

the n-of-many-ones: fueling community science with personal data

quantifying myself

back in 2011

23andMe is Celebrating DNA Day a
Little Early With a Sale!

April 11, 2011 By [Shwu](#) under [announcements](#)

back in 2011

23andMe is Celebrating DNA Day a Little Early With a Sale!

April 11, 2011 By [Shwu](#) under [announcements](#)



back in 2011

23andMe is Celebrating DNA Day a Little Early With a Sale!

April 11, 2011 By [Shwu](#) under [announcements](#)



Julia Reda ✓

@Senficon

Following

Der [@gedankenstuecke](#) darf sich freuen, ich habe soeben brav seine Speichelprobe an Big Brother verschickt. [#spackeria](#)
[#23andme](#)

Translated from German by  Microsoft

He [@gedankenstuecke](#) can rejoice, I have just sent His Saliva sample To Big Brother.
[#spackeria](#) [#23andme](#)

4:16 AM - 15 Apr 2011



since then...



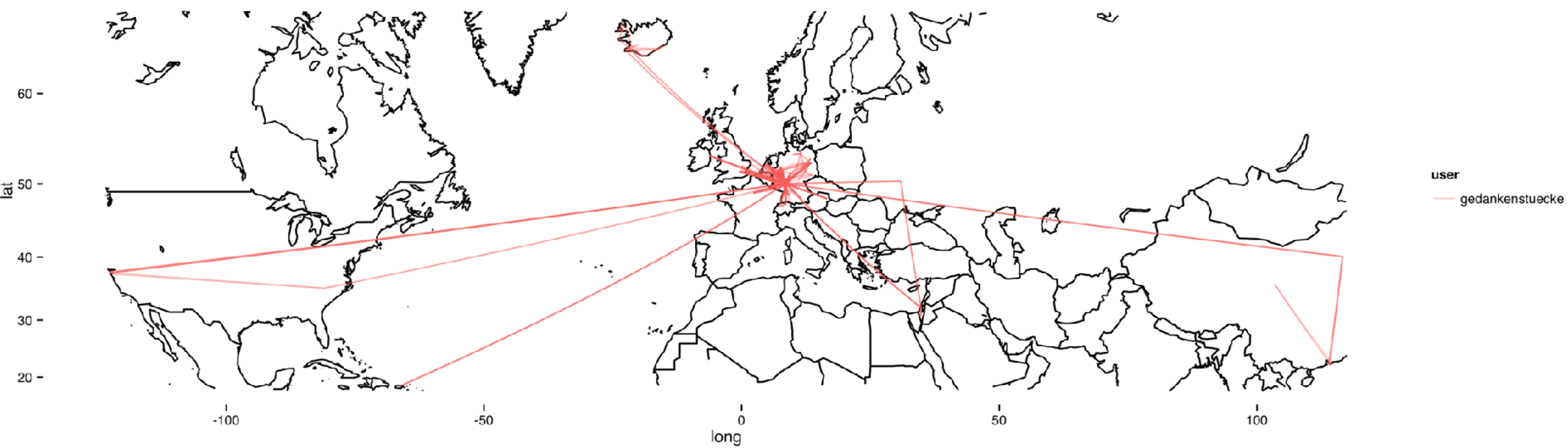
since then...



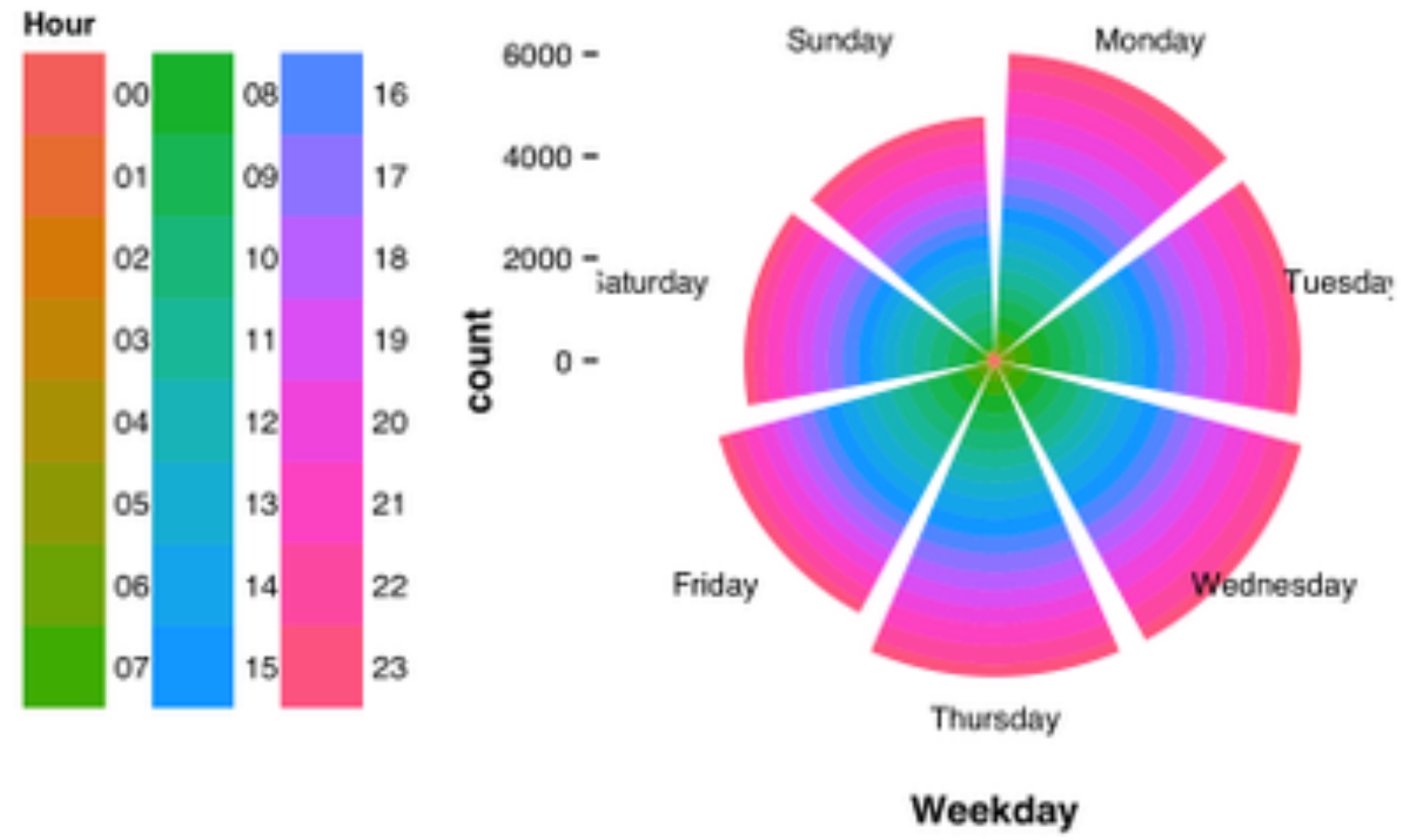
OURA



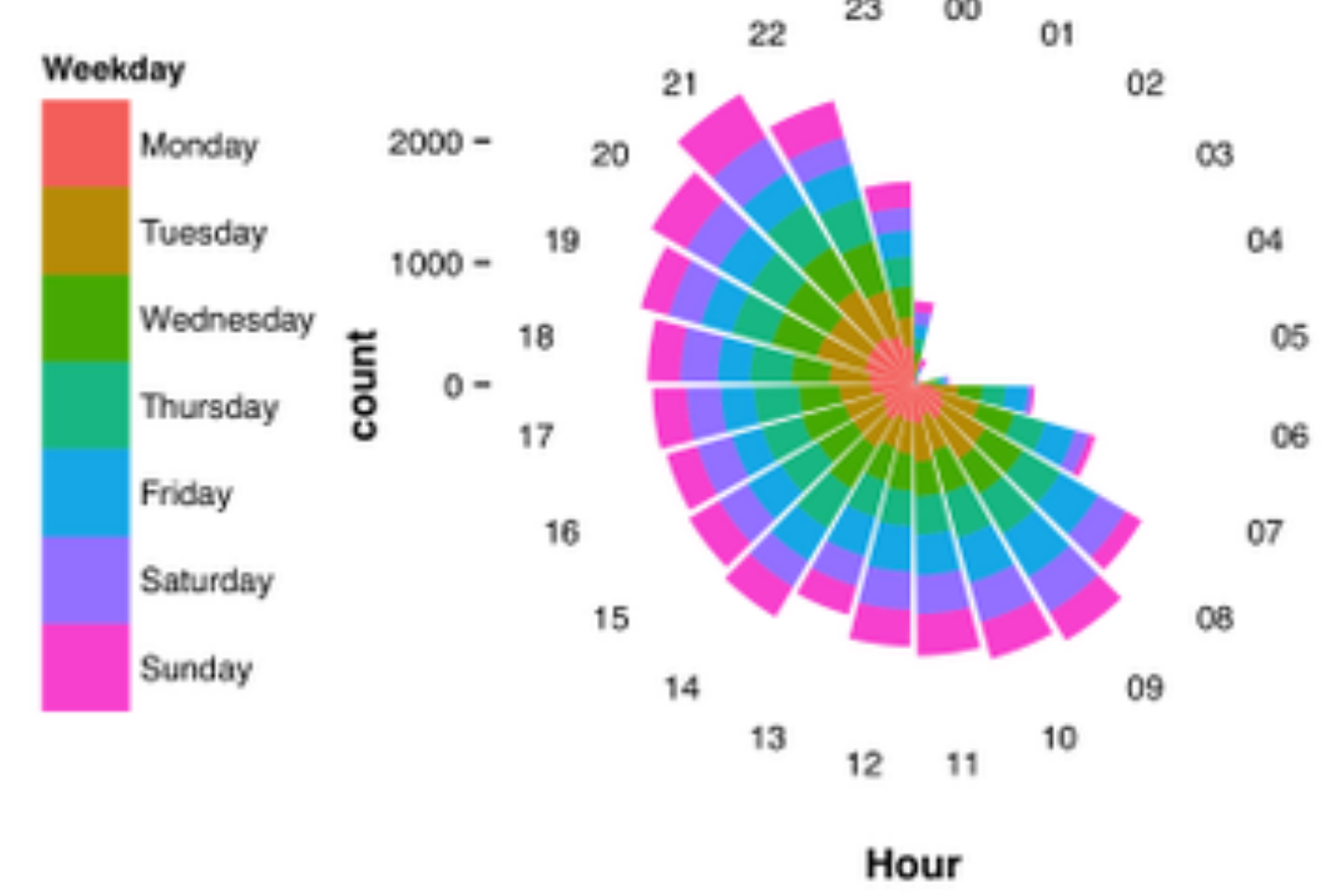
that's collecting a ton of data



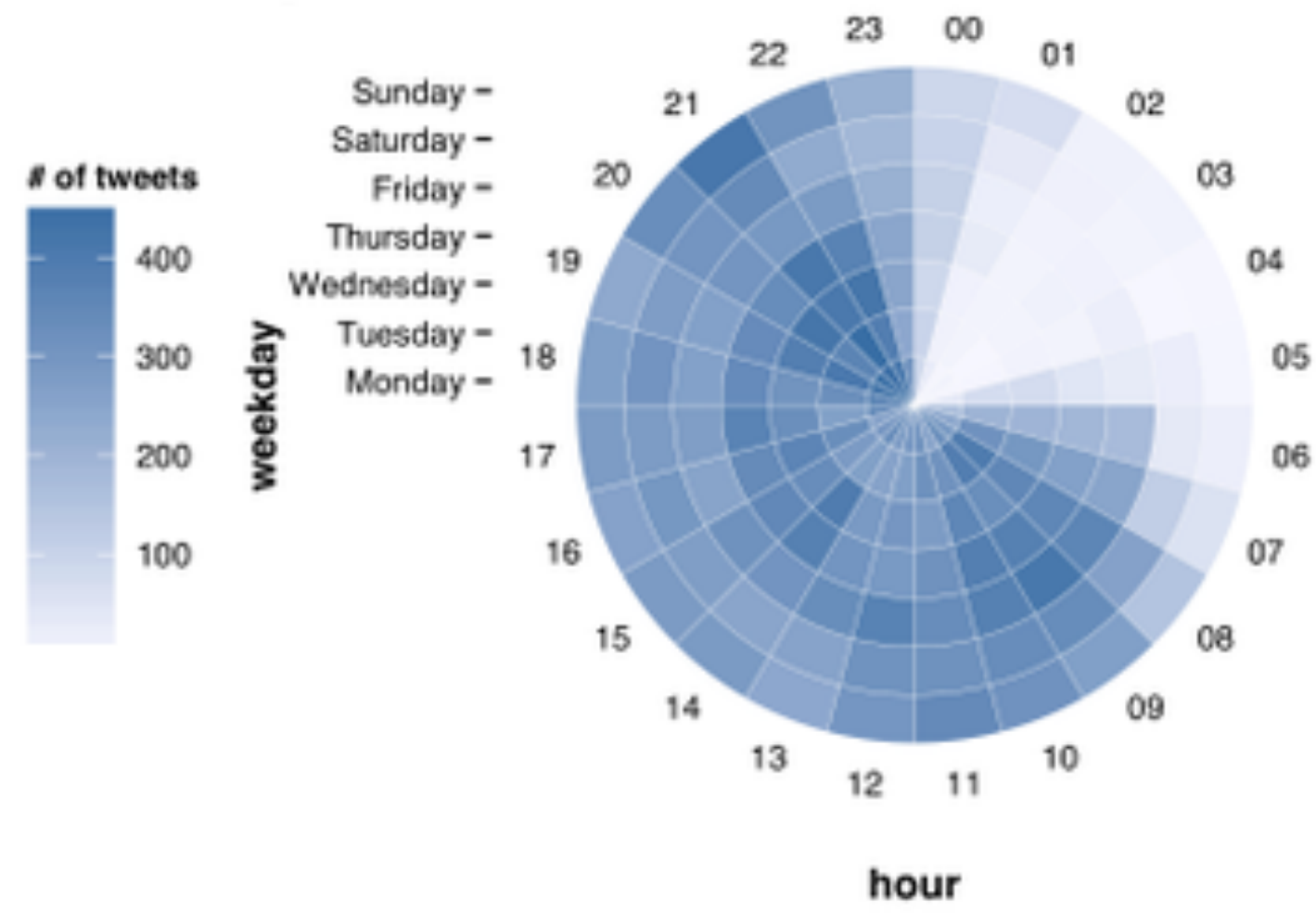
tweets/day



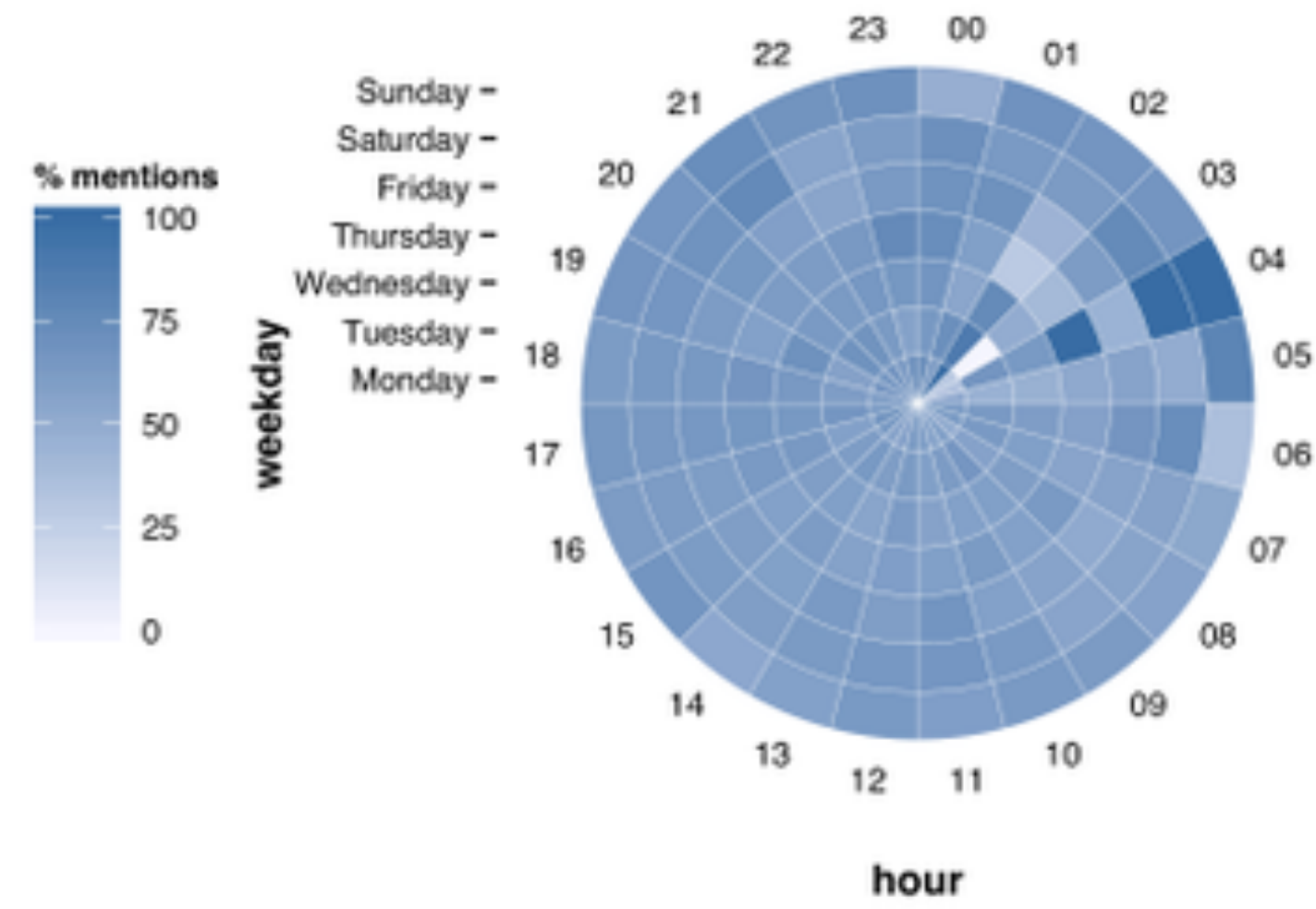
tweets/hour



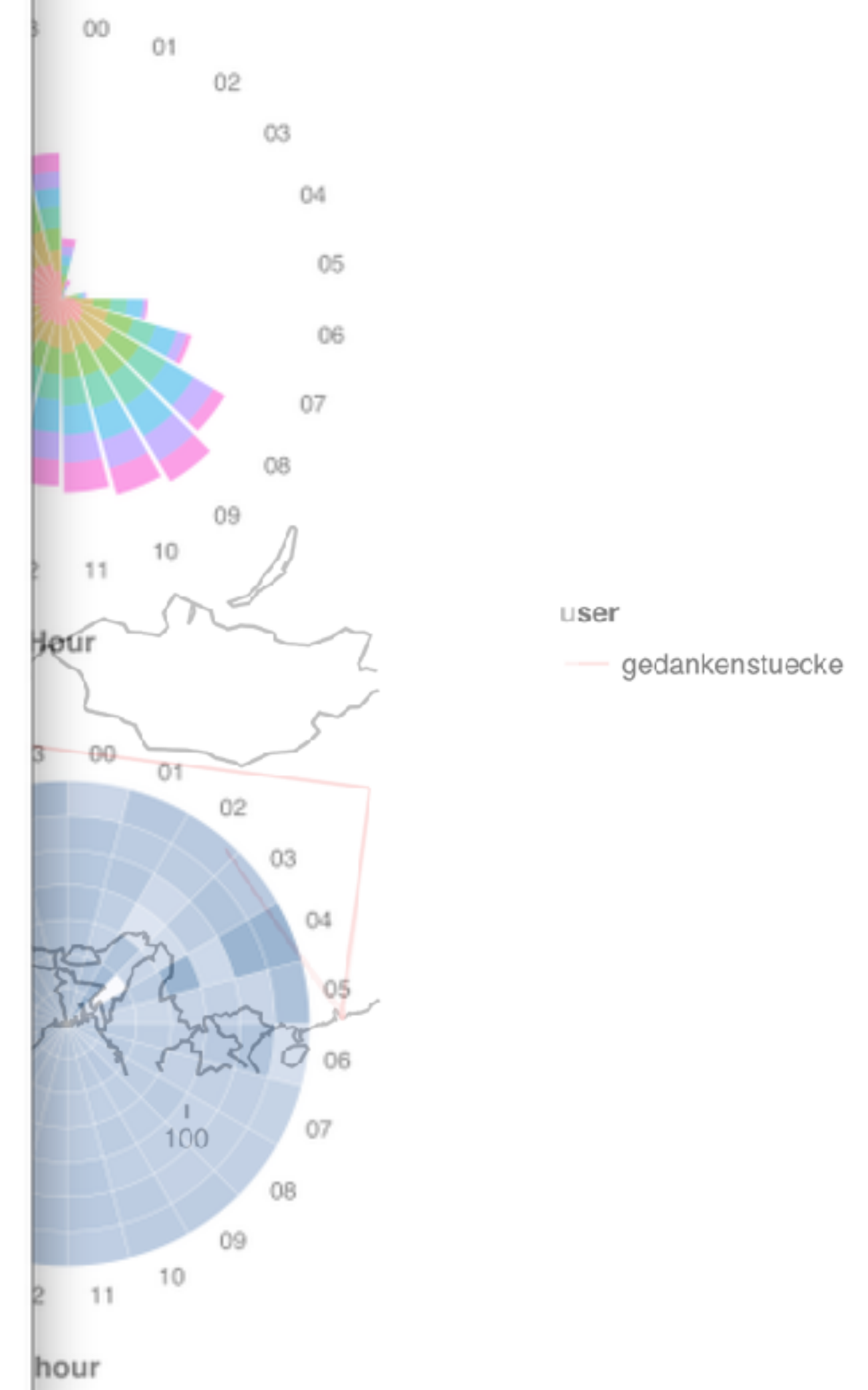
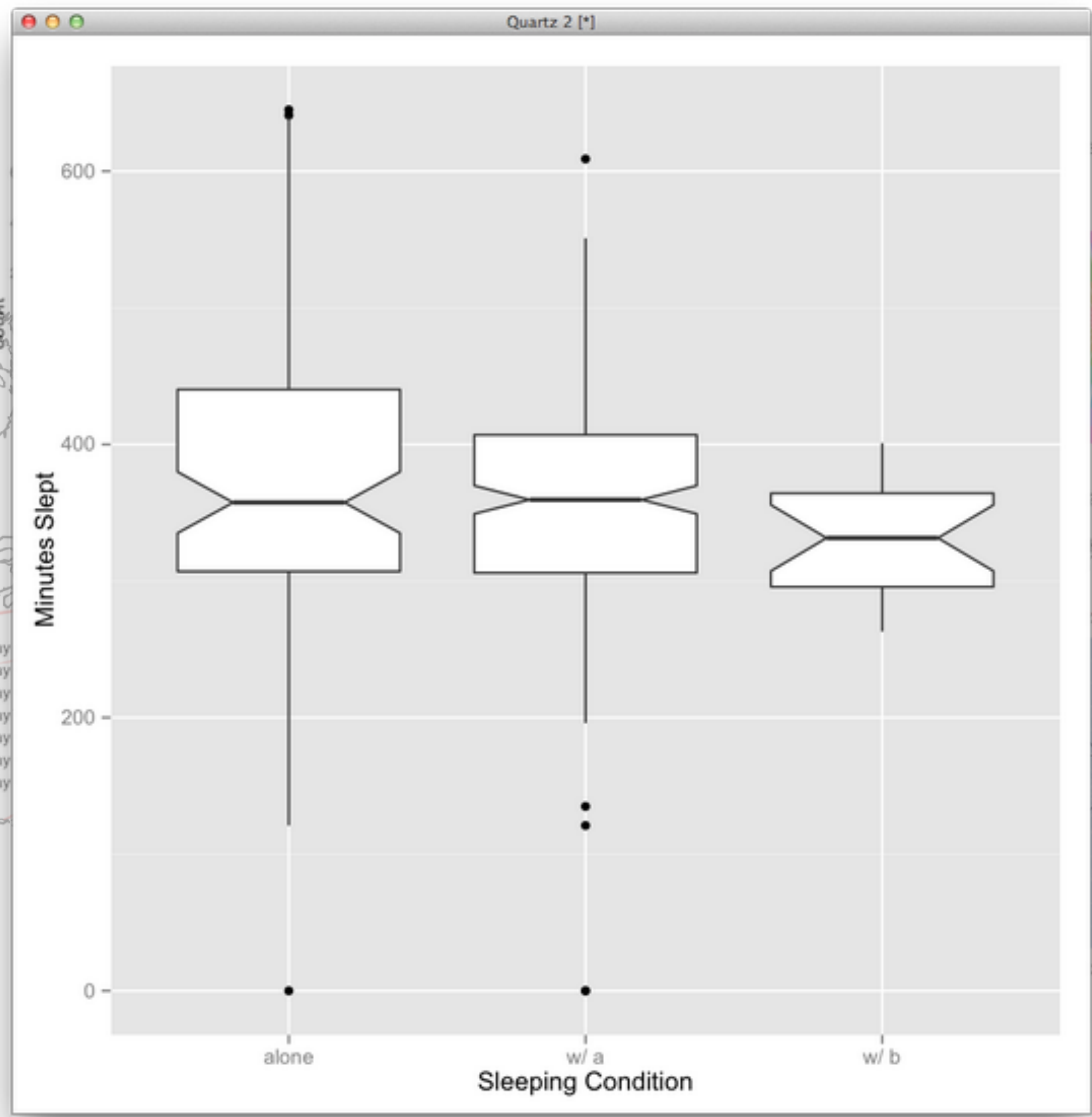
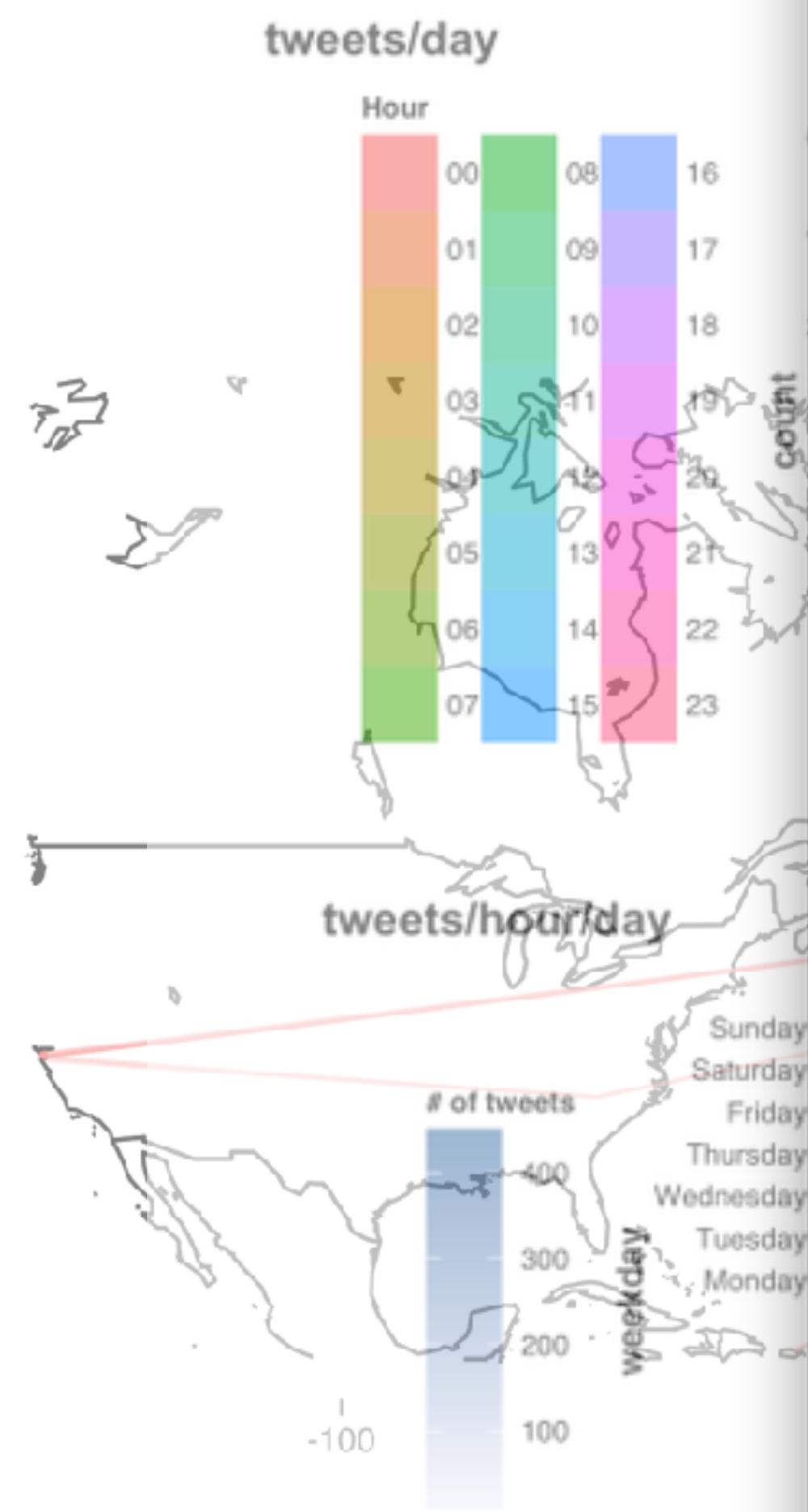
tweets/hour/day



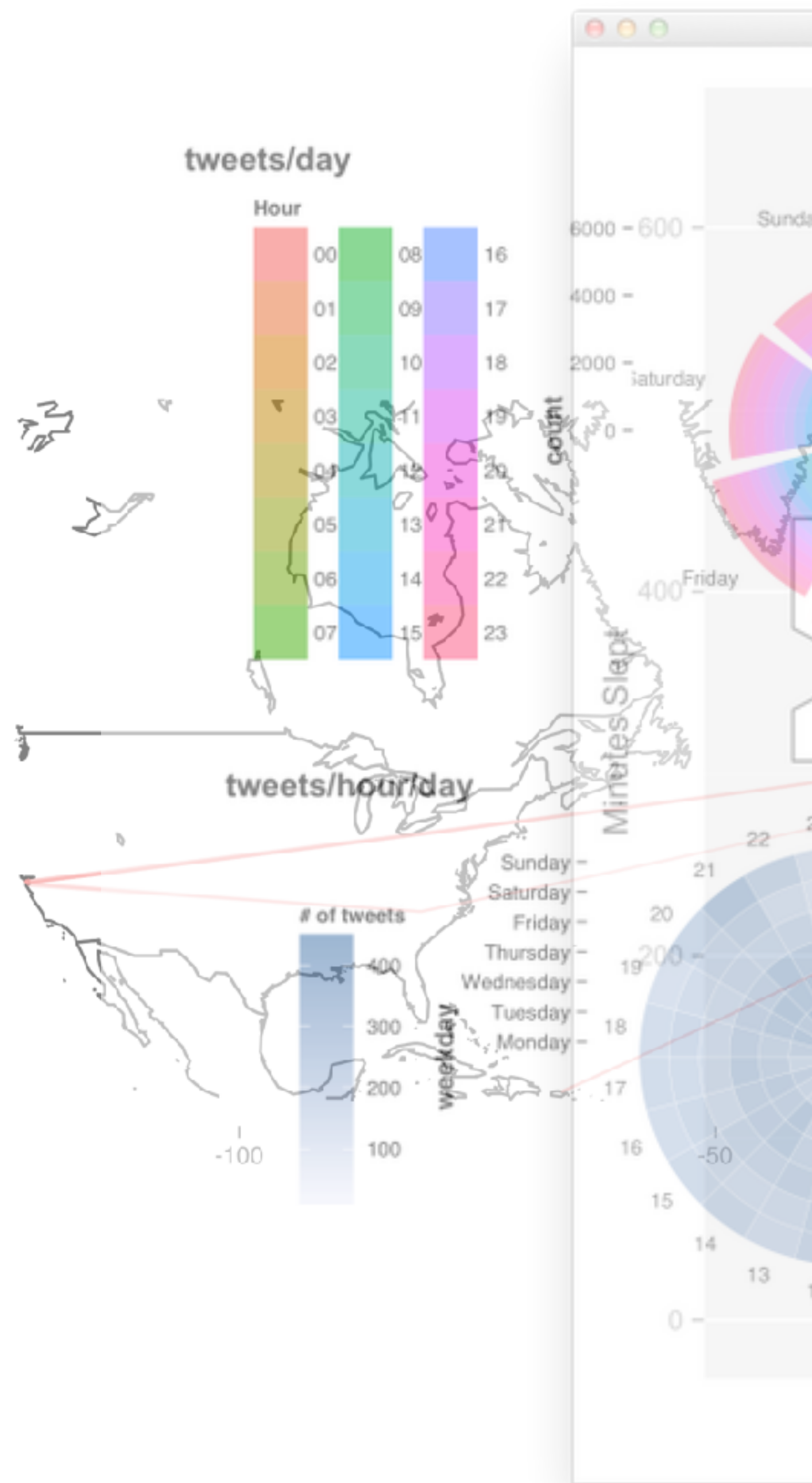
% mentions



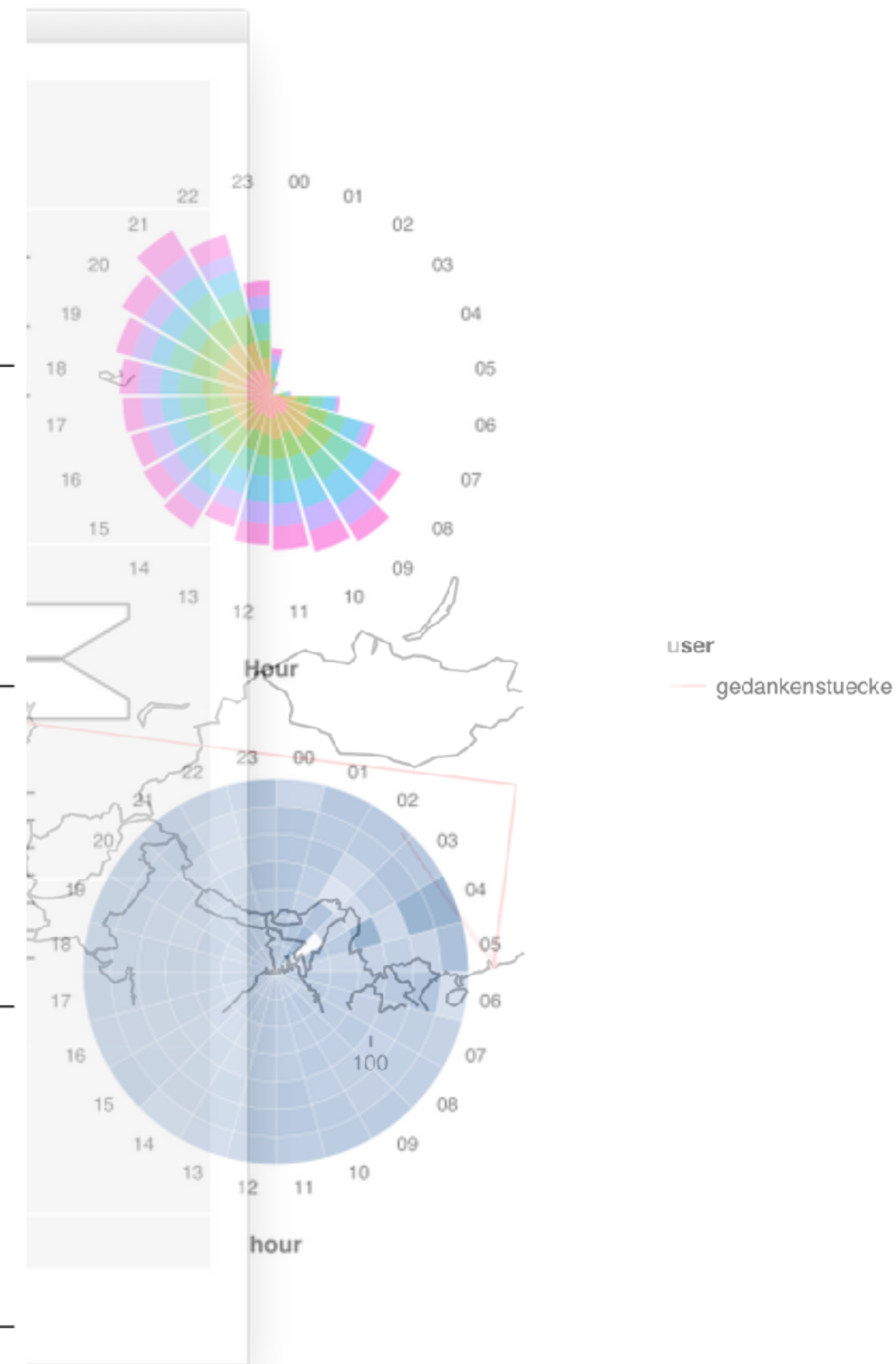
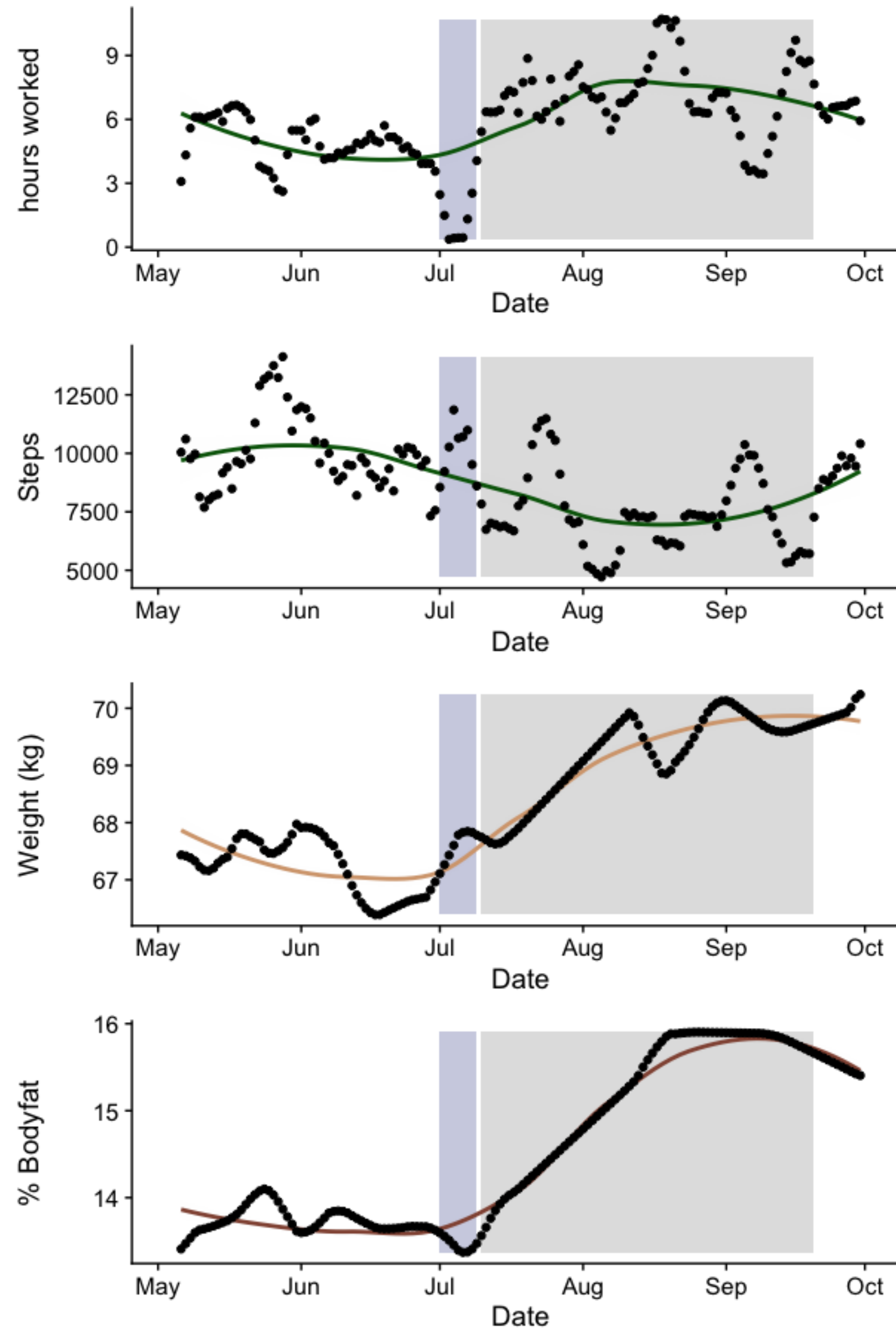
lat
60 -
50 -
40 -
30 -
20 -



lat
60 -
50 -
40 -
30 -
20 -



before, during (gray box) and after PhD thesis writing
(points are rolling average over 7 days)



that's fun, but...

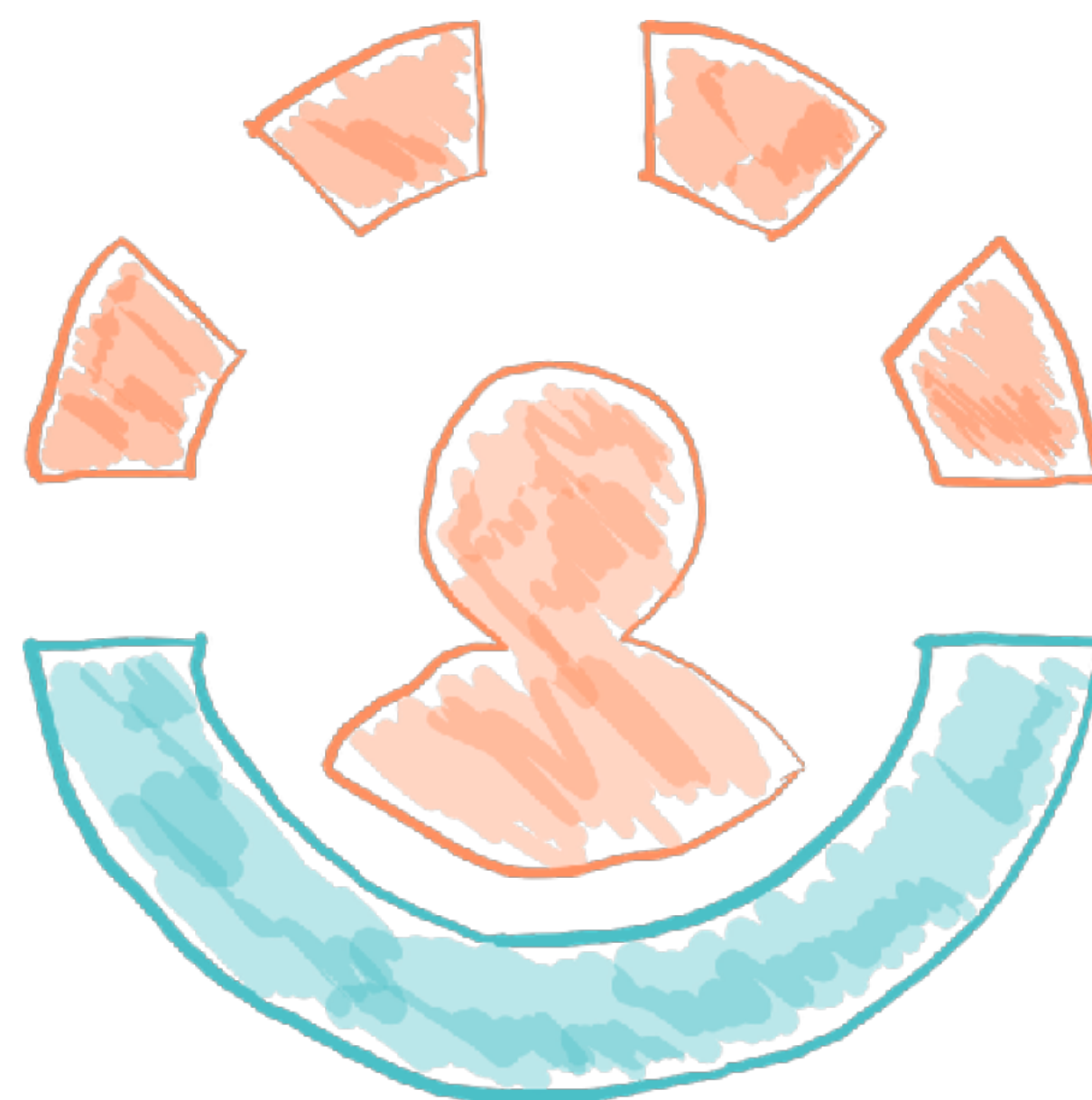
that's fun, but...

- gotta learn ***all*** the APIs of the different services to liberate your data
- inherently one-off projects (don't ask me to reproduce most of those figures)
- super limited utility for a larger community

what if...

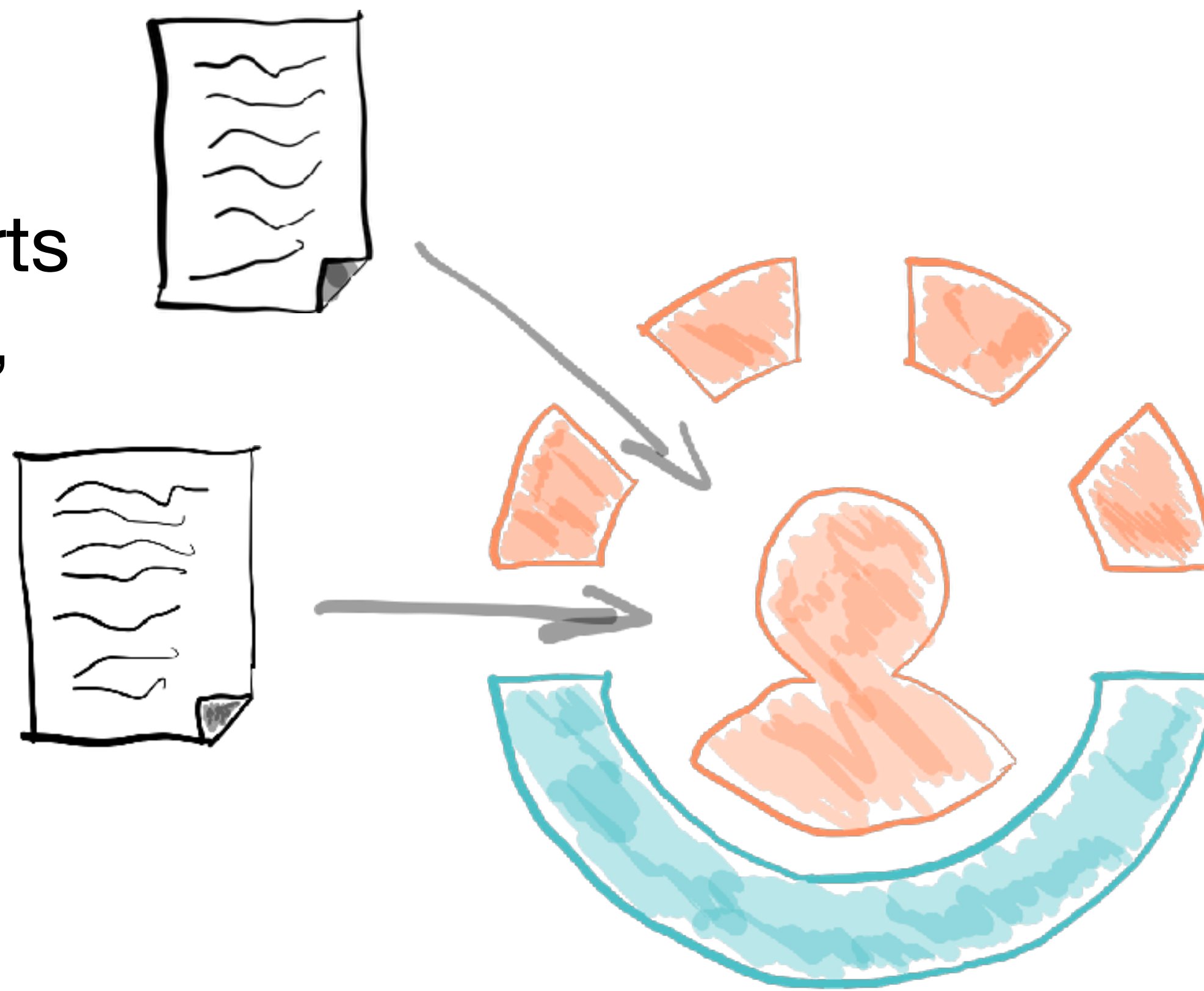
what if...

- we could make accessing & aggregating your data less painful?
- make n-of-1 data visualizations easily reproducible?
- even allow for sharing some of those data?



OPEN HUMANS

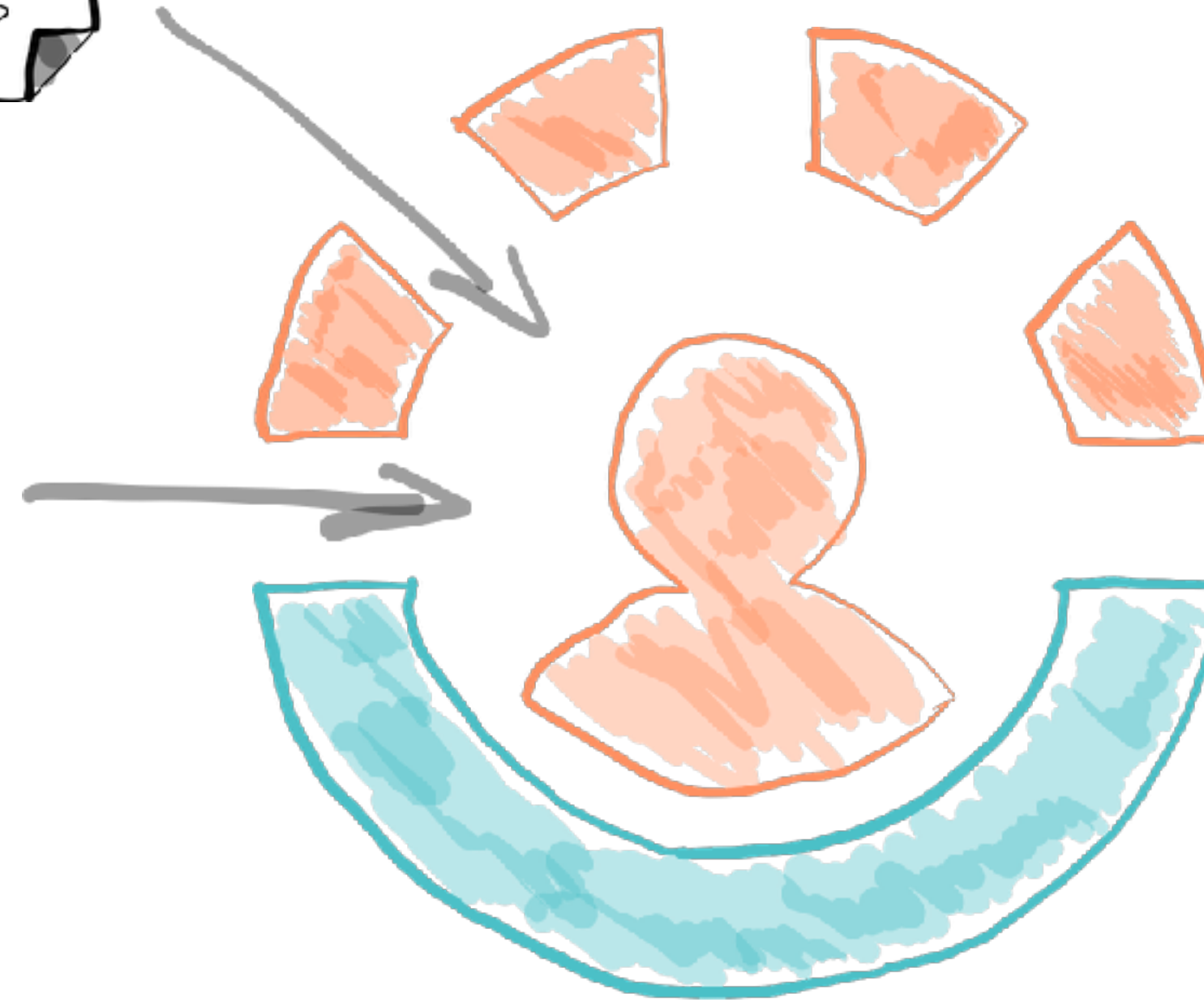
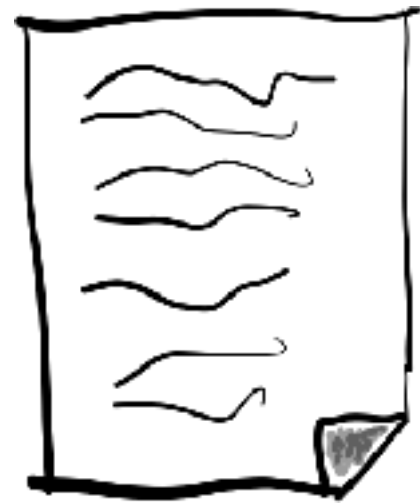
individual imports
personal data,



OPEN HUMANS



individual imports
personal data,



OPEN HUMANS

data sources



23andMe Upload

Open Humans
Open Humans Foundation
Connected by **1152** members

23andMe is a direct-to-consumer genetic testing company that tests about one million genetic locations.

[Learn more](#)



Twitter Archive Analyzer

Bastian Greshake Tzovaras
Connected by **519** members

The TwArxiv is a Twitter Archive Analyzer. Upload your Twitter archive and get new insights.

[Learn more](#)

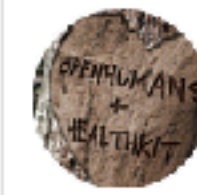


AncestryDNA Upload

Open Humans
Open Humans Foundation
Connected by **411** members

Ancestry.com's AncestryDNA is a direct-to-consumer genetic testing product that tests about 700,000 genetic locations.

[Learn more](#)



Open Humans Healthkit Integration

James Turner
self
Connected by **113** members

Install this third-party app on your iPhone or iPad to upload HealthKit data to Open Humans.

[Learn more](#)



Data Selfies

Open Humans
Open Humans Foundation
Connected by **174** members

Upload Data Selfies

[Learn more](#)



Nokia Health (Withings) Connection

Open Humans
Open Humans Foundation
Connected by **101** members

Add your Nokia Health (Withings) data to Open Humans

[Learn more](#)



Gencove

Joseph K. Pickrell
Gencove, Inc.

Connected by **295** members

Your genome app - get your ancestry, microbiome, and more! Contribute your data to Open Humans.

[Learn more](#)



Fitbit Connection

Open Humans
Open Humans Foundation
Connected by **398** members

Connect your Fitbit account to add data from your Fitbit activity trackers and other Fitbit devices.

[Learn more](#)



openSNP

Bastian Greshake Tzovaras
openSNP

Connected by **290** members

openSNP allows you to put your genetic and phenotypic data into the public domain. Connect your openSNP account to Open Humans.

[Learn more](#)



uBiome Upload

Open Humans
Open Humans Foundation
Connected by **69** members

uBiome is a company based in San Francisco that performs microbiome sequencing.

[Learn more](#)



FamilyTreeDNA integration

Bastian Greshake Tzovaras
Connected by **74** members

Upload your FamilyTreeDNA data to Open Humans

[Learn more](#)



Imputer

Kevin Arvai
Connected by **55** members

Imputer will perform imputation on user's genotype data to estimate genotypes at missing loci.

[Learn more](#)



Runkeeper connection

Open Humans
Open Humans Foundation
Connected by **210** members

RunKeeper is a free smartphone app for GPS fitness-tracking. You can use it to record GPS timopoint data for runs, walks, bicycling, and oth

[Learn more](#)



Nightscout Data Transfer

James Wedding, P.E.
The Nightscout Foundation
Connected by **280** members

A tool to easily enable upload of data from individual Nightscout databases to the Open Humans platform

[Learn more](#)



Genome/Exome Upload

Open Humans
Open Humans Foundation
Connected by **209** members

Do you have genome or exome data? You can upload genome, exome, and genotyping data in VCF format.



Google Location History Upload

Open Humans
Open Humans Foundation
Connected by **46** members

Upload your Google Location History archive

[Learn more](#)



Spotify integration

Open Humans
Open Humans Foundation
Connected by **47** members

Collect the songs you listen to on Spotify

[Learn more](#)



Oura Connect

Open Humans
Open Humans Foundation
Connected by **38** members

Connect your Oura to Open Humans

[Learn more](#)



RescueTime connection

Bastian Greshake Tzovaras
Connected by **31** members

RescueTime is a web service tracks your productivity and computer usage over time.

[Learn more](#)



Overland connection

Open Humans
Open Humans Foundation
Connected by **15** members

Import GPS data from Overland

[Learn more](#)



Google search history analyzer

Athina Tzovara
Connected by **12** members

Google search history analyzer

data sources

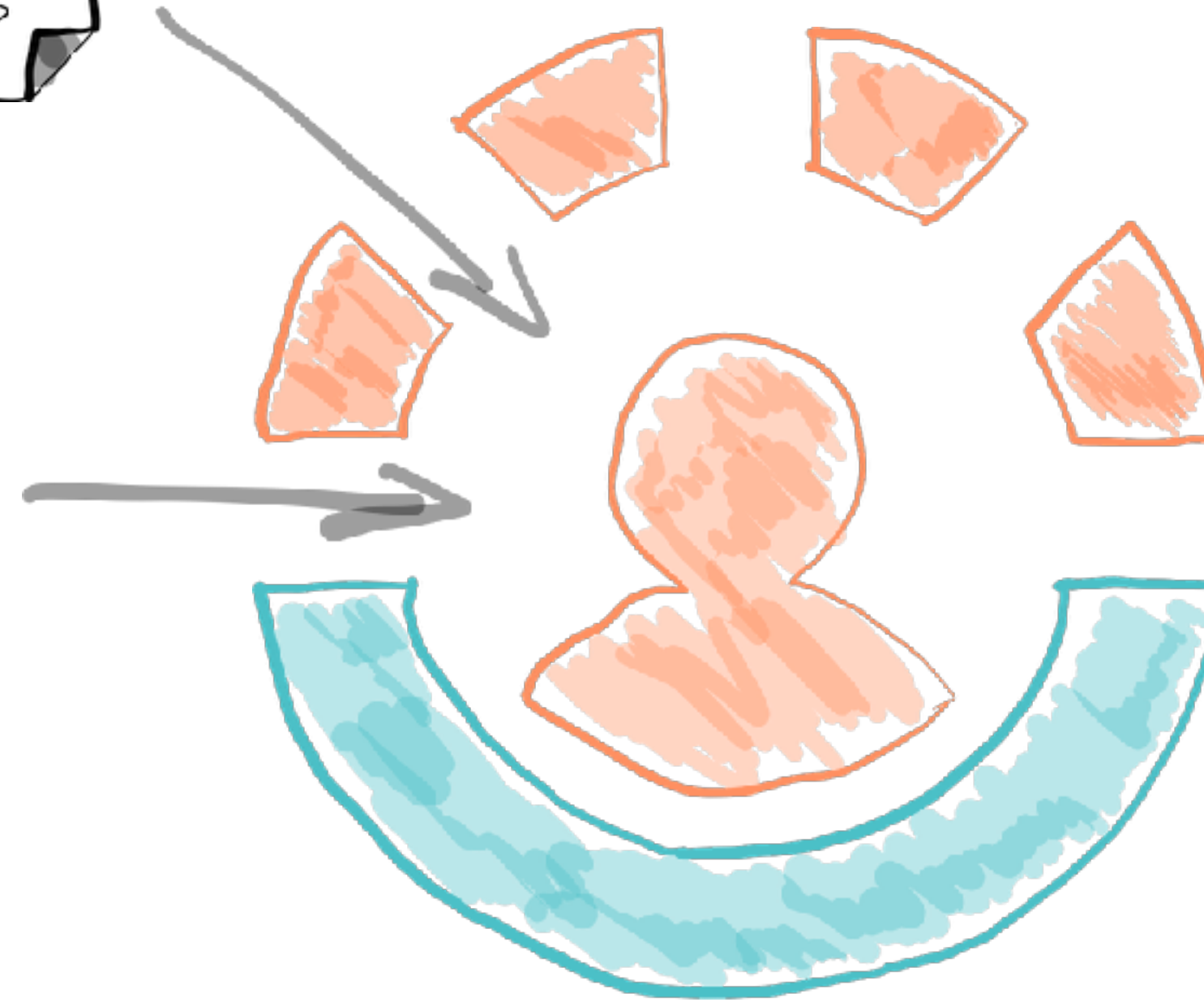
- upload of files (e.g. Google Takeouts, Twitter Archives, genetic data)
- continuous API integrations (e.g. Fitbit, continuous glucose monitors, device usage data, Spotify, ...)
- data storage standardized by data source



individual imports
personal data,

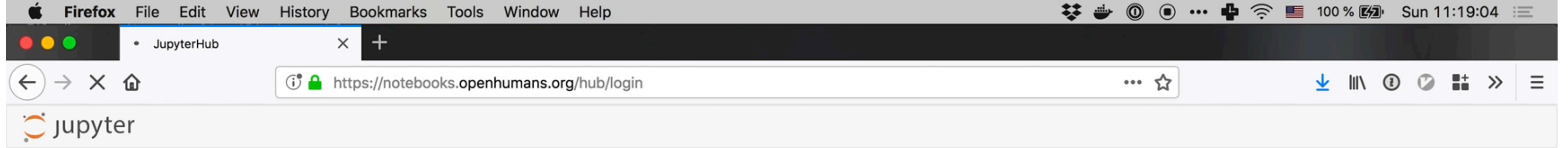


stores data
privately & can
explore it,



OPEN HUMANS

Sign in with OpenHumans



[Sign in with OpenHumans](#)

access all of our personal data through one API!

Using RescueTime to analyze your productivity and computer usage

RescueTime is a service that monitors how you're using your computer. Split into 5 minute intervals you can't only see what you did with your computer, but it also classifies these activities automatically in productive/unproductive time.

Let's use this data to see

- How our productivity changes over time
- How we're spending most of our computer time

and if we're also counting steps with Fitbit and/or Apple HealthKit to see

- How our computer usage influences our activity

To get started let's find out which data sources we have on Open Humans and identify & load the data from RescueTime

```
In [1]: import json
import os
import requests
import pandas as pd
from datetime import datetime
import numpy as np
import seaborn as sns
response = requests.get("https://www.openhumans.org/api/direct-sharing/project/exchange-member/?access_token={}".format(
fileurl = ''
for entry in response.json()['data']:
    if entry['source'] == "direct-sharing-149":
        fileurl = entry['download_url']
        break

rescuetime_data = requests.get(fileurl).json()
```

We can now see how the data from RescueTime looks like:

Using RescueTime to analyze your productivity and computer usage

RescueTime is a service that monitors how you're using your computer. Split into 5 minute intervals you can't only see what you did with your computer, but it also classifies these activities automatically in productive/unproductive time.

Let's use this data to see

- How our productivity changes over time
- How we're spending most of our computer time

and if we're also counting steps with Fitbit and/or Apple HealthKit to see

- How our computer usage influences our activity

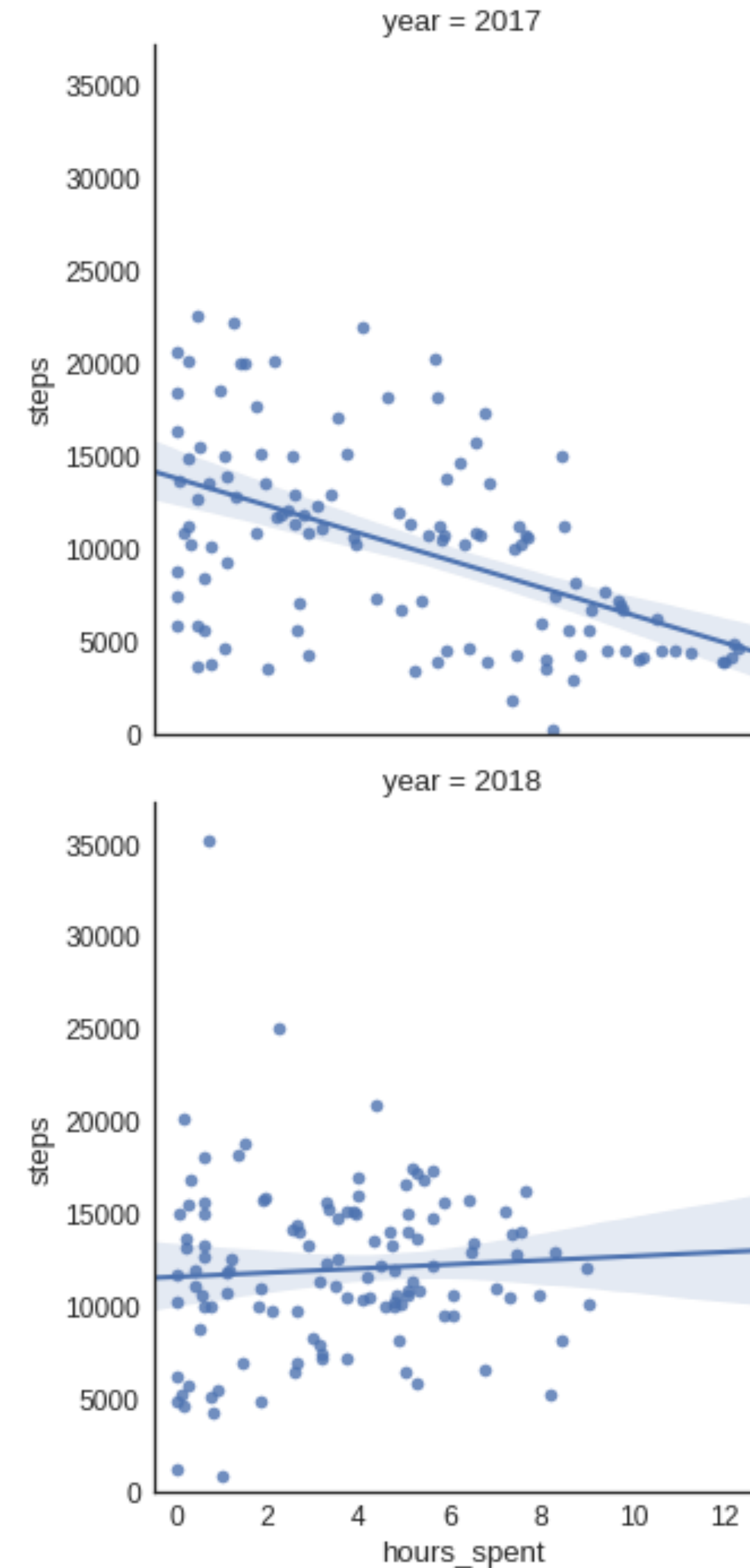
To get started let's find out which data sources we have on Open Humans and identify & load the data from RescueTime

```
In [1]: import json
import os
import requests
import pandas as pd
from datetime import datetime
import numpy as np
import seaborn as sns
response = requests.get("https://www.openhumans.org/api/direct-sharing/project/exchange-member/?access_token={}".format(
fileurl = ''
for entry in response.json()['data']:
    if entry['source'] == "direct-sharing-149":
        fileurl = entry['download_url']
        break

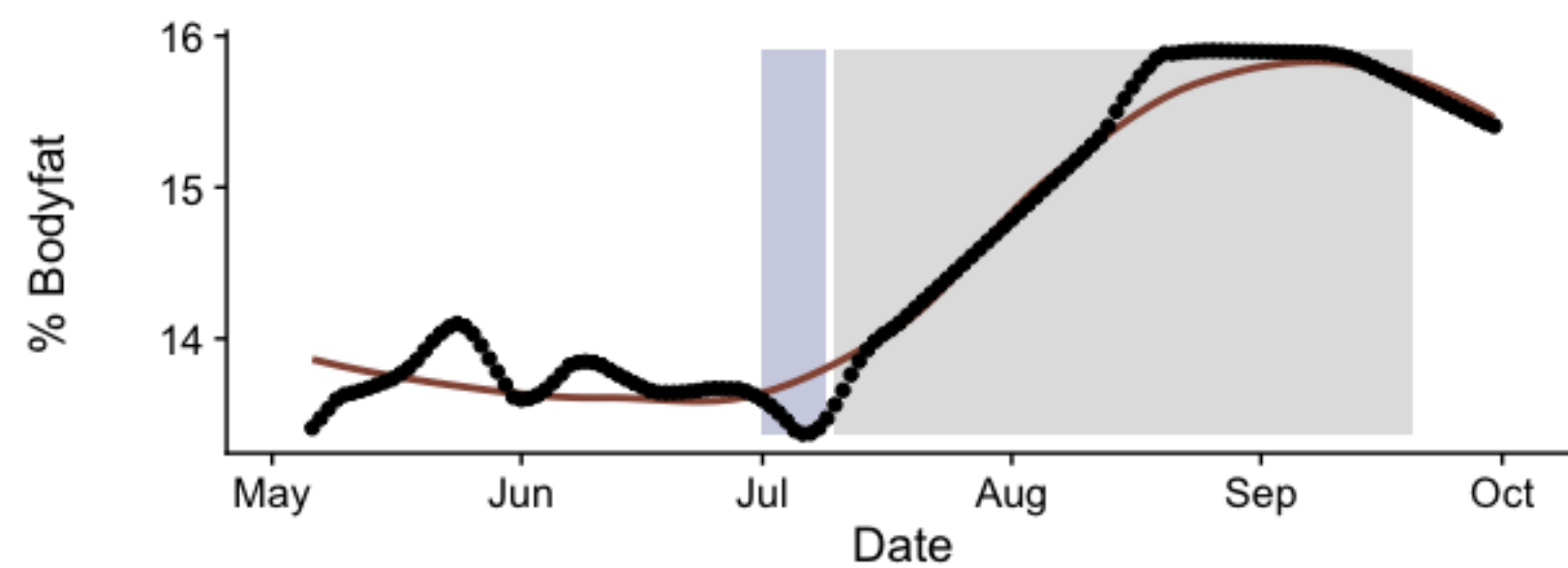
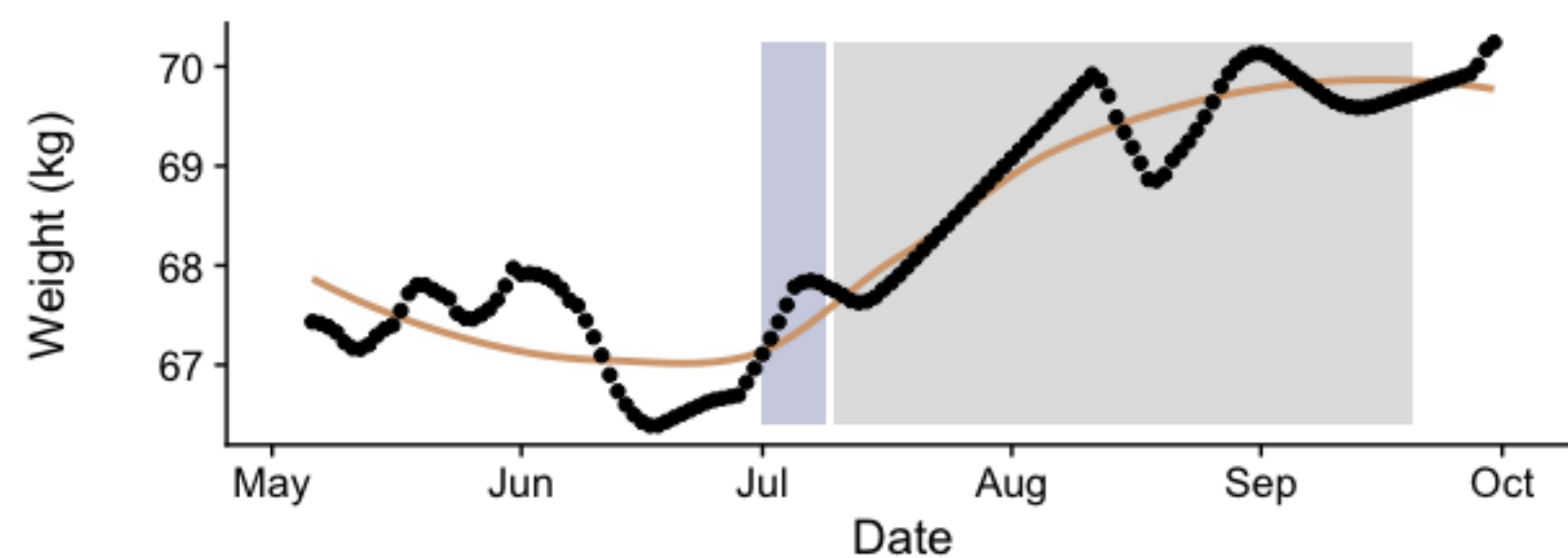
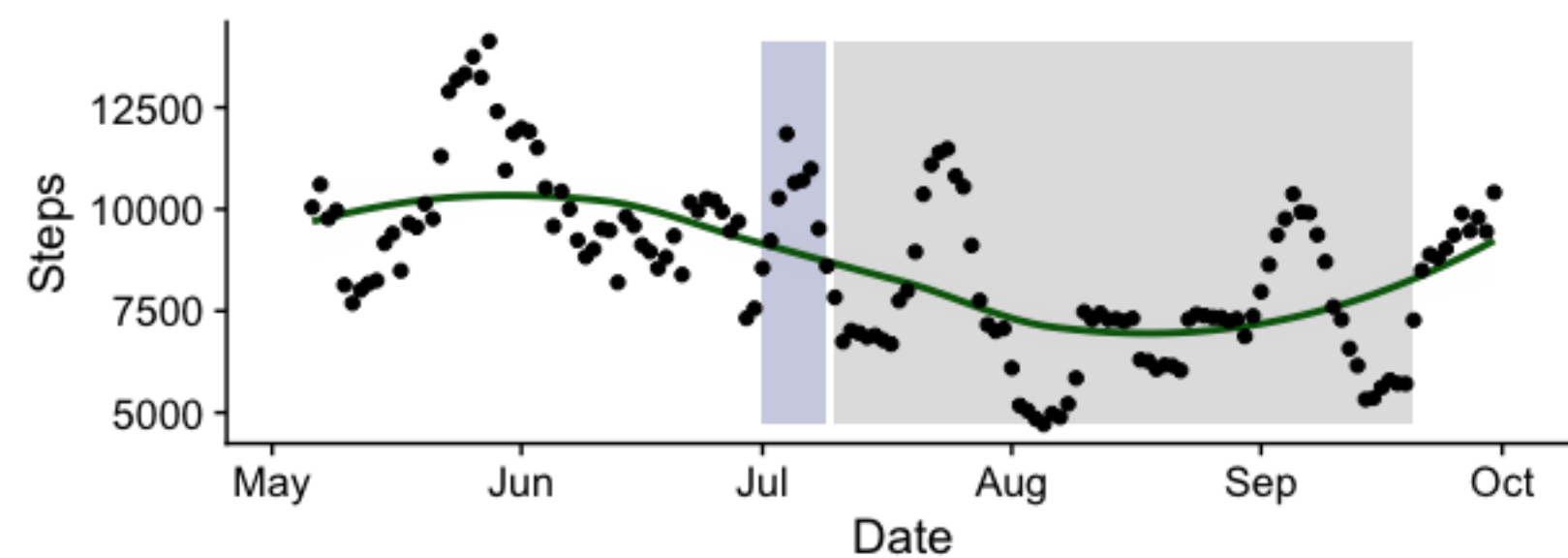
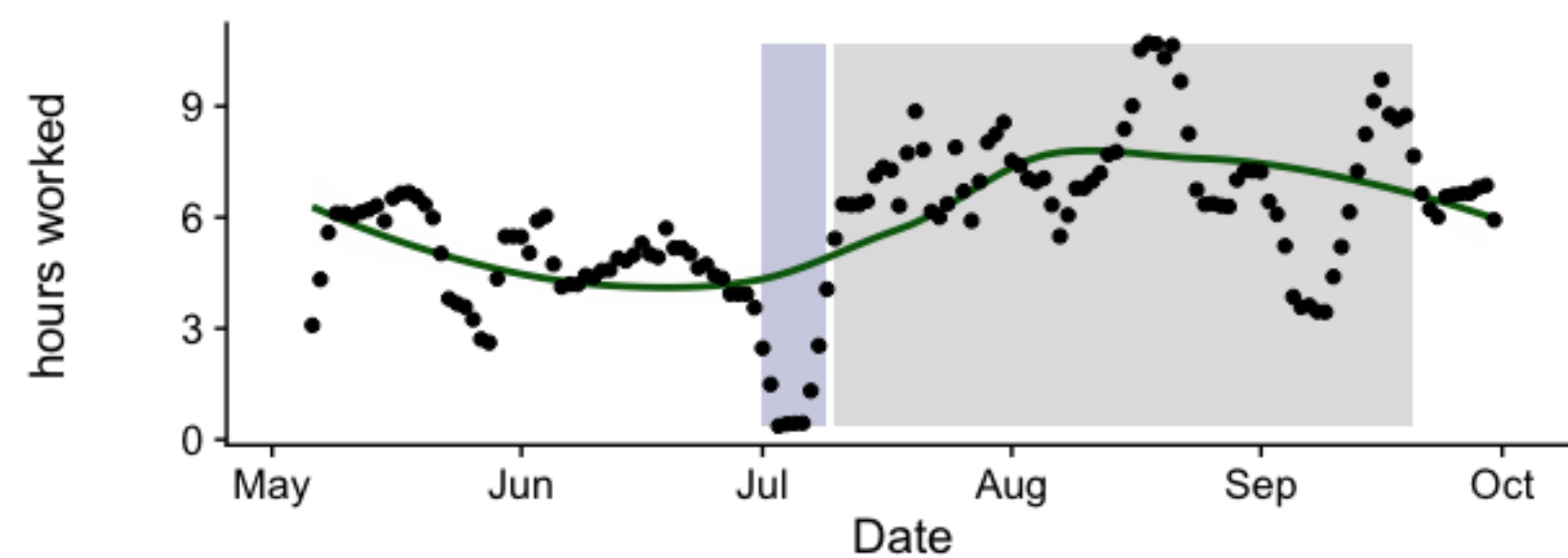
rescuetime_data = requests.get(fileurl).json()
```

We can now see how the data from RescueTime looks like:

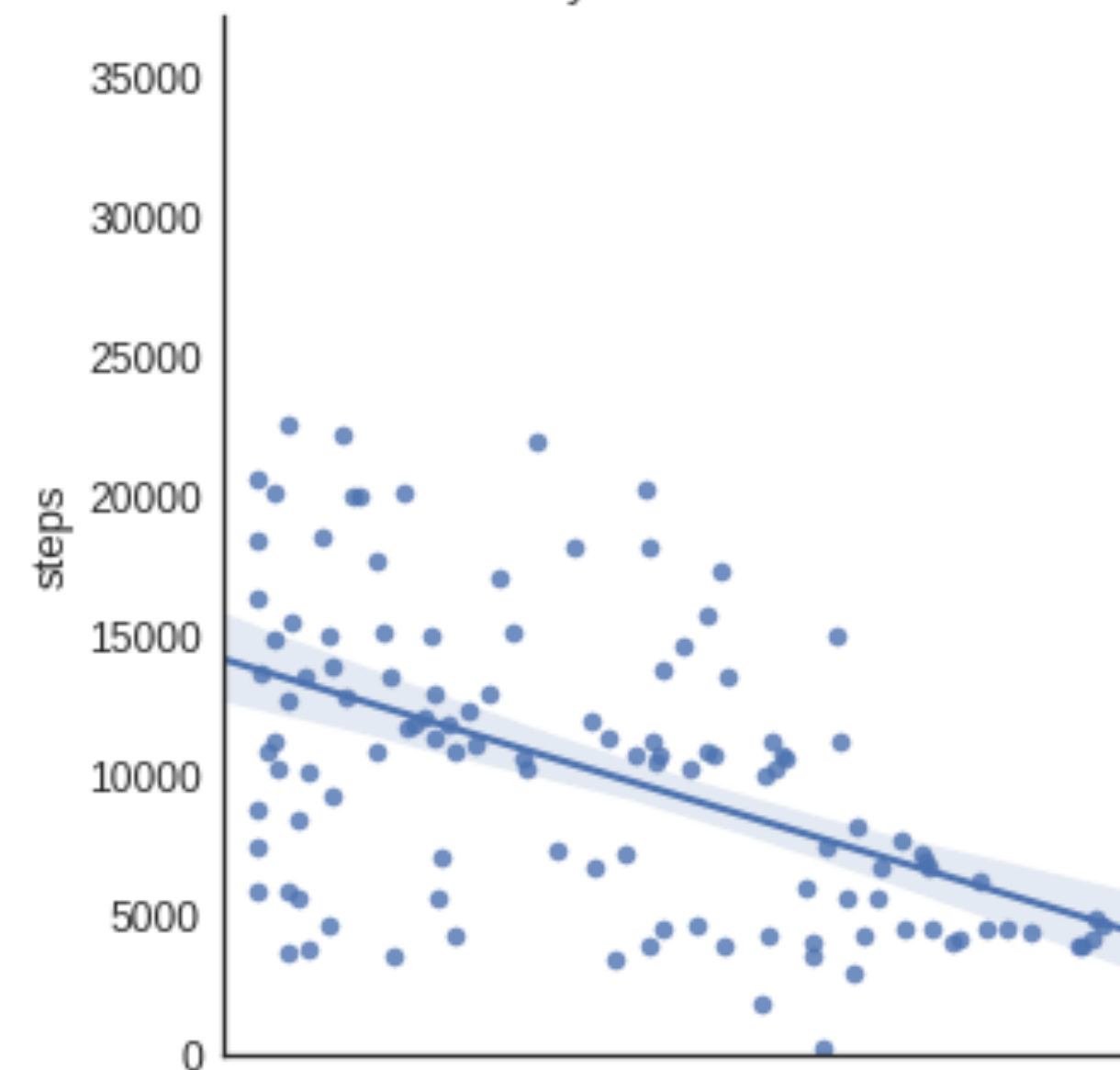
**we can easily correlate
across different data
sets!**



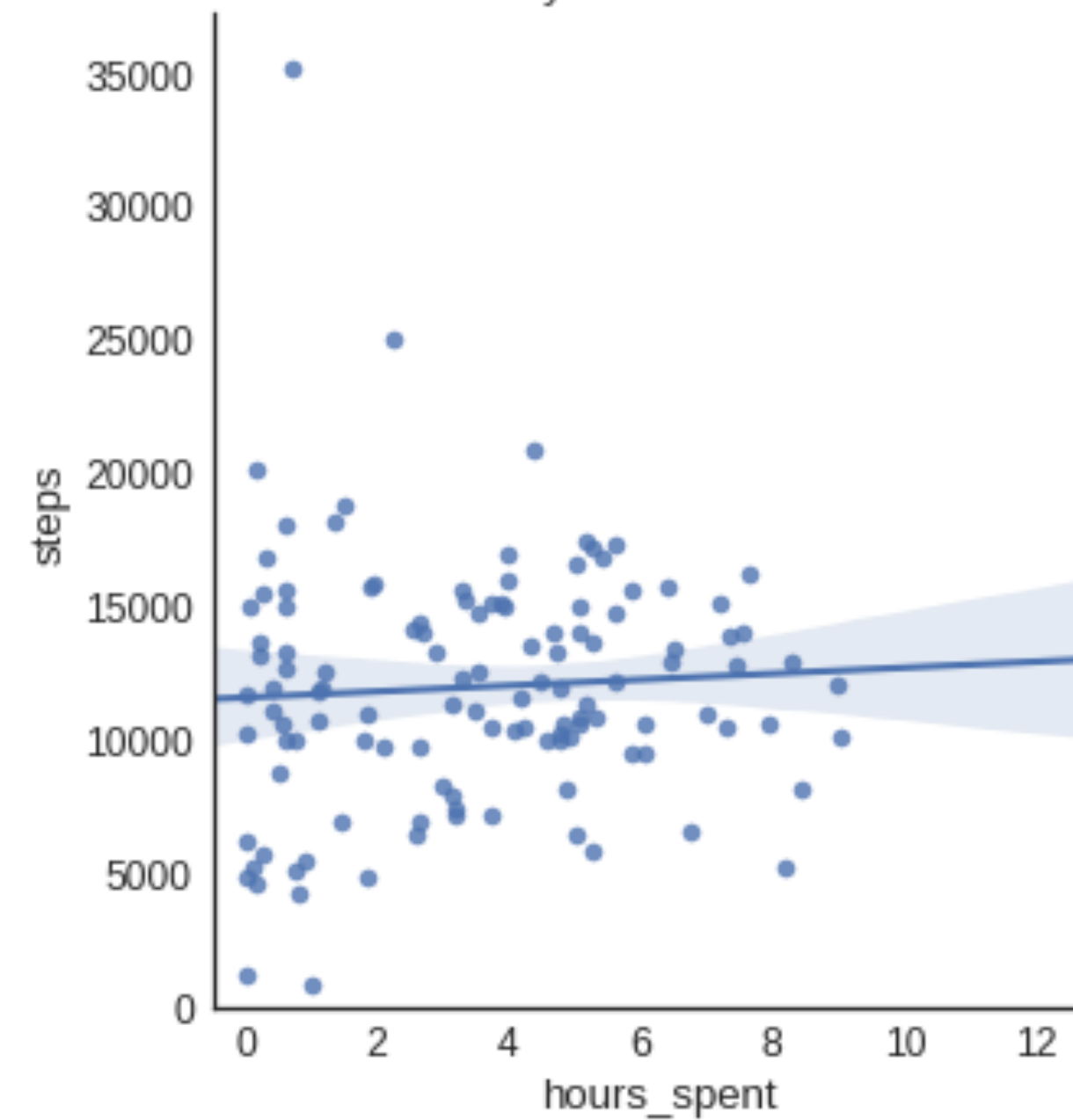
before, during (gray box) and after PhD thesis writing
(points are rolling average over 7 days)



year = 2017



year = 2018



rescuetime-vs-step-counts-2019-05-01-22

+

← → ↺ 🏠

https://notebooks.openhumans.org/user/gedankenstuecke/notebooks/rescuetime-vs-step-counts-2019-05-01-22

⋮ ☆

📄 ⓘ ⌂ ⚙️ ⏪ ⏩ ≡

jupyter rescuetime-vs-step-counts-2019-05-01-22:14:17 (autosaved)

Logout Control Panel

File Edit View Insert Cell Kernel Widgets Help

Not Trusted | Python 3

New Notebook

Open...

Make a Copy...

Rename...

Save and Checkpoint

Revert to Checkpoint

Print Preview

Download as

Share Notebook

Trust Notebook

Close and Halt

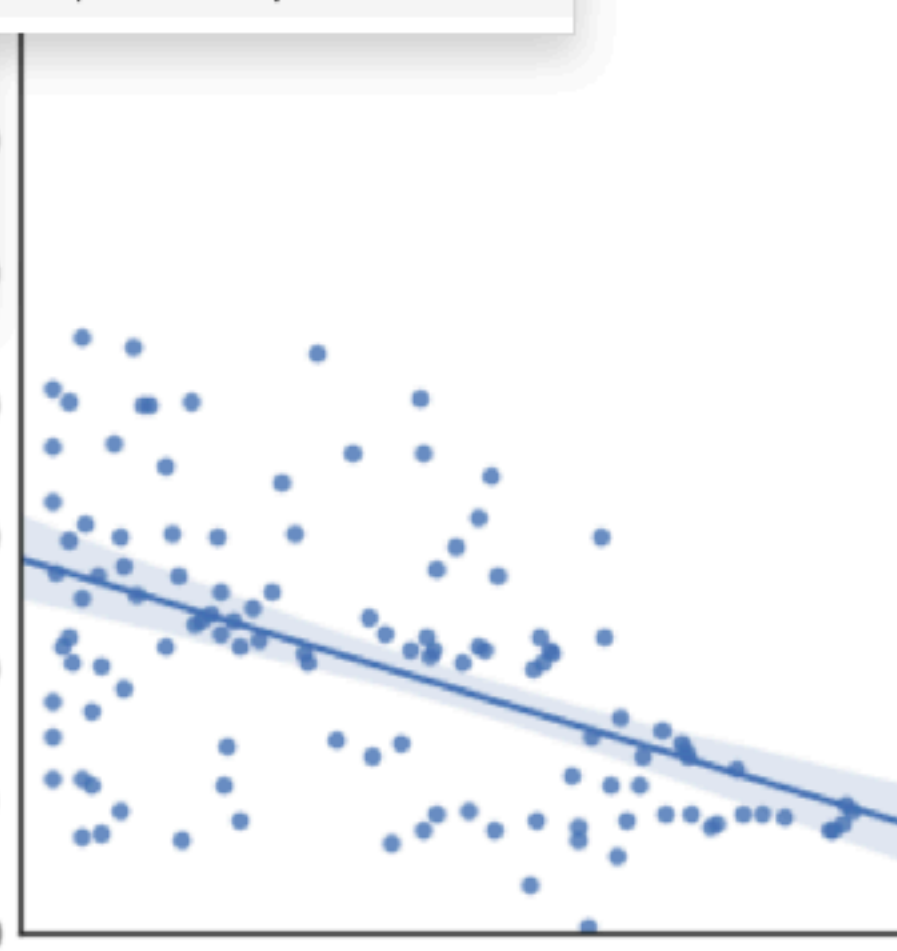
⬆ ⬇ ⏪ Run ⬆ ⬇ ⏩ Code

it done we can plot our data:


```
plot(data=joined_data,
      y='steps',
      x='hours_spent',
      row='year')
t.ylim(-0.5, None)
t.xlim(-0.5, None)

13.042686893113581)
```

year = 2017



year = 2018



https://notebooks.openhumans.org/user/gedankenstuecke/notebooks/rescuetime-vs-step-counts-2019-05-01-22:14:17.ipynb#

✕ 🔍 day

⬆ ⬇

Highlight All Match Case Whole Words 4 of 19 matches

rescuetime-vs-step-counts-2019-05-01-22

+

← → ↺ 🏠

https://notebooks.openhumans.org/user/gedankenstuecke/notebooks/rescuetime-vs-step-counts-2019-05-01-22

⋮ ☆

📄 ⓘ ⌂ ⚙️ ⏪ ⏩ ≡

jupyter rescuetime-vs-step-counts-2019-05-01-22:14:17 (autosaved)

Logout Control Panel

File Edit View Insert Cell Kernel Widgets Help

Not Trusted | Python 3

New Notebook

Open...

Make a Copy...

Rename...

Save and Checkpoint

Revert to Checkpoint

Print Preview

Download as

Share Notebook

Trust Notebook

Close and Halt

↑ ↓ ▶ Run ■ ↺ ▶▶ Code

it done we can plot our data:

```
plot(data=joined_data,
      y='steps',
      x='hours_spent',
      row='year')
t.ylim(-0.5, None)
t.xlim(-0.5, None)
```

13.042686893113581)

year = 2017

Upload to Open Humans

steps

20000

15000

10000

5000

0

year = 2018

35000

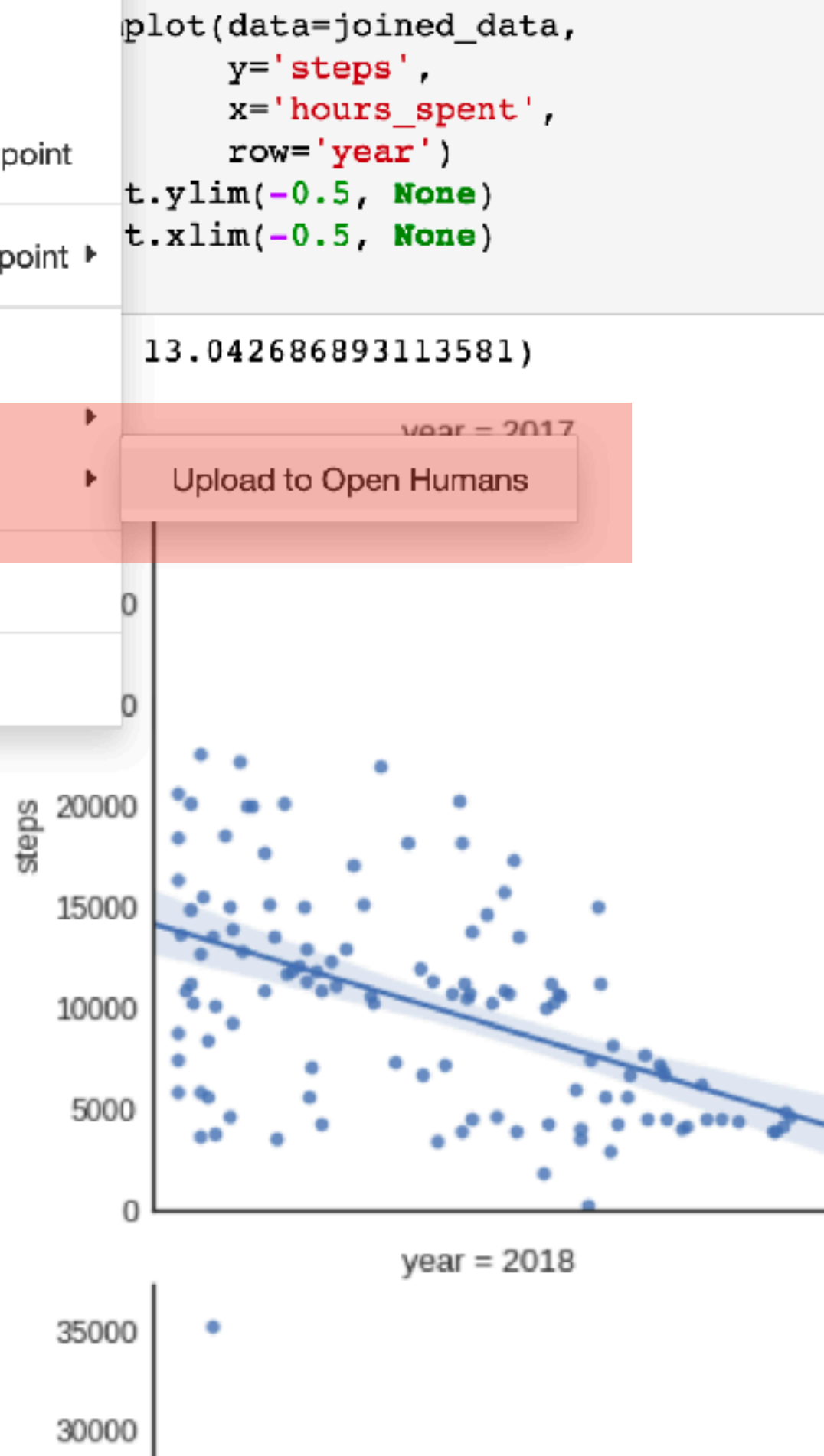
30000

https://notebooks.openhumans.org/user/gedankenstuecke/notebooks/rescuetime-vs-step-counts-2019-05-01-22:14:17.ipynb#

✕ 🔍 day

⬆ ⬇

Highlight All Match Case Whole Words 4 of 19 matches



Juno's Personal Data Exploratory

https://exploratory.openhumans.org/notebook/3/

120%

Log In

Juno's Exploratory

Explore

Search Notebook

Submit

About

Details for **rescuetime-vs-step-counts.ipynb**

Open

Published by [gedankenstuecke](#)

Description

How do your productivity and screen time influence your physical activity? Find out by looking at your **RescueTime** and **Fitbit** or **Apple Healthkit** data.
Important: You only need one of the activity trackers to use this notebook.

101

1

Tags & Data Sources

activity tracking

productivity tracking

work life balance

RescueTime connection

Fitbit Connection

Open Humans Healthkit Integration

Open in Personal Data Notebooks

Comments

Please [log in to comment](#).

Notebook

Last updated 10 months, 4 weeks ago

Using **RescueTime** to analyze your productivity and computer usage

RescueTime is a service that monitors how you're using your computer. Split into 5 minute intervals you can't only see what you did with your computer, but it also classifies these activities automatically in productive/unproductive time.
Let's use this data to see

- How our productivity changes over time
- How we're spending most of our computer time

and if we're also counting steps with **Fitbit** and/or **Apple HealthKit** to see

- How our computer usage influences our activity

To get started let's find out which data sources we have on *Open Humans* and identify & load the data from **RescueTime**
In [1]:

```
import json
import os
```

day

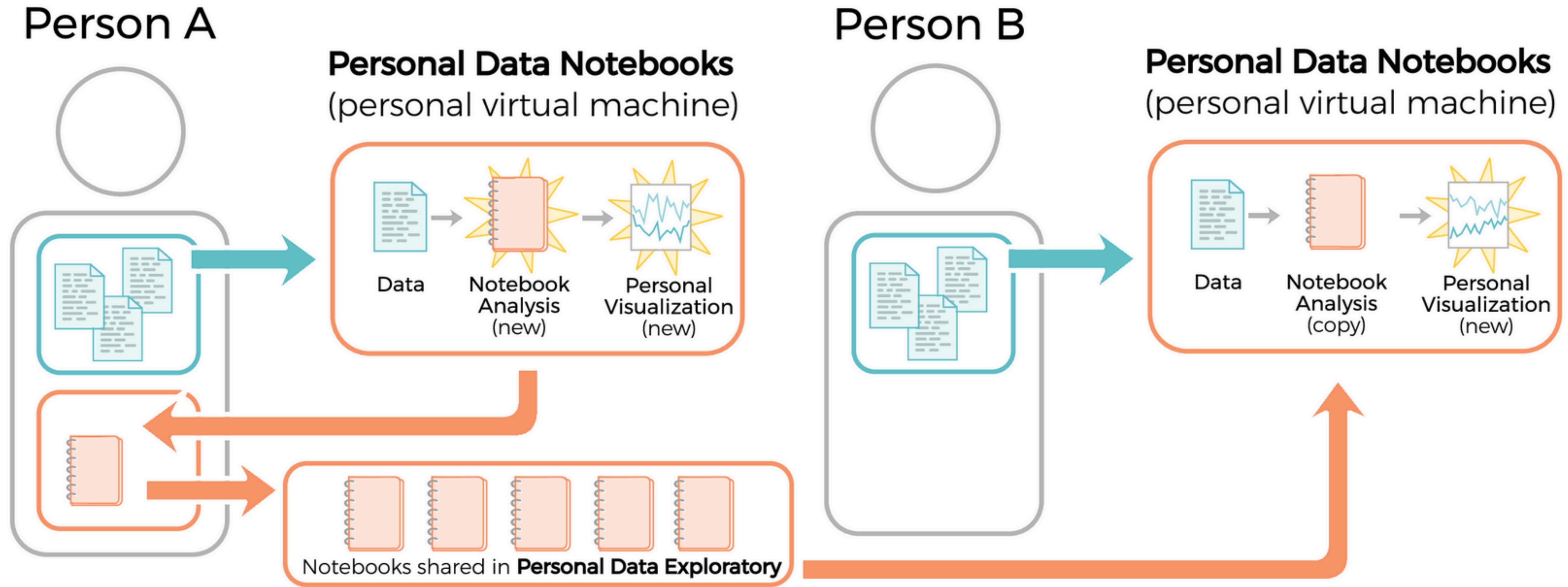
Highlight All

Match Case

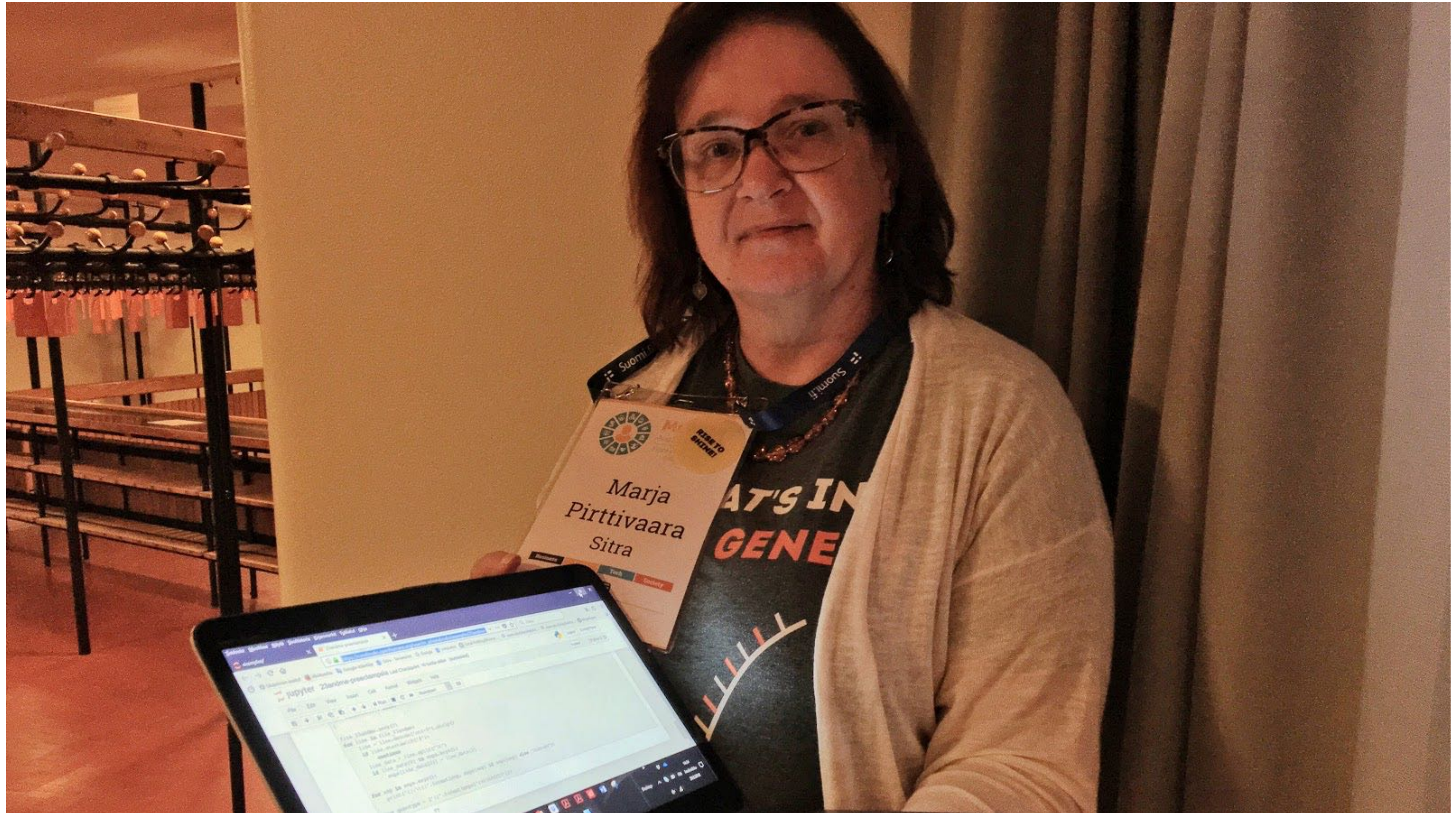
Whole Words

4 of 19 matches

sharing notebooks

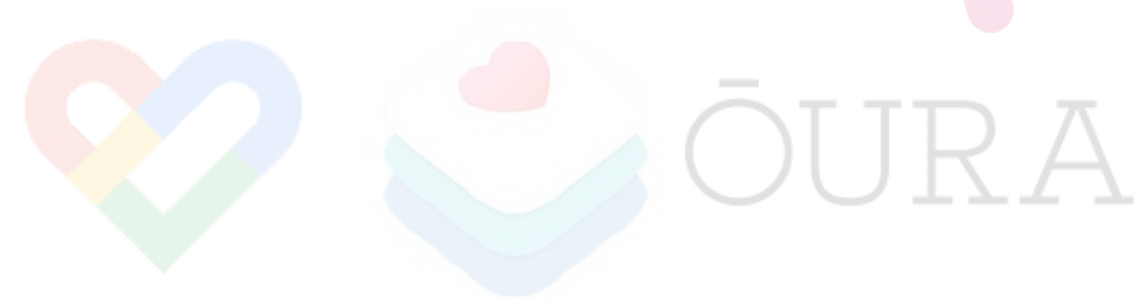


data analyses over 🍺🍺



personal data notebooks

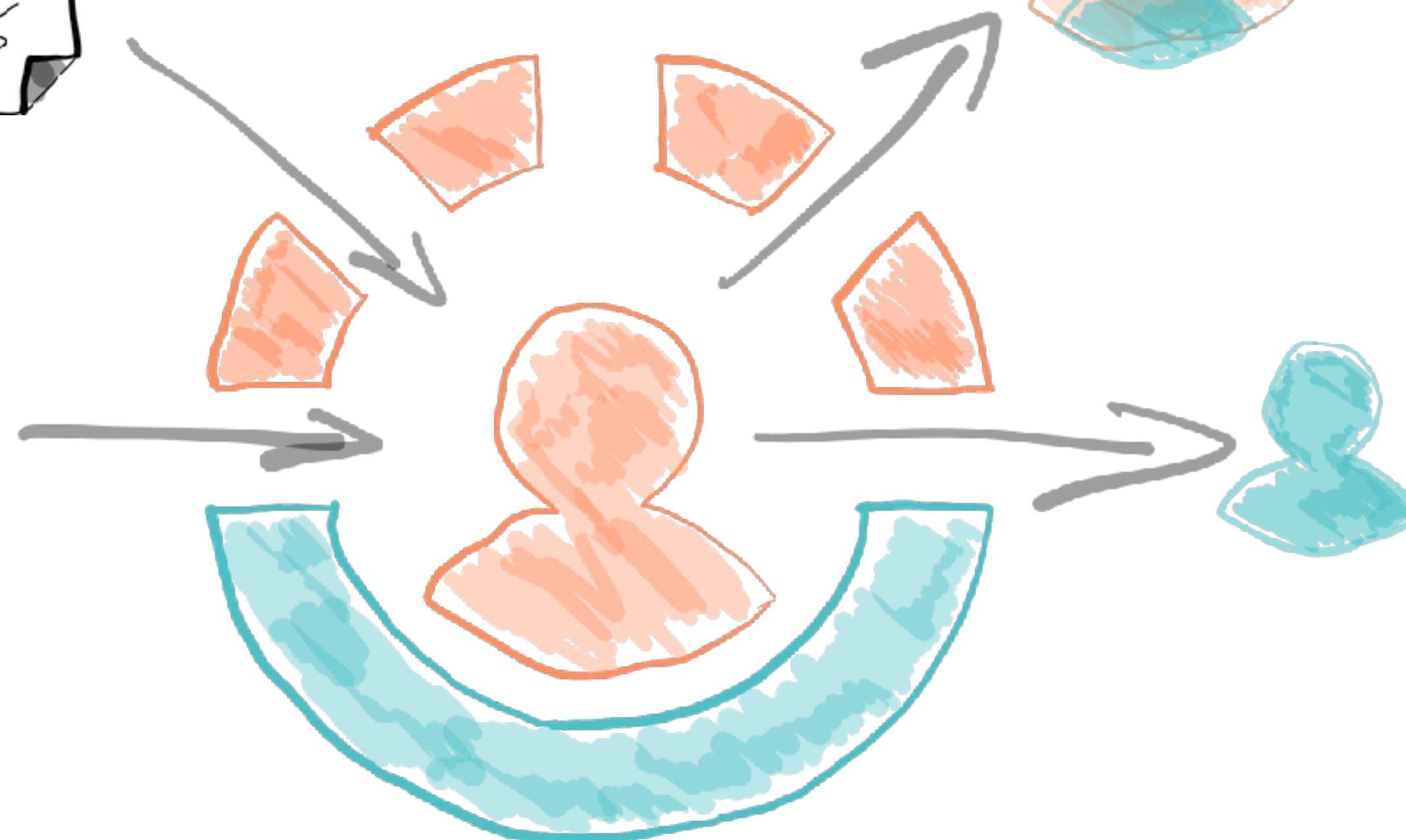
- write & run your data analysis
 - without having to download your data sets
 - without having to code against multiple APIs
- share analyses to be re-used by others



individual imports
personal data,



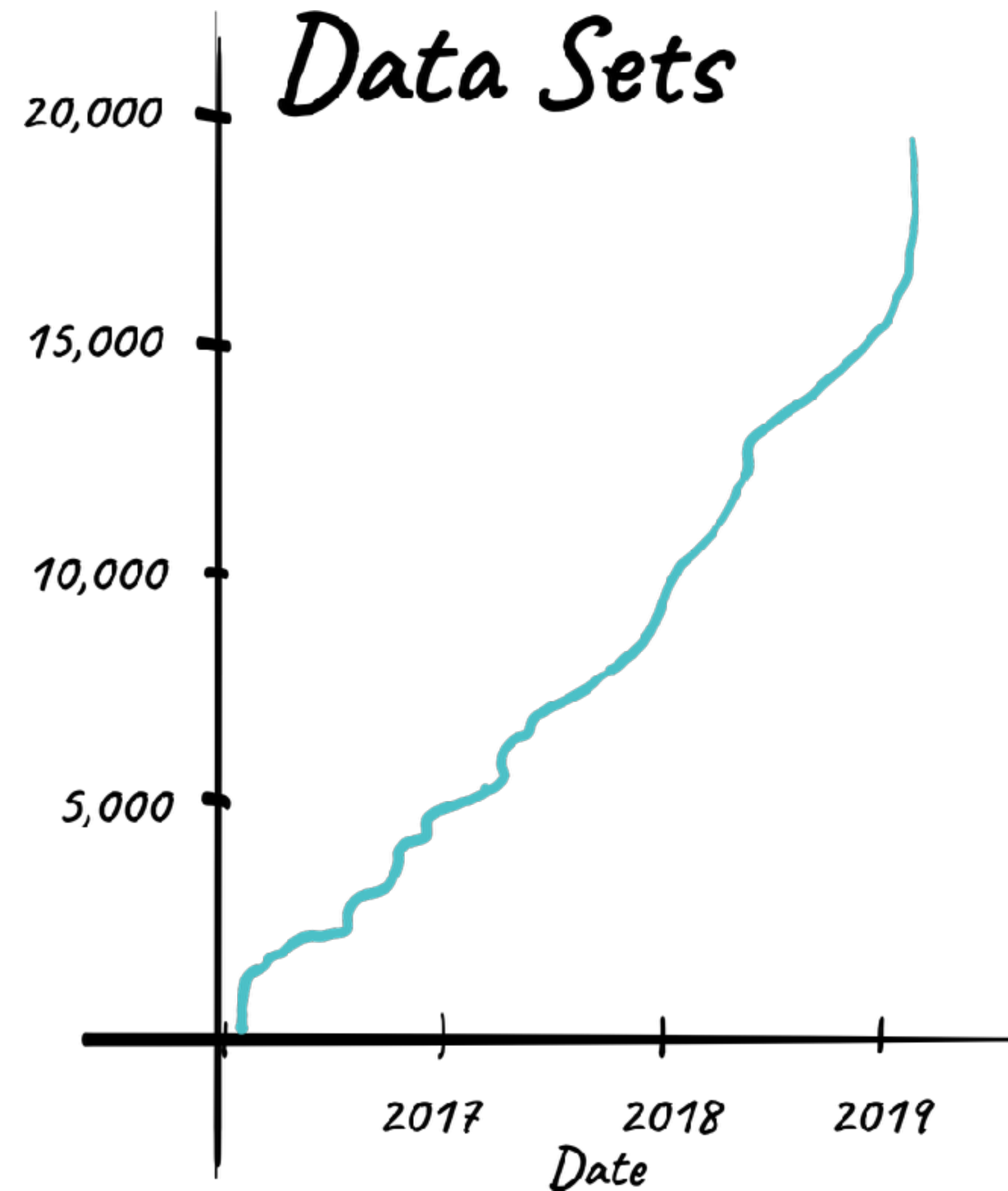
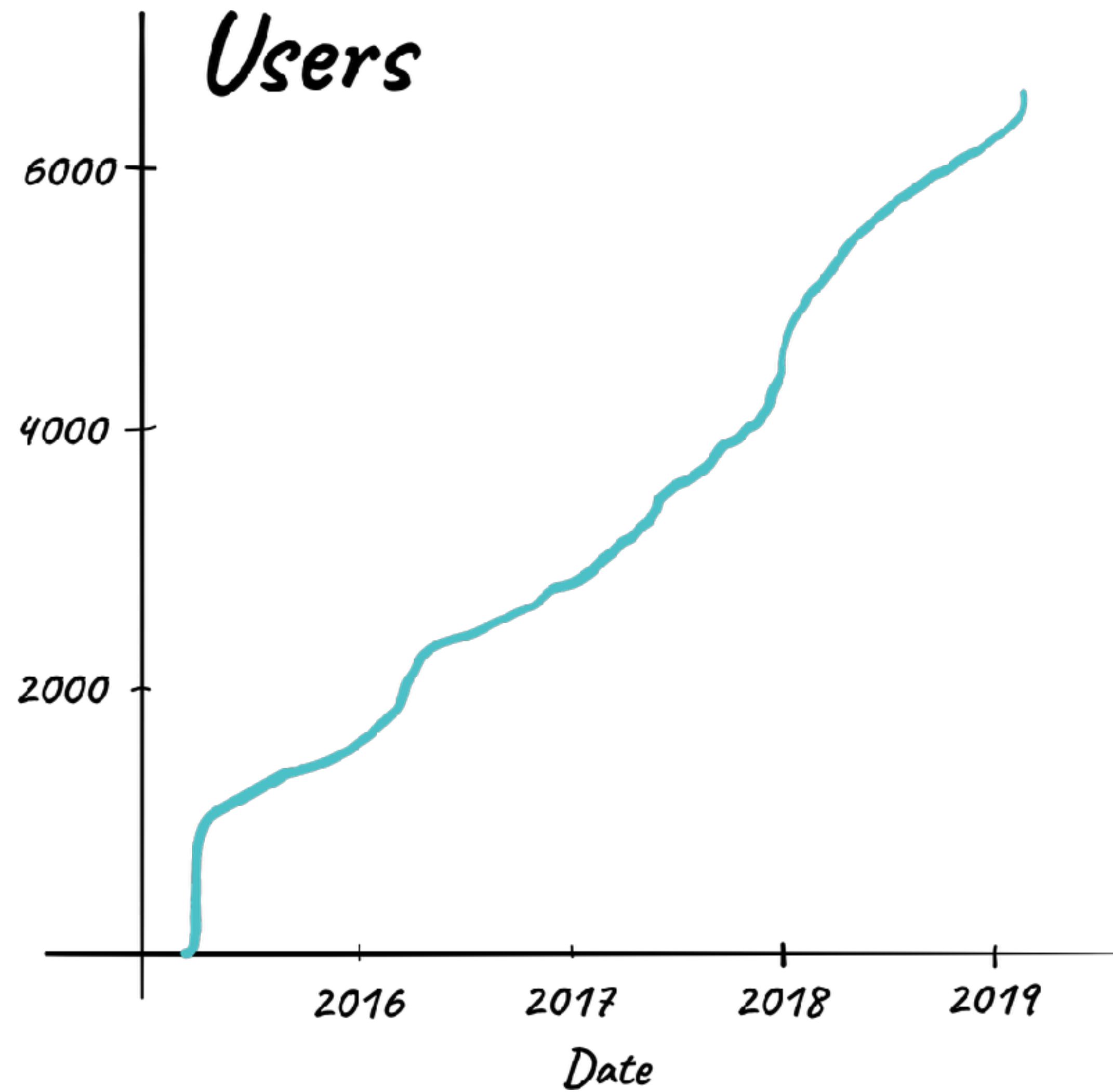
stores data
privately & can
explore it,



OPEN HUMANS

can share with
projects (academic
& grassroots)

Member Community



**can we leverage this community to do
participatory research?**



This is **Dana Lewis** and her
Open Artificial Pancreas System.

CGM data = continuous glucose monitor

#OpenAPS



NIGHTSCOUT
#WeAreNotWaiting



Welcome to QCycle!

QCycle is a participant-led research (PLR) project focusing on physiological change across the ovulatory cycle. We're a mix of researchers, computer scientists, engineers, and self-trackers looking to learn about ourselves, predict our LH surges, investigate links between our cycles and mood, HRV, productivity and more.

This is our private website for project resources. You can come here to find:

[Background information](#) on our project.

[Instructions](#) on how to use your devices and test kits.

[Important links.](#)

[Iterative Data Visualizations.](#)

individual-centric personal data sharing

- get individuals to actively consent to your study
 - data access is automatically granted after consent
- reduces friction, as researchers can use preexisting data

Projects using Data



Genevieve Genome Report

Mad Ball

Connected by **807** members

Genevieve matches your genome against public variant data, and invites you to contribute to shared notes.

[Learn more](#)



Twitter Archive Analyzer

Bastian Greshake Tzovaras

Connected by **519** members

The TwArxiv is a Twitter Archive Analyzer. Upload your Twitter archive and get new insights.

[Learn more](#)



Keeping Pace

Dr. Rumi Chunara
New York University

Joined by **400** members

Keeping Pace seeks to study data about how we move around, to understand how seasons and local environment influence our movement patterns.

[Learn more](#)



Juno's Personal Data Exploratory

Bastian Greshake Tzovaras
Open Humans Foundation

Connected by **96** members

Showcase and share your Personal Data Notebooks with others

[Learn more](#)



UbiQomix Microbiome Exploration

Orit Shaer, Mad Price Ball,
Oded Nov
Wellesley College, Open Humans,
NYU

Joined by **123** members

This research survey study aims to gain insight into user needs as they engage with their personal and social microbiome data.

[Learn more](#)



Gut Instinct

Vineet Pandey
UCSD Design Lab

Connected by **93** members

Learn about current science of the gut microbiome with expert-curated learning materials.

[Learn more](#)



Circles

Cliff Tabin and Abigail Wark
Harvard Medical School

Joined by **316** members

Circles is a research study that aims to discover the genetic basis for a mysterious and remarkable human trait: the areola.

[Learn more](#)



Personal Data Notebooks

Bastian Greshake Tzovaras
Open Humans Foundation

Connected by **466** members

Play with code for analyzing personal data! Your data stays private on a personal virtual machine.

[Learn more](#)



GenomX Genome Exploration

Orit Shaer & Oded Nov
Wellesley College & NYU

Joined by **356** members

Is your 23andMe data public? Register interest in a study of how folks interact with their genome data using GenomX, our visualization tool

[Learn more](#)



Imputer

Kevin Arvai

Connected by **55** members

Imputer will perform imputation on user's genotype data to estimate genotypes at missing loci.

[Learn more](#)



Galileo

Vineet Pandey
UCSD Design Lab

Connected by **54** members

Learn about current science of the gut microbiome with expert-curated learning materials.

[Learn more](#)



Nightscout Data Commons

The Nightscout Data Commons Committee
The Nightscout Foundation

Joined by **66** members

The Nightscout Data Commons collects data from Nightscout users and facilitates research in partnership with the DIY diabetes community.

[Learn more](#)



Data sharing and ethical oversight

Prof. Dr. Effy Vayena
ETH Zurich

Joined by **210** members

Through a quick survey we aim at understanding what the Open Humans community thinks about ethical oversight.

[Learn more](#)



openSNP

Bastian Greshake Tzovaras
openSNP

Connected by **290** members

openSNP allows you to put your genetic and phenotypic data into the public domain. Connect your openSNP account to Open Humans.

[Learn more](#)



OpenAPS Data Commons

Dana Lewis
OpenAPS

Joined by **128** members

The OpenAPS Data Commons collects data from DIY closed loopers and facilitates research in partnership with the DIY closed loop community.

[Learn more](#)



lineage

Andrew Riha
Scedastic Software, LLC

Connected by **35** members

lineage provides tools for genetic genealogy - find shared DNA and genes between individuals, merge raw data files, and re-map SNPs

[Learn more](#)



Resilience Project Study

Jason Bobe
Icahn School of Medicine at Mount Sinai

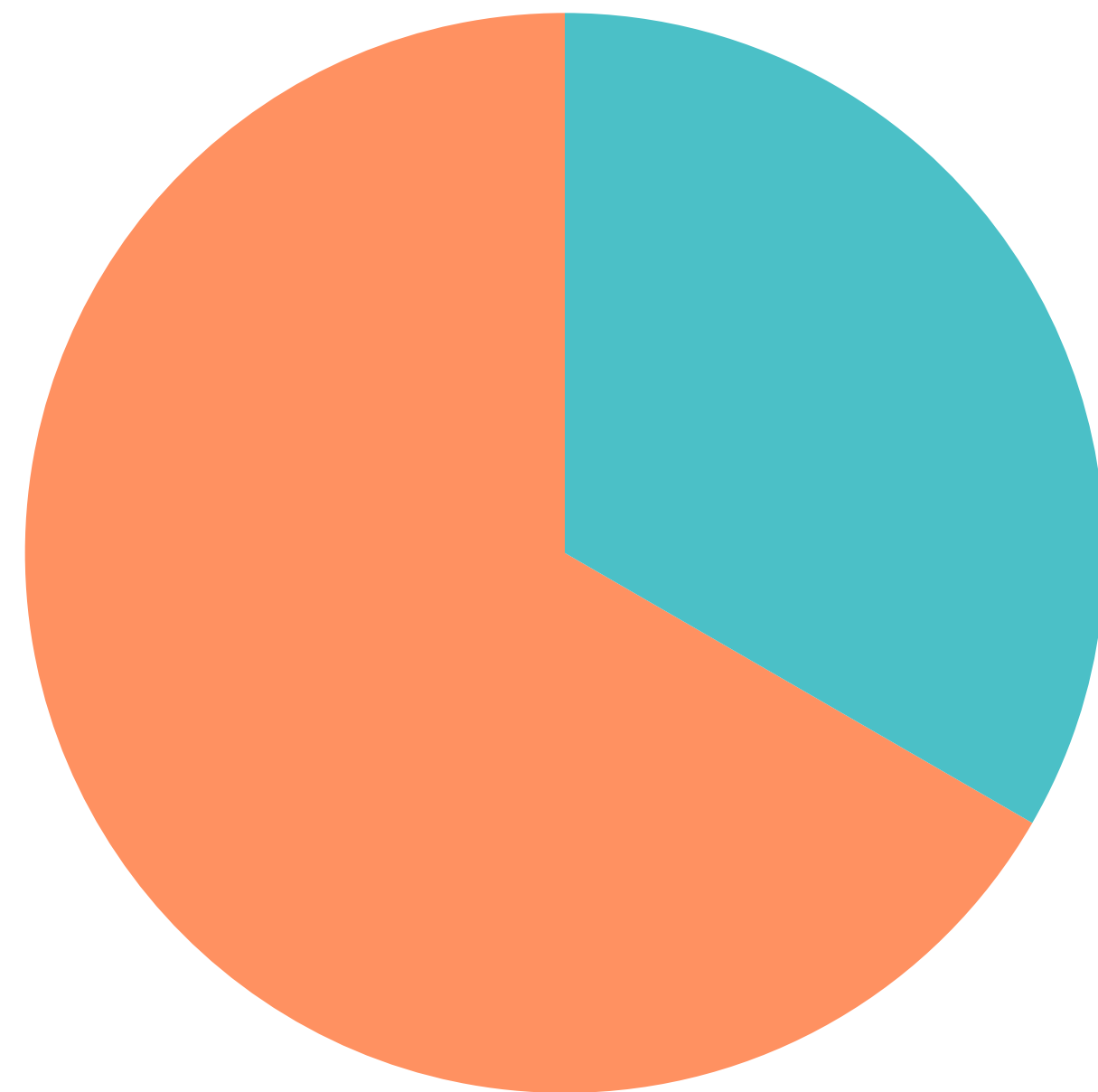
Connected by **6** members

The Resilience Project aims to discover why some people are more able to overcome challenges to their health and escape disease.

Governance: Building **trust** &
community

Governance: Building **trust** & **community**

OHF Board of Directors: Self-nominations invited



● Community-Elected
● Board-Elected

I'm also delighted to share that we have restructured our organization's governance to create "community seats". One of the three seats this round will be chosen by Open Humans members!

Anyone may apply to our board. The process involves a self-nomination, and nominees should be seconded by a current member of the Board of Directors.

Being a director of this organization is a position of trust. While I manage day-to-day operations, the board is our highest tier of governance – our ultimate decision-making authority. Many board members also contribute to the organization as officers, e.g. as secretary or treasurer. As Executive Director, I serve at the board's behest.

Governance: Building **trust** & **community**

Approval for Resilience Project Study

 Project Reviews approved



madprime 

1  23d

I'm posting a review request for the [Resilience Project Study](#) ⁴, which is applying for project approval on Open Humans. This project is part of an IRB-approved study led by Jason Bobe at the Icahn School of Medicine at Mount Sinai, and we've received documentation of this IRB approval.

The project would like to launch their app in the Apple app store and would like a quick review turnaround, if possible.

Should this project be visible and available for all Open Humans members to join?

Please vote **Approve** or **Deny**, and/or comment.

Governance: Building **trust** & **community**



beau

23d

Is the documentation of the IRB approval something that can be shared on the project page? (I recall that there are factors that sometimes make sharing anything related to the actual IRB documents tricky...)

I'm posting a review request for the [Resilience Project Study](#), which is applying for project approval on Open Humans. This project is part of an IRB-approved study led by Jason Bobe at the Icahn School of Medicine at Mount Sinai, and we've received documentation of this IRB approval.

The project would like to launch their app in the Apple app store and would like a quick review turnaround, if possible.

Should this project be visible and available for all Open Humans members to join?

Please vote **Approve** or **Deny**, and/or comment.

Governance: Building **trust** & **community**



beau

23d

Is the documentation of the IRB approval something that can be shared on the project page? (I recall that there are factors that sometimes make sharing anything related to the actual IRB documents tricky...)



jasonbobe

22d

@beau Here is the most recent approval letter from the IRB.



Icahn School of Medicine at Mount Sinai
Mount Sinai Beth Israel
Mount Sinai Brooklyn
The Mount Sinai Hospital
Mount Sinai Queens
New York Eye and Ear Infirmary
of Mount Sinai
Mount Sinai St. Luke's
Mount Sinai West

**Program for the Protection
of Human Subjects**
Institutional Review Boards
Mount Sinai Health System
One Gustave L. Levy Place, Box 1081
New York, NY 10029-6574
T 212-824-8200
F 212-876-6789
irb@mssm.edu
icahn.mssm.edu/pphs

APPROVAL OF RESEARCH



- import your own data, 📱
- explore it and share your analyses notebooks, 📊
- share your data with cool research projects. 🎉🔬

Acknowledgements



Mad Ball
Mairi Dulaney
Mike Escalante
Tim Head
Rosy Gupta
Manaswini Das
Jasmine Tamak
Tarannum Khan



Chris Mungall
Nomi Harris
Deepak Unni

Ext. Collaborators: Samuel Gershman (Harvard)
Marcel Salathé (EPFL)
John Wilbanks (Sage Bionetworks)
Effy Vayena (ETH Zurich)

With support from:



Google
Summer of Code

