Data visualization

Thijmen Jeroense

Department of Sociology/ICS

01-06-2022

Radboud University

What is a data visualization?

A way of transforming data into something that is perceivable by the human mind.

We can define data in a **broad** or **narrow** way.

Examples:

1. Figure or graph
2. Visual of theoretical model or process
3. Visual abstract (even a video)

PNAS

Exposure to opposing views on social media can increase political polarization

Christopher A. Bail[a,1], Lisa P. Argyle[b], Taylor W. Brown[a], John P. Bumpus[a], Haohan Chen[c], M. B. Fallin Hunzaker[d], Jaemin Lee[a], Marcus Mann[a], Friedolin Merhout[a], and Alexander Volfovsky[e]

[a]Department of Sociology, Duke University, Durham, NC 27708; [b]Department of Political Science, Brigham Young University, Provo, UT 84602; [c]Department of Political Science, Duke University, Durham, NC 27708; [d]Department of Sociology, New York University, New York, NY 10012; and [e]Department of Statistical Science, Duke University, Durham, NC 27708

Edited by Peter S. Bearman, Columbia University, New York, NY, and approved August 9, 2018 (received for review March 20, 2018)

Visualizing data

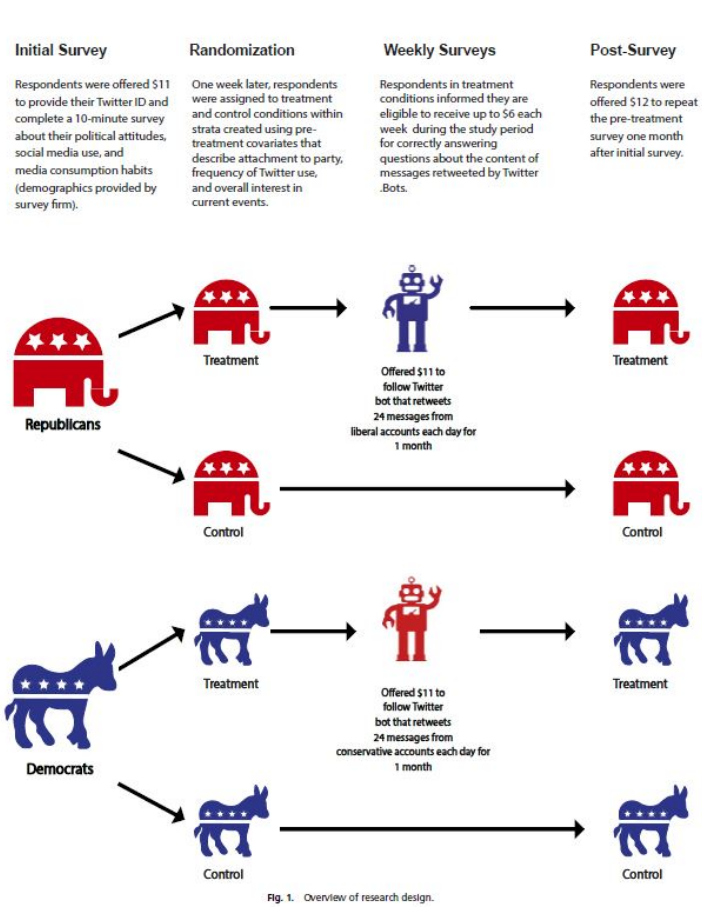


Fig. 1. Overview of research design.

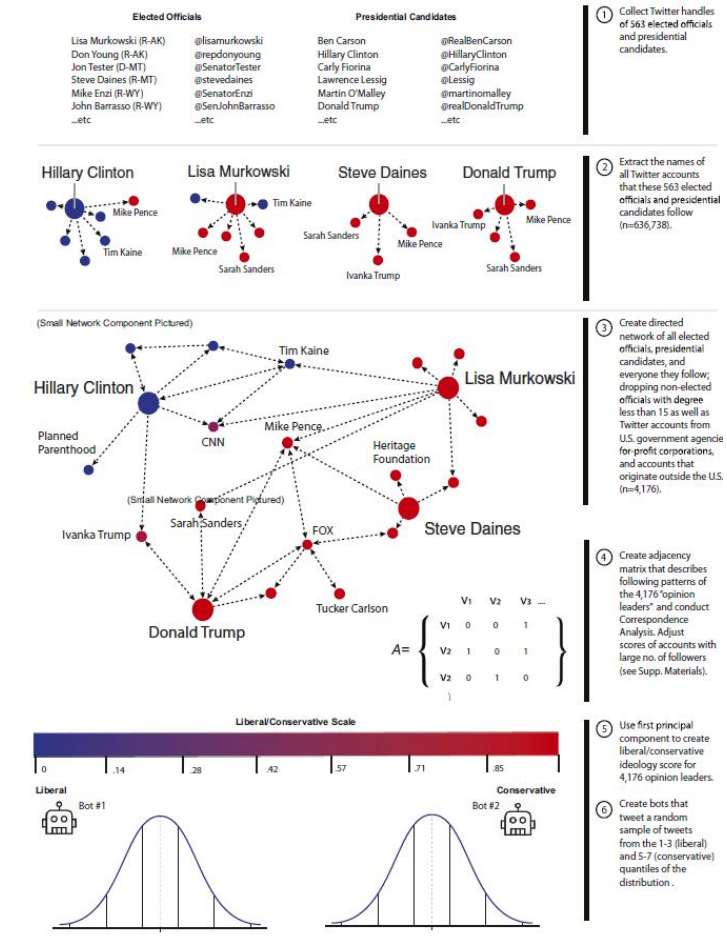


Fig. 2. Design of study's Twitter bots.

Why do we need data visualizations?

1. Communicate knowledge and research more effectively
2. Visual aid to clarify your story
3. Attract a bigger and broader audience for your research
4. Career prospects (publications and grants)

Data visualization = storytelling

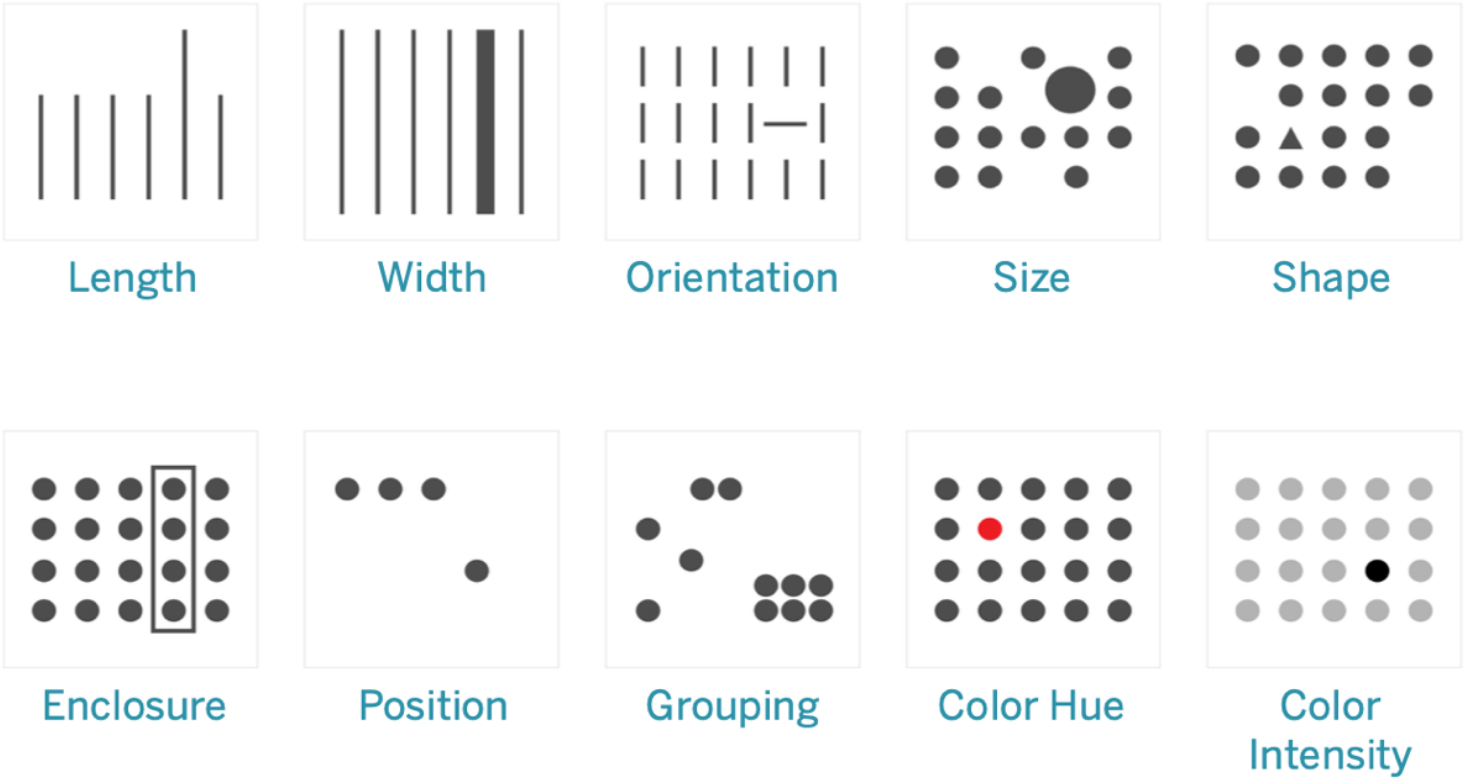
Things to keep in mind when writing your blog:

- What is my story?
- Who is my audience?
- What is my goal?

Not so different for creating (data)visuals:

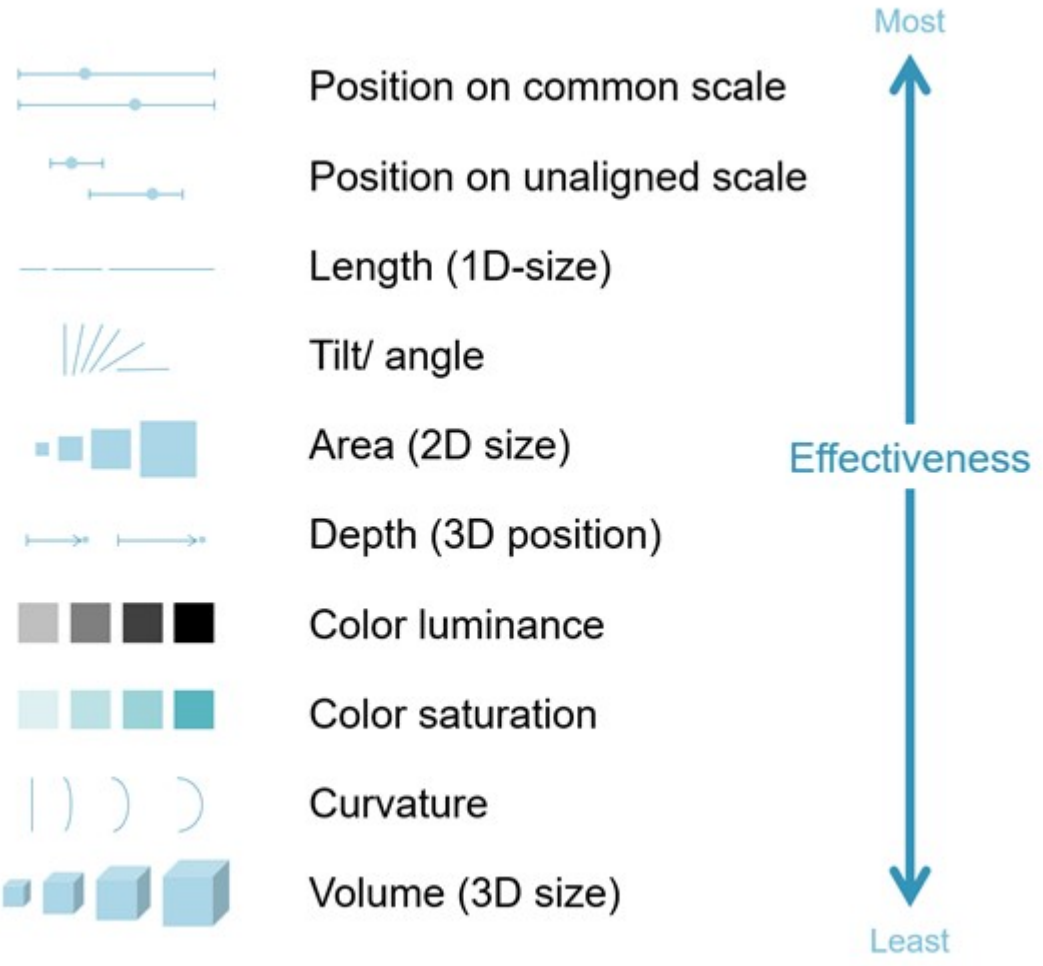
- Message
- Informative
- Goal

Preattentive attributes of visual perception



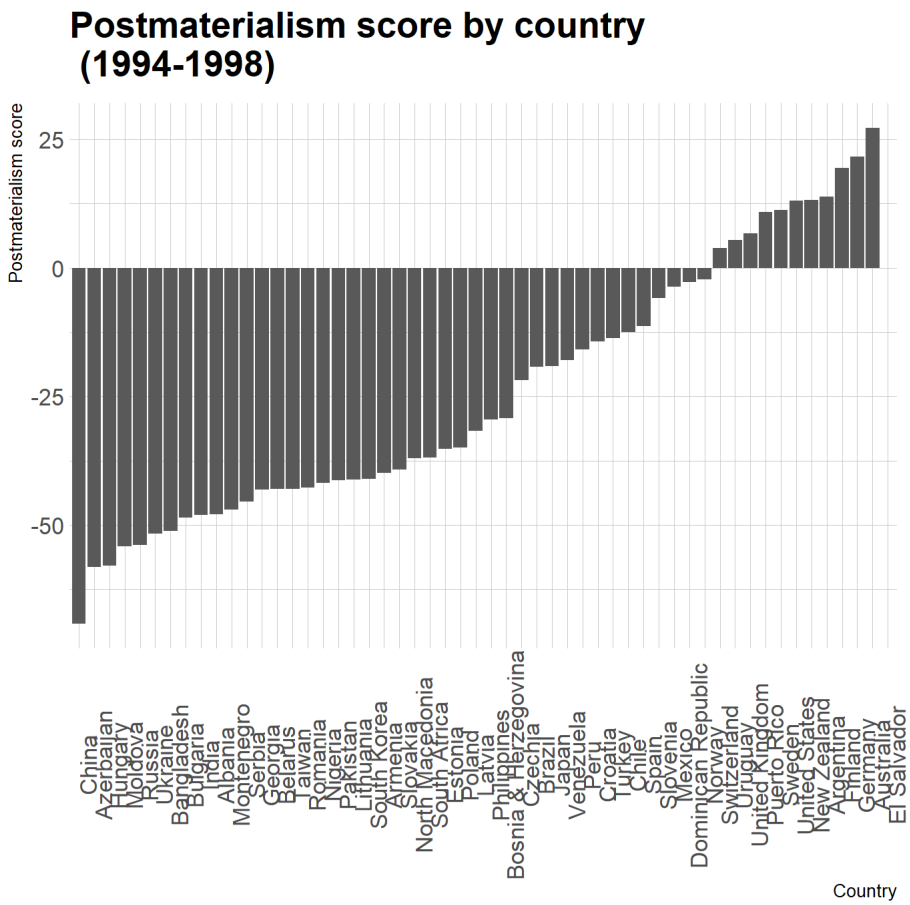
source: https://help.tableau.com/current/blueprint/en-us/bp_why_visual_analytics.htm

Preattentive attributes of visual perception

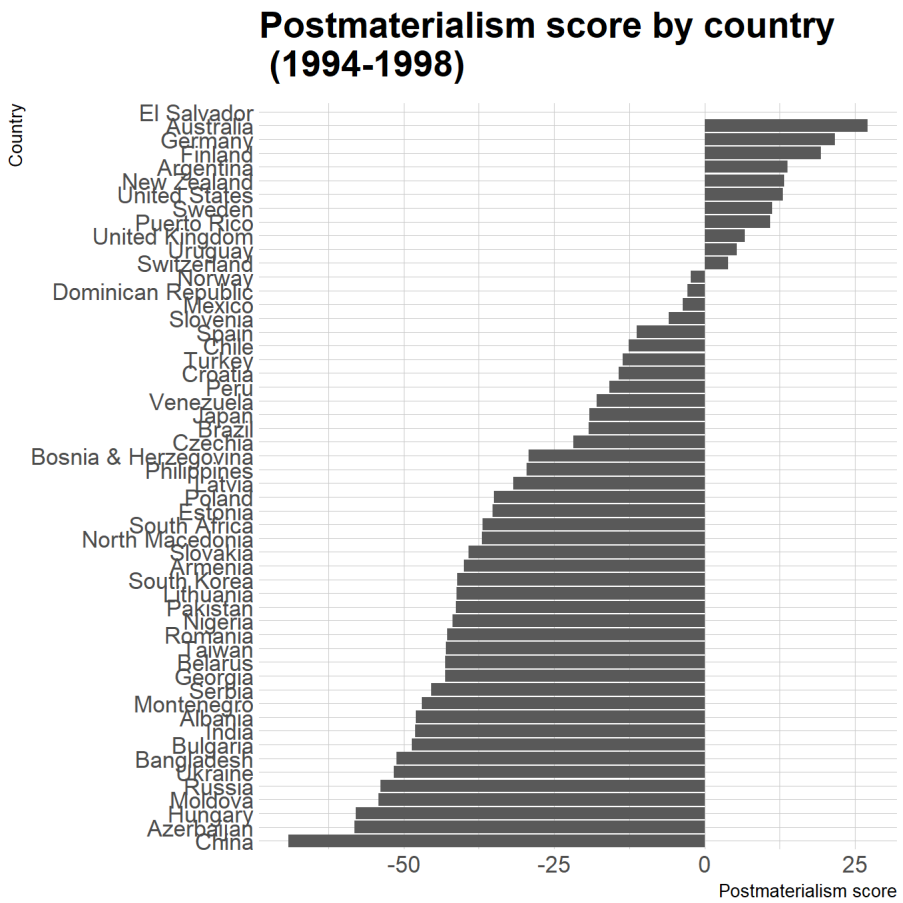


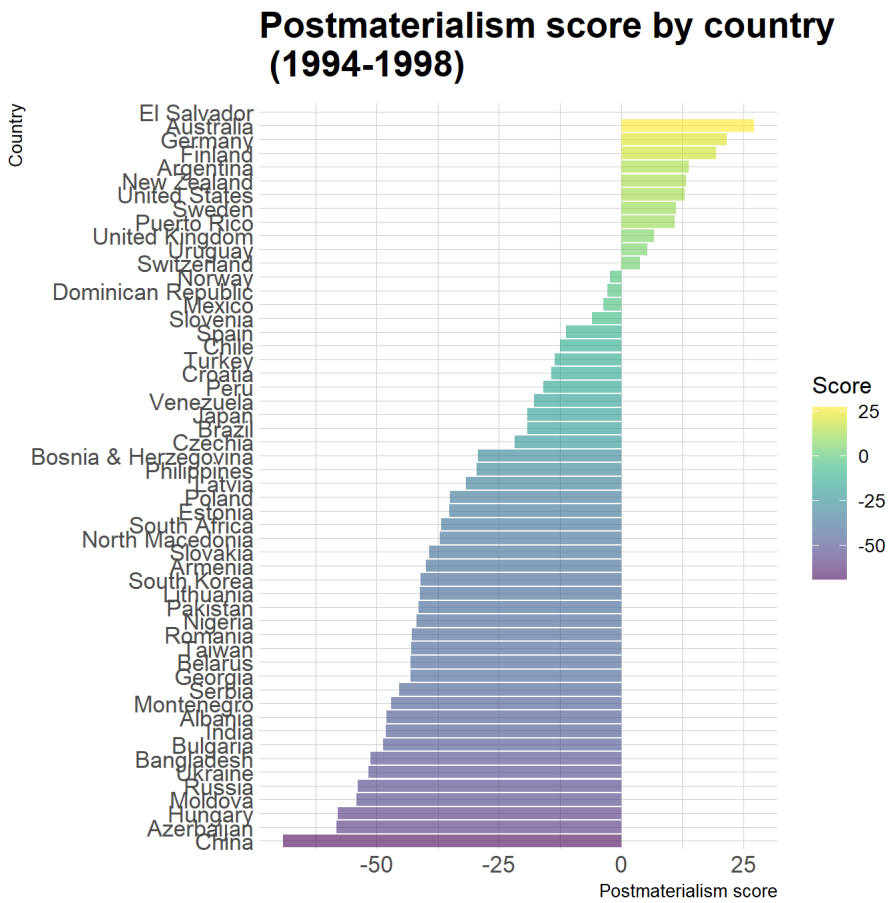
source: https://www.statstories.com

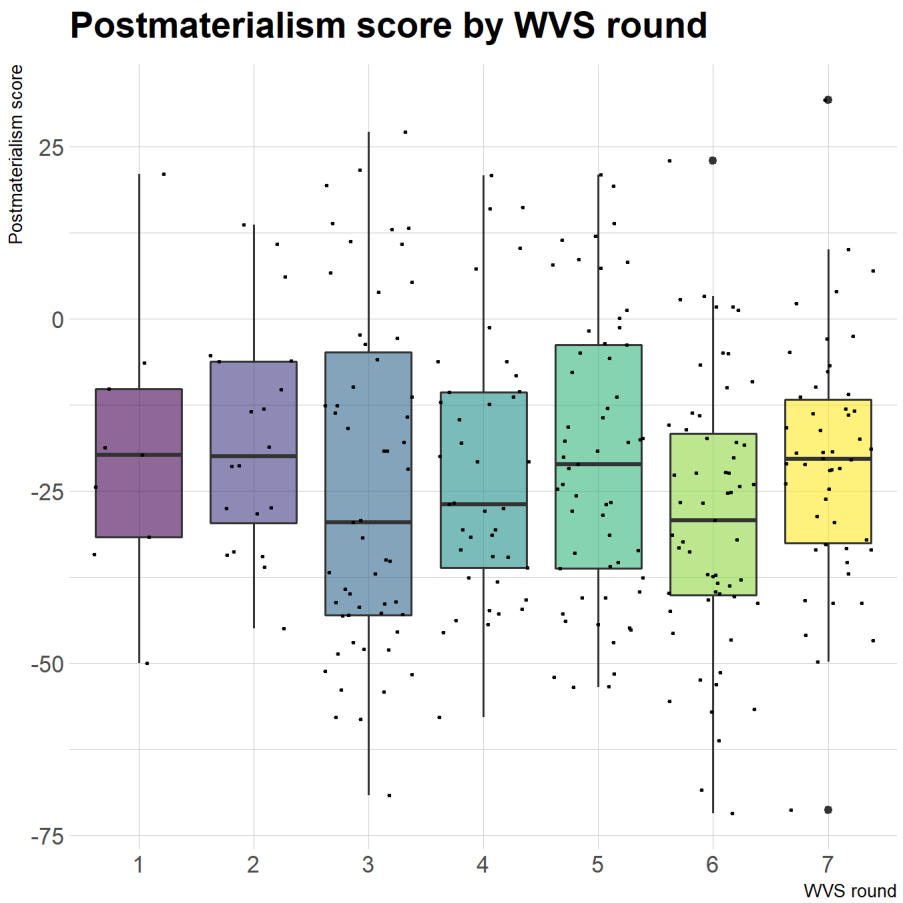
Types of data visualizations with the WVS data


Postmaterialism score by country
(1994-1998)
Postmaterialism score
25
0
-25
-50
China
Azerbaijan
Hungary
Moldova
Russia
Ukraine
Bangladesh
Bulgaria
India
Albania
Montenegro
Serbia
Georgia
Belarus
Taiwan
Romania
Nigeria
Pakistan
Lithuania
South Korea
Armenia
Slovakia
North Macedonia
South Africa
Estonia
Poland
Latvia
Philippines
Bosnia & Herzegovina
Czechia
Brazil
Japan
Venezuela
Peru
Croatia
Turkey
Chile
Spain
Slovenia
Mexico
Dominican Republic
Norway
Switzerland
Uruguay
United Kingdom
Puerto Rico
Sweden
United States
New Zealand
Argentina
Finland
Germany
Australia
El Salvador
Country

Comparisons: Rotate the damn plot!




Postmaterialism score by country
(1994-1998)
Country
El Salvador
Australia
Germany
Finland
Argentina
New Zealand
United States
Sweden
Puerto Rico
United Kingdom
Uruguay
Switzerland
Norway
Dominican Republic
Mexico
Slovenia
Spain
Chile
Turkey
Croatia
Peru
Venezuela
Japan
Brazil
Czechia
Bosnia & Herzegovina
Philippines
Latvia
Poland
Estonia
South Africa
North Macedonia
Slovakia
Armenia
South Korea
Lithuania
Pakistan
Nigeria
Romania
Taiwan
Belarus
Georgia
Serbia
Montenegro
Albania
India
Bulgaria
Bangladesh
Ukraine
Russia
Moldova
Hungary
Azerbaijan
China
-50
-25
0
25
Postmaterialism score
Score
25
0
-25
-50


Postmaterialism score by WVS round
Postmaterialism score
25
0
-25
-50
-75
1
2
3
4
5
6
7
WVS round

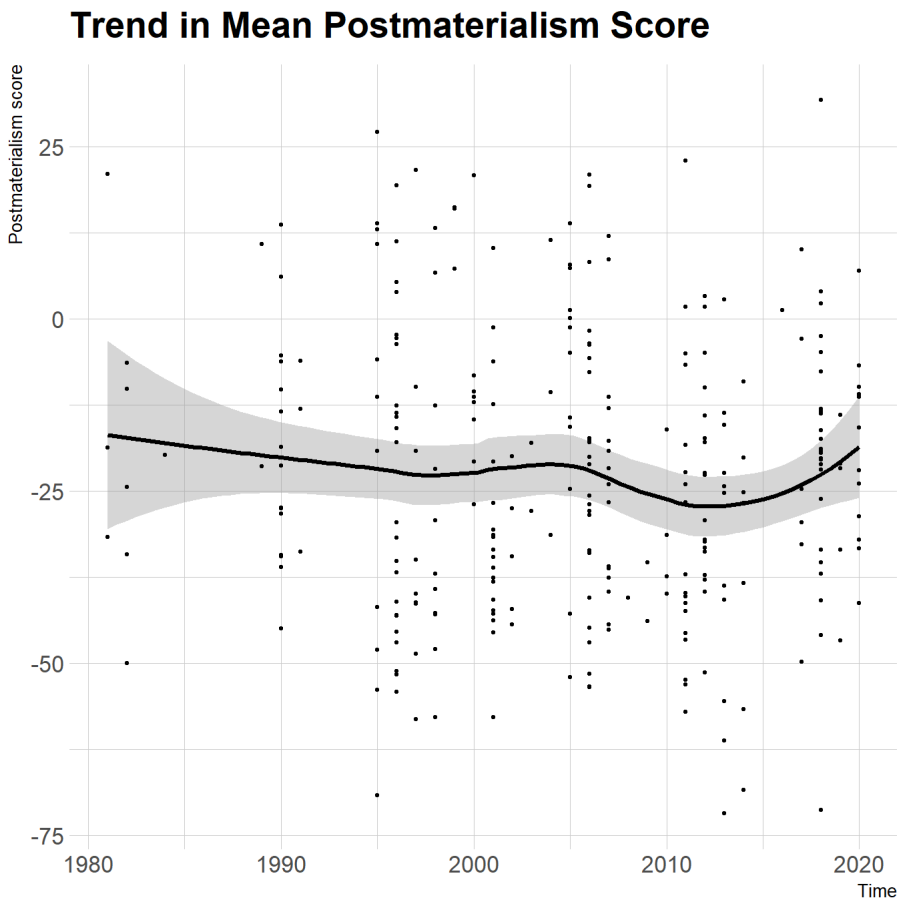
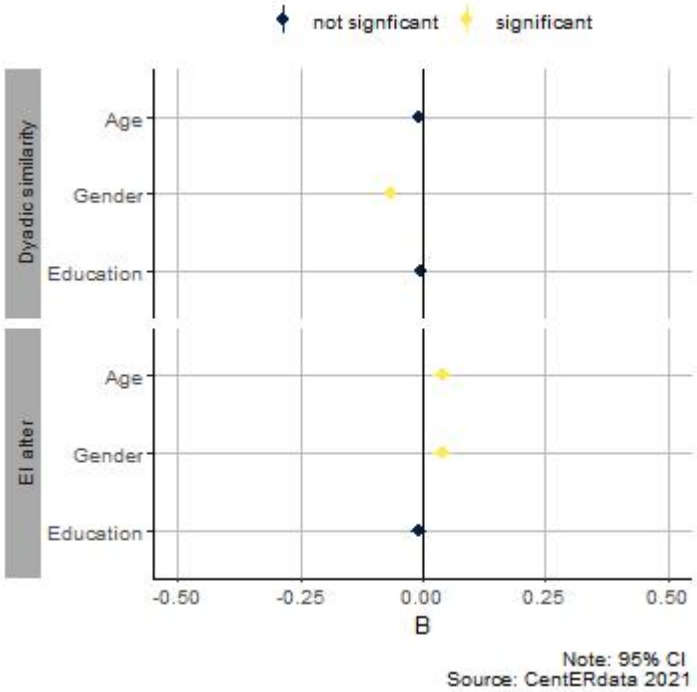
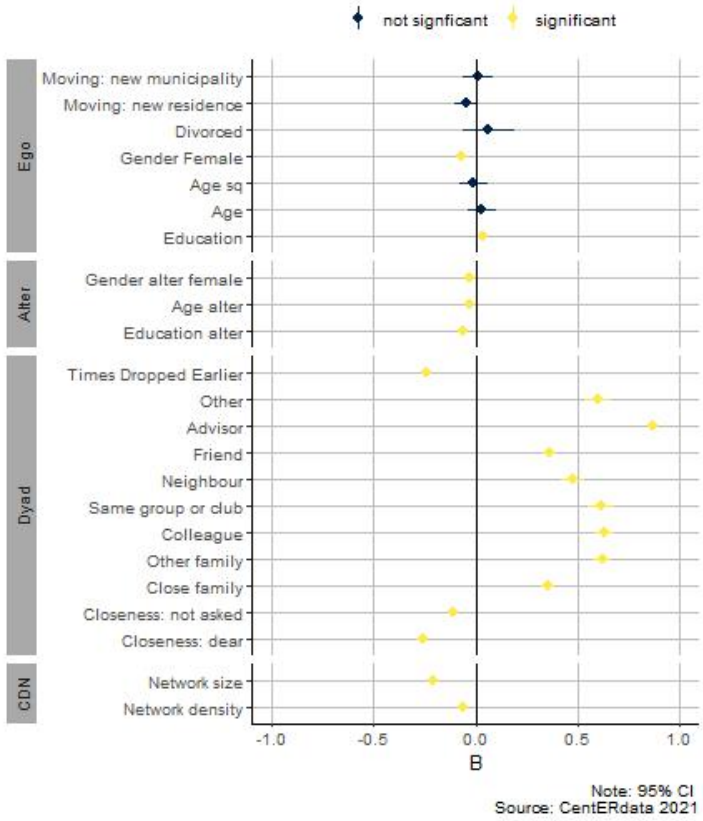

