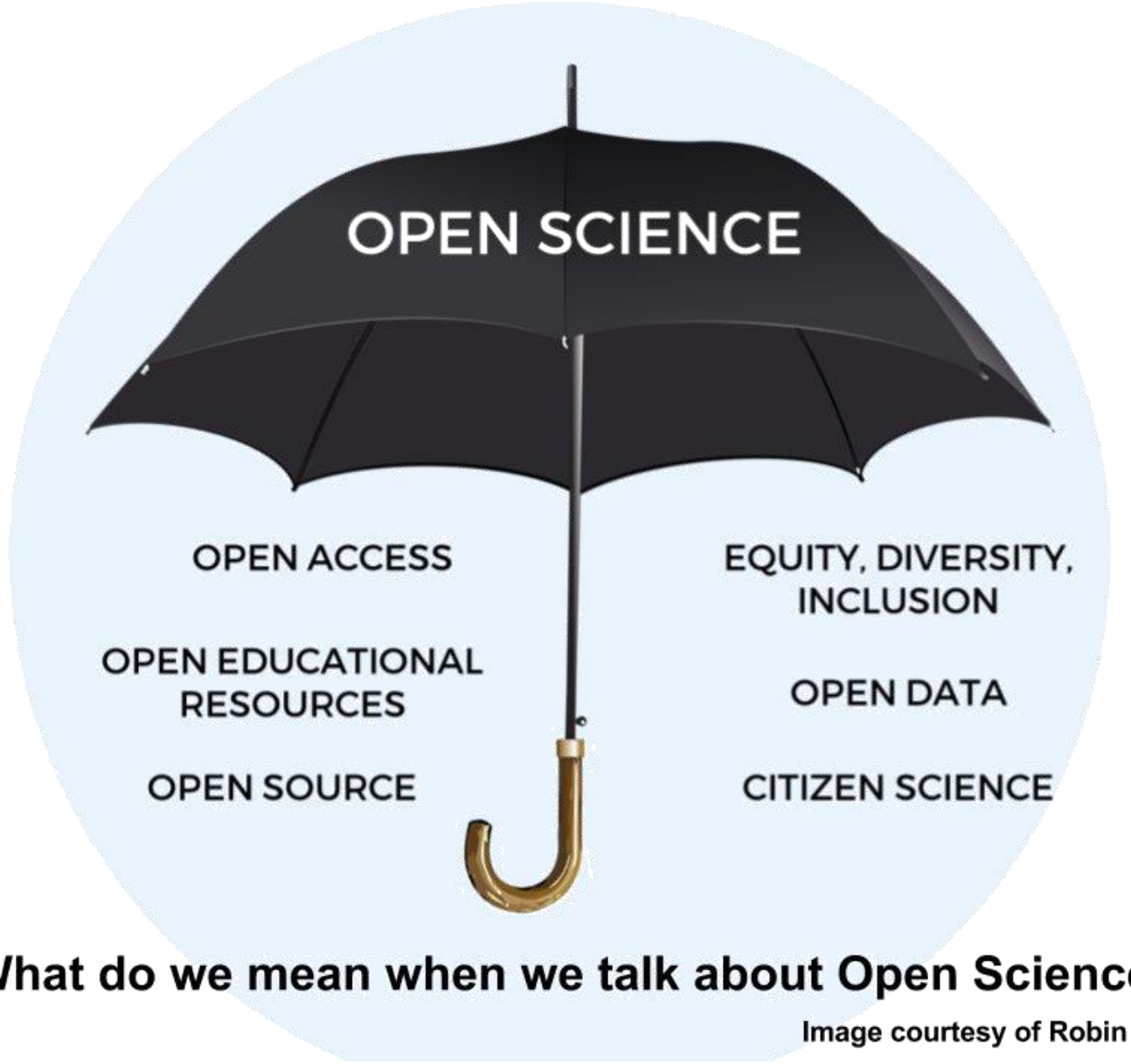
https://twitter.com/kcranstn/status/370914072511791104

@kirstie_j

doi: https://dx.doi.org/10.6084/m9.figshare.5914444

Reproducible <> Open
(and that's fine)

@kirstie_j
doi: https://dx.doi.org/10.6084/m9.figshare.5914444



What do we mean when we talk about Open Science?

Image courtesy of Robin Champieux

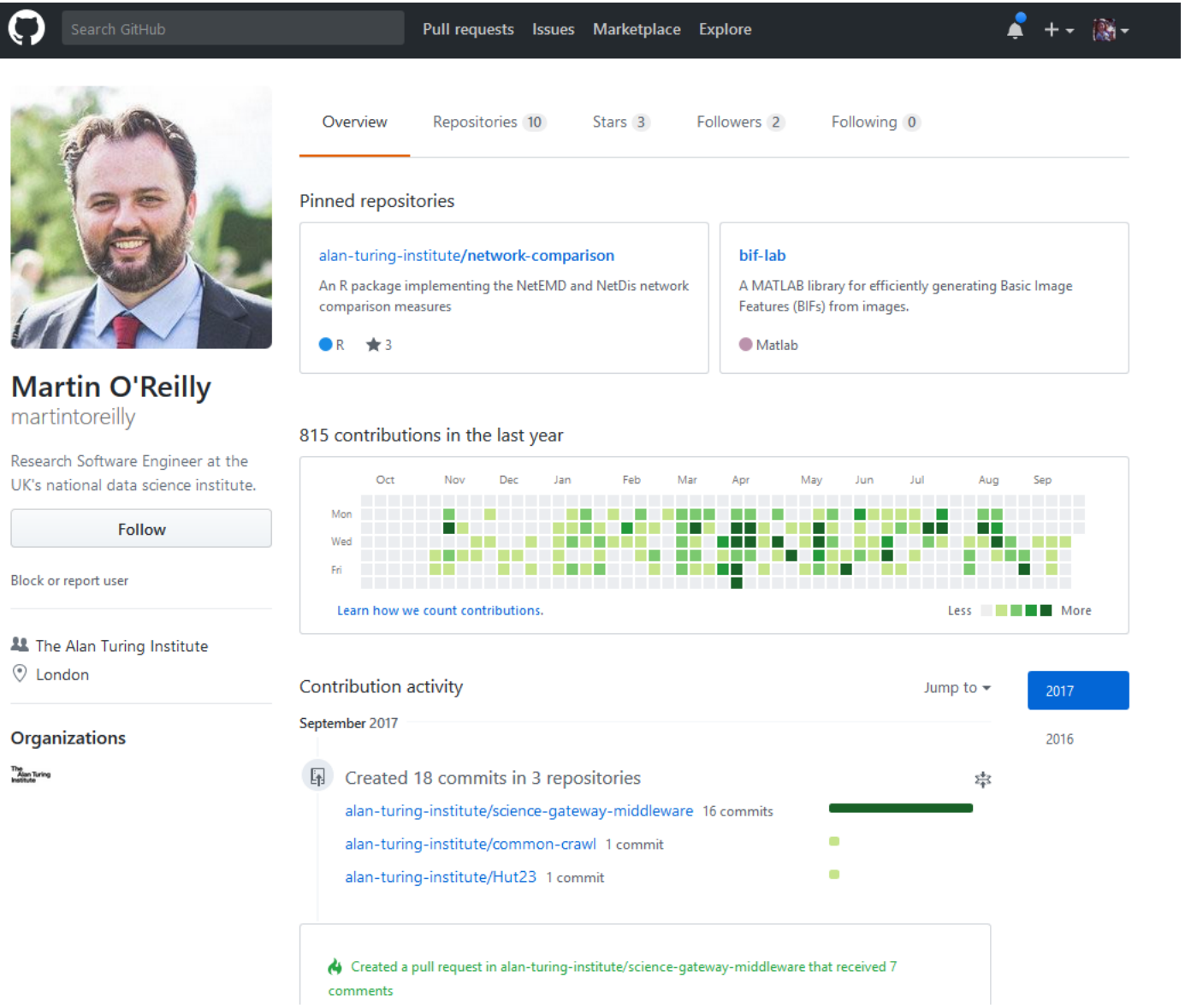
@kirstie_j

doi: https://dx.doi.org/10.6084/m9.figshare.5914444

Reproducible research at the Turing

@kirstie_j
doi: https://dx.doi.org/10.6084/m9.figshare.5914444

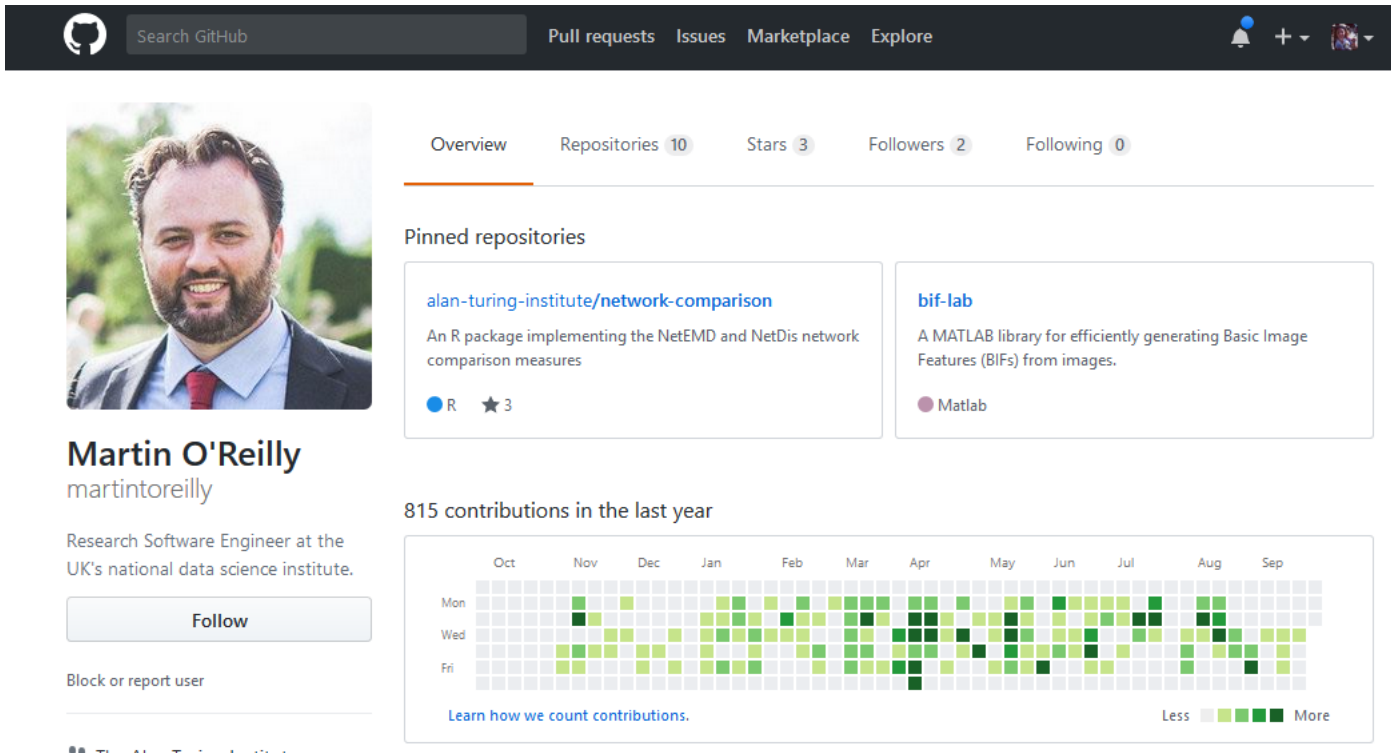
The Alan Turing Institute



@kirstie_j

doi: https://dx.doi.org/10.6084/m9.figshare.5914444

The Alan Turing Institute



My colleagues in the research software engineering team and I **seek to position The Alan Turing Institute as a world leader for reproducible research** by embedding a culture of "reproducible by default" at the institute and providing training and tools to make reproducible research "too easy not to do".

@kirstie_j

doi: https://dx.doi.org/10.6084/m9.figshare.5914444

Reproducible research at the Turing

- Training & support for incoming students
- Software development support
  - From the beginning if possible
  - Posthoc for now
- Turing Reproducibility Champions
- Institutional buy in

@kirstie_j
doi: https://dx.doi.org/10.6084/m9.figshare.5914444

Reproducible research at the Turing

- Training & support for incoming students
- Software development support
  - From the beginning if possible
  - Posthoc for now
- Turing Reproducibility Champions
- Institutional buy in
- "The Turing Way"

Attr: edhiggins, CC BY 3.0, via Wikimedia Commons

Find what works for you

@kirstie_j
doi: https://dx.doi.org/10.6084/m9.figshare.5914444

Find what works for you

Every little helps

@kirstie_j
doi: https://dx.doi.org/10.6084/m9.figshare.5914444

DATA CARPENTRY

software carpentry

You can do this!

https://www.mozillascience.org

http://software-carpentry.org

http://data-carpentry.org

https://www.coursera.org/specializations/jhu-data-science

https://www.coursera.org/learn/python/home/info

@kirstie_j

doi: https://dx.doi.org/10.6084/m9.figshare.5914444

DATA CARPENTRY

software carpentry

WE can do this!

https://www.mozillascience.org

http://software-carpentry.org

http://data-carpentry.org

https://www.coursera.org/specializations/jhu-data-science

https://www.coursera.org/learn/python/home/info

@kirstie_j

doi: https://dx.doi.org/10.6084/m9.figshare.5914444

The Alan Turing Institute

UNIVERSITY OF CAMBRIDGE

Thank you!

WhitakerLab

@kirstie_j

doi: 10.6084/m9.figshare.5914444

Start small

@kirstie_j
doi: https://dx.doi.org/10.6084/m9.figshare.5914444

Protocols.io

@kirstie_j
doi: https://dx.doi.org/10.6084/m9.figshare.5914444

UCHANGE structural neuroimaging pipeline

Kirstie Whitaker [1]

[1] Brain Mapping Unit, Department of Psychiatry, University of Cambridge

10.17504/protocols.io.eu5bey6

Steps

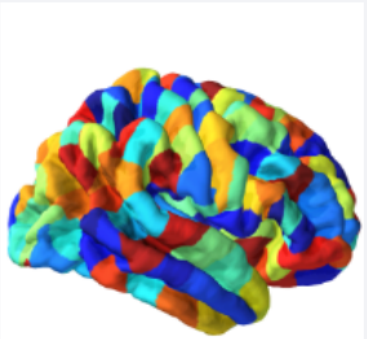
1 Check or complete software installation

You must have the following software libraries installed:

- **Freesurfer**: https://surfer.nmr.mgh.harvard.edu/fswiki/DownloadAndInstall
- **FSL**: http://fsl.fmrib.ox.ac.uk/fsl/fslwiki/FslInstallation
- **Anaconda**: https://www.continuum.io/downloads

From within **Anaconda** install the following python packages following the command below:

- VTK
- Mayavi
- Nibabel
- pysurfer



CONTACT
Kirstie Whitaker

UCHANGE structural neuroimaging pipeline

Kirstie Whitaker [1]

[1] Brain Mapping Unit, Department of Psychiatry, University of Cambridge
10.17504/protocols.io.eu5bey6

Steps

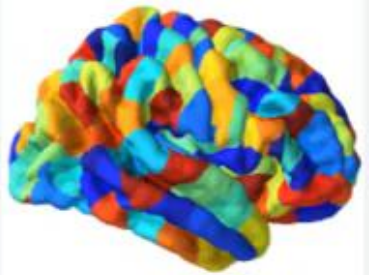
1 Check or complete software installation

You must have the following software libraries installed:

- **Freesurfer**: https://surfer.nmr.mgh.harvard.edu/fswiki/DownloadAndInstall
- **FSL**: http://fsl.fmrib.ox.ac.uk/fsl/fslwiki/FslInstallation
- **Anaconda**: https://www.continuum.io/downloads

From within **Anaconda** install the following python packages following the command below:

- VTK
- Mayavi
- Nibabel
- pysurfer



CONTACT
Kirstie Whitaker

Asking for help: StackOverflow & MWEs

http://stackoverflow.com

@kirstie_j
doi: https://dx.doi.org/10.6084/m9.figshare.5914444

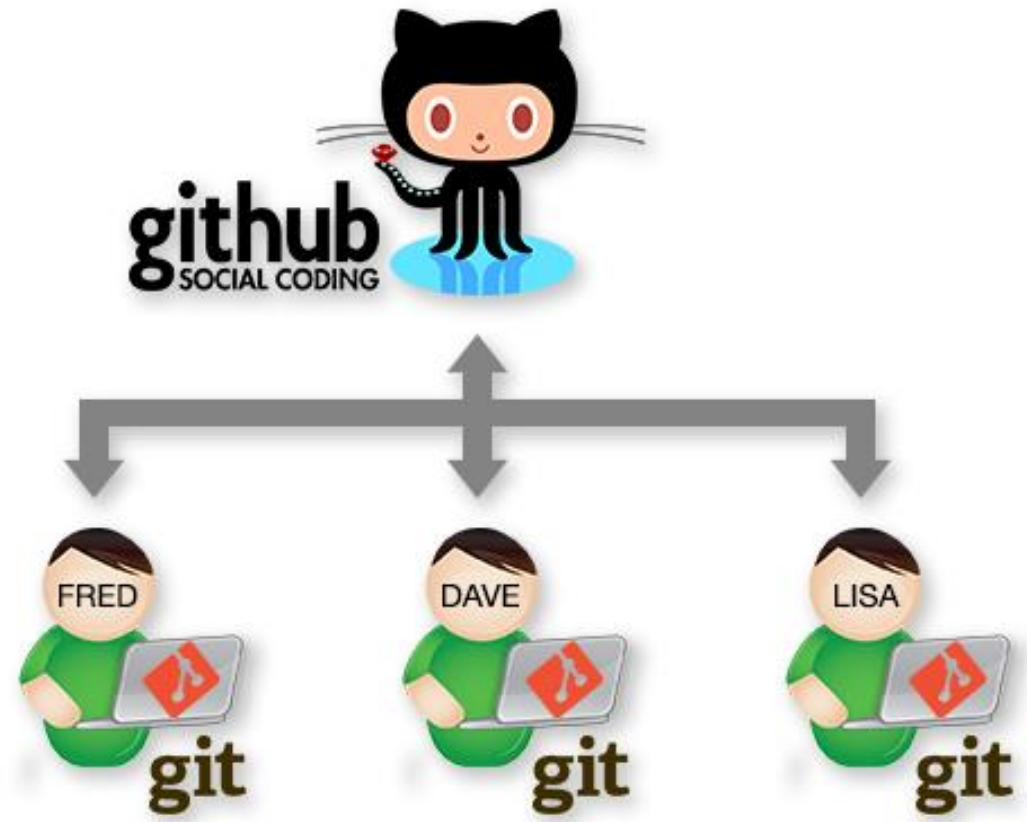
https://mrcolley.com/2014/07/01/
talk-to-the-duck-debugging-and-resilience

@kirstie_j
doi: https://dx.doi.org/10.6084/m9.figshare.5914444

Version Control

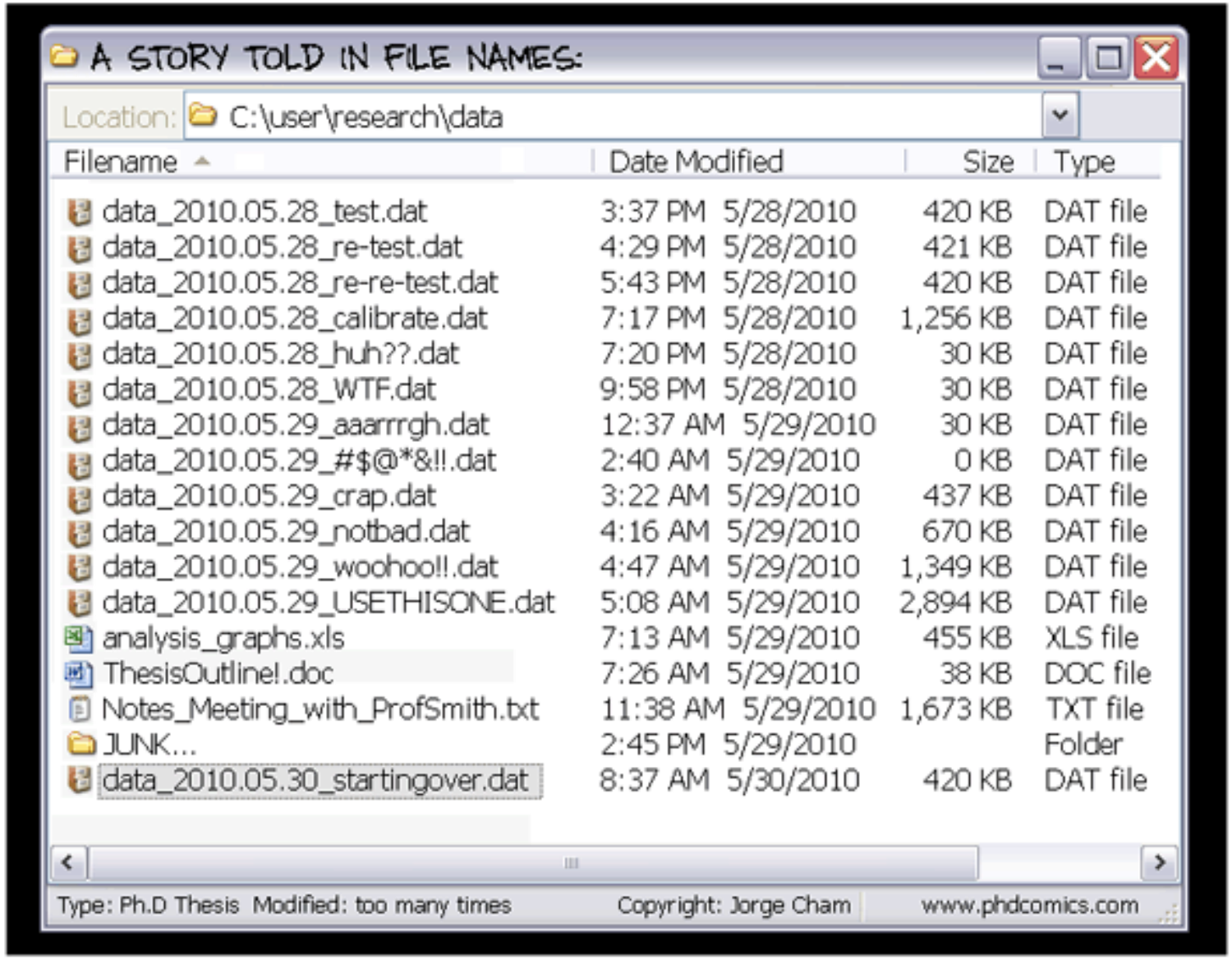
@kirstie_j
doi: https://dx.doi.org/10.6084/m9.figshare.5914444

