

R Workshop

Charles Crabtree (Dartmouth College)

January 18, 2022

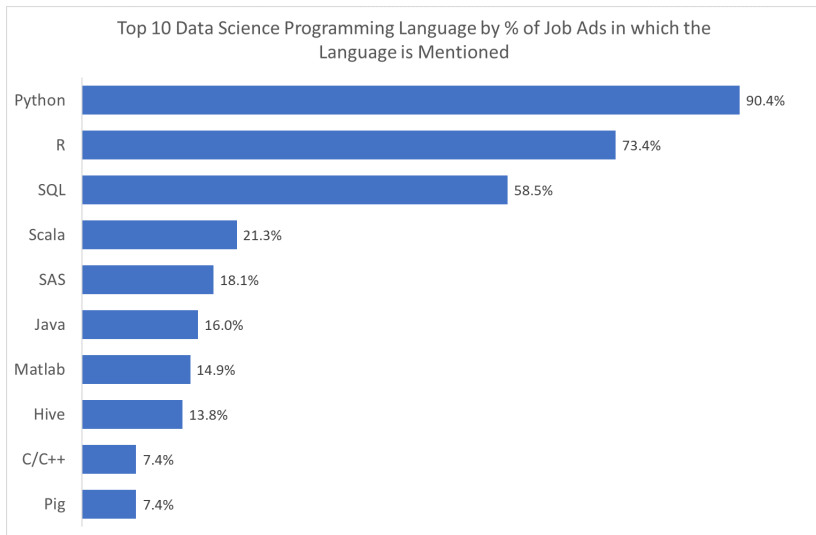
Statistical software



Why R and not something else (that is maybe easier)?

- ▶ Free and open-source
- ▶ Large community of users
- ▶ Flexible, powerful, command-based (no point-and-click)
- ▶ Fun

Statistical software



Statistical software



- ▶ Integrated development environment (IDE) for R
- ▶ Makes R much easier to use
- ▶ Facilitates good coding practices

What are data?

Used to seeing it in this format:

	A	B	C	D	E	F	G	H
1		country_name	country_code	year	totch	totir	totgr	post_cw
2	1	Afghanistan	700	1923	2	2	1.5	0
3	2	Afghanistan	700	1931	4.5	0.75	2	0
4	3	Afghanistan	700	1964	6.5	0.5	1	0
5	4	Afghanistan	700	1977	7.5	0.5	1.5	0
6	5	Afghanistan	700	1979	0	0	0	0
7	6	Afghanistan	700	1980	4.5	5.75	4	0
8	7	Afghanistan	700	1987	6.5	5.25	6	0
9	8	Afghanistan	700	1990	8	5.75	5.75	0
10	9	Afghanistan	700	2001	5.5	0.5	0.5	1
11	10	Afghanistan	700	2004	9	1.25	8.25	1
12	11	Albania	339	1914	2	3.5	6.5	0
13	12	Albania	339	1920	0	0	0	0
14	13	Albania	339	1925	4.5	2.5	0.5	0
15	14	Albania	339	1928	9	2.5	1.5	0
16	15	Albania	339	1939	3	1	1	0
17	16	Albania	339	1946	3	8.25	7	0
18	17	Albania	339	1976	5	4	7	0
19	18	Albania	339	1991	0.5	2	2	0
20	19	Albania	339	1998	7.5	12	8.75	1
21	20	Albania	339	2012	2	11	11	1

What are data?

Two dimensions:

- ▶ Observations (rows)
- ▶ Variables (columns)

Example

Observations:

- ▶ Data on constitutional revisions

Variables:

- ▶ Country name
- ▶ Country code
- ▶ Year
- ▶ Cultural homogeneity index: Measuring the extent to which written constitutions subscribe to an exclusionary nationalism and institutionalize majority privileges.
- ▶ Group minority rights index: Measuring the degree to which individuals are protected from discrimination and entitled to exercise equal rights in articulating their ethnicity, language or religion.
- ▶ Individual minority rights index: Measuring the level of the state's commitment to put ethnic, linguistic, religious, and indigenous groups on equal footing and promote a pluralistic vision of the nation.

Intro to R

github.com/cdcrabtree/uva-2022

1. Download and extract files
2. Put them in your working directory
3. Open RStudio

Intro to R

R needs a working directory* – a place where it looks for files (like data) and where it puts things (like graphs you make).

What *is* a working directory? A folder on your computer.

Step 1: Create a folder for R stuff in this class. i.e. “class”

* with Chromebook, working directory is the rstudio cloud

Intro to R

Step 2: Get the filepath to that folder. For example, if your class folder was in your Dropbox in a folder called “UVA-2022”, your filepath would be:

`~/Dropbox/UVA-2022`

`C:/Users/cdcrabtree/Dropbox/UVA-2022`

depending on if you have Mac/Linux or Windows