Trend in Mean Postmaterialism Score
Postmaterialism score
25
0
-25
-50
-75
1980
1990
2000
2010
2020
Time

Table 2 Binomial logistic regression on political social media use. *Source*: LISS core study (CentERdata 2018) and Work and Politics (Lehr 2016)

	M1: Political news on social media				M2: Follows politician				M3: Reacted to political issue on social media			
	Simple		Multivariate		Simple		Multivariate		Simple		Multivariate	
	B	SE	*B*	SE	*B*	SE	*B*	SE	*B*	SE	*B*	SE
Independent variables												
Postmaterial attitudes	−0.036	(0.061)	0.034	0.069	0.168	(0.104)	0.049	0.111	0.279**	(0.093)	0.247*	0.099
Populist attitudes	−0.252*	(0.109)	0.055	0.125	−0.019	(0.184)	0.280	0.207	0.210	(0.163)	0.464**	0.180
Control variables												
Age			−0.049***	0.005			−0.013	0.008			−0.016*	0.007
Female			0.059	0.132			−0.696**	0.215			−0.353[a]	0.188
Gross personal income			0.003	0.004			−0.004	0.007			−0.010[a]	0.006
Employment status												
Employed			Ref.				Ref.				Ref.	
Self-employed			0.048	0.256			0.214	0.372			0.496	0.327
Unemployed			0.807**	0.257			0.379	0.394			0.477	0.340
Not active			0.582***	0.176			−0.149	0.301			0.121	0.254
Education												
Primary			0.185	0.383			0.121	0.807			0.542	0.511
Lower secondary			Ref.				Ref.				Ref.	
Upper secondary			0.358	0.237			0.606	0.462			0.159	0.362
Vocational			0.194	0.196			0.585	0.399			0.273	0.301
First-stage tertiary			0.506*	0.205			1.083**	0.399			0.667*	0.307
Second-stage tertiary			0.481*	0.244			1.713***	0.423			0.970**	0.344
Constant			−0.147	0.404			−3.043***	0.699			−3.015***	0.592
Null deviance			1752.000				800.730				954.500	

Presenting statistical models





Data types

Of course there are many different chart types out there!

Take a look at:

- https://datavizproject.com/
- https://datavizcatalogue.com/

Explore different options and play around with the data.

But remember: keep your story and audience in mind!

Important tips

1. Readable labels
2. Data to ink ratio
3. Colour blindness
4. Careful use of colours
5. Always visualize error and uncertainty

Now it's your turn

Step 1: **Think!** Formulate a clear story for your blog and think quickly about visual aids you can use. (figure or a visual)

Step 2: **Explore!** Get inspired by charts and visuals that other authors have used:

Look at the data page of the NYT for inspiration and also the SDG atlas from the worldbank.

Step 3: **Sketch!** Get out your notepads and sketch different chart types for the same information.

Step 4: **Create!** Use a software program to create the chart you decided. All the information you could possibly need is on the internet.

Now it's your turn

All code used for this presentation are on my github.

Presentation made in R with the Xaringan package.