Quantitative and Computational Methods in the Social Sciences

Week 4: Research Methods

Why study quantification?

… a recent twitter brawl

one tweet…

Felix Chopra
@FelixChopra

···

1/ Qualitative interviews offer unparalleled richness but are rarely used in economics. Let's change that! New WP with @Ingar30 uses an AI-driven approach to conducting q qualitative interviews, making them scalable, cheap, and ripe for both qualitative and quantitative analysis!

Traduzir post

Conducting Qualitative Interviews with AI

Felix Chopra   Ingar Haaland

This version: September 15, 2023

Abstract

Qualitative interviews are one of the fundamental tools of empirical social science research and give individuals the opportunity to explain how they understand and interpret the world, allowing researchers to capture detailed and nuanced insights into complex phenomena. However, qualitative interviews are seldom used in economics and other disciplines inclined toward quantitative data analysis, likely due to concerns about limited scalability, high costs, and low generalizability. In this paper, we introduce an AI-assisted method to conduct semi-structured interviews. This approach retains the depth of traditional qualitative research while enabling large-scale, cost-effective data collection suitable for quantitative analysis. We demonstrate the feasibility of this approach through a large-scale data collection to understand the stock market participation puzzle. Our 395 interviews allow for quantitative analysis that we demonstrate yields richer and more robust conclusions compared to qualitative interviews with traditional sample sizes as well as to survey responses to a single open-ended question. We also demonstrate high interviewee satisfaction with the AI-assisted interviews. In fact, a majority of respondents indicate a strict preference for AI-assisted interviews over human-led interviews. Our novel AI-assisted approach bridges the divide between qualitative and quantitative data analysis and substantially lowers the barriers and costs of conducting qualitative interviews at scale.

10:07 PM · 15 de set de 2023 · **1,1 mi** Visualizações

50   489   1.130   1.043

… and different "types" of responses

Heba Gowayed هبة جويد @hebagowayed · 16 de set
Qualitative research, but make it robot.

Abolish Econ.

Bradley 🩷💚 @Be_Radiculous · 16 de set
What do economists think "qualitative" means

Kirstin Munro @kmunro_econ · 17 de set ···
Organ failure couldn't take me out, but this thread going to be the thing that kills me, isn't it

Marco Dehnert @MarcoDehnert · 16 de set ···
Qual research is not interested in becoming scalable, generalizable, or ripe for quantitative analysis. That's a conflation with positivistic research and misunderstands the many ways in which qual research is based on different principles and has different goals than quant…

Aanchal Saraf @aanch__ · 17 de set ···
These are not qualitative interviews, much less semi structured omg. Nor are your conclusions born of qualitative analysis. This is just fancy Qualtrics, ie quantitative surveys that you are analyzing quantitatively! Words mean things!!

💬 1   🔁 1   ♡ 65   ılıl 2.169   ⬆

Alex Kindel @alextkindel · 18 de set ···
The sample selection in the application study makes me worry about ascertainment bias: the research question is about "reasons for stock market non-participation" but only people who are already non-participants are selected. So it's like asking Oscar winners why they won.

💬 2   🔁   ♡ 2   ılıl 494   ⬆

Laura K. Nelson | @LauraNelson@mastodon.social ···
@LauraK_Nelson

I am once again asking sociologists to be curious and open to new methods. We need to stake our ground: we have strong methodological traditions that we are building on and should continue to do so. And surely we can keep improving and updating what we do.
Traduzir post

Alex Kindel @alextkindel · 18 de set ···
The comparison to qualitative methods per se is limited to coding results. I would have liked to see a comparison to interviews conducted by a person.

💬 1   🔁 1   ♡ 4   ılıl 701   ⬆

5:38 PM · 18 de set de 2023 · **10,8 mil** Visualizações

Why study quantification?

...to learn how to do it
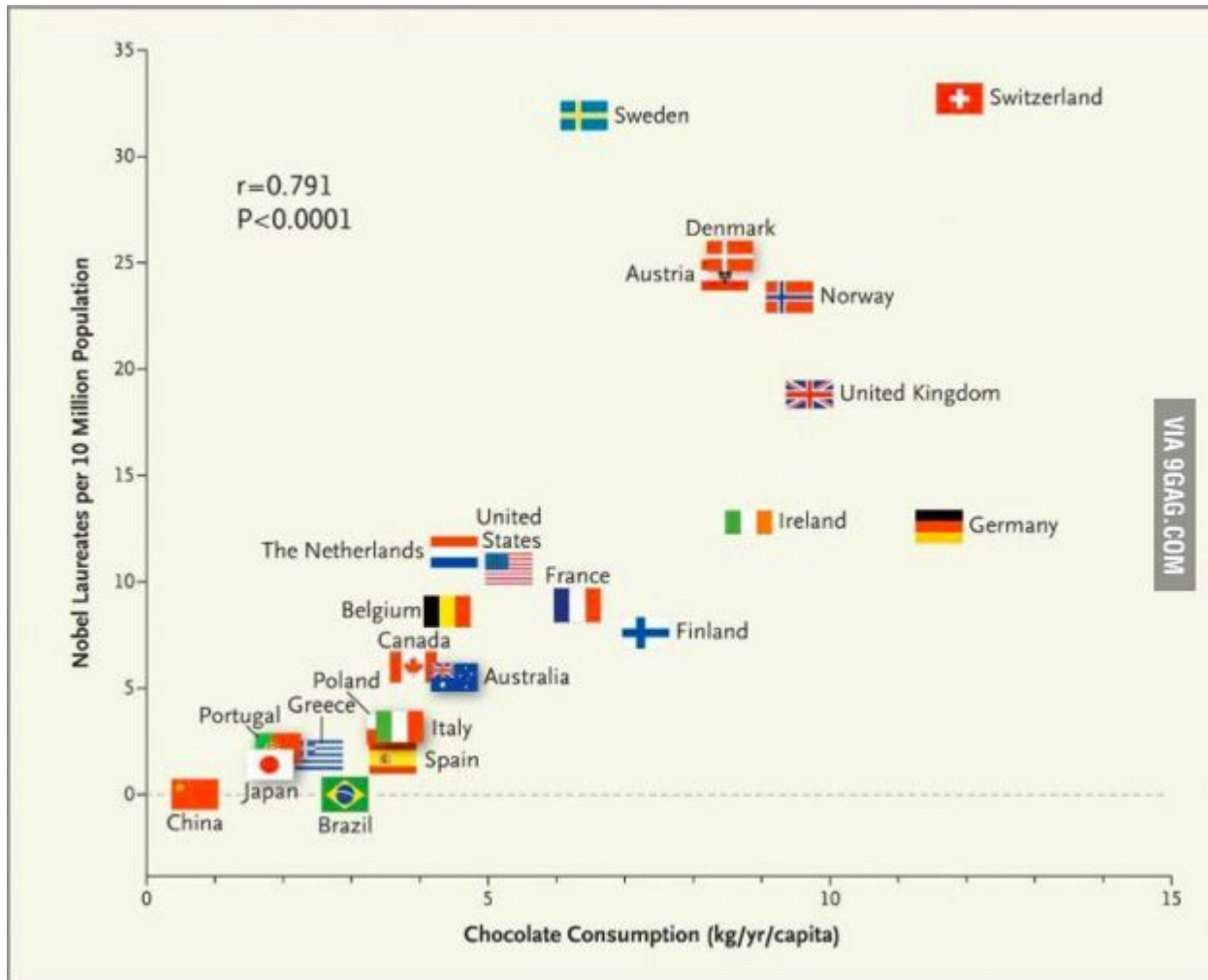
…to dialogue with other types of research

…to analyze it critically or substantially

…. but above all…
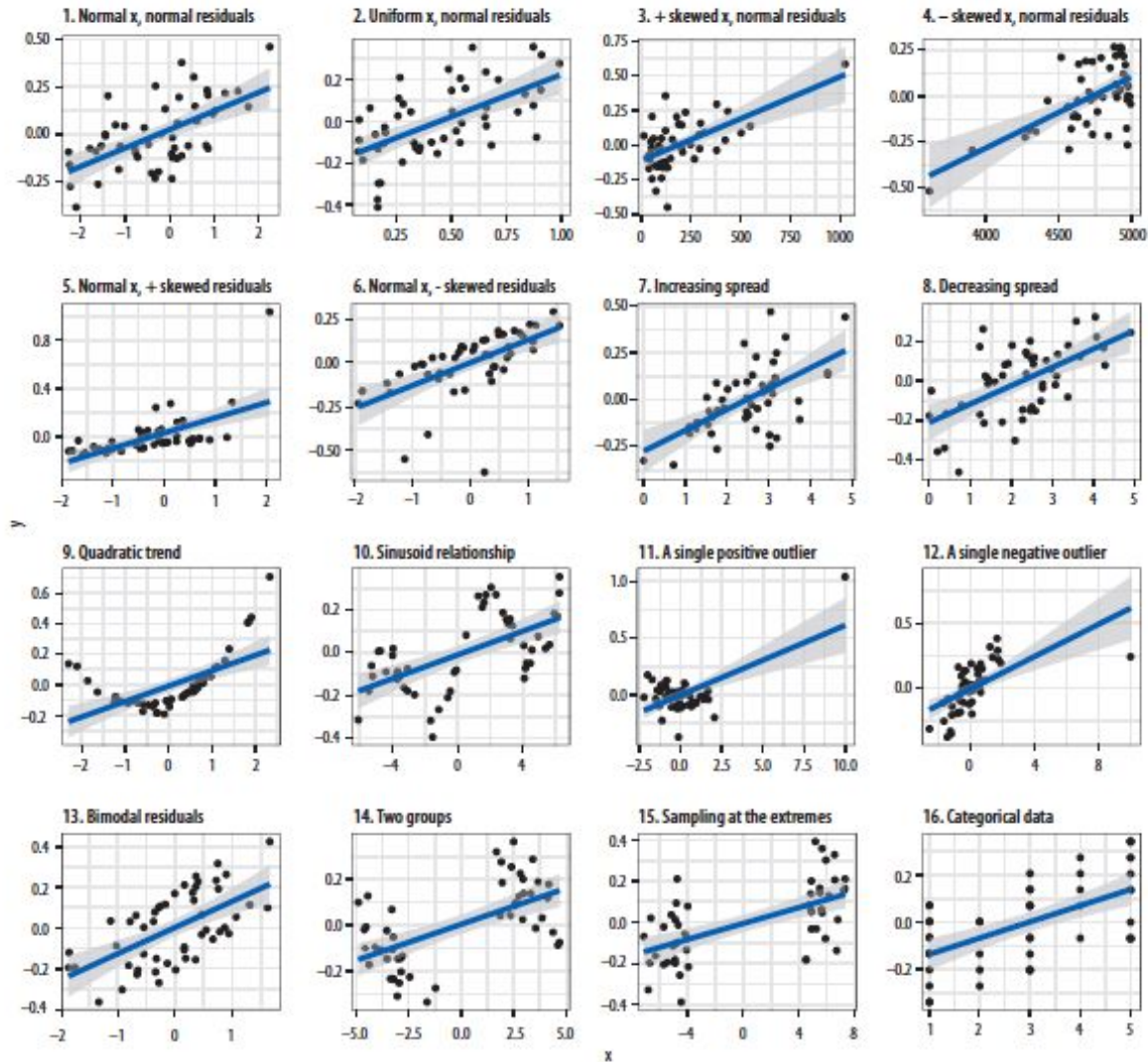
… ANSO folks should know a bit of quantification because of…

NUMERICAL LITERACY!

Correlations….

...next level!

1. Normal x, normal residuals
2. Uniform x, normal residuals
3. + skewed x, normal residuals
4. – skewed x, normal residuals
5. Normal x, + skewed residuals
6. Normal x, - skewed residuals
7. Increasing spread
8. Decreasing spread
9. Quadratic trend
10. Sinusoid relationship
11. A single positive outlier
12. A single negative outlier
13. Bimodal residuals
14. Two groups
15. Sampling at the extremes
16. Categorical data

How to spot bullshit?
(from http://callingbullshit.org/)

- Too good (simple?) to be true
- Multiple working hypothesis
- Think about orders of magnitude
- Beware of unfair comparisons

We start slow…

1. **Traditional Quantitative Methods:** *when do elites support redistribution?*

2. **Computational Social Sciences:** *how has the Amazon been constructed as a problem?*
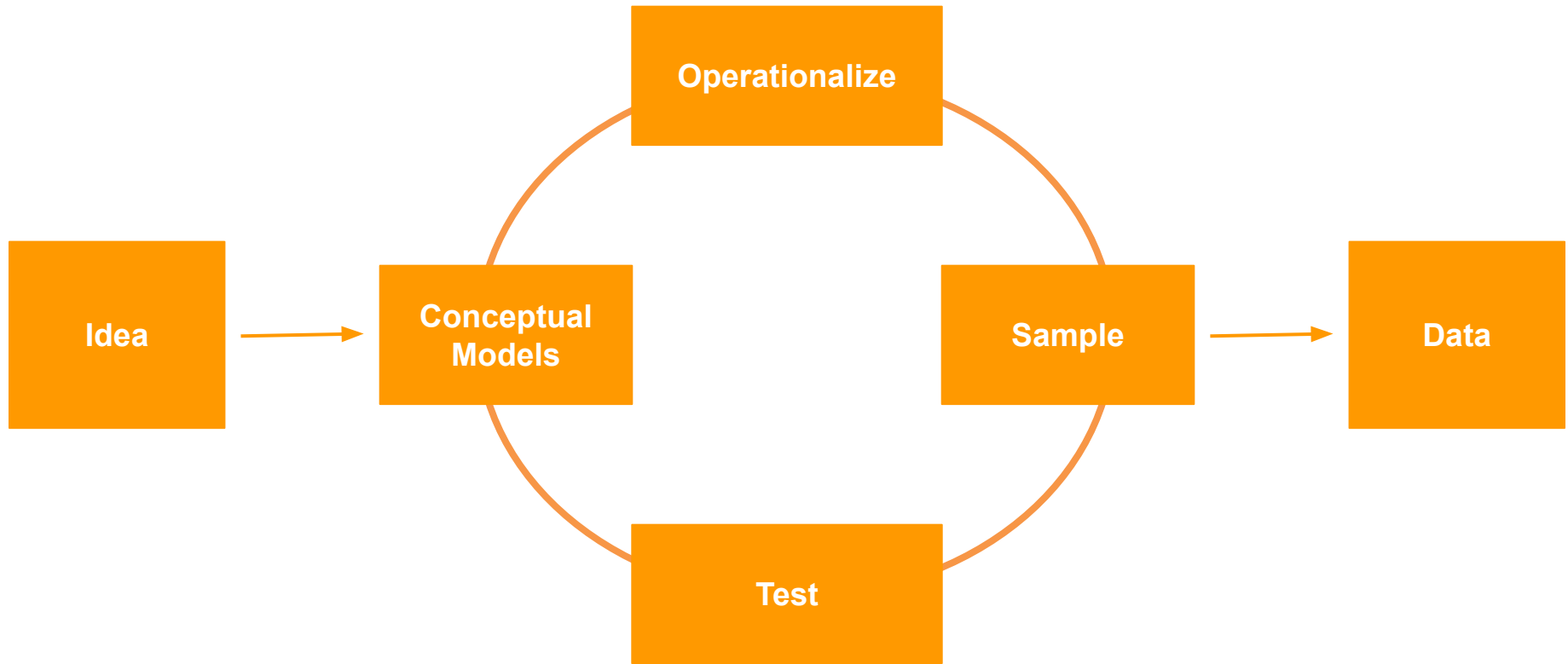
3. Assignment for next week.

From Conceptual to Empirical Model
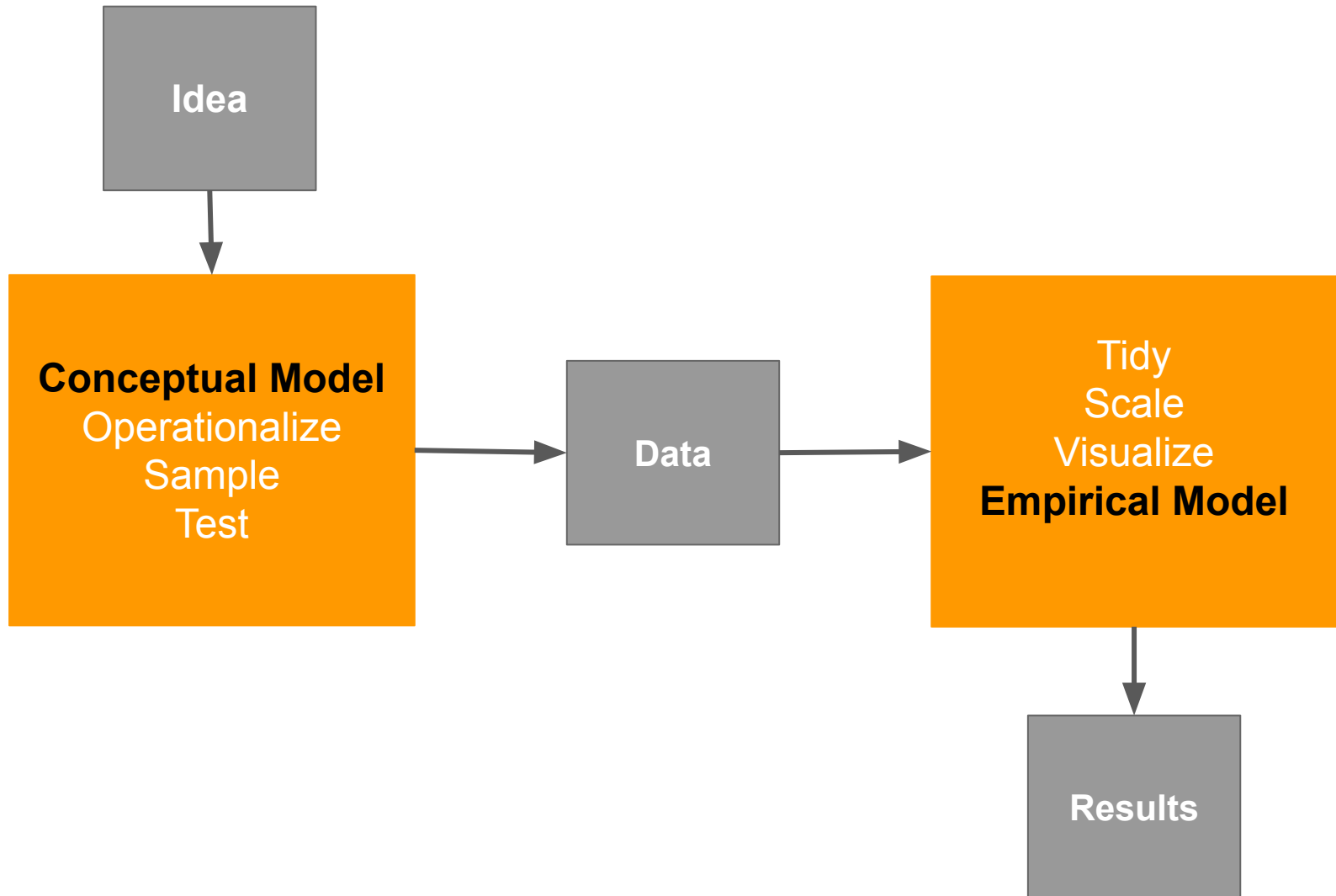
Let's think about elites...

When do elites support redistributive policies?

A survey workflow...

Conceptualizing

Consequences of inequality → **Mediators of support for redistribution** → **Support for (what) redistribution**

- Generally
  - Main consequences of inequality?

- Politics
  - Perception of threat
  - Linked to inequality?

- Health
  - Perception of threat
  - Linked to inequality?

- Violence
  - Perception of threat
  - Linked to inequality?

- Ideologies
  - Left or Right?
  - Neo-liberal or developmentalist?
  - Fatalist or Meritocratic?
  - Colorblind or racialized?

- Knowledge of the poor and poverty
- Trust in institutions
- Trust in others
- Responsibility to redistribute
- Willingness to pay
- Possibilities of opting-out

- At the bottom:
  - cash-transfers
- At the middle:
  - social nets
- At the top:
  - wealth tax
- Categorical:
  - race and gender

Socio-economic background

Sampling I

Elites because they occupy positions of power		
Economic power	Legislative power	Executive power
CEOs, CFOs Super-wealthy	Congressperson Senators	Top career civil servants Ministers

Sampling II

Sector	Legislative		Executive	Judiciary	Economic			Total
	Congress and Senate	State Assemblies	**Top-civil servants**	2nd and 3rd Instance Judges	**CEO, CFO, or President of the board**	Super Rich (Forbes)	Super Donors (TSE)	
Population (N)	594	420	626	602	500	238	687	3667
Contacted	150	90	120	42	150	20	60	632
Estimated Sample (n)	75	45	60	21	75	10	30	316

Priority Groups

Operationalizing I

16. In your opinion, income redistribution policies:

Must be avoided	1
Are necessary	2
*Do not know (**Do not read**)*	*88*
*Do not want to answer (**Do not read**)*	*99*

Operationalizing I

Redistribution Binary Question

In your opinion, income redistribution policies (1) should be avoided, or (2) are necessary?
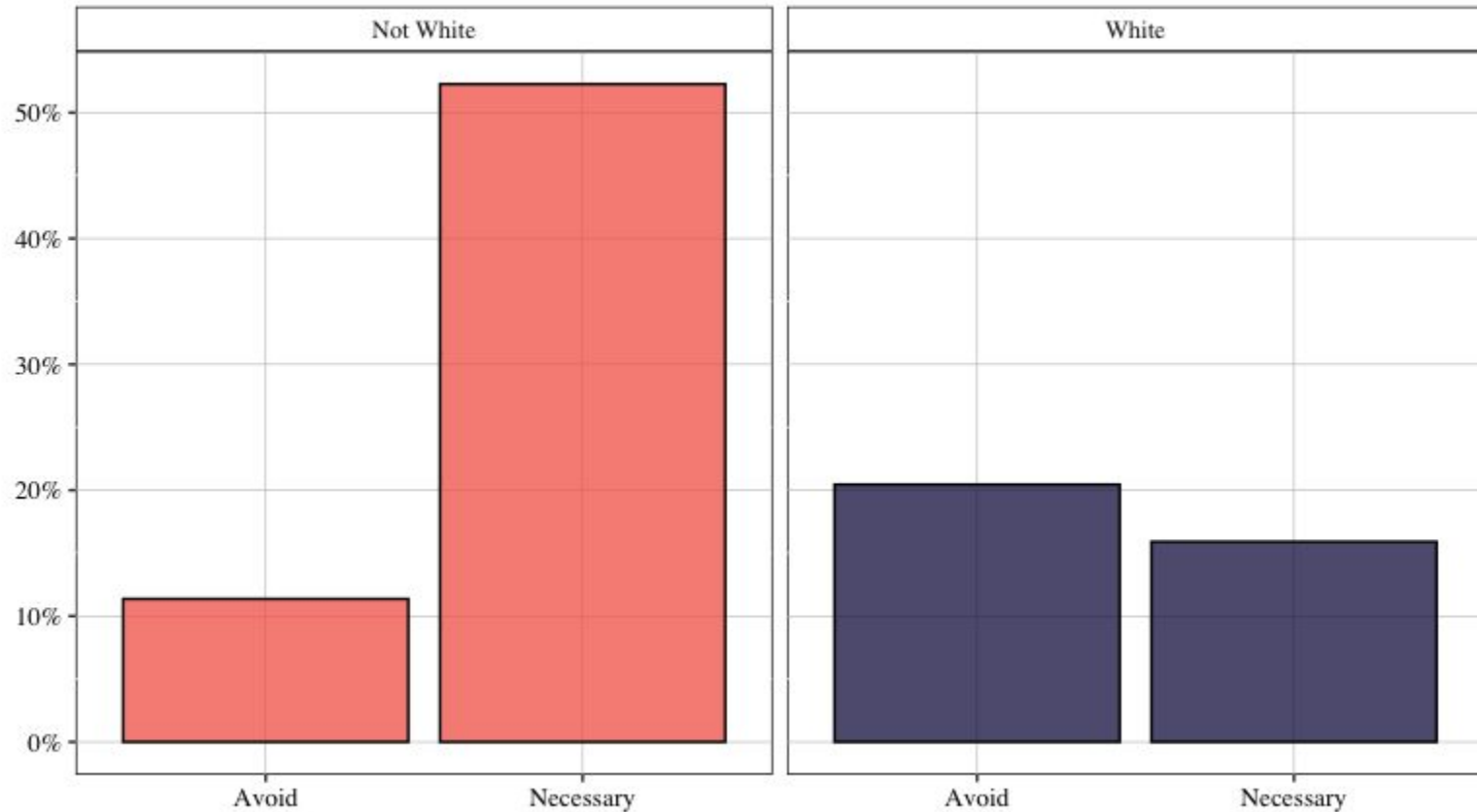
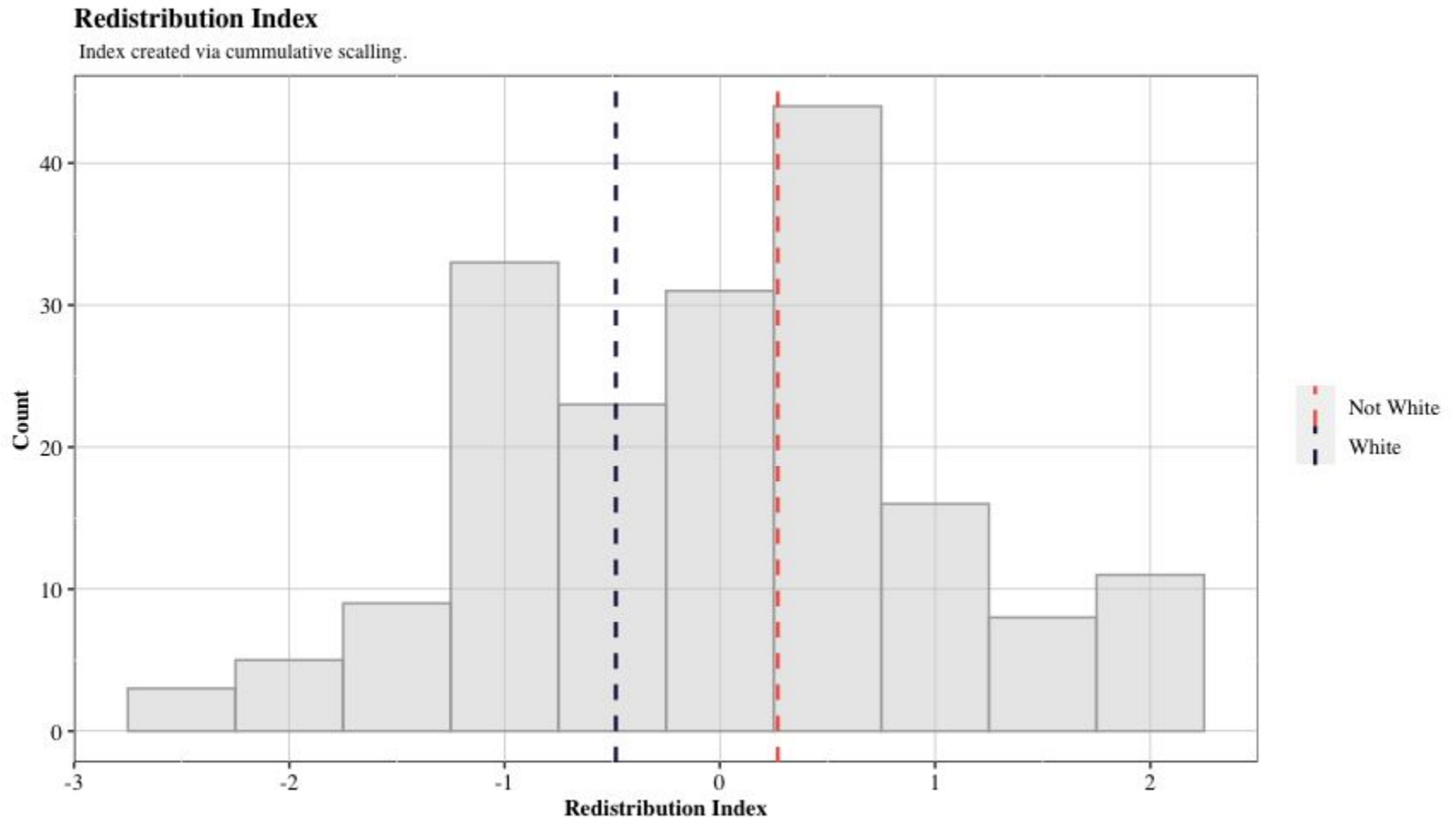Source: Unequal democracies surveys Brazil and South Africa 2022

Operationalizing II

18. Now I am going to mention a series of political policies, and I would like you to indicate if you consider them desirable and also regard them as viable.

1. Yes   88. Do not know (*Do not read*)
2. No   99. Do not want to answer (*Do not read*)

	Is it desirable?				Is it viable?			
	Yes	No	Do not know (*Do not read*)	Do not want to answer (*Do not read*)	Yes	No	Do not know (*Do not read*)	Do not want to answer (*Do not read*)
1 Prioritise disadvantaged student access to higher education	1	2	88	99	1	2	88	99
2 Ensure free universal access to health services	1	2	88	99	1	2	88	99
3 Guarantee unemployment insurance	1	2	88	99	1	2	88	99
4 Improve the quality of primary school education	1	2	88	99	1	2	88	99
5 Maintain preferential admissions policies in universities for previously disadvantaged groups	1	2	88	99	1	2	88	99
6. Maintain preferential hiring policies for previously disadvantaged groups	1	2	88	99	1	2	88	99
7. Continue Black Economic Empowerment programmes	1	2	88	99	1	2	88	99
6 Decrease income disparity between high-level occupations and others	1	2	88	99	1	2	88	99
7 Implement a basic income grant for everyone over 25	1	2	88	99	1	2	88	99
8 Guarantee free university education	1	2	88	99	1	2	88	99
9 Maintain a food distribution programme for the poor	1	2	88	99	1	2	88	99
10 Guarantee shelter through government housing	1	2	88	99	1	2	88	99
11. Deal with land claims issues	1	2	88	99	1	2	88	99

Operationalizing II

Redistribution Index
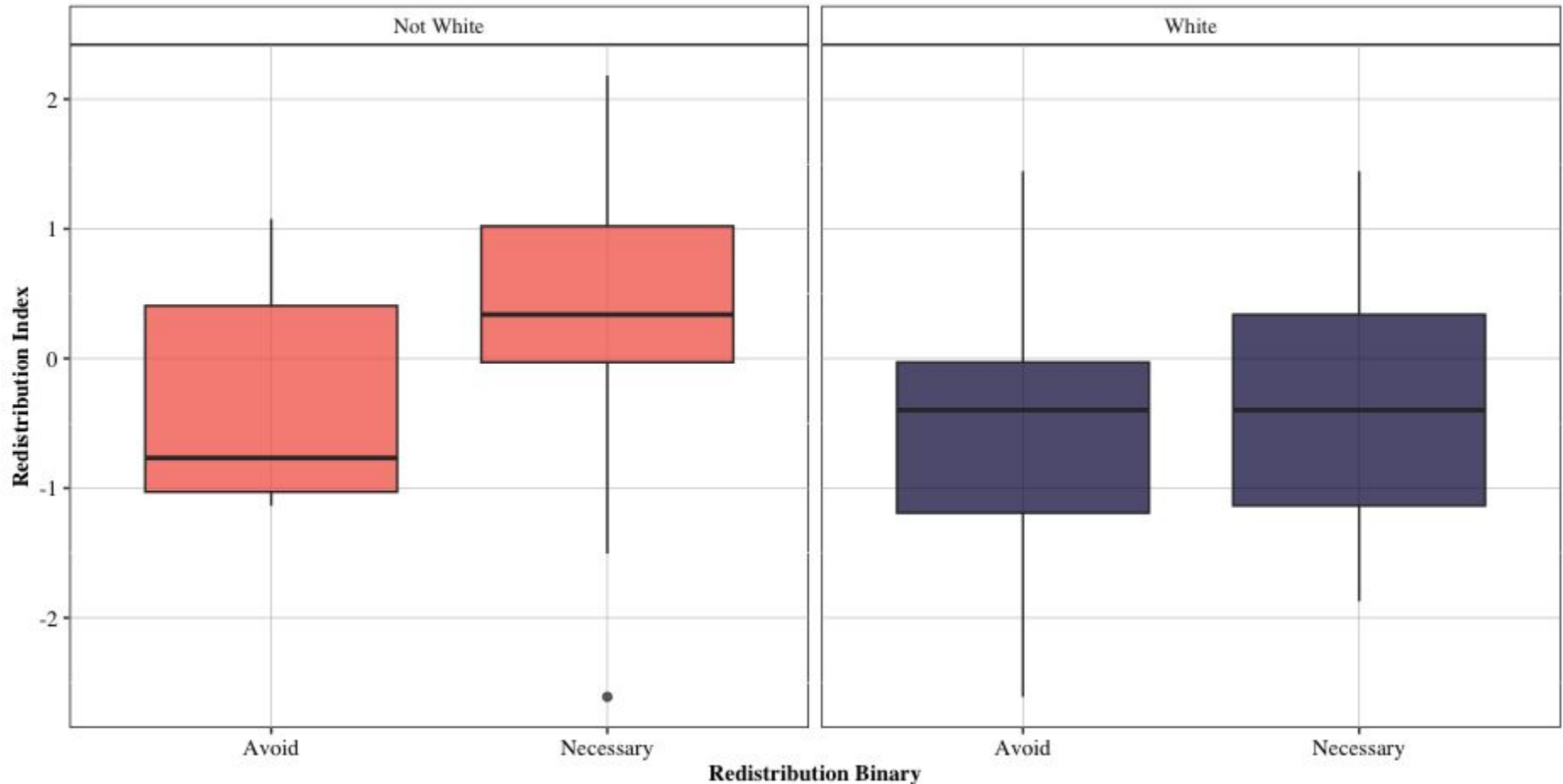
Index created via cummulative scalling.

Source: Unequal democracies surveys Brazil and South Africa 2022, n=180.

Operationalizing II

Is the redistribution binary a good measure?

Sample composed of 180 elites

From quantification to qualification

Qualitative environment → Support for Redistribution → Quantitative environment

Go hang out with parliamentarians	What's your take on the South African Social Grant?	Are viable and/or desirable?	Should be avoided **versus** are necessary
You produce a thick description	You produce different narratives	You produce different levels of support	You produce different categories: either support or not

... (between first and second column) ... (between second and third column)

Ethnography   Open-ended Interviews   Surveys

Behaviour   Perceptions

Empirical Environment

Testing

From Conceptual to Empirical Model

data…

A survey workflow (cont.)

Modelling I

Consequences of inequality → **Mediators of support** → **Support for (what) redistribution**

Modelling II

Table 4. Effect of perceptions of the poor on elite support for redistribution in Brazil

	Model 1		Model 2		Model 3		Model 4	
	Coef	SE	Coef	SE	Coef	SE	Coef	SE
Poor are irrational	−0.191***	0.035	−0.192***	0.035	−0.129***	0.034	−0.134***	0.036
Externalities			.018	0.037	−0.002	0.033	0.014	0.034
Business elite					−0.152***	0.037	−0.140***	0.042
Political elite					0.076*	0.040	0.081*	0.044
Female							−0.108	0.079
Father education							0.001	0.012
Mother education							−0.003	0.014
Race							−0.055	0.042
Age							−0.002	0.002
Ideology							−0.053	0.036
Perception of self as responsible							−0.051	0.048
Willingness to pay more taxes							0.052^	0.032
Intercept	0.741***	0.028	0.729***	0.038	0.732***	0.043	0.872***	0.117
R^2	0.130		0.170		0.341		0.419	
N	147		147		147		135	

^$P < 0.11$; *$P < 0.10$; **$P < 0.05$; ***$P < 0.01$.

Computational Social Sciences!

(un)supervised machine learning
web-scraping techniques
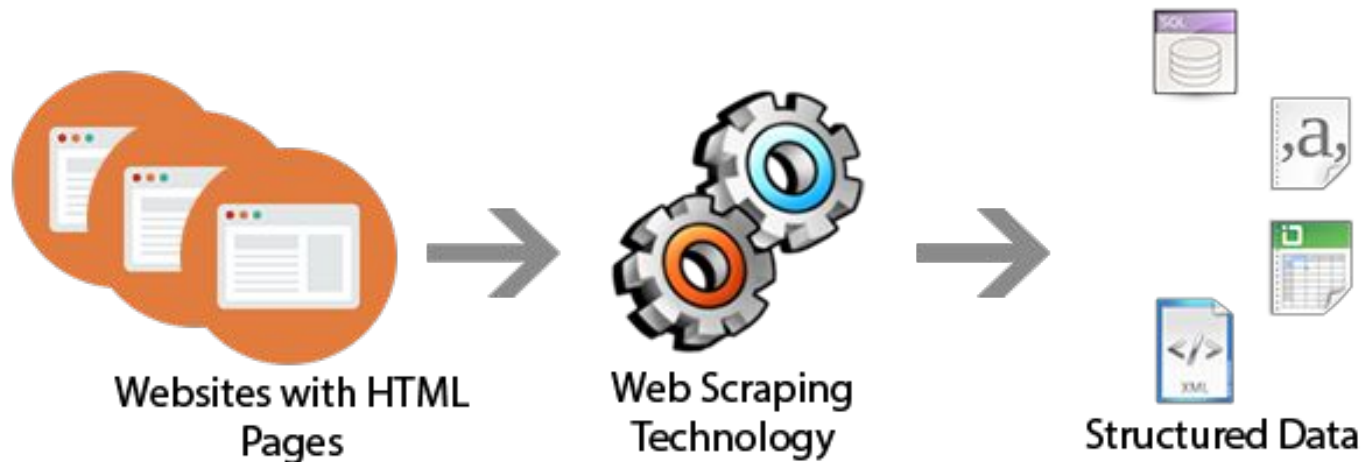large language models
causal inference
text-as-data
python
R
…

what are and how can we think about 'newer' quantification tools?
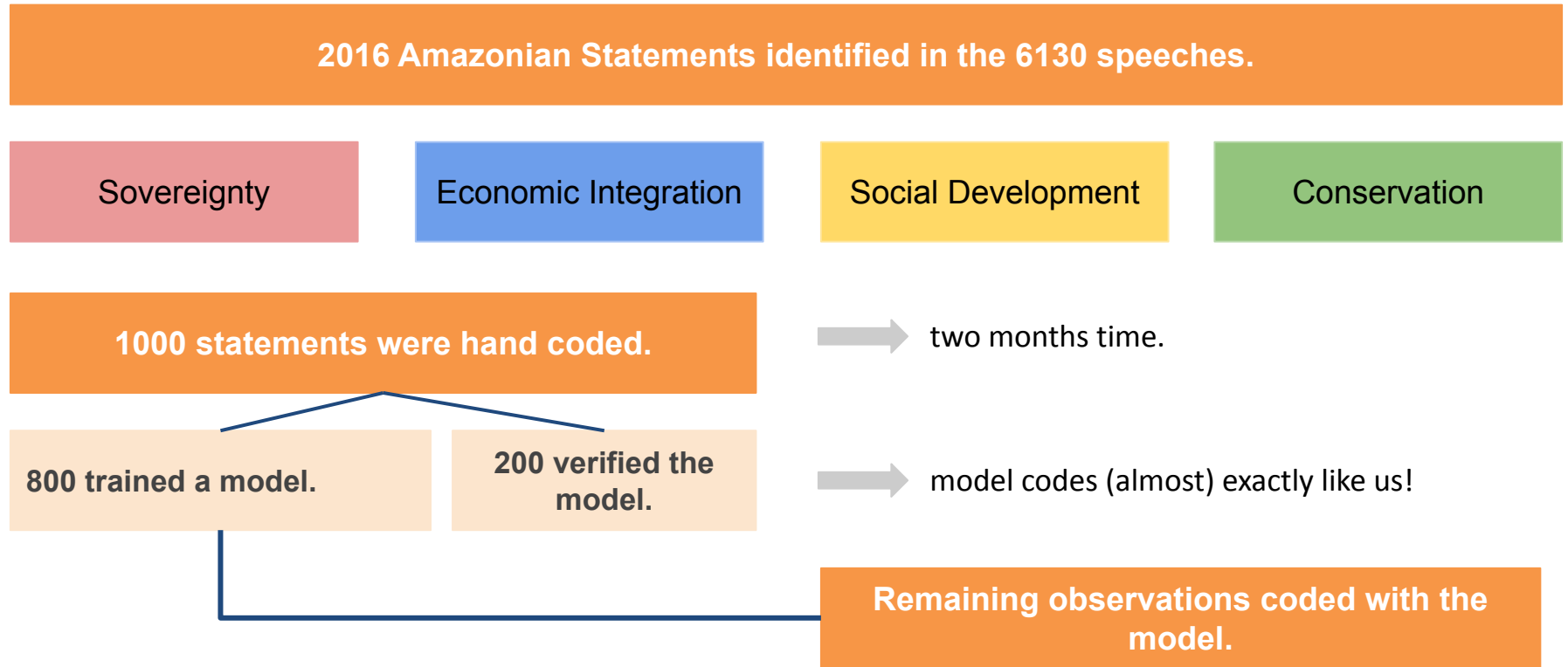
Let's think about discourse…

How has the Brazilian Amazon been constructed as a problem?

Web scraping: building a dataset with all presidential speeches since 1985

When we scrappe, we often get more than we expected…

Text-as-data:supervised learning

2016 Amazonian Statements identified in the 6130 speeches.

Sovereignty	Economic Integration	Social Development	Conservation

1000 statements were hand coded. ⟶ two months time.

800 trained a model.   **200 verified the model.** ⟶ model codes (almost) exactly like us!

**Remaining observations coded with the model.**

When does such an approach make sense?

…when codebook is limited and defined.

Simple description is already a contribution…

Share of pure-types problem-constructions by year

Sample composed by 1924 Amazonian statements in presidential speeches since 1985

Economic Integration

Environmental Conservation

Social Development

National Sovereignty
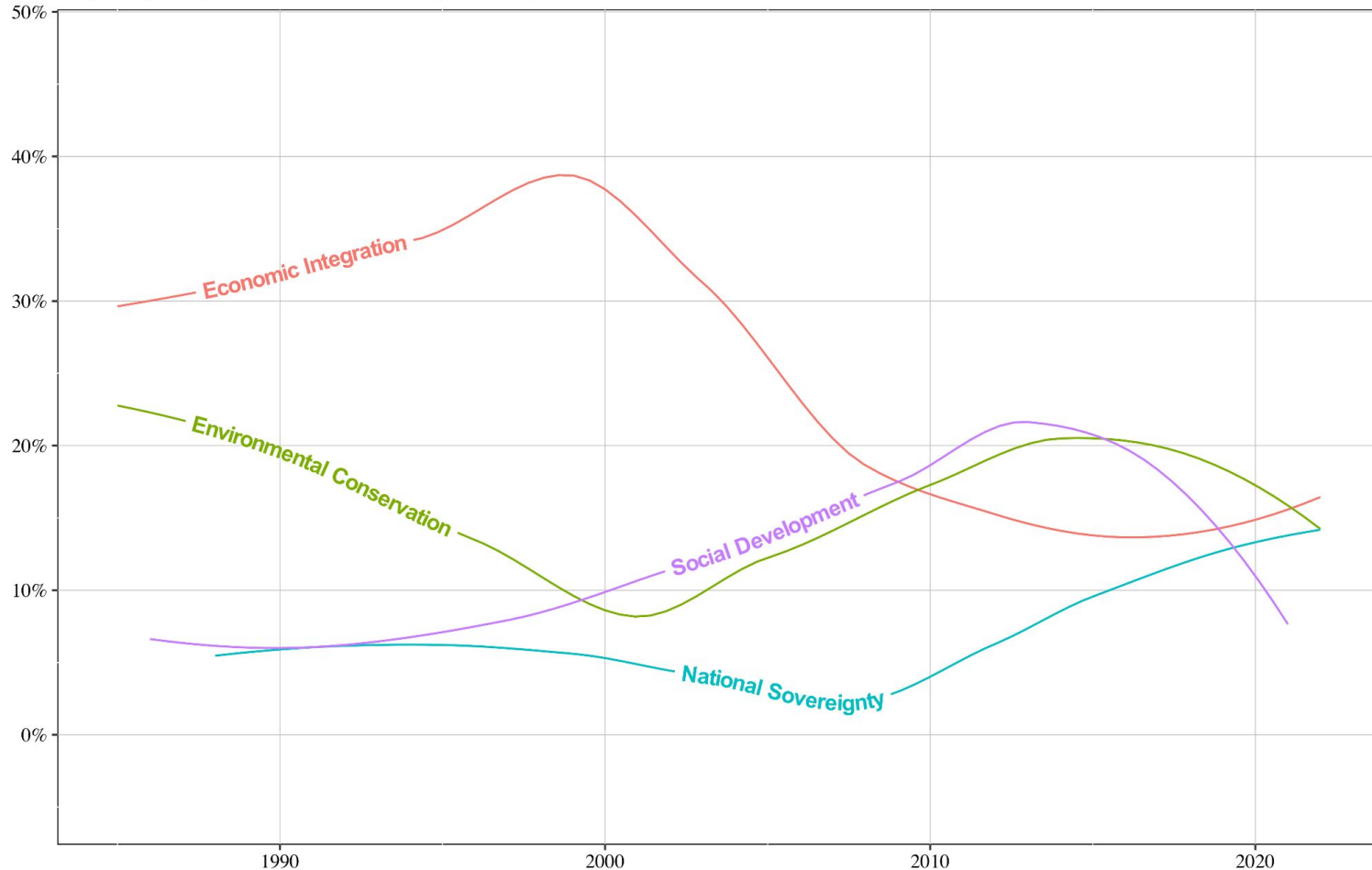
Curves in the plot were smoothed using loess method.

Modelling III

Table 1. Log odds for fixed-effects logistic regressions indexed by president.

	Amazon Speech (1)	Environmental Conservation (2)	Economic Integration (3)	Social Development (4)	National Sovereignty (5)
Amazonian Countries	−0.167*** (0.033)	0.012 (0.032)	0.091** (0.039)	−0.097*** (0.028)	−0.004 (0.022)
Non- Amazonian States	−0.403*** (0.025)	0.079*** (0.018)	−0.039* (0.022)	−0.013 (0.016)	0.028** (0.012)
Non-Amazonian Countries	−0.430*** (0.026)	0.202*** (0.031)	−0.056 (0.037)	−0.057** (0.027)	−0.004 (0.021)
Deforestation	−0.075*** (0.025)	−0.098** (0.047)	0.303*** (0.056)	−0.045 (0.041)	−0.030 (0.031)
Inflation	0.232*** (0.049)	0.164** (0.071)	−0.291*** (0.086)	0.072 (0.063)	0.102** (0.048)
Election Year	−0.010 (0.012)	0.041* (0.022)	0.011 (0.026)	0.008 (0.019)	−0.016 (0.014)
Observations	6,167	1,924	1,924	1,924	1,924
R^2	0.066	0.035	0.031	0.010	0.007
Adjusted R^2	0.064	0.028	0.025	0.004	0.0001
F Statistic	72.132*** (df = 6; 6153)	11.499*** (df = 6; 1910)	10.249*** (df = 6; 1910)	3.320*** (df = 6; 1910)	2.197*** (df = 6; 1910)

Note: *$p<0.1$; **$p<0.05$; ***$p<0.01$.

Computational Social Sciences
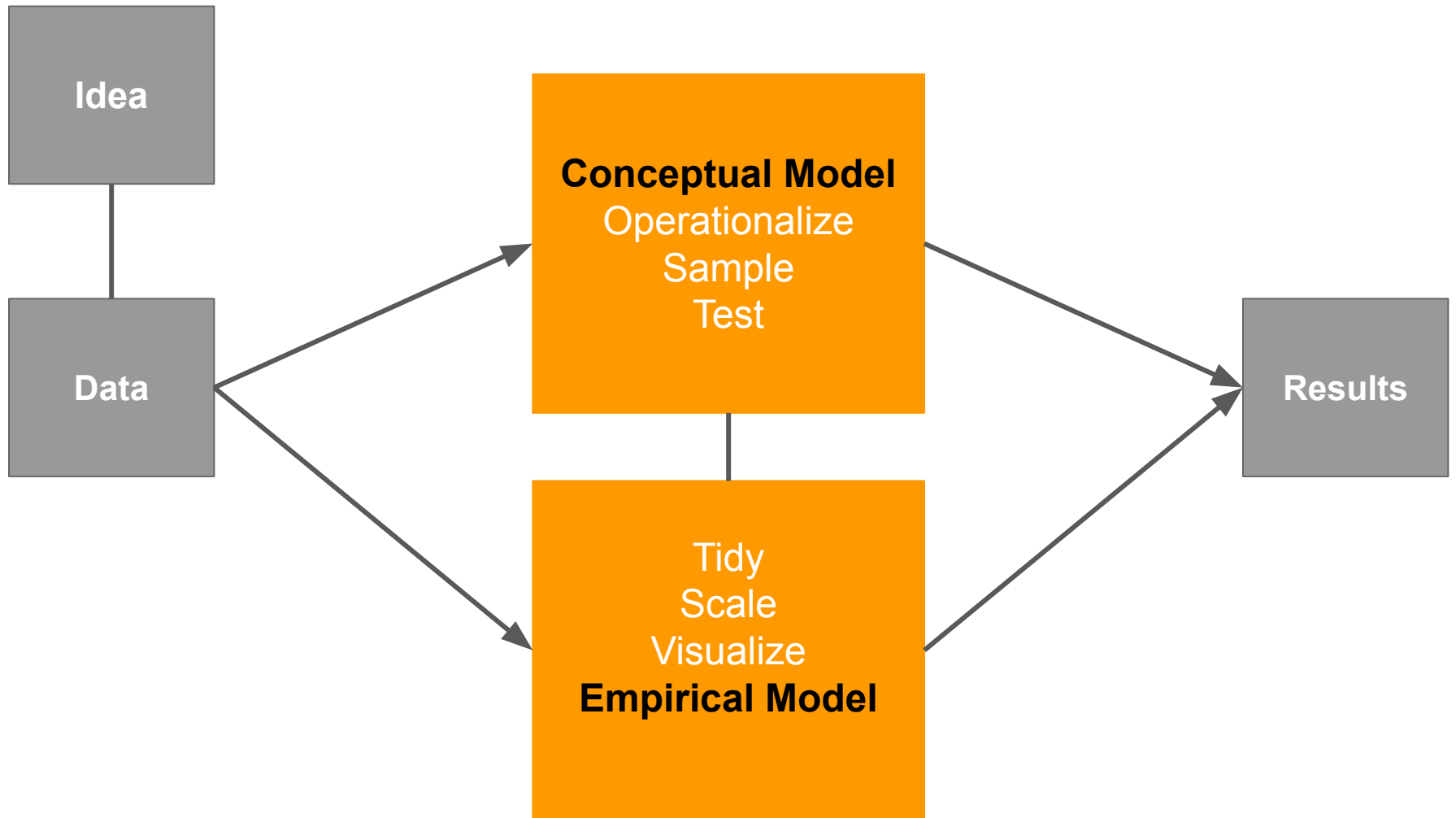
'newer' quantification tools…

…rearrange the logics of induction, deduction and abduction.

Qualitative? Quantitative? Computational?

Laura K. Nelson | @LauraNelson@mastodon.social
@LauraK_Nelson

This is the second job I've seen at the intersection of qual and comp methods 🔥❤️

"Bridging the gap between [qual] and [quant] social research is more important than ever, as is addressing new ethical challenges in research and teaching in the age of artificial intelligence."

Traduzir post

> **Marc Keuschnigg** @MarcKeuschnigg · 18 de set
> Leipzig sociology is looking for a postdoc at the intersection of qualitative methods and computational social science
> uni–leipzig.de/en/university/…
> Fans of @LauraK_Nelson, @achimedelmann, @BerlinerLufti, @simon_lindgren, @eollion etc, please get in touch with me. Deadline Oct 15

9:32 PM · 18 de set de 2023 · **5.165** Visualizações

💬 1   🔁 13   ♡ 36   🔖 2   ↥

Poste sua resposta!   **Responder**

Mario L. Small @MarioLuisSmall · 20 de set
Fantastic. Our research group at @UBremen is focused on similar sets of issues

💬   🔁 1   ♡ 1   📊 374   ↥

Computational Grounded Theory: A Methodological Framework

Laura K. Nelson (iD) ✉ View all authors and affiliations

The Extended Computational Case Method: A Framework for Research Design

Juan Pablo Pardo-Guerra (iD) ✉ and Prithviraj Pahwa (iD) View all authors and affiliations

The promises of computational ethnography: Improving transparency, replicability, and validity for realist approaches to ethnographic analysis

Corey M. Abramson ✉, Jacqueline Joslyn, [...], and Daniel Dohan (+2) View all authors and affiliations

…'newer' quantification tools increase (change?) the range and types of questions we can answer.

Assignment

You will form four groups and each group will read one of the following papers. Next week, you will seat within your group and discuss the questions in the next slide. By <mark>date</mark>, you should submit a 1200 words reading response.

- Andrew M. Guess et al. , **How do social media feed algorithms affect attitudes and behavior in an election campaign?**. Science 381,398-404(2023). DOI:10.1126/science.abp9364

- Matias Spektor, Umberto Mignozzetti, Guilherme N. Fasolin (2022); **Nationalist Backlash Against Foreign Climate Shaming.** Global Environmental Politics; 22 (1): 139–158. doi: https://doi.org/10.1162/glep_a_00644

- Paxton, P., Velasco, K., & Ressler, R. W. (2020). **Does Use of Emotion Increase Donations and Volunteers for Nonprofits?** American Sociological Review, 85(6), 1051–1083. https://doi.org/10.1177/0003122420960104

- Jorgenson, A. K., Clark, B., Thombs, R. P., Kentor, J., Givens, J. E., Huang, X., El Tinay, H., Auerbach, D., & Mahutga, M. C. (2023). **Guns versus Climate: How Militarization Amplifies the Effect of Economic Growth on Carbon Emissions.** American Sociological Review, 88(3), 418-453. https://doi.org/10.1177/00031224231169790

<mark>**You'll receive a google sheet, please fill it in by tomorrow 14:00 and we will answer with your group.**</mark>

Preparing your reading response

- What is quantified in this paper? How is the sampling and operationalization done?

- What type of data does the paper relies on?

- What type of statistical analysis does the paper relies on (descriptive, inferential, etc)? Are there any causal claims?

- What types of conclusions are taken from the statistical analysis? Do you buy it?

What next?

- this is crazy **(in a bad way)**… how do I do the exercise?

- this is crazy **(in a good way)**… how do I learn more?

What next?

- this is crazy (in a bad way)… **how do I do the exercise?**
  1. Bring your questions next week, we are here to help.
  2. Check support material and the internet.
  3. Drop me an e-mail, we can find time to discuss it.

- this is crazy (in a good way)… **how do I learn more?**
  – MINT338: Fundamentals of R (with myself & Henrique)
  – MINT411: Intermediate R (with myself & Henrique)
  – MINT009: Advanced Quant Methods (with Prof. Vandewalle)
  – IRPS093: Networks (with Prof. Hollway)
  – MINT388: Data and Text Analysis with Python (w/ Prof. Noailly)
  – Civica Data Science Online School
  – Journal Sociological Science
  – And you can come see us for books, references, ways forwards…