

# Logic 2: Controversies about the quantitative approach

## LQRPS

Frederik Hjorth  
fh@ifs.ku.dk  
fghjorth.github.io  
@fghjorth

Department of Political Science  
University of Copenhagen

February 6<sup>th</sup>, 2017

## 1 Motivating example 1

## 2 The DA-RT movement

## 3 The counter-DA-RT movement

## 4 Motivating example 2

## 5 Introduction to R

- 1 Motivating example 1
- 2 The DA-RT movement
- 3 The counter-DA-RT movement
- 4 Motivating example 2
- 5 Introduction to R

- 1 Motivating example 1
- 2 The DA-RT movement
- 3 The counter-DA-RT movement
- 4 Motivating example 2
- 5 Introduction to R

- 1 Motivating example 1
- 2 The DA-RT movement
- 3 The counter-DA-RT movement
- 4 Motivating example 2
- 5 Introduction to R

- 1 Motivating example 1
- 2 The DA-RT movement
- 3 The counter-DA-RT movement
- 4 Motivating example 2
- 5 Introduction to R

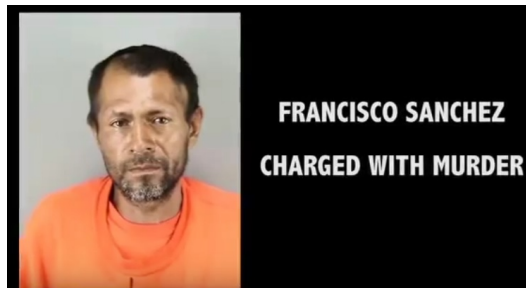
- 1 Motivating example 1
- 2 The DA-RT movement
- 3 The counter-DA-RT movement
- 4 Motivating example 2
- 5 Introduction to R



**Willie Horton**

<https://youtu.be/Io9KMSSEZ0Y>

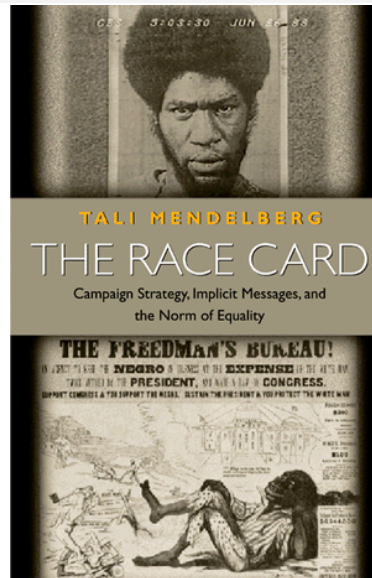




<https://youtu.be/yD4wnQvLhyI>

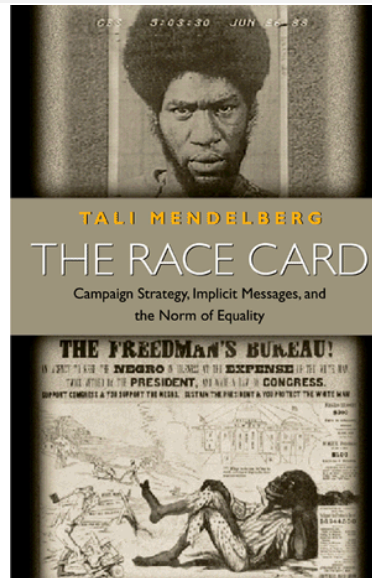
»The most important and underplayed lesson of the Horton message is that (...) [w]hen a society has repudiated racism, yet racial conflict persists, candidates can win by playing the race card only through implicit racial appeals. The implicit nature of these appeals allows them to prime racial stereotypes, fears, and resentments while appearing not to do so. When an implicit appeal is rendered explicit—when other elites bring the racial meaning of the appeal to voters' attention—it appears to violate the norm of racial equality. It then loses its ability to prime white voters' racial predispositions.«

→ the “IE model”



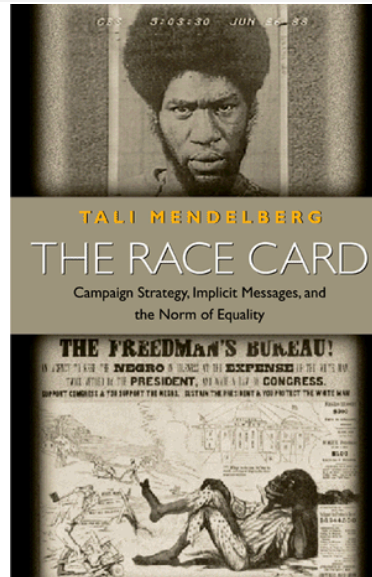
»The most important and underplayed lesson of the Horton message is that (...) [w]hen a society has repudiated racism, yet racial conflict persists, candidates can win by playing the race card only through implicit racial appeals. The implicit nature of these appeals allows them to prime racial stereotypes, fears, and resentments while appearing not to do so. When an implicit appeal is rendered explicit—when other elites bring the racial meaning of the appeal to voters' attention—it appears to violate the norm of racial equality. It then loses its ability to prime white voters' racial predispositions.«

→ the “IE model”



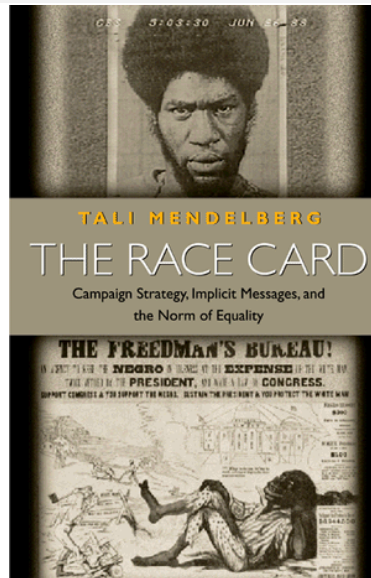
»The most important and underplayed lesson of the Horton message is that (...) [w]hen a society has repudiated racism, yet racial conflict persists, candidates can win by playing the race card only through implicit racial appeals. The implicit nature of these appeals allows them to prime racial stereotypes, fears, and resentments while appearing not to do so. When an implicit appeal is rendered explicit—when other elites bring the racial meaning of the appeal to voters' attention—it appears to violate the norm of racial equality. It then loses its ability to prime white voters' racial predispositions.«

→ the “IE model”



»The most important and underplayed lesson of the Horton message is that (...) [w]hen a society has repudiated racism, yet racial conflict persists, candidates can win by playing the race card only through implicit racial appeals. The implicit nature of these appeals allows them to prime racial stereotypes, fears, and resentments while appearing not to do so. When an implicit appeal is rendered explicit—when other elites bring the racial meaning of the appeal to voters' attention—it appears to violate the norm of racial equality. It then loses its ability to prime white voters' racial predispositions.«

→ the “IE model”



But: Huber & Lapinski (AJPS 2006) in large-scale replication find no support for IE model

TABLE 4 Baseline and Incremental Effects of Antiblack Predispositions on Policy Views, by Message Style

Message Style	Decrease Spending		Strengthen Welfare Work Requirements		Oppose Aid to Blacks		Decrease Affirmative Action	
	Experiment A	Experiment B	Experiment A	Experiment B	Experiment A	Experiment B	Experiment A	Experiment B
Implicit Racial (Baseline effect)	0.732*** (3.96)	1.094*** (6.04)	1.780*** (10.18)	1.577*** (8.96)	3.508*** (9.69)	3.019*** (8.17)	2.184*** (9.74)	1.635*** (7.39)
Explicit Racial (Incremental effect)	0.186 (0.72)	-0.335 (1.36)	-0.192 (0.82)	0.157 (0.66)	-0.208 (0.43)	0.340 (0.66)	-0.112 (0.39)	0.269 (0.93)
Control/GOTV (Incremental effect)	0.064 (0.25)	—	-0.322 (1.29)	—	-0.352 (0.72)	—	-0.136 (0.45)	—
Counter-Stereotypical (Incremental effect)	—	0.075 (0.29)	—	-0.079 (0.32)	—	0.127 (0.25)	—	0.057 (0.21)
Observations	1852	1995	1991	2138	1749	1859	1986	2091
$\chi^2$	440.22***	445.89***	454.69***	430.29***	397.19***	309.50***	419.85***	336.66***

Robust Z-statistics in parentheses. \*\*\*denotes  $p < .01$ , \*\*denotes  $p < .05$ , \*denotes  $p < .10$ . Functional form is ordered probit, except for oppose aid to blacks item, for which functional form is probit.

## Why the discrepancy?

»Mendelberg's experimental work, while innovative, relies on a sample of about 200 respondents, most with relatively low levels of racial resentment. (...) in estimating the differential effects of racial messages on these high-resentment individuals, Mendelberg's statistical analysis can rely on at most about 11 individuals for each of her three treatments. In contrast, in experiment A, we have more than 900 respondents meeting these conditions, and a minimum of 295 in each of the three treatments. Thus, Mendelberg's earlier results may simply be due to the vagaries of small samples...«

→ original experiment suffers from severe power problems (we will return to this on Thursday+Friday)

## Why the discrepancy?

»Mendelberg's experimental work, while innovative, relies on a sample of about 200 respondents, most with relatively low levels of racial resentment. (...) in estimating the differential effects of racial messages on these high-resentment individuals, Mendelberg's statistical analysis can rely on at most about 11 individuals for each of her three treatments. In contrast, in experiment A, we have more than 900 respondents meeting these conditions, and a minimum of 295 in each of the three treatments. Thus, Mendelberg's earlier results may simply be due to the vagaries of small samples...«

→ original experiment suffers from severe power problems (we will return to this on Thursday+Friday)



## Why the discrepancy?

»Mendelberg's experimental work, while innovative, relies on a sample of about 200 respondents, most with relatively low levels of racial resentment. (...) in estimating the differential effects of racial messages on these high-resentment individuals, Mendelberg's statistical analysis can rely on at most about 11 individuals for each of her three treatments. In contrast, in experiment A, we have more than 900 respondents meeting these conditions, and a minimum of 295 in each of the three treatments. Thus, Mendelberg's earlier results may simply be due to the vagaries of small samples...«

→ original experiment suffers from severe power problems (we will return to this on Thursday+Friday)

## Why the discrepancy?

»Mendelberg's experimental work, while innovative, relies on a sample of about 200 respondents, most with relatively low levels of racial resentment. (...) in estimating the differential effects of racial messages on these high-resentment individuals, Mendelberg's statistical analysis can rely on at most about 11 individuals for each of her three treatments. In contrast, in experiment A, we have more than 900 respondents meeting these conditions, and a minimum of 295 in each of the three treatments. Thus, Mendelberg's earlier results may simply be due to the vagaries of small samples...«

→ original experiment suffers from severe power problems (we will return to this on Thursday+Friday)

## Why the discrepancy?

»Mendelberg's experimental work, while innovative, relies on a sample of about 200 respondents, most with relatively low levels of racial resentment. (...) in estimating the differential effects of racial messages on these high-resentment individuals, Mendelberg's statistical analysis can rely on at most about 11 individuals for each of her three treatments. In contrast, in experiment A, we have more than 900 respondents meeting these conditions, and a minimum of 295 in each of the three treatments. Thus, Mendelberg's earlier results may simply be due to the vagaries of small samples...«

→ original experiment suffers from severe power problems (we will return to this on Thursday+Friday)

## Why the discrepancy?

»Mendelberg's experimental work, while innovative, relies on a sample of about 200 respondents, most with relatively low levels of racial resentment. (...) in estimating the differential effects of racial messages on these high-resentment individuals, Mendelberg's statistical analysis can rely on at most about 11 individuals for each of her three treatments. In contrast, in experiment A, we have more than 900 respondents meeting these conditions, and a minimum of 295 in each of the three treatments. Thus, Mendelberg's earlier results may simply be due to the vagaries of small samples...«

→ original experiment suffers from severe power problems (we will return to this on Thursday+Friday)

## Replicability standards in mid/late 00's:

»We provided Professor Mendelberg with the data (and our Stata code) from our experiments in April of 2006. At that time, we requested the data underlying her experiments. It was not provided to us. We again requested those data in June of 2007. In July of 2007, Professor Mendelberg reported to us that the data were unavailable because they were stored in a cardboard box that was damaged due to flooding in her home in April of 2007.«

Replicability standards in mid/late 00's:

»We provided Professor Mendelberg with the data (and our Stata code) from our experiments in April of 2006. At that time, we requested the data underlying her experiments. It was not provided to us. We again requested those data in June of 2007. In July of 2007, Professor Mendelberg reported to us that the data were unavailable because they were stored in a cardboard box that was damaged due to flooding in her home in April of 2007.«

Replicability standards in mid/late 00's:

»We provided Professor Mendelberg with the data (and our Stata code) from our experiments in April of 2006. At that time, we requested the data underlying her experiments. It was not provided to us. We again requested those data in June of 2007. In July of 2007, Professor Mendelberg reported to us that the data were unavailable because they were stored in a cardboard box that was damaged due to flooding in her home in April of 2007.«

Replicability standards in mid/late 00's:

»We provided Professor Mendelberg with the data (and our Stata code) from our experiments in April of 2006. At that time, we requested the data underlying her experiments. It was not provided to us. We again requested those data in June of 2007. In July of 2007, Professor Mendelberg reported to us that the data were unavailable because they were stored in a cardboard box that was damaged due to flooding in her home in April of 2007.«



Replicability standards in mid/late 00's:

»We provided Professor Mendelberg with the data (and our Stata code) from our experiments in April of 2006. At that time, we requested the data underlying her experiments. It was not provided to us. We again requested those data in June of 2007. In July of 2007, Professor Mendelberg reported to us that the data were unavailable because they were stored in a cardboard box that was damaged due to flooding in her home in April of 2007.«

At that point, replicability as an idea was well-established (but tacitly ignored):

## Replication, Replication

Gary King,<sup>1</sup> *Harvard University*

**P**olitical science is a community enterprise; the community of empirical political scientists needs access to the body of data necessary to replicate existing studies to understand, evaluate, and especially build on this work. Unfortunately, the norms we have in place now do not encourage, or in some cases even permit, this aim. Following are suggestions that would facilitate replication and are easy to implement—by teachers, students, dissertation writers, graduate programs, authors, reviewers, funding agencies, and journal and book editors.

### Problems in Empirical Political Science

How were the respondents selected? Who did the interviewing? What was the question order? How did you decide which informants to interview or villages to visit? How long did you spend in each community? Did you speak to people in their language or through an interpreter? Which version of the ICPSR file did you extract information from? How knowledgeable were the coders? How frequently did the coders agree? Exactly what codes were originally generated and what were all the recodes performed? Precisely which measure of unemployment was used? What were the exact rules used for conducting the content analysis? When did the time series begin and end? What countries were included in your study and how were they cho-

*standard* holds that sufficient information exists with which to understand, evaluate, and build upon a prior work if a third party could replicate the results without any additional information from the author. The replication standard does not actually require anyone to replicate the results of an article or book. It only requires sufficient information to be provided—in the article or book or in some other publicly accessible form—so that the results could in principle be replicated. Since many believe that research standards should be applied equally to quantitative and qualitative analyses (King, Keohane, and Verba 1994), the replication standard is also appropriate for qualitative research, although the rich complexity of the data often

- 1 Motivating example 1
- 2 The DA-RT movement
- 3 The counter-DA-RT movement
- 4 Motivating example 2
- 5 Introduction to R

## Journals adhering to Data Access and Research Transparency (DA-RT) standards:

- ① Require authors to ensure that **cited data are available at the time of publication** through a trusted digital repository.
- ② Require authors to **delineate clearly the analytic procedures upon which their published claims rely**, and where possible to provide access to all relevant analytic materials.
- ③ Maintain a **consistent data citation policy** to increase the credit that data creators and suppliers receive for their work.
- ④ Ensure that journal style guides, codes of ethics, publication manuals, and other forms of guidance are updated and expanded to include improved data access and research transparency requirements.

## Journals adhering to Data Access and Research Transparency (DA-RT) standards:

- ① Require authors to ensure that **cited data are available at the time of publication** through a trusted digital repository.
- ② Require authors to **delineate clearly the analytic procedures upon which their published claims rely**, and where possible to provide access to all relevant analytic materials.
- ③ Maintain a **consistent data citation policy** to increase the credit that data creators and suppliers receive for their work.
- ④ Ensure that journal style guides, codes of ethics, publication manuals, and other forms of guidance are updated and expanded to include improved data access and research transparency requirements.

Journals adhering to Data Access and Research Transparency (DA-RT) standards:

- ① Require authors to ensure that **cited data are available at the time of publication** through a trusted digital repository.
- ② Require authors to **delineate clearly the analytic procedures upon which their published claims rely**, and where possible to provide access to all relevant analytic materials.
- ③ Maintain a **consistent data citation policy** to increase the credit that data creators and suppliers receive for their work.
- ④ Ensure that journal style guides, codes of ethics, publication manuals, and other forms of guidance are updated and expanded to include improved data access and research transparency requirements.

Journals adhering to Data Access and Research Transparency (DA-RT) standards:

- ① Require authors to ensure that **cited data are available at the time of publication** through a trusted digital repository.
- ② Require authors to **delineate clearly the analytic procedures upon which their published claims rely**, and where possible to provide access to all relevant analytic materials.
- ③ Maintain a **consistent data citation policy** to increase the credit that data creators and suppliers receive for their work.
- ④ Ensure that journal style guides, codes of ethics, publication manuals, and other forms of guidance are updated and expanded to include improved data access and research transparency requirements.

Journals adhering to Data Access and Research Transparency (DA-RT) standards:

- ① Require authors to ensure that **cited data are available at the time of publication** through a trusted digital repository.
- ② Require authors to **delineate clearly the analytic procedures upon which their published claims rely**, and where possible to provide access to all relevant analytic materials.
- ③ Maintain a **consistent data citation policy** to increase the credit that data creators and suppliers receive for their work.
- ④ Ensure that journal style guides, codes of ethics, publication manuals, and other forms of guidance are updated and expanded to include improved data access and research transparency requirements.



## Signatories:

*American Journal of Political Science, American Political Science Review, American Politics Research, British Journal of Political Science, Comparative Political Studies, Conflict Management and Peace Science, Cooperation and Conflict, European Journal of Political Research, European Political Science, European Union Politics, International Interactions, International Security, Journal of Conflict Resolution, Journal of Elections, Public Opinion, and Parties, Journal of Experimental Political Science, Journal of European Public Policy, Journal of Peace Research, Journal of Theoretical Politics, Quarterly Journal of Political Science, Party Politics, Political Analysis, Political Behavior, Political Communication, Political Science Research and Methods, Research and Politics, Rivista Italiana di Scienza Politica, State Politics and Policy Quarterly, The Journal of Politics, The Political Methodologist*

A more ambitious standard: since 2005, QJPS has required authors to submit a data+code 'package', subject to in-house replication

»Given that the QJPS review is relatively basic, however, one might ask whether it is even worth the considerable time the QJPS invests. Experience has shown the answer is an unambiguous 'yes'. Of the 24 empirical papers subject to in-house replication review since September 2012, only 4 packages required no modifications. Of the remaining 20 papers, 13 had code that would not execute without errors, 8 failed to include code for results that appeared in the paper, and 7 failed to include installation directions for software dependencies. Most troubling, however, 14 (58 percent) had results in the paper that differed from those generated by the author's own code.«

Source: *A Decade of Replications: Lessons from the Quarterly Journal of Political Science*

A more ambitious standard: since 2005, QJPS has required authors to submit a data+code 'package', subject to in-house replication

»Given that the QJPS review is relatively basic, however, one might ask whether it is even worth the considerable time the QJPS invests. Experience has shown the answer is an unambiguous 'yes'. Of the 24 empirical papers subject to in-house replication review since September 2012, only 4 packages required no modifications. Of the remaining 20 papers, 13 had code that would not execute without errors, 8 failed to include code for results that appeared in the paper, and 7 failed to include installation directions for software dependencies. Most troubling, however, 14 (58 percent) had results in the paper that differed from those generated by the author's own code.«

Source: *A Decade of Replications: Lessons from the Quarterly Journal of Political Science*

A more ambitious standard: since 2005, QJPS has required authors to submit a data+code 'package', subject to in-house replication

»Given that the QJPS review is relatively basic, however, one might ask whether it is even worth the considerable time the QJPS invests. Experience has shown the answer is an unambiguous 'yes'. Of the 24 empirical papers subject to in-house replication review since September 2012, only 4 packages required no modifications. Of the remaining 20 papers, 13 had code that would not execute without errors, 8 failed to include code for results that appeared in the paper, and 7 failed to include installation directions for software dependencies. Most troubling, however, 14 (58 percent) had results in the paper that differed from those generated by the author's own code.«

Source: *A Decade of Replications: Lessons from the Quarterly Journal of Political Science*

A more ambitious standard: since 2005, QJPS has required authors to submit a data+code 'package', subject to in-house replication

»Given that the QJPS review is relatively basic, however, one might ask whether it is even worth the considerable time the QJPS invests. Experience has shown the answer is an unambiguous 'yes'. Of the 24 empirical papers subject to in-house replication review since September 2012, only 4 packages required no modifications. Of the remaining 20 papers, 13 had code that would not execute without errors, 8 failed to include code for results that appeared in the paper, and 7 failed to include installation directions for software dependencies. Most troubling, however, 14 (58 percent) had results in the paper that differed from those generated by the author's own code.«

Source: *A Decade of Replications: Lessons from the Quarterly Journal of Political Science*

A more ambitious standard: since 2005, QJPS has required authors to submit a data+code 'package', subject to in-house replication

»Given that the QJPS review is relatively basic, however, one might ask whether it is even worth the considerable time the QJPS invests. Experience has shown the answer is an unambiguous 'yes'. Of the 24 empirical papers subject to in-house replication review since September 2012, only 4 packages required no modifications. Of the remaining 20 papers, 13 had code that would not execute without errors, 8 failed to include code for results that appeared in the paper, and 7 failed to include installation directions for software dependencies. Most troubling, however, 14 (58 percent) had results in the paper that differed from those generated by the author's own code.«

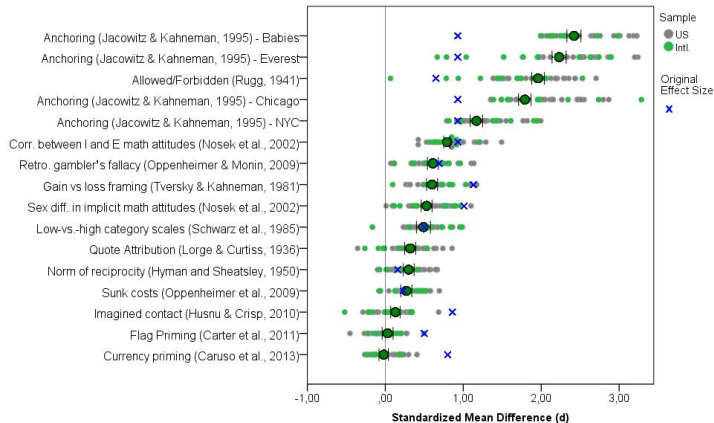
Source: *A Decade of Replications: Lessons from the Quarterly Journal of Political Science*

A more ambitious standard: since 2005, QJPS has required authors to submit a data+code 'package', subject to in-house replication

»Given that the QJPS review is relatively basic, however, one might ask whether it is even worth the considerable time the QJPS invests. Experience has shown the answer is an unambiguous 'yes'. Of the 24 empirical papers subject to in-house replication review since September 2012, only 4 packages required no modifications. Of the remaining 20 papers, 13 had code that would not execute without errors, 8 failed to include code for results that appeared in the paper, and 7 failed to include installation directions for software dependencies. Most troubling, however, 14 (58 percent) had results in the paper that differed from those generated by the author's own code.«

Source: *A Decade of Replications: Lessons from the Quarterly Journal of Political Science*

## Context: ongoing replication crisis in (mostly social) psychology




Source: *Investigating Variation in Replicability: A “Many Labs” Replication Project*



Carter et al. (PS 2011)

## A Single Exposure to the American Flag Shifts Support Toward Republicanism up to 8 Months Later

Psychological Science  
22(8) 1011–1018  
© The Author(s) 2011  
Reprints and permission:  
sagepub.com/journalsPermissions.nav  
DOI: 10.1177/0956797611414726  
<http://pss.sagepub.com>  
 SAGE

**Travis J. Carter<sup>1</sup>, Melissa J. Ferguson<sup>2</sup>, and Ran R. Hassin<sup>3</sup>**

<sup>1</sup>Center for Decision Research, Booth School of Business, University of Chicago; <sup>2</sup>Department of Psychology, Cornell University; and <sup>3</sup>Department of Psychology and The Center for the Study of Rationality, Hebrew University

### Abstract

There is scant evidence that incidental cues in the environment significantly alter people's political judgments and behavior in a durable way. We report that a brief exposure to the American flag led to a shift toward Republican beliefs, attitudes, and voting behavior among both Republican and Democratic participants, despite their overwhelming belief that exposure to the flag would not influence their behavior. In Experiment 1, which was conducted online during the 2008 U.S. presidential election, a single exposure to an American flag resulted in a significant increase in participants' Republican voting intentions, voting behavior, political beliefs, and implicit and explicit attitudes, with some effects lasting 8 months after the exposure to the prime. In Experiment 2, we replicated the findings more than a year into the current Democratic presidential term. These results constitute the first evidence that nonconscious priming effects from exposure to a national flag can bias the citizenry toward one political party and can have considerable durability.

Source: <http://journals.sagepub.com/doi/pdf/10.1177/0956797611414726>

- 1 Motivating example 1
- 2 The DA-RT movement
- 3 The counter-DA-RT movement**
- 4 Motivating example 2
- 5 Introduction to R

## Argument 1: DA-RT promotes a “rather narrow conception of political science relevance” (277)

»The conception of political science reflected in the DA-RT initiative (...) is a modernizing vision centered on the revolutionary scholarly and communicative possibilities conferred by the new digital technologies. In this sense, it is novel and indeed pioneering. But at its heart, this vision articulates the old positivist idea that social science, like natural science, is all about employing the most up-to-date methods and techniques to generate and then disseminate veridical truths about the world. On this view we must work as hard as possible to eliminate the perspectival and interpretive dimensions of science, and support methods and procedures to ensure that every piece of research is purged of partiality and oriented toward achieving an accurate representation of reality that anyone else, anywhere, can potentially ‘replicate’ and assess.« (278)

→ does DA-RT promote a narrow conception of polisci? If so, it that problematic?

## Argument 1: DA-RT promotes a “rather narrow conception of political science relevance” (277)

»The conception of political science reflected in the DA-RT initiative (...) is a modernizing vision centered on the revolutionary scholarly and communicative possibilities conferred by the new digital technologies. In this sense, it is novel and indeed pioneering. But at its heart, this vision articulates the old positivist idea that social science, like natural science, is all about employing the most up-to-date methods and techniques to generate and then disseminate veridical truths about the world. On this view we must work as hard as possible to eliminate the perspectival and interpretive dimensions of science, and support methods and procedures to ensure that every piece of research is purged of partiality and oriented toward achieving an accurate representation of reality that anyone else, anywhere, can potentially ‘replicate’ and assess.« (278)

→ does DA-RT promote a narrow conception of polisci? If so, is that problematic?

## Argument 1: DA-RT promotes a “rather narrow conception of political science relevance” (277)

»The conception of political science reflected in the DA-RT initiative (...) is a modernizing vision centered on the revolutionary scholarly and communicative possibilities conferred by the new digital technologies. In this sense, it is novel and indeed pioneering. But at its heart, this vision articulates the old positivist idea that social science, like natural science, is all about employing the most up-to-date methods and techniques to generate and then disseminate veridical truths about the world. On this view we must work as hard as possible to eliminate the perspectival and interpretive dimensions of science, and support methods and procedures to ensure that every piece of research is purged of partiality and oriented toward achieving an accurate representation of reality that anyone else, anywhere, can potentially ‘replicate’ and assess.« (278)

→ does DA-RT promote a narrow conception of polisci? If so, is that problematic?

## Argument 1: DA-RT promotes a “rather narrow conception of political science relevance” (277)

»The conception of political science reflected in the DA-RT initiative (...) is a modernizing vision centered on the revolutionary scholarly and communicative possibilities conferred by the new digital technologies. In this sense, it is novel and indeed pioneering. But at its heart, this vision articulates the old positivist idea that social science, like natural science, is all about employing the most up-to-date methods and techniques to generate and then disseminate veridical truths about the world. On this view we must work as hard as possible to eliminate the perspectival and interpretive dimensions of science, and support methods and procedures to ensure that every piece of research is purged of partiality and oriented toward achieving an accurate representation of reality that anyone else, anywhere, can potentially ‘replicate’ and assess.« (278)

→ does DA-RT promote a narrow conception of polisci? If so, it that problematic?

## Argument 1: DA-RT promotes a “rather narrow conception of political science relevance” (277)

»The conception of political science reflected in the DA-RT initiative (...) is a modernizing vision centered on the revolutionary scholarly and communicative possibilities conferred by the new digital technologies. In this sense, it is novel and indeed pioneering. But at its heart, this vision articulates the old positivist idea that social science, like natural science, is all about employing the most up-to-date methods and techniques to generate and then disseminate veridical truths about the world. On this view we must work as hard as possible to eliminate the perspectival and interpretive dimensions of science, and support methods and procedures to ensure that every piece of research is purged of partiality and oriented toward achieving an accurate representation of reality that anyone else, anywhere, can potentially ‘replicate’ and assess.« (278)

→ does DA-RT promote a narrow conception of polisci? If so, is that problematic?

## Argument 2: DA-RT tries to solve a non-existent problem

»Have major arguments about any particular domain of inquiry or any broad theoretical perspective—historical institutionalism? rational choice institutionalism? theory of democratization? theory of civil war onset and duration?—been damaged by the lack of sufficient standards of data transparency? Have the standard practices of political science publishing—prepublication double-blind peer reviewing, demanding processes of revision and resubmission and further review, and ongoing post-publication processes of scholarly critique—failed to ‘remedy’ weaknesses that have been discerned in particular pieces of research?« (275)

»I am unaware of any major breach of scholarly ethics in our discipline in recent years (one of my colleagues at IU, himself a quantitative researcher, has described DA-RT to me as “a solution in search of a problem”).« (278)

→ is DA-RT a solution in search of a problem?



## Argument 2: DA-RT tries to solve a non-existent problem

»Have major arguments about any particular domain of inquiry or any broad theoretical perspective—historical institutionalism? rational choice institutionalism? theory of democratization? theory of civil war onset and duration?—been damaged by the lack of sufficient standards of data transparency? Have the standard practices of political science publishing—prepublication double-blind peer reviewing, demanding processes of revision and resubmission and further review, and ongoing post-publication processes of scholarly critique—failed to ‘remedy’ weaknesses that have been discerned in particular pieces of research?« (275)

»I am unaware of any major breach of scholarly ethics in our discipline in recent years (one of my colleagues at IU, himself a quantitative researcher, has described DA-RT to me as “a solution in search of a problem”).« (278)

→ is DA-RT a solution in search of a problem?

## Argument 2: DA-RT tries to solve a non-existent problem

»Have major arguments about any particular domain of inquiry or any broad theoretical perspective—historical institutionalism? rational choice institutionalism? theory of democratization? theory of civil war onset and duration?—been damaged by the lack of sufficient standards of data transparency? Have the standard practices of political science publishing—prepublication double-blind peer reviewing, demanding processes of revision and resubmission and further review, and ongoing post-publication processes of scholarly critique—failed to ‘remedy’ weaknesses that have been discerned in particular pieces of research?« (275)

»I am unaware of any major breach of scholarly ethics in our discipline in recent years (one of my colleagues at IU, himself a quantitative researcher, has described DA-RT to me as “a solution in search of a problem”).« (278)

→ is DA-RT a solution in search of a problem?

## Argument 2: DA-RT tries to solve a non-existent problem

»Have major arguments about any particular domain of inquiry or any broad theoretical perspective—historical institutionalism? rational choice institutionalism? theory of democratization? theory of civil war onset and duration?—been damaged by the lack of sufficient standards of data transparency? Have the standard practices of political science publishing—prepublication double-blind peer reviewing, demanding processes of revision and resubmission and further review, and ongoing post-publication processes of scholarly critique—failed to ‘remedy’ weaknesses that have been discerned in particular pieces of research?« (275)

»I am unaware of any major breach of scholarly ethics in our discipline in recent years (one of my colleagues at IU, himself a quantitative researcher, has described DA-RT to me as “a solution in search of a problem”).« (278)

→ is DA-RT a solution in search of a problem?

## Argument 2: DA-RT tries to solve a non-existent problem

»Have major arguments about any particular domain of inquiry or any broad theoretical perspective—historical institutionalism? rational choice institutionalism? theory of democratization? theory of civil war onset and duration?—been damaged by the lack of sufficient standards of data transparency? Have the standard practices of political science publishing—prepublication double-blind peer reviewing, demanding processes of revision and resubmission and further review, and ongoing post-publication processes of scholarly critique—failed to ‘remedy’ weaknesses that have been discerned in particular pieces of research?« (275)

»I am unaware of any major breach of scholarly ethics in our discipline in recent years (one of my colleagues at IU, himself a quantitative researcher, has described DA-RT to me as “a solution in search of a problem”).« (278)

→ is DA-RT a solution in search of a problem?

## November 2015: *Petition to Delay DA-RT Implementation*

→ exercise: can you find familiar names and/or colleagues?

## November 2015: *Petition to Delay DA-RT Implementation*

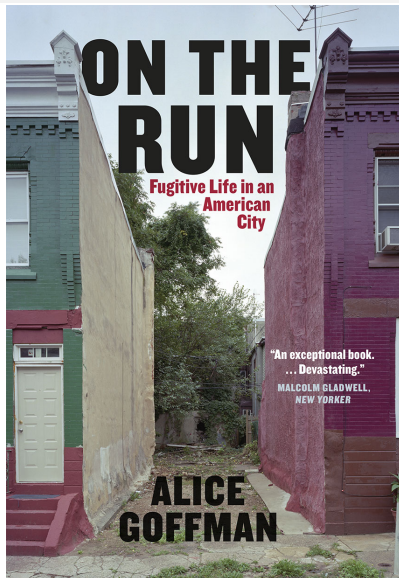
→ exercise: can you find familiar names and/or colleagues?

November 2015: *Petition to Delay DA-RT Implementation*

→ exercise: can you find familiar names and/or colleagues?

- 1 Motivating example 1
- 2 The DA-RT movement
- 3 The counter-DA-RT movement
- 4 Motivating example 2**
- 5 Introduction to R





## On the Run: Wanted Men in a Philadelphia Ghetto

Alice Goffman  
Princeton University

*Although recent increases in imprisonment are concentrated in poor Black communities, we know little about how daily life within these neighborhoods is affected. Almost all ethnographic work in poor minority neighborhoods was written before the expansion of the criminal justice system, and the bulk of research on “mass imprisonment” relies on survey data, field experiments, or interviews, conceptualizing its impact in terms of current or former felons and their families. Drawing on six years of fieldwork in Philadelphia, this article shifts the focus from imprisonment and criminal records to the increase in policing and supervision in poor Black neighborhoods, and what this has meant for a growing status group of wanted people. For many young men, avoiding jail has become a daily preoccupation: they have warrants out for minor infractions, like failing to pay court fees or breaking curfew, and will be detained if they are identified. Such threat of imprisonment transforms social relations by undermining already tenuous attachments to family, work, and community. But young men also rely on their precarious legal standing to explain failures that would have occurred anyway, while girlfriends and neighbors exploit their wanted status as an instrument of social control. I discuss the implications of my ethnographic observations relative to prior treatments of the poor and policing, and with regard to broader sociological questions about punishment and surveillance in the modern era.*

## Alice Goffman TED talk

## Oft-repeated criticism: virtually zero transparency in the field work behind *On The Run*

»I do not know if Goffman's editors and dissertation committee held her to a journalist's standard of fact checking. There is no footnote for the hospital incident in *On the Run*, and her dissertation is not available from the Princeton library. Alas, it is now too late to obtain any additional documentation, because Goffman shredded all of her field notes and disposed of her hard drive. Her reason, as she explained to the Philadelphia Inquirer, was to remove "the threat of being subpoenaed" for the identities of her subjects, many of whom had discussed or committed crimes in her presence.«

Source: *Ethics On The Run*

→ is this science? If yes, what makes it so?

Oft-repeated criticism: virtually zero transparency in the field work behind *On The Run*

»I do not know if Goffman's editors and dissertation committee held her to a journalist's standard of fact checking. There is no footnote for the hospital incident in *On the Run*, and her dissertation is not available from the Princeton library. Alas, it is now too late to obtain any additional documentation, because Goffman shredded all of her field notes and disposed of her hard drive. Her reason, as she explained to the Philadelphia Inquirer, was to remove "the threat of being subpoenaed" for the identities of her subjects, many of whom had discussed or committed crimes in her presence.«

Source: *Ethics On The Run*

→ is this science? If yes, what makes it so?

Oft-repeated criticism: virtually zero transparency in the field work behind *On The Run*

»I do not know if Goffman's editors and dissertation committee held her to a journalist's standard of fact checking. There is no footnote for the hospital incident in *On the Run*, and her dissertation is not available from the Princeton library. Alas, it is now too late to obtain any additional documentation, because Goffman shredded all of her field notes and disposed of her hard drive. Her reason, as she explained to the Philadelphia Inquirer, was to remove "the threat of being subpoenaed" for the identities of her subjects, many of whom had discussed or committed crimes in her presence.«

Source: *Ethics On The Run*

→ is this science? If yes, what makes it so?

Oft-repeated criticism: virtually zero transparency in the field work behind *On The Run*

»I do not know if Goffman's editors and dissertation committee held her to a journalist's standard of fact checking. There is no footnote for the hospital incident in *On the Run*, and her dissertation is not available from the Princeton library. Alas, it is now too late to obtain any additional documentation, because **Goffman shredded all of her field notes and disposed of her hard drive**. Her reason, as she explained to the *Philadelphia Inquirer*, was to remove "the threat of being subpoenaed" for the identities of her subjects, many of whom had discussed or committed crimes in her presence.«

Source: *Ethics On The Run*

→ is this science? If yes, what makes it so?

Oft-repeated criticism: virtually zero transparency in the field work behind *On The Run*

»I do not know if Goffman's editors and dissertation committee held her to a journalist's standard of fact checking. There is no footnote for the hospital incident in *On the Run*, and her dissertation is not available from the Princeton library. Alas, it is now too late to obtain any additional documentation, because **Goffman shredded all of her field notes and disposed of her hard drive**. Her reason, as she explained to the Philadelphia Inquirer, was to remove "the threat of being subpoenaed" for the identities of her subjects, many of whom had discussed or committed crimes in her presence.«

Source: *Ethics On The Run*

→ is this science? If yes, what makes it so?

Oft-repeated criticism: virtually zero transparency in the field work behind *On The Run*

»I do not know if Goffman's editors and dissertation committee held her to a journalist's standard of fact checking. There is no footnote for the hospital incident in *On the Run*, and her dissertation is not available from the Princeton library. Alas, it is now too late to obtain any additional documentation, because **Goffman shredded all of her field notes and disposed of her hard drive**. Her reason, as she explained to the Philadelphia Inquirer, was to remove "the threat of being subpoenaed" for the identities of her subjects, many of whom had discussed or committed crimes in her presence.«

Source: *Ethics On The Run*

→ is this science? If yes, what makes it so?



Oft-repeated criticism: virtually zero transparency in the field work behind *On The Run*

»I do not know if Goffman's editors and dissertation committee held her to a journalist's standard of fact checking. There is no footnote for the hospital incident in *On the Run*, and her dissertation is not available from the Princeton library. Alas, it is now too late to obtain any additional documentation, because **Goffman shredded all of her field notes and disposed of her hard drive**. Her reason, as she explained to the Philadelphia Inquirer, was to remove "the threat of being subpoenaed" for the identities of her subjects, many of whom had discussed or committed crimes in her presence.«

Source: *Ethics On The Run*

→ is this science? If yes, what makes it so?

- 1 Motivating example 1
- 2 The DA-RT movement
- 3 The counter-DA-RT movement
- 4 Motivating example 2
- 5 Introduction to R

In recent years, two revolutions have vastly increased availability and accessibility of quantitative data

- ① 'data revolution': vast quantities of data much more assessible
- ② 'computational revolution': exponential increases in processing power

(Note: a parallel trend, closely connected to causal inference, typically labeled 'credibility revolution')

→ implication: mere familiarity with traditional structured data will not suffice

»It is simply not sufficient to achieve 'statistical literacy' by learning about common statistical concepts and methods. Instead, all students in the social sciences should acquire basic data analysis skills so that they can exploit ample opportunities to learn from data (...)« (3)

In recent years, two revolutions have vastly increased availability and accessibility of quantitative data

- ① 'data revolution': vast quantities of data much more assessible
- ② 'computational revolution': exponential increases in processing power

(Note: a parallel trend, closely connected to causal inference, typically labeled 'credibility revolution')

→ implication: mere familiarity with traditional structured data will not suffice

»It is simply not sufficient to achieve 'statistical literacy' by learning about common statistical concepts and methods. Instead, all students in the social sciences should acquire basic data analysis skills so that they can exploit ample opportunities to learn from data (...)« (3)

In recent years, two revolutions have vastly increased availability and accessibility of quantitative data

- ① 'data revolution': vast quantities of data much more assessible
- ② 'computational revolution': exponential increases in processing power

(Note: a parallel trend, closely connected to causal inference, typically labeled 'credibility revolution')

→ implication: mere familiarity with traditional structured data will not suffice

»It is simply not sufficient to achieve 'statistical literacy' by learning about common statistical concepts and methods. Instead, all students in the social sciences should acquire basic data analysis skills so that they can exploit ample opportunities to learn from data (...)« (3)

In recent years, two revolutions have vastly increased availability and accessibility of quantitative data

- ① 'data revolution': vast quantities of data much more assessible
- ② 'computational revolution': exponential increases in processing power

(Note: a parallel trend, closely connected to causal inference, typically labeled 'credibility revolution')

→ implication: mere familiarity with traditional structured data will not suffice

»It is simply not sufficient to achieve 'statistical literacy' by learning about common statistical concepts and methods. Instead, all students in the social sciences should acquire basic data analysis skills so that they can exploit ample opportunities to learn from data (...)« (3)

In recent years, two revolutions have vastly increased availability and accessibility of quantitative data

- ① 'data revolution': vast quantities of data much more assessible
- ② 'computational revolution': exponential increases in processing power

(Note: a parallel trend, closely connected to causal inference, typically labeled 'credibility revolution')

→ implication: mere familiarity with traditional structured data will not suffice

»It is simply not sufficient to achieve 'statistical literacy' by learning about common statistical concepts and methods. Instead, all students in the social sciences should acquire basic data analysis skills so that they can exploit ample opportunities to learn from data (...)« (3)

In recent years, two revolutions have vastly increased availability and accessibility of quantitative data

- ① 'data revolution': vast quantities of data much more assessible
- ② 'computational revolution': exponential increases in processing power

(Note: a parallel trend, closely connected to causal inference, typically labeled 'credibility revolution')

→ implication: mere familiarity with traditional structured data will not suffice

»It is simply not sufficient to achieve 'statistical literacy' by learning about common statistical concepts and methods. Instead, all students in the social sciences should acquire basic data analysis skills so that they can exploit ample opportunities to learn from data (...)« (3)



In recent years, two revolutions have vastly increased availability and accessibility of quantitative data

- ① 'data revolution': vast quantities of data much more assessible
- ② 'computational revolution': exponential increases in processing power

(Note: a parallel trend, closely connected to causal inference, typically labeled 'credibility revolution')

→ implication: mere familiarity with traditional structured data will not suffice

»It is simply not sufficient to achieve 'statistical literacy' by learning about common statistical concepts and methods. Instead, all students in the social sciences should acquire basic data analysis skills so that they can exploit ample opportunities to learn from data (...)« (3)

## What is R?

- a program for conducting statistical programming
- a programming language (like C++, Python, Perl, etc.)
- object-oriented (ctr. e.g. Stata)
- open source
- invented by Ross Ihaka & Ross Gentleman
- extension of the programming language S
- first stable beta: 2000

## What is R?

- a program for conducting statistical programming
- a programming language (like C++, Python, Perl, etc.)
- object-oriented (ctr. e.g. Stata)
- open source
- invented by Ross Ihaka & Ross Gentleman
- extension of the programming language S
- first stable beta: 2000

## What is R?

- a program for conducting statistical programming
- a programming language (like C++, Python, Perl, etc.)
- object-oriented (ctr. e.g. Stata)
- open source
- invented by Ross Ihaka & Ross Gentleman
- extension of the programming language S
- first stable beta: 2000

## What is R?

- a program for conducting statistical programming
- a programming language (like C++, Python, Perl, etc.)
- object-oriented (ctr. e.g. Stata)
- open source
- invented by Ross Ihaka & Ross Gentleman
- extension of the programming language S
- first stable beta: 2000

## What is R?

- a program for conducting statistical programming
- a programming language (like C++, Python, Perl, etc.)
- object-oriented (ctr. e.g. Stata)
- open source
- invented by Ross Ihaka & Ross Gentleman
- extension of the programming language S
- first stable beta: 2000

## What is R?

- a program for conducting statistical programming
- a programming language (like C++, Python, Perl, etc.)
- object-oriented (ctr. e.g. Stata)
- open source
- invented by Ross Ihaka & Ross Gentleman
- extension of the programming language S
- first stable beta: 2000

## What is R?

- a program for conducting statistical programming
- a programming language (like C++, Python, Perl, etc.)
- object-oriented (ctr. e.g. Stata)
- open source
- invented by Ross Ihaka & Ross Gentleman
- extension of the programming language S
- first stable beta: 2000



## What is R?

- a program for conducting statistical programming
- a programming language (like C++, Python, Perl, etc.)
- object-oriented (ctr. e.g. Stata)
- open source
- invented by Ross Ihaka & Ross Gentleman
- extension of the programming language S
- first stable beta: 2000

## What is R?

- a program for conducting statistical programming
- a programming language (like C++, Python, Perl, etc.)
- object-oriented (ctr. e.g. Stata)
- open source
- invented by Ross Ihaka & Ross Gentleman
- extension of the programming language S
- first stable beta: 2000

## Why R?

- highly modular: almost infinitely applicable
- reproducible workflow
- in most industrial data science, predominant standard
- free as in free speech and free beer
- clearly superior software for data visualization

## Why R?

- highly modular: almost infinitely applicable
- reproducible workflow
- in most industrial data science, predominant standard
- free as in free speech and free beer
- clearly superior software for data visualization

## Why R?

- highly modular: almost infinitely applicable
- reproducible workflow
- in most industrial data science, predominant standard
- free as in free speech and free beer
- clearly superior software for data visualization

## Why R?

- highly modular: almost infinitely applicable
- reproducible workflow
- in most industrial data science, predominant standard
- free as in free speech and free beer
- clearly superior software for data visualization

## Why R?

- highly modular: almost infinitely applicable
- reproducible workflow
- in most industrial data science, predominant standard
- free as in free speech and free beer
- clearly superior software for data visualization

## Why R?

- highly modular: almost infinitely applicable
- reproducible workflow
- in most industrial data science, predominant standard
- free as in free speech and free beer
- clearly superior software for data visualization



## Why R?

- highly modular: almost infinitely applicable
- reproducible workflow
- in most industrial data science, predominant standard
- free as in free speech and free beer
- clearly superior software for data visualization

Company  
value

In billions

100 —

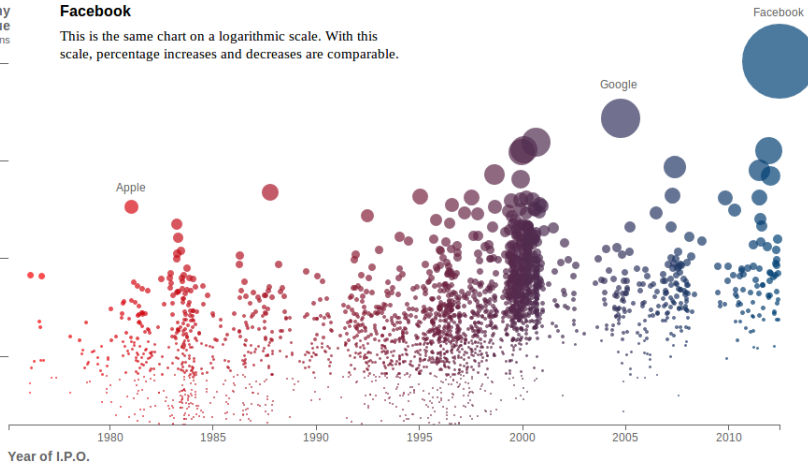
10 —

1 —

0.1 —

**Facebook**

This is the same chart on a logarithmic scale. With this scale, percentage increases and decreases are comparable.



The New York Times

## Mapping America: Every City, Every Block

Browse local data from the Census Bureau's American Community Survey, based on samples from 2005 to 2009.

Find something interesting? Share this view on [Twitter](#) or [Facebook](#)

### Distribution of racial and ethnic groups

[View More Maps](#)

Address, ZIP code or city

Go

Go

Go

Go

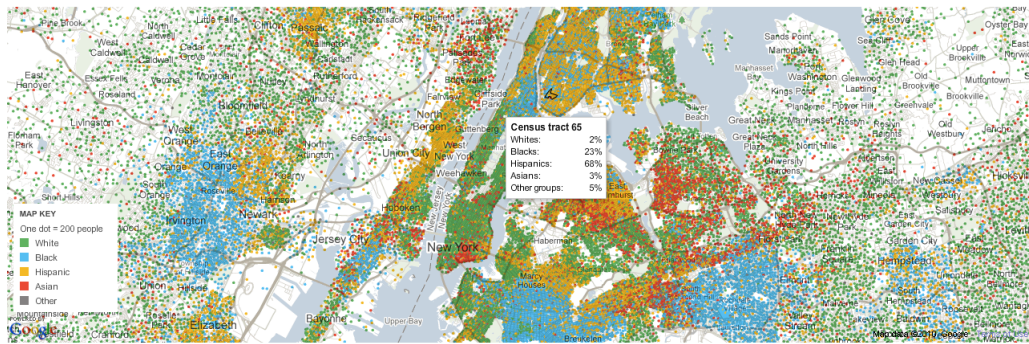
Go

Go

Go

Go

Go



By MATTHEW BLOCH and SHAN CARTER | Source: 2005-9 American Community Survey, Census Bureau; [socialexplorer.com](#)

## Why *not* R?

- no graphical user interface (GUI)
- highly modular: constant need to import extra packages
- almost zero tolerance for error
- indecipherable error messages

## Why *not* R?

- no graphical user interface (GUI)
- highly modular: constant need to import extra packages
- almost zero tolerance for error
- indecipherable error messages

## Why *not* R?

- no graphical user interface (GUI)
- highly modular: constant need to import extra packages
- almost zero tolerance for error
- indecipherable error messages

## Why *not* R?

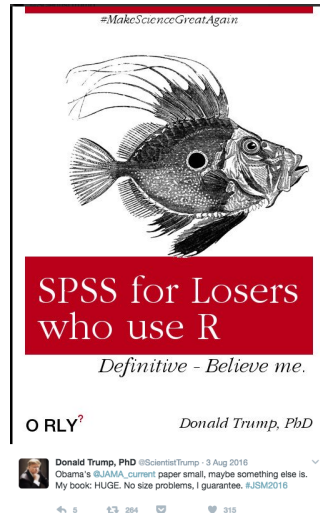
- no graphical user interface (GUI)
- highly modular: constant need to import extra packages
- almost zero tolerance for error
- indecipherable error messages

## Why *not* R?

- no graphical user interface (GUI)
- highly modular: constant need to import extra packages
- almost zero tolerance for error
- indecipherable error messages



R has some very prominent critics:



But:



## I Was Intimidated by Coding Until I Learned This Secret Strategy: Googling

"You don't need to go to grad school. Save your money. I'll teach you how to code."  
Seven years ago, in a bar near downtown Los Angeles, I was sharing ...

SLATE.COM

## Concepts introduced here:

- working directory
- objects
- vectors
- functions
- data frames
- export/import

## Concepts introduced here:

- working directory
- objects
- vectors
- functions
- data frames
- export/import

## Concepts introduced here:

- working directory
- objects
- vectors
- functions
- data frames
- export/import

## Concepts introduced here:

- working directory
- objects
- vectors
- functions
- data frames
- export/import

## Concepts introduced here:

- working directory
- objects
- vectors
- functions
- data frames
- export/import

## Concepts introduced here:

- working directory
- objects
- vectors
- functions
- data frames
- export/import



## Concepts introduced here:

- working directory
- objects
- vectors
- functions
- data frames
- export/import

See you tomorrow!