How to get into R

Alexander Matrunich
18 May 2017, R-Ladies Tbilisi

What are we going to do

See details in the original post by Maëlle Salmon and a description in David Smith's post.

- Who am I?
- Why R?
- What is R?
  - Short R history
  - Basic elements of R installation
  - RStudio IDE
- How to R?
- Contacts

Alexander (Alex) Matrunich

Specialist degree in sociology at Pskov Volny Institute

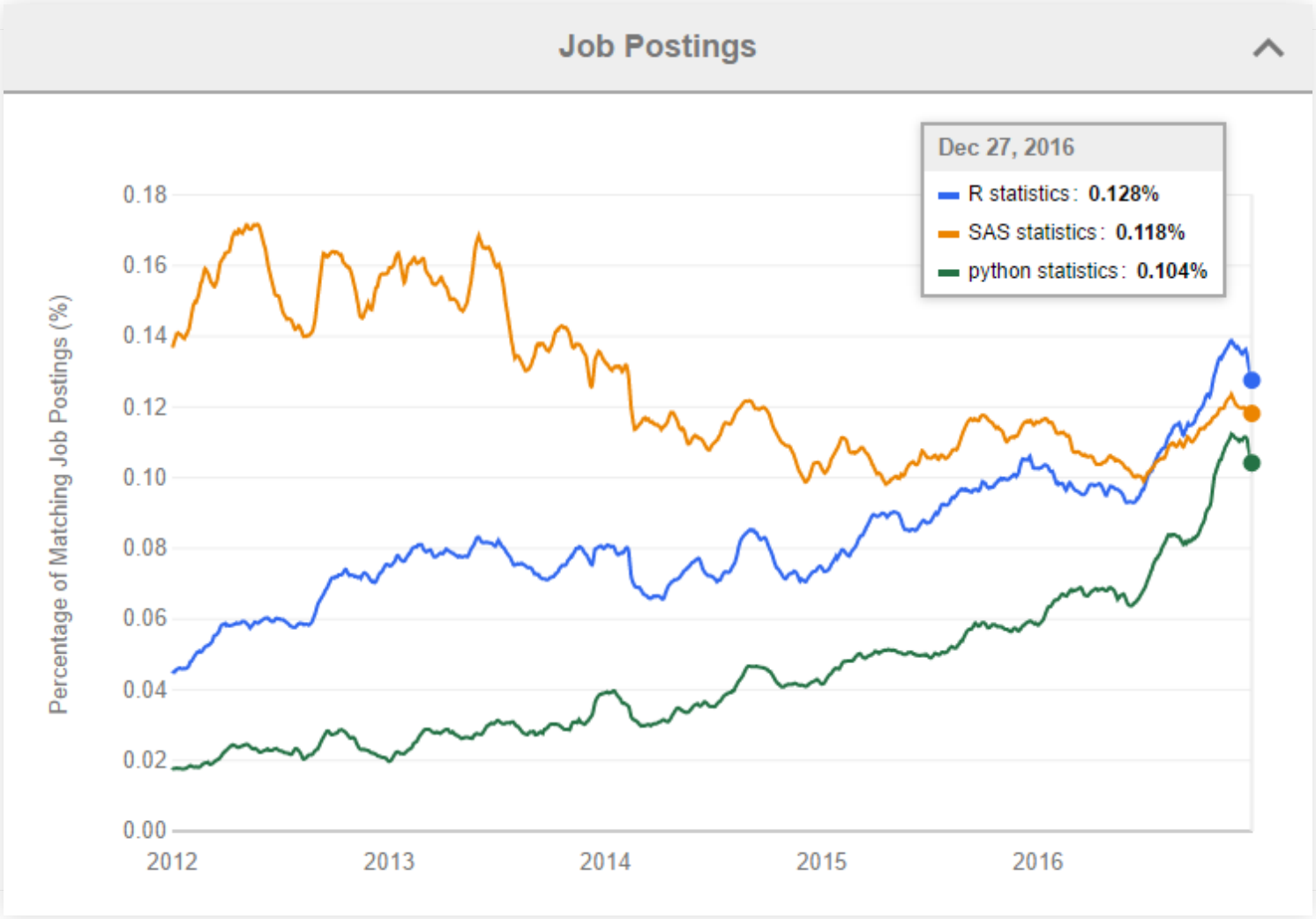
Free open source enthusiast (GNU/Linux, LibreOffice, etc.)

Open data contributor (Wikipedia, OpenStreetMap)

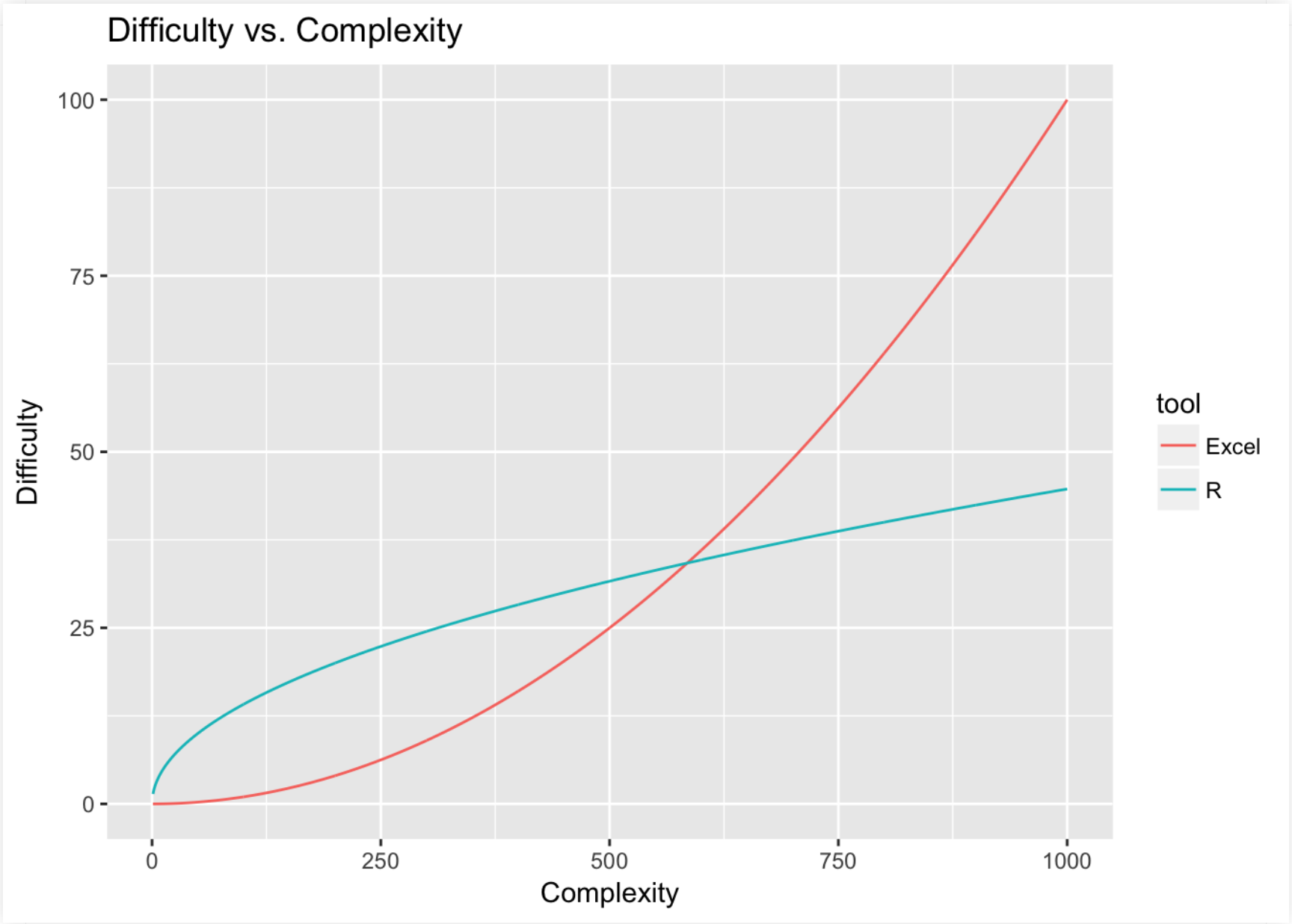
Notable projects with R:

- Opinion poll data processing and report generation
- Exit-poll real-time data processing and report generation
- R training for the UN FAO staff
- World trade data analysis for the UN FAO

Why R?



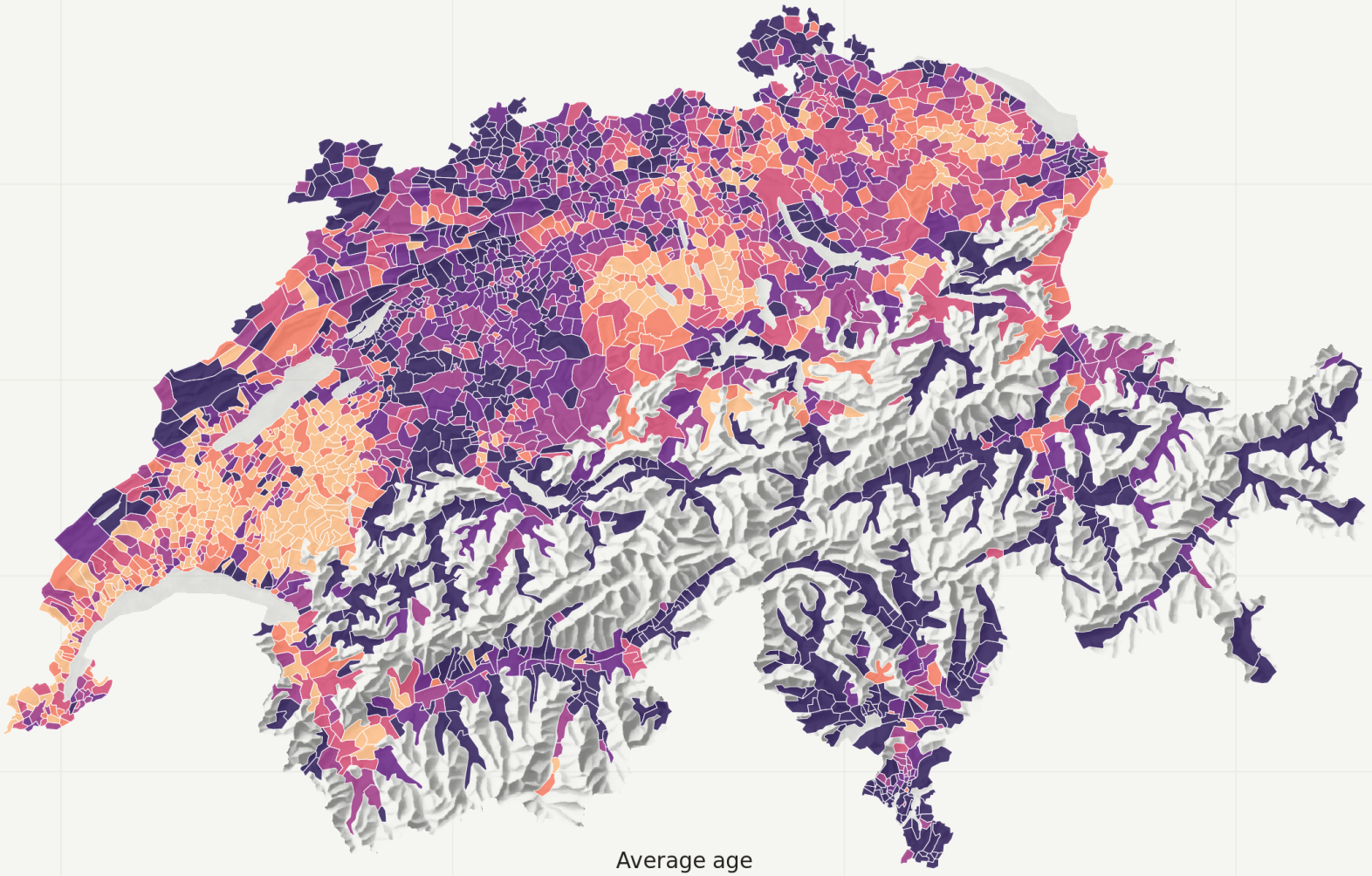
Screenshoot by David Smith from Indeed Job Trends



Idea and plot by **Gordon Shotwell**

Switzerland's regional demographics

Average age in Swiss municipalities, 2015



Map CC-BY-SA; Author: Timo Grossenbacher (@grssnbchr), Geometries: ThemaKart, BFS; Data: BFS, 2016; Relief: swisstopo, 2016

Plot by **Timo Grossenbacher**. See notes by **David Smith**

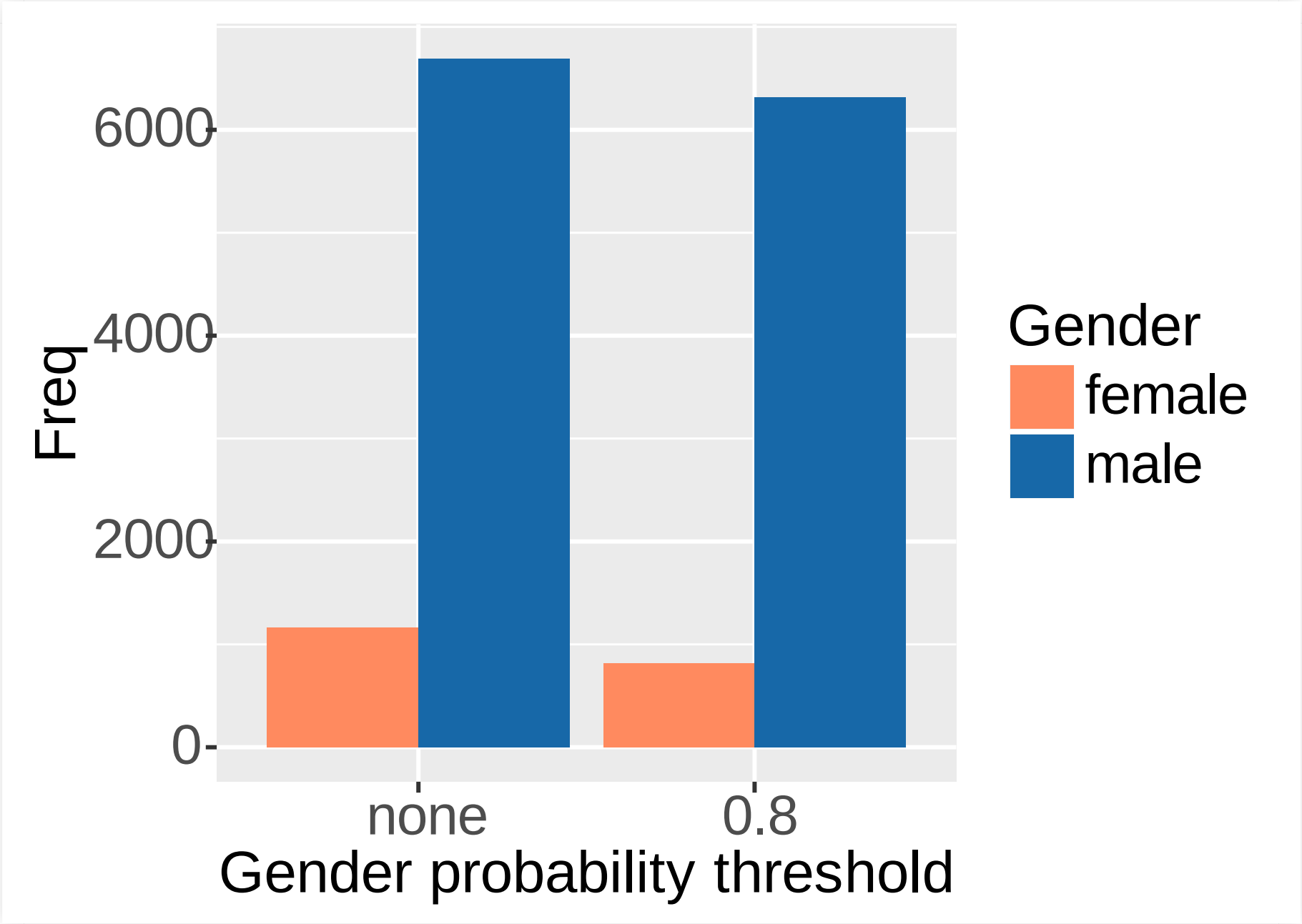
See details in **David Smith's post**

Where Europe lives, in 14 lines of R code

```
read_csv('GEOSTAT_grid_POP_1K_2011_V2_0_1.csv') %>%
  rbind(read_csv('JRC-GHSL_AIT-grid-POP_1K_2011.csv') %>%
          mutate(TOT_P_CON_DT='')) %>%
  mutate(lat = as.numeric(gsub('.*N([0-9]+)[EW].*', '\\1', GRD_ID))/100,
         lng = as.numeric(gsub('.*[EW]([0-9]+)', '\\1', GRD_ID)) *
           ifelse(gsub('.*([EW]).*', '\\1', GRD_ID) == 'W', -1, 1) / 100)
%>%
  filter(lng > 25, lng < 60) %>%
  group_by(lat = round(lat, 1), lng = round(lng, 1)) %>%
  summarize(value = sum(TOT_P, na.rm = TRUE))  %>%
  ungroup() %>%
  complete(lat, lng) %>%
  ggplot(aes(lng, lat + 5*(value/max(value, na.rm = TRUE)))) +
  geom_line(size = 0.4, alpha = 0.8, color='#5A3E37', aes(group=lat),
na.rm=TRUE) +
  ggthemes::theme_map() +
  coord_equal(0.9)
```

See details in **David Smith's post**

About ladies... R-ladies



See slides **R Foundation Women in R Taskforce**

What is R?

History of R

1976 - S language in Bell Labs by John Chambers

1993 - R - open source implementation of S by Ross Ihaka and Robert Gentleman at the University of Auckland

1997 - CRAN; part of the GNU Project

2000 - Version 1.0, suitable for production use

2004 - Version 2.0

2013 - Version 3.0. Values 2^{31}

2016 - R-Ladies Tbilisi founded

Base R Software

1. Base R (written in C, Fortran and R)
2. An R package:
    - Collection of R functions
    - Documentation (manuals and optionally vignettes)
    - Optionally data sets
    - Optionally inclusions of other programming languages (C++, Python, Java)
3. CRAN - Comprehensive R Archive Network (10K packages January 2017)
4. Every R user has:
    1. Base R installation
    2. System library of packages
    3. User's library of packages

Task views

Integrated Development Environments for R

Free open source **RStudio IDE**:

- Syntax highlighting, code completion, and auto indentation
- Execute R code directly from the source editor
- Quickly jump to function definitions
- Easily manage multiple working directories using projects
- Integrated R help and documentation
- Interactive debugger to diagnose and fix errors quickly
- Extensive package development tools

How to R?

How to start

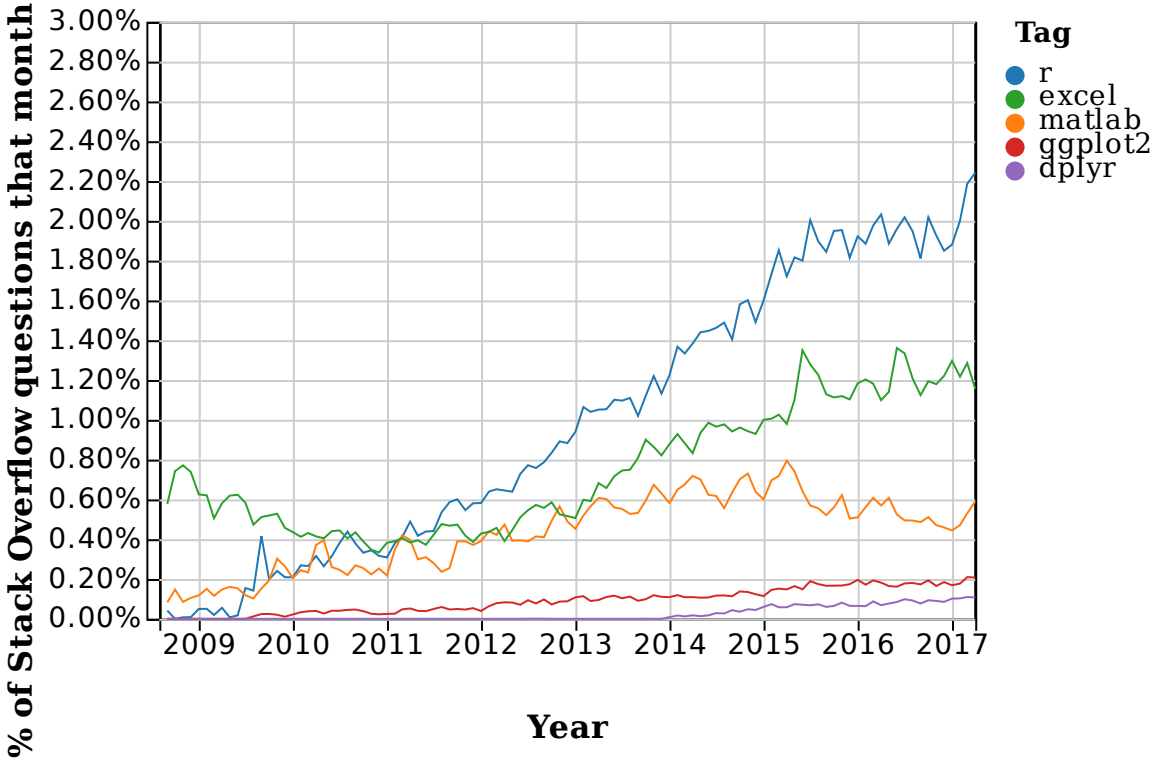
- Work in RStudio (if you are not an Emacs/Vim/… user)
- Every job in separate directory (projects in RStudio)
- Don't save and don't restore workspace image (settings in RStudio)
- Follow code style
- Restart R-session to avoid traces (Ctrl+Shift+F10 in RStudio)
- Comment your code: why but not what

Errors

- Read them!
- Search in English
- Use `r`, `rstat`, `rlang`, `cran` keywords
- Short reporducible examples when asking for help

Storages of R-knowledge

https://insights.stackoverflow.com/trends?tags=r%2Cstatistics



- Stackoverflow
- Twitter: #rstat, Hadley Wickham, David Smith
- Blogs: R-Bloggers

Contacts

Alexander (Alex) Matrunich

a@matrunich.com

Twitter: **@matrunich**