@kirstie_j
doi: https://dx.doi.org/10.6084/m9.figshare.5914444



@kirstie_j

doi: https://dx.doi.org/10.6084/m9.figshare.5914444

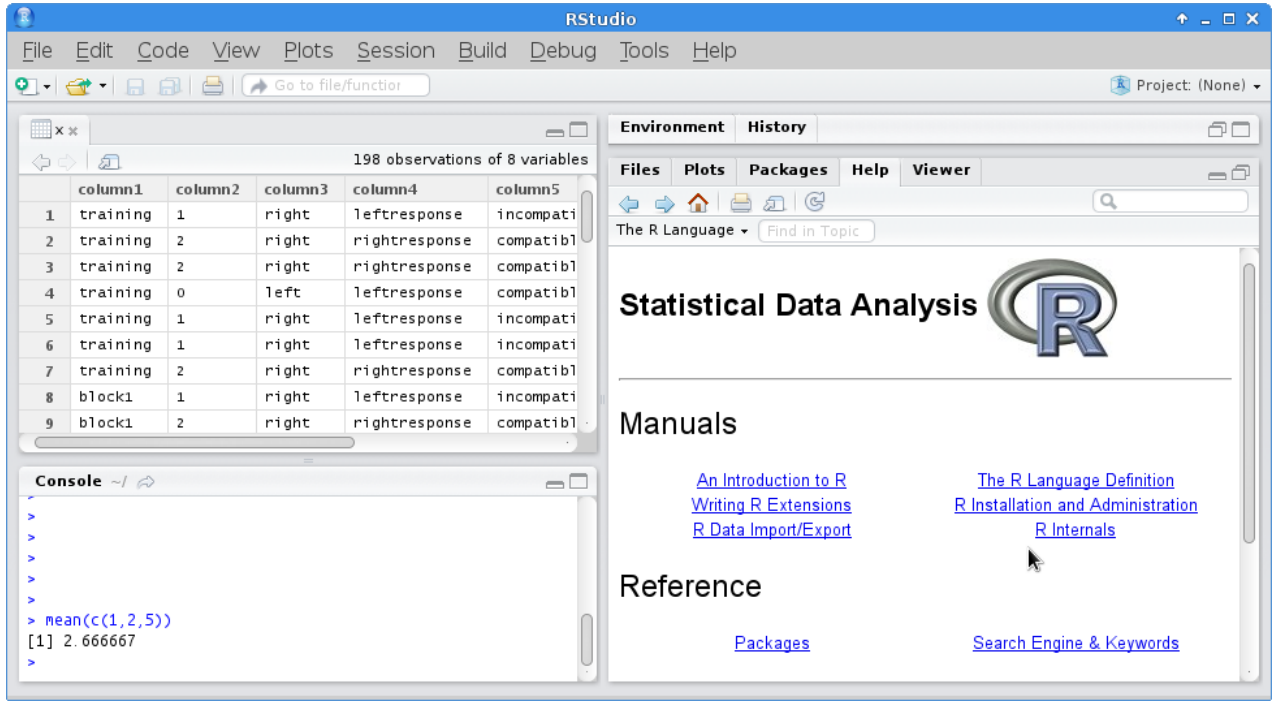


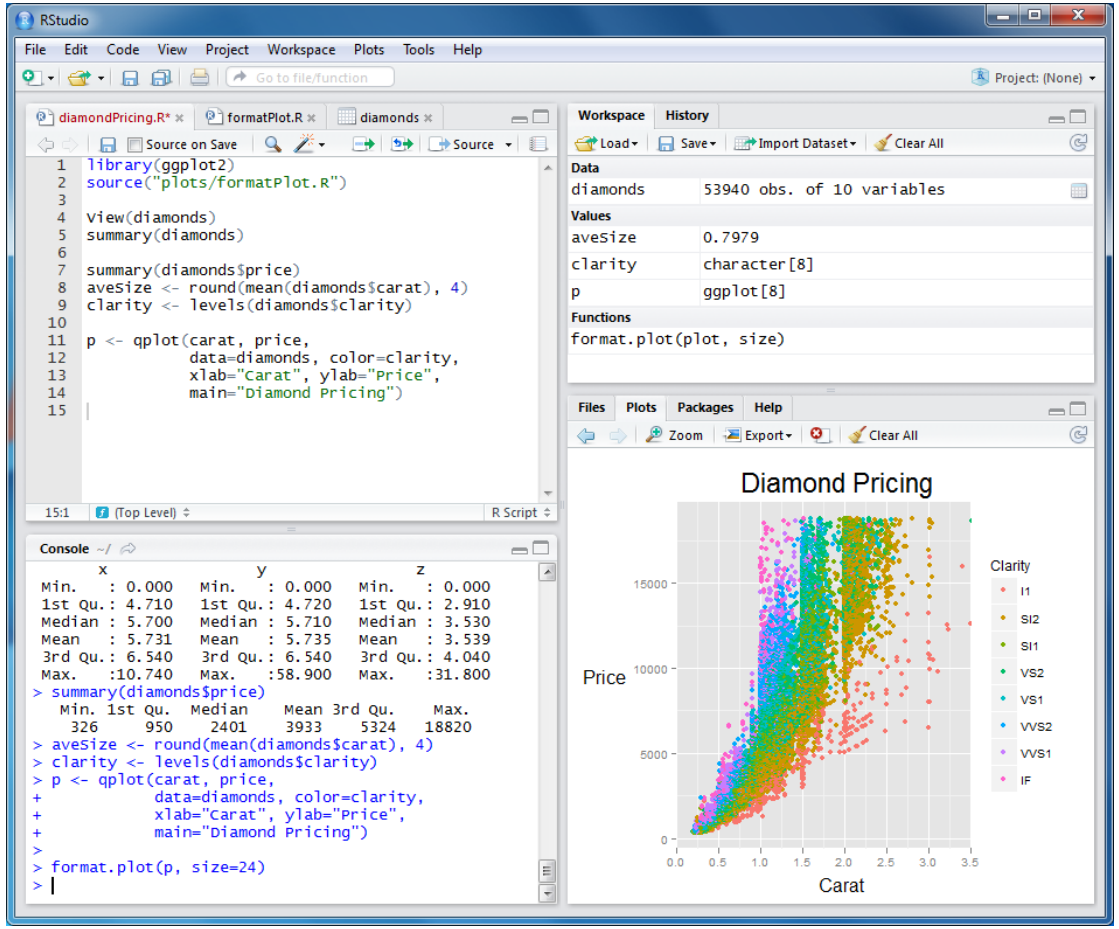
@kirstie_j
doi: https://dx.doi.org/10.6084/m9.figshare.5914444

Python vs R

(vs Matlab vs STATA etc…)

@kirstie_j
doi: https://dx.doi.org/10.6084/m9.figshare.5914444





@kirstie_j
doi: https://dx.doi.org/10.6084/m9.figshare.5914444

Shiny by RStudio

Search

OVERVIEW

TUTORIAL

ARTICLES

GALLERY

REFERENCE

DEPLOY

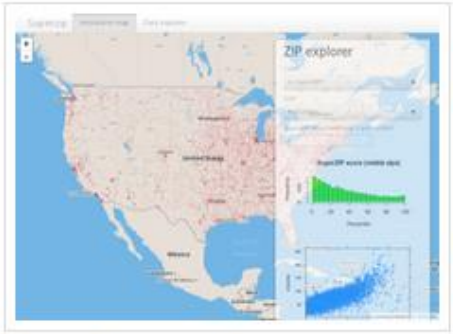
HELP

Gallery

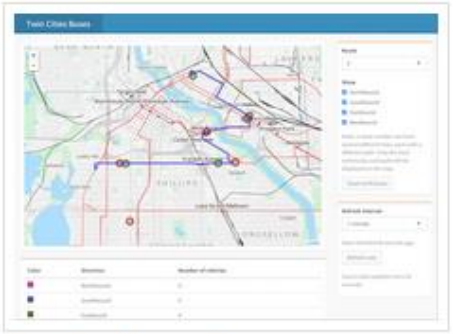
This gallery contains useful examples to learn from. Visit the Shiny User Showcase to see an inspiring set of sophisticated apps.

Interactive visualizations

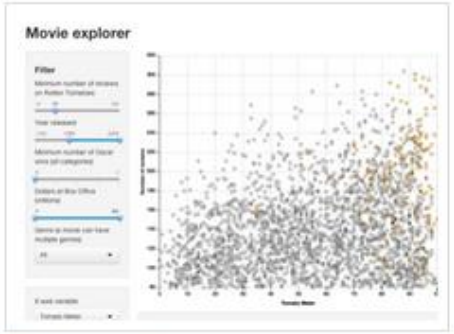
Shiny is designed for fully interactive visualization, using JavaScript libraries like d3, Leaflet, and Google Charts.



SuperZip example



Bus dashboard



Movie explorer

Google Charts

Shiny gallery

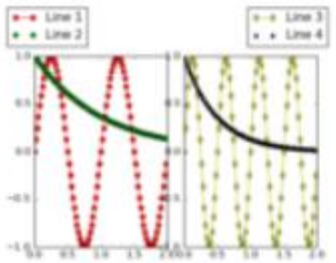
@kirstie_j

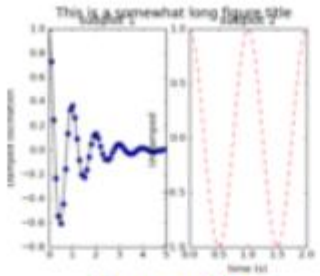
doi: https://dx.doi.org/10.6084/m9.figshare.5914444

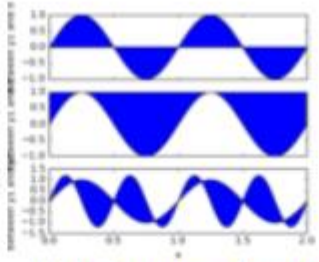
@kirstie_j
doi: https://dx.doi.org/10.6084/m9.figshare.5914444

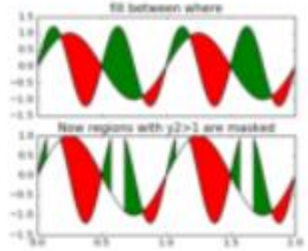
Matplotlib gallery

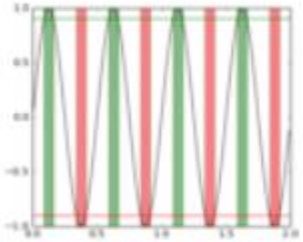
Click each example to see source code

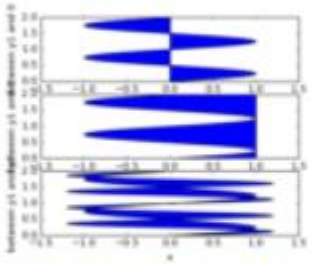

figlegend_demo

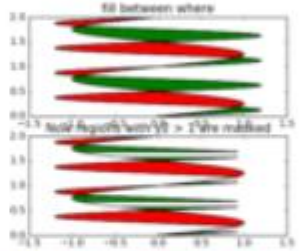

figure_title

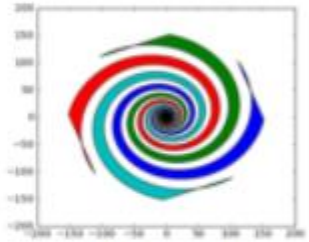

fill_between_demo


fill_between_demo

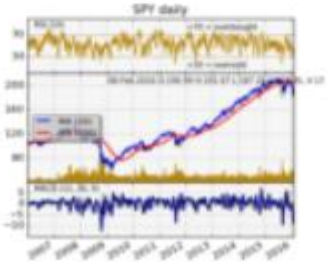

fill_between_demo

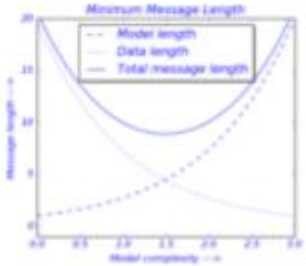

fill_betweenx_demo

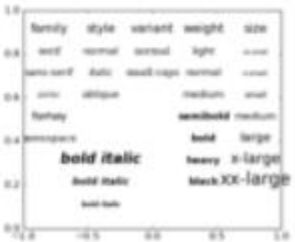

fill_betweenx_demo


fill_spiral


finance_demo

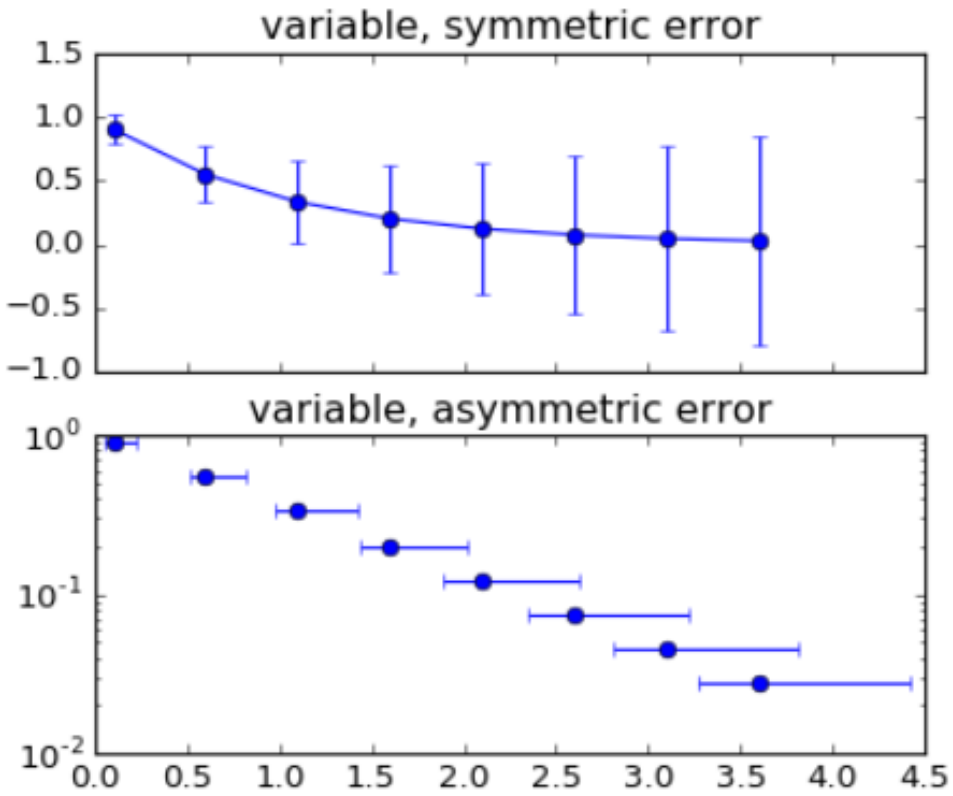

finance_work2


findobj_demo


fonts_demo

statistics example code: errorbar_demo_features.py

(Source code, png, hires.png, pdf)




```
"""
Demo of errorbar function with different ways of specifying error bars.

Errors can be specified as a constant value (as shown in `errorbar_demo.py`),
or as demonstrated in this example, they can be specified by an N x 1 or 2 x N,
where N is the number of data points.

N x 1:
    Error varies for each point, but the error values are symmetric (i.e. the
    lower and upper values are equal).

2 x N:
    Error varies for each point, and the lower and upper limits (in that order)
    are different (asymmetric case)

In addition, this example demonstrates how to use log scale with errorbar.
"""
import numpy as np
import matplotlib.pyplot as plt

# example data
x = np.arange(0.1, 4, 0.5)
y = np.exp(-x)
# example error bar values that vary with x-position
error = 0.1 + 0.2 * x
# error bar values w/ different -/+ errors
lower_error = 0.4 * error
upper_error = error
asymmetric_error = [lower_error, upper_error]

fig, (ax0, ax1) = plt.subplots(nrows=2, sharex=True)
ax0.errorbar(x, y, yerr=error, fmt='-o')
ax0.set_title('variable, symmetric error')

ax1.errorbar(x, y, xerr=asymmetric_error, fmt='o')
ax1.set_title('variable, asymmetric error')
ax1.set_yscale('log')
plt.show()
```

http://matplotlib.org/examples/statistics/errorbar_demo_features.html

@kirstie_j

doi: https://dx.doi.org/10.6084/m9.figshare.5914444

Jupyter Notebook

http://jupyter.org

@kirstie_j
doi: https://dx.doi.org/10.6084/m9.figshare.5914444

SIGNAL PROCESSING WITH GW150914 OPEN DATA

Welcome! This ipython notebook (or associated python script GW150914_tutorial.py) will go through some typical signal processing tasks on strain time-series data associated with the LIGO GW150914 data release from the LIGO Open Science Center (LOSC):

- https://losc.ligo.org/events/GW150914/
- View the tutorial as a web page - https://losc.ligo.org/s/events/GW150914/
- Download the tutorial as a python script - https://losc.ligo.org/s/events/GW1
- Download the tutorial as iPython Notebook - https://losc.ligo.org/s/events/G

To begin, download the ipython notebook, readligo.py, and the data files listed b
can run the python script GW150914_tutorial.py. You will need the python packa

On Windows, or if you prefer, you can use a python development environment s
/why-anaconda) or Enthought Canopy (https://www.enthought.com/products/ca

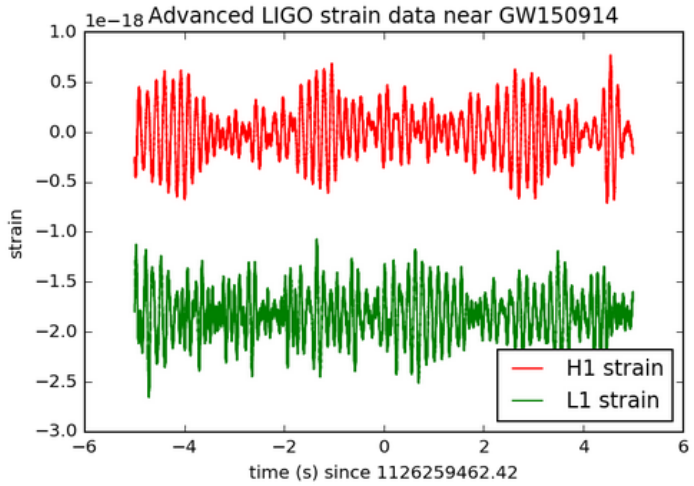
Questions, comments, suggestions, corrections, etc: email losc@ligo.org

v20160208b

https://losc.ligo.org/s/events/GW150914/GW150914_tutorial.html

In [6]:
```
# plot +- 5 seconds around the event:
tevent = 1126259462.422         # Mon Sep 14 09:50:45 GMT 2015
deltat = 5.                     # seconds around the event
# index into the strain time series for this time interval:
indxt = np.where((time_H1 >= tevent-deltat) & (time_H1 < tevent+deltat))

plt.figure()
plt.plot(time_H1[indxt]-tevent,strain_H1[indxt],'r',label='H1 strain')
plt.plot(time_L1[indxt]-tevent,strain_L1[indxt],'g',label='L1 strain')
plt.xlabel('time (s) since '+str(tevent))
plt.ylabel('strain')
plt.legend(loc='lower right')
plt.title('Advanced LIGO strain data near GW150914')
plt.savefig('GW150914_strain.png')
```



Observation of Gravitational Waves from a Binary Black Hole Merger

B. P. Abbott *et al.*[*]
(LIGO Scientific Collaboration and Virgo Collaboration)
(Received 21 January 2016; published 11 February 2016)

Repository

Clone

Pull request

Branch

Version control

Some jargon busting

Merge

Issues

Release

Commit

Fork

Markdown

@kirstie_j

doi: https://dx.doi.org/10.6084/m9.figshare.5914444

Submit your first pull request!

Inspired by: https://yourfirstpr.github.io

@kirstie_j
doi: https://dx.doi.org/10.6084/m9.figshare.5914444

ReproducibleResearch

This repository supports Kirstie's presentation on tips and tricks for making your research reproducible.

The goal is to build a directory of useful links, and a jargon busting glossary.

Guide for contributors

See our guidelines for how to contribute to the project.

Code of conduct

Everyone is welcome to join this project, particularly people who have not used GitHub before and are feeling unsure of how to begin!

Please follow our code of conduct in all your on and offline interactions.

@kirstie_j
doi: https://dx.doi.org/10.6084/m9.figshare.5914444

Aim for 40% comments in your code

@kirstie_j
doi: https://dx.doi.org/10.6084/m9.figshare.5914444

Share your comments with the original author

@kirstie_j
doi: https://dx.doi.org/10.6084/m9.figshare.5914444

(which is almost always going to be YOU!)

@kirstie_j
doi: https://dx.doi.org/10.6084/m9.figshare.5914444

What if I can't share my code until I'm published?

@kirstie_j
doi: https://dx.doi.org/10.6084/m9.figshare.5914444

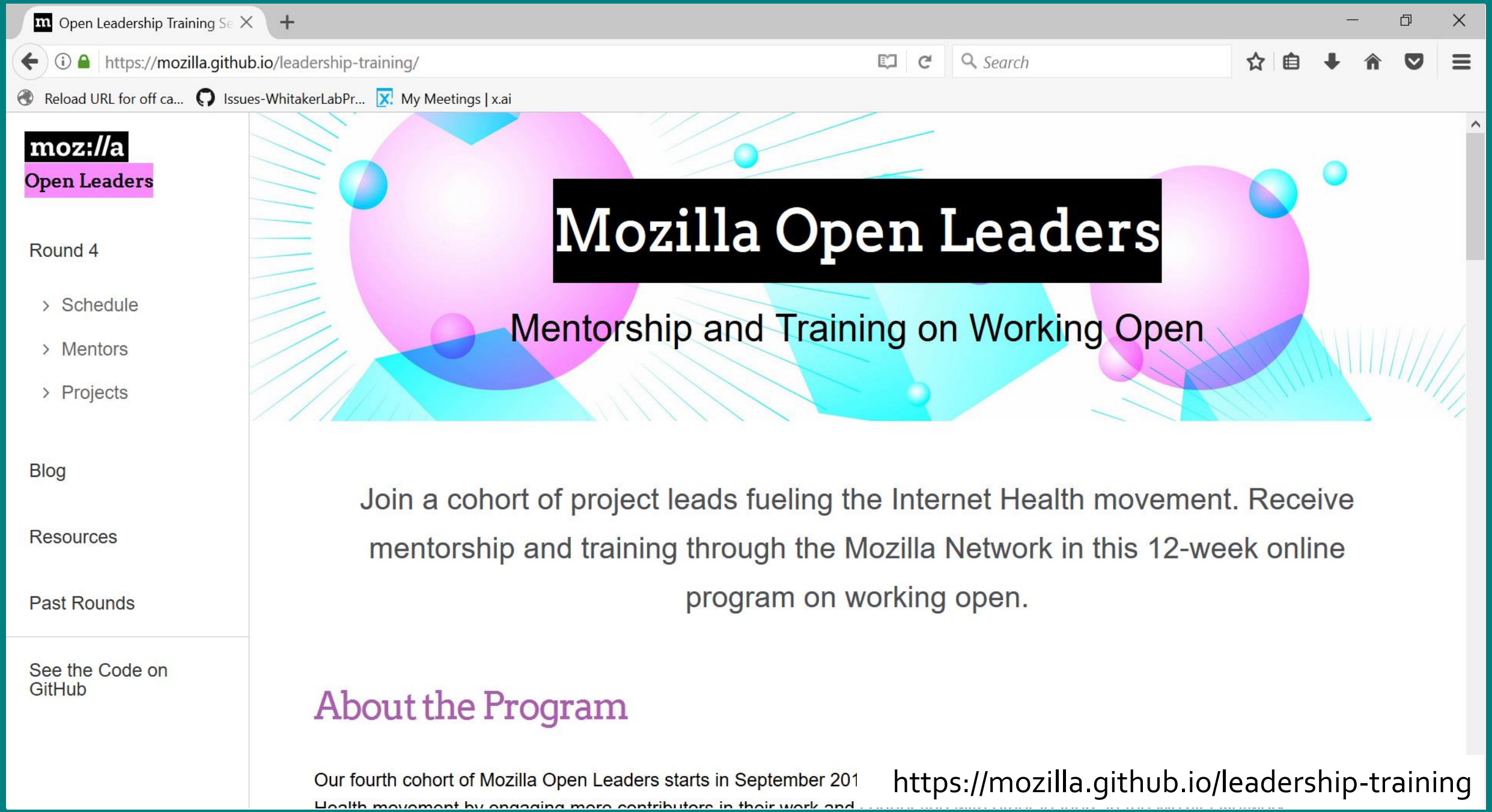
https://education.github.com

@kirstie_j

doi: https://dx.doi.org/10.6084/m9.figshare.5914444

Reproducible <> Open
(and that's fine)

@kirstie_j
doi: https://dx.doi.org/10.6084/m9.figshare.5914444


Open Leadership Training Se
https://mozilla.github.io/leadership-training/
Search
Reload URL for off ca...
Issues-WhitakerLabPr...
My Meetings | x.ai
moz://a
Open Leaders
Round 4
Schedule
Mentors
Projects
Blog
Resources
Past Rounds
See the Code on GitHub
Mozilla Open Leaders
Mentorship and Training on Working Open
Join a cohort of project leads fueling the Internet Health movement. Receive mentorship and training through the Mozilla Network in this 12-week online program on working open.
About the Program
Our fourth cohort of Mozilla Open Leaders starts in September 201
https://mozilla.github.io/leadership-training

What if I can't (ever) share my data?

@kirstie_j
doi: https://dx.doi.org/10.6084/m9.figshare.5914444

AUTHOR
creates pseudo IDs, and deposits data, metadata and text into Data Repository under permanent embargo BEFORE PAPER IS ACCEPTED

REPOSITORY
generates unique DOI number (Digital Online Identifier)

AUTHOR
adds DOI to manuscript with standard text: "Data for this specific paper has been uploaded to the Cambridge Data Repository ([INSERT DOI LINK]) and password protected. Our participants did not give informed consent for their data to be made publicly available, and it is possible that they could be identified from this data set. Access to the data supporting the analyses presented in this paper will be made available to researchers with a reasonable request to OpenNSPN@medschl.cam.ac.uk."

OPEN NSPN
keeps record of DOIs

REQUESTER
Either 'Requests a copy' at the data repository OR sends an email to Open NSPN.

REPOSITORY
forwards request to Open NSPN

OPEN NSPN
forwards request to First Author, Last Author and Data Management Chair for decision

YES
External

OPEN NSPN
sends Data Access Agreement (DAA) to Requester

REQUESTER
returns DAA

OPEN NSPN
Authorises data access request

YES
NSPN team member

NO

OPEN NSPN
sends refusal to Requester

END

REPOSITORY
Emails data to requester (happens automatically)

22nd May 2017
version 0.2

New Results

Previous Next

Adolescent Tuning Of Association Cortex In Human Structural Brain Networks

František Váša, Jakob Seidlitz, Rafael Romero-Garcia, Kirstie J. Whitaker, Gideon Rosenthal, Petra E. Vértes, Maxwell Shinn, Aaron Alexander-Bloch, Peter Fonagy, Raymond J. Dolan, Peter B. Jones, Ian M. Goodyer, The NSPN Consortium, Olaf Sporns, Edward T. Bullmore

doi: https://doi.org/10.1101/126920

Posted September 15, 2017.

Abstract

How does human brain organization change over the course of adolescence? Motivated by prior data on local cortical shrinkage and intracortical myelination, we predicted age-related changes in topological organisation of cortical structural networks. We estimated the structural correlation matrix from magnetic resonance imaging (MRI) measures of cortical thickness at 308 regions in a sample of N=297 healthy participants, aged 14-24 years (inclusive). We used

Subject Area

Neuroscience

Availability of data and code

Data for this specific paper has been uploaded to the Cambridge Data Repository (https://doi.org/10.17863/CAM.8856) and password protected. Our participants did not give informed consent for their questionnaire measures to be made publicly available, and it is possible that they could be identified from this data set. Access to the data supporting the analyses presented in this paper will be made available to researchers with a reasonable request to NSPNdata@medschl.cam.ac.uk. The code used to conduct analyses is available from FV's github: https://github.com/frantisekvasa/structural_network_development (DOI: 10.5281/zenodo.528674).

Abstract

How does
prior data
changes in topological organisation of cortical structural networks. We estimated the structural correlation matrix from magnetic resonance imaging (MRI) measures of cortical thickness at 308 regions in a sample of N=297 healthy participants, aged 14-24 years (inclusive). We used

UNIVERSITY OF CAMBRIDGE

Apollo

Browse

All of Apollo

- Communities & Collections
- Authors
- Titles
- Keywords
- Type

This Collection

- Authors
- Titles
- Keywords
- Type

Statistics

Data supporting NSPN publication "Adolescent tuning of association cortex in human structural brain networks"

View / Open Files

str.net.dev.data.RData (Unknown, 7Mb)

data-processing-and-description.docx (Microsoft Word 2007, 106Kb)

Authors

Vasa, Frantisek

Citation

Vasa, F., Seidlitz, J., Romero Garcia, R., Whitaker, K. J., Rosenthal, G., Vertes, P. E., Shinn, M., et al. *Data supporting NSPN publication "Adolescent tuning of association cortex in human structural brain networks"* [Dataset]. https://doi.org/10.17863/CAM.8856

Description

There are two files supporting this publication. The first is an RData file containing all variables necessary to reproduce the main findings of the publication. The second is a Microsoft Word document describing 1) how the data was collected and processed and 2) all the variables stored in the RData file.

Software

Custom scripts written in R, available from Frantisek Vasa's github page: https://github.com/frantisekvasa/structural_network_development (DOI: 10.5281/zenodo.528674)

Keywords

Váša et al, 2017

UNIVERSITY OF CAMBRIDGE

Apollo

Browse

All of Apollo

- Communities & Collections
- Authors
- Titles
- Keywords
- Type

This Collection

- Authors
- Titles
- Keywords
- Type

Statistics

Data supporting NSPN publication "Adolescent tuning of association cortex in human structural brain networks"

View / Open Files

str.net.dev.data.RData (Unknown, 7Mb)

data-processing-and-description.docx (Microsoft Word 2007, 106Kb)

Authors

Vasa, Frantisek

Citation

Vasa, F., Seidlitz, J., Romero Garcia, R., Whitaker, K. J., Rosenthal, G., Vertes, P. E., Shinn, M., et al. *Data supporting NSPN publication "Adolescent tuning of association cortex in human structural brain networks"* [Dataset]. https://doi.org/10.17863/CAM.8856

Description

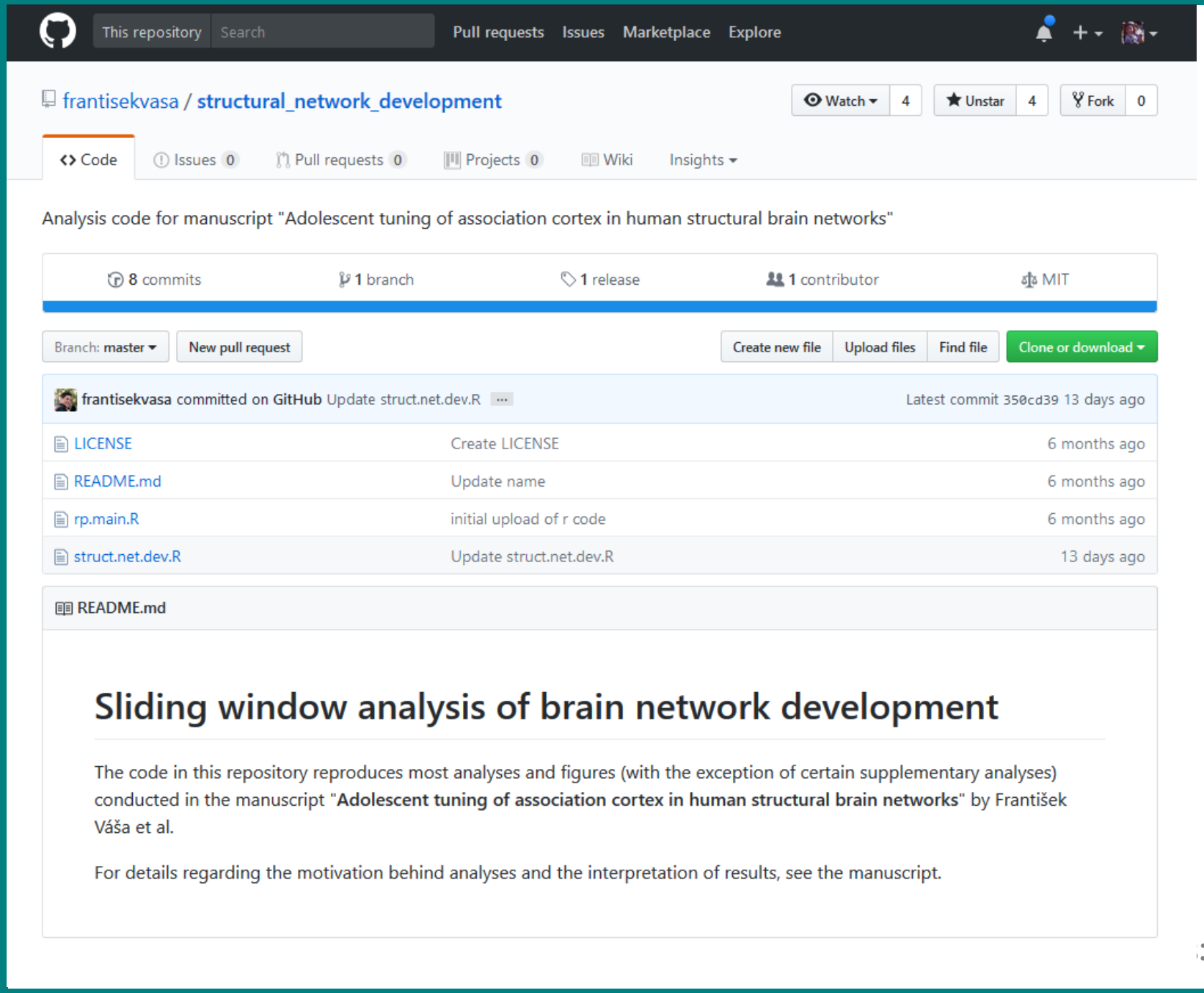
There are two files supporting this publication. The first is an RData file containing all variables necessary to reproduce the main findings of the publication. The second is a Microsoft Word document describing 1) how the data was collected and processed and 2) all the variables stored in the RData file.

Software

Custom scripts written in R, available from Frantisek Vasa's github page: https://github.com/frantisekvasa/structural_network_development (DOI: 10.5281/zenodo.528674)

Keywords

Váša et al, 2017



@kirstie_j

://dx.doi.org/10.6084/m9.figshare.5914444

This repository Search Pull requests Issues Marketplace Explore

frantisekvasa / structural_network_development

Watch 4 Unstar 4 Fork 0

Code Issues 0 Pull requests 0 Projects 0 Wiki Insights

Analysis code for manuscript "Adolescent tuning of association corte

8 commits 1 branch 1

Branch: master New pull request

frantisekvasa committed on GitHub Update struct.net.dev.R

LICENSE	Create LICENSE
README.md	Update name
rp.main.R	initial upload of r co
struct.net.dev.R	Update struct.net.d

README.md

Sliding window analysis of bra

The code in this repository reproduces most analyses and figures (with the exception of certain supplementary analyses) conducted in the manuscript "**Adolescent tuning of association cortex in human structural brain networks**" by František Váša et al.

For details regarding the motivation behind analyses and the interpretation of results, see the manuscript.

zenodo Search Upload Communities Log in Sign up

April 11, 2017

Software Open Access

frantisekvasa/structural_network_development: Initial release

František Váša

Initial release of the code supporting the manuscript "Adolescent tuning of association cortex in human structural brain networks" by František Váša et al.

Preview

structural_network_development-v0.1.zip

frantisekvasa-structural_network_development-fcf901d

LICENSE	1.1 kB
README.md	422 Bytes
rp.main.R	681 Bytes
struct.net.dev.R	103.3 kB

Available in GitHub

Publication date:
April 11, 2017

DOI:
DOI 10.5281/zenodo.528674

Related identifiers:
Supplement to:
https://github.com/frantisekvasa/structural_network_development/tree/v0.1

License (for files):
Other (Open)

@kirstie_j

://dx.doi.org/10.6084/m9.figshare.5914444