

Maths and computer science for social good



David Sumpter
Uppsala University
Sweden

Ida-Maria Sintorn
Uppsala University
Sweden



UPPSALA
UNIVERSITET

Tools and data for doing social good

- A case study from this week...
- Using tools.
- Data sources.
- A few thoughts on ambition level

A case study...

PLAYLIST

The most popular talks of all time

Are schools killing creativity? What makes a great leader? How can I find happiness? These 25 talks are the ones that you and your fellow TED fans just can't stop sharing.



Watch now



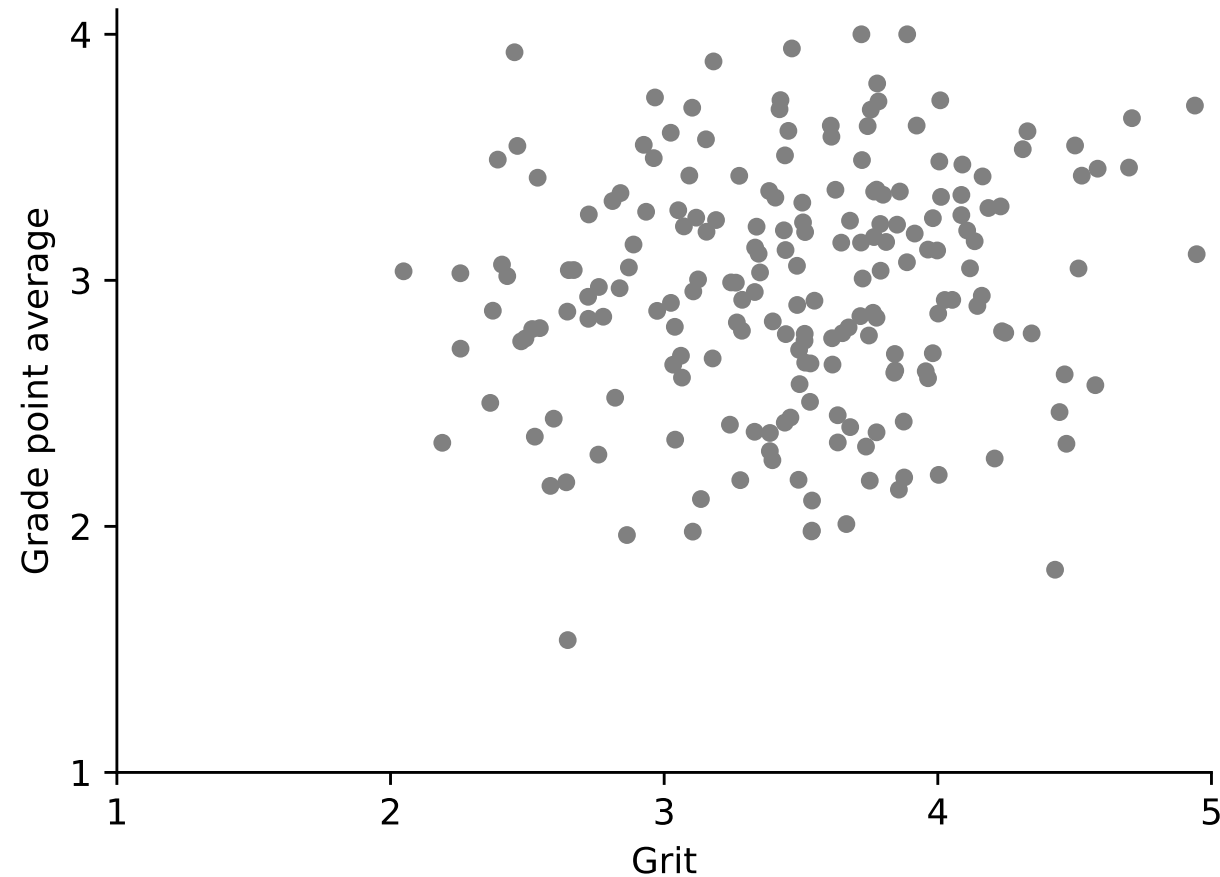
Add to list



ANGELA LEE DUCKWORTH

Grit: The power of passion and perseverance

What does 4% variance explained look like?



EKONOMI

För mycket fritid kan göra dig olycklig

UPPDATERAD 2021-09-21 PUBLICERAD 2021-09-19



Having Too Little or Too Much Time Is Linked to Lower Subjective Well-Being

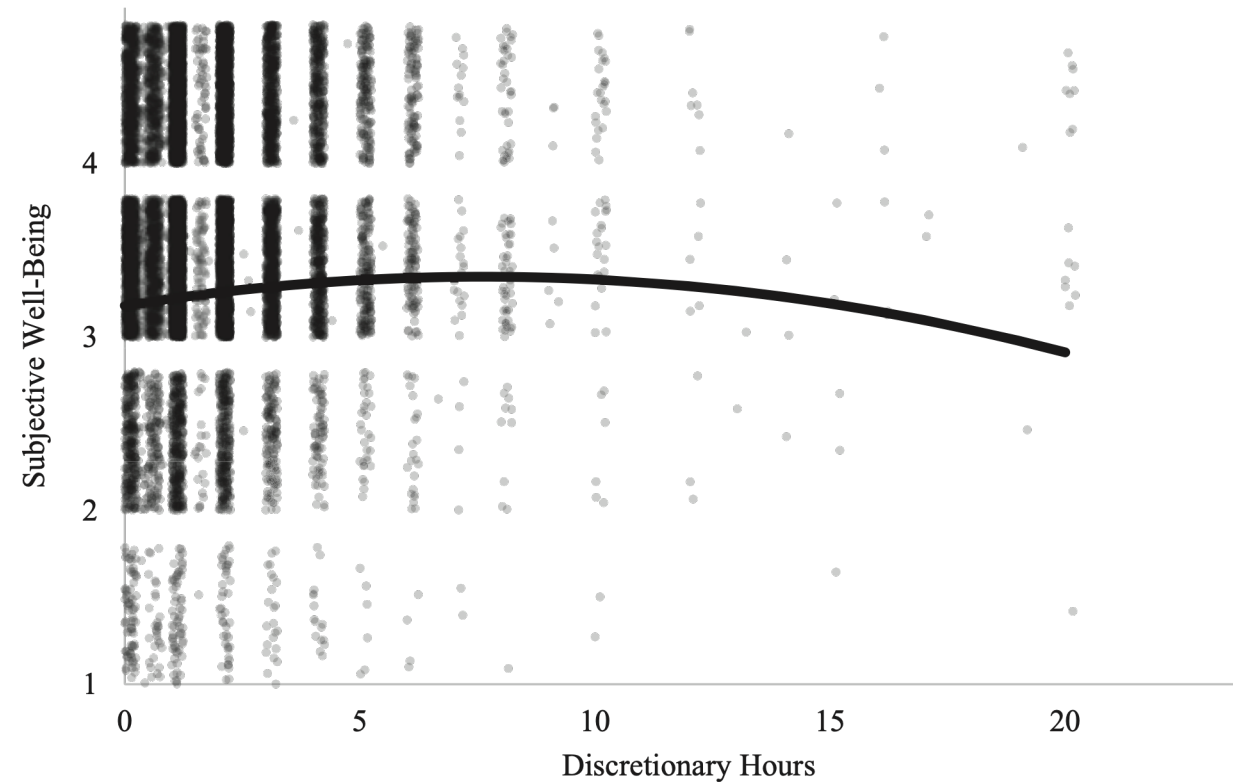
Marissa A. Sharif¹, Cassie Mogilner², and Hal E. Hershfield²

¹ Department of Marketing, The Wharton School, University of Pennsylvania

² Marketing and Behavioral Decision Making Areas, Anderson School of Management, University of California, Los Angeles

Figure 1

Scatterplot of the Relationship Between Discretionary Time and Subjective Well-Being in Study 1



Note. For ease of visualization, a jitter was added to subjective well-being scores and discretionary hours. Fit line represents the negative curvilinear relationship between discretionary time and subjective well-being.

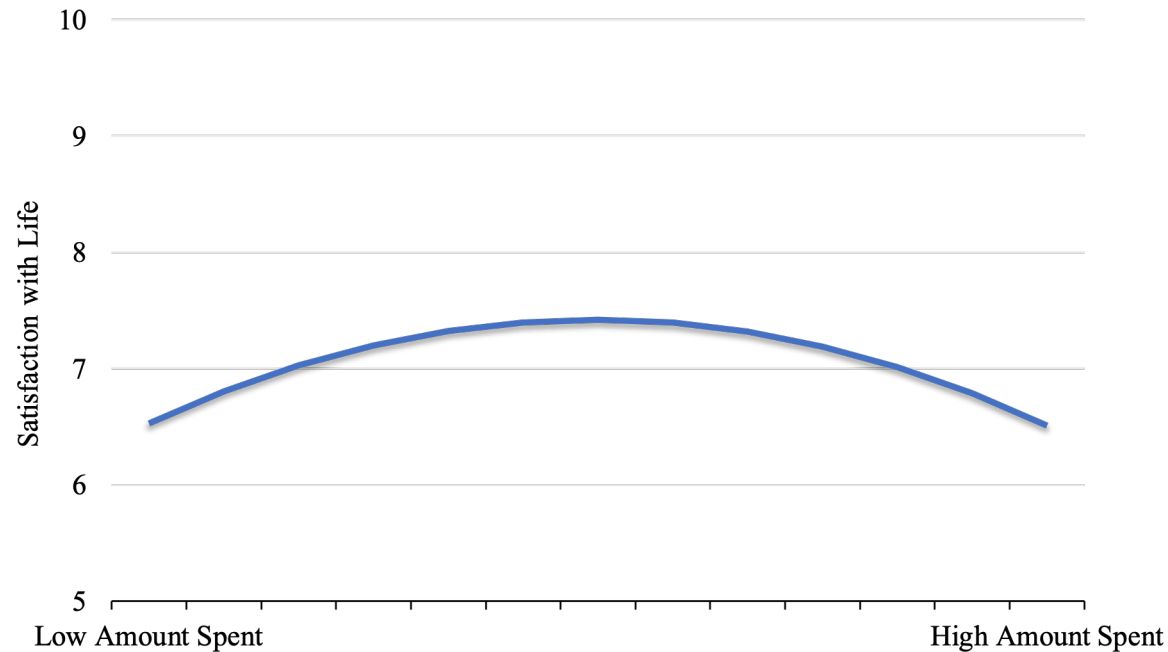
Buying time promotes happiness

Ashley V. Whillans^{a,1}, Elizabeth W. Dunn^b, Paul Smeets^c, Rene Bekkers^d, and Michael I. Norton^a

^aHarvard Business School, Harvard University, Cambridge, MA 02163; ^bDepartment of Psychology, University of British Columbia, Vancouver, BC, Canada V6T 1Z4; ^cDepartment of Finance, Maastricht University, 6200 MD Maastricht, The Netherlands; and ^dCenter for Philanthropic Studies, Vrije Universiteit Amsterdam, 1081 HV Amsterdam, The Netherlands

Figure S1

The meta-analytic quadratic effect of amount spent on time-saving purchases on life satisfaction across Studies 1-6



Note. The midpoint corresponds to spending approx. \$101 to \$200 USD to outsource disliked tasks per month. The endpoints depict ± 1 SD = \$71-\$80 USD spent to outsource per month.

Supporting Material for
Buying time promotes happiness

Ashley V. Whillans[†], Elizabeth W. Dunn,
Paul Smeets, Rene Bekkers, & Michael I. Norton

Does volunteering improve well-being?

Ashley V. Whillans^a, Scott C. Seider^b, Lihan Chen^a, Ryan J. Dwyer^a, Sarah Novick^b, Kathryn J. Gramigna^b, Brittany A. Mitchell^b, Victoria Savalei^a, Sally S. Dickerson^c and Elizabeth W. Dunn^a

^aDepartment of Psychology, University of British Columbia, Vancouver, BC, Canada; ^bSchool of Education, Boston University, Boston, MA, USA; ^cDepartment of Psychology, Pace University, New York, NY, USA

ABSTRACT

Does volunteering causally improve well-being? To empirically test this question, we examined one instantiation of volunteering that is common at post-secondary institutions across North America: community service learning (CSL). CSL is a form of experiential learning that combines volunteer work with intentional learning goals and active reflection. We partnered with an academic program that randomly assigns interested students to participate in a CSL program or to a wait-list. As part of this CSL program, students are required to engage in 10–12 h of formal volunteering each week in addition to completing related coursework. To assess the well-being benefits of formal volunteering through CSL participation, we examined the subjective well-being (SWB) of students from both groups over a six-month period. Using Bayesian statistics, and comparing a null model to a model specifying a small to moderate benefit of CSL participation, we found conclusive evidence in support of the null model. These findings diverge from previous correlational research in this area by providing no evidence for the causal benefits of volunteering on SWB. These findings highlight the critical importance of using experimental methodology to establish the causal benefits of volunteer work, such as the experiences provided by CSL programs, on SWB.

KEYWORDS

Prosocial behavior; education; well-being; college students; Bayesian statistics

When Giving to Others is Most Likely to Increase Well-Being

Behaving generously can increase happiness—but this effect is not inevitable. Instead, research has identified several key ingredients that seem to be important for turning good deeds into good feelings. Specifically, people are more likely to derive joy from helping others when:

- (1) they feel free to choose whether or how to help.
- (2) they feel connected to the people they are helping.
- (3) they can see how their help is making a difference.

10 Scientifically Proven Ways to Be Incredibly Happy

Try one. Try them all. They work.
Science says so. 

<https://www.inc.com/jeff-haden/10-scientifically-proven-ways-to-be-incredibly-happy-wed.html>

Spending Money on Others Promotes Happiness

Elizabeth W. Dunn,^{1*} Lara B. Aknin,¹ Michael I. Norton²

summed to create an index of prosocial spending ($M = \$145.96$, $SD = 306.06$). Entering the personal and prosocial spending indices simultaneously into a regression predicting general happiness revealed that personal spending was unrelated to happiness (standardized regression coefficient $\beta = -0.02$, NS), but higher prosocial spending was associated with significantly greater happiness ($\beta = 0.11$, $P < 0.01$). When we included income

Contact authors with concrete question/idea

My question is why in your work you don't tend to give (to the non-expert) an easy way of measuring the effects of the experiments with respect to themselves? For example, in the your experiment on giving to others the effect is measured by beta. This does give some measure of the effect size, but it is difficult to interpret in the context of one person.

There are two ways I can see of giving an effect size that is easier to understand. To make statements like:

- There was a X point increase in life satisfaction (on a scale of 0-10) for people in treatment A compared to those in treatment B.
- There was a X point increase in life satisfaction for paper Y cent spent on helping other people/saving time

Or, and this is my favourite, "if we chose two people at random from our study what is the probability that the person who had the greatest X also had the greatest Y?"

I realise that you are sometimes controlling for other variables, but there are ways to do this and retain the spirit of the measure.

I am asking this, because (as an applied mathematician) I find it an interesting challenge to think of ways of making statistics easier to understand. I also ask because in your writing (for example I enjoyed the article for the World Happiness immensely, with its careful discussion of causation) you emphasise the fact these studies are most relevant to the population. But I do feel that a clear interpretation of the effect sizes (in ways people can relate to in their own lives) is missing.

Paper	Studies	Link
<i>Prosocial Spending</i>		
Aknin, Barrington-Leigh, et al. (2013)	2a, 3	http://bit.ly/30hRnUX
Aknin et al. (2015)	1, 2	http://bit.ly/32aVHHD
Aknin, Dunn, Sandstrom, et al. (2013)	3	http://bit.ly/2FZBhHT
Aknin et al. (2008)	Holiday Study, Mother's Day Study, Father's Day Study	http://bit.ly/30jeheB
Aknin, Dunn, Whillans, et al. (2013)	1,2	http://bit.ly/2Xx0OTv
Aknin et al. (2014)	1	http://bit.ly/2xDzZO4
Aknin et al. (2012)	Main study	http://bit.ly/2xwmVu6
Aknin et al. (2009)	1, 2	http://bit.ly/2JjRoSK
Aknin, Van Boven, and Graham-Johnson (2015)	2	http://bit.ly/2YG5Jhu
Anik et al. (2013)	2a/b	http://bit.ly/2LI69QW
Dunn et al. (2008)	1, 3	http://bit.ly/2Jt4Hz3
Hanniball et al. (2019)	1, 2, 3, 4	https://osf.io/x3ypv/
Whillans, Aknin, et al. (2019)	1	https://osf.io/kvrhs/
<i>Buying Time</i>		
Smeets et al. (2019)	1	https://osf.io/vndmt/
Whillans, Dunn, et al. (2017)	1–8	https://osf.io/vr9pa/
Whillans, Weidman, and Dunn (2016)	1–4	https://osf.io/3zdv7/

Action plan

- I will go through the 2019 review paper data (once I have managed to open an SPSS file in Python!) and work out ways of putting 'personal happiness' measurements on the outcome of each of the studies. Liz point was very valid that, I also have to think about (if I give a percentage) how others will interpret it. It is very interesting that Cohen's d makes most sense for you. Maybe one option is to work with that idea and make it easier to interpret (visualised with overlapping distributions for example). Anyway, I will think about this.

- When it comes to feedbacks and interactions, I hadn't thought this through properly before we spoke. But I wanted to share two of my papers which are relevant. One of them is on clapping behaviour. We found that the proportion of people clapping was a very good predictor of the rate people started clapping. I suppose clapping is somehow prosocial. But I thought of this with regard to what we said about feedback. Watching a setting where you do something nice, I do something nice and so on. In the paper we used a Bayesian approach to identify social feedback.

The second paper is an experiment we did in a classroom. We went to schools across Sweden and got kids to play a game where they had to manipulate yellow and blue avatars according to certain rules. The whole point of this game is to show how small preferences on the individual level lead to large changes on the level of the group. I thought of figure 4 in particular, where we are able to measure the response of people to each other. It might be possible to set up a giving game of that sort?

So.... what I will also try to do is create a simulation, based on the measurements in your paper, of how we might expect these feedbacks to manifest themselves. In both the papers I attach (and in general in my collective behaviour research) what we do is build social interaction simulation models in parallel of the experiments in order to build up intuition.

Tools you need...

All I use now is...



But...

- Python notebooks are typically not good for public communication.
- R/Shiny is better combination for data science to app.
- Still not really good solution for quickly getting Python-based app online (correct me if I am wrong)
- And my prediction is this will change rapidly. (Powerbi)

The screenshot shows the CRUSH Covid dashboard. The header is dark purple with the logo 'CRUSH Covid' and a hamburger menu icon. A dark blue sidebar on the left contains a list of navigation items: 'Välkommen', 'Veckorapport', 'Topplistor', 'Vaccinationkartor', 'Vaccinationstäckning', 'Vaccinationbubbel', 'Trender', 'Smittspårning', 'Testkartor', 'ECDC kartor', 'Avlopp', 'Åldersfördelning', '1177 och 112', and 'Rörelsemönster'. The main content area has a light pink background and is divided into three columns. The first column, titled 'Välkommen', contains a welcome message, a description of the project, and update information. The second column, titled 'Länkar och mer information', lists links to COVID-19 information, the project's website, social media, and informational videos. The third column, titled 'Undantag och detaljer', provides details about the data, including age groups and the date of the latest update.

CRUSH Covid ☰

Välkommen

Välkommen till CRUSH Covids dashboard med interaktiv grafik. På vänster sida ser du menyn med de olika analyserna. På den här sidan kan du följa utvecklingen av covid-19 pandemin i Uppsala Län.

CRUSH Covid är ett innovativt tvärvetenskapligt forskningsprojekt i samverkan mellan Region Uppsala och forskare från fem olika institutioner vid Uppsala universitet. Syftet med projektet är att kartlägga och försöka dämpa ökad smittspridning samt kraftiga lokala utbrott av covid-19 i Uppsala län.

Webbplatsen uppdateras med de senaste tillgängliga uppgifterna varje torsdag.

Senast data uppdaterad: 2021-09-23
Nästa schemalagda uppdatering: torsdag 2021-09-30

Länkar och mer information

- Information om covid-19 och provtagning: [1177](#)
- Forskningsprojektets hemsida: [CRUSH COVID](#)
- Uppdateringar om forskningsprojektet: CRUSH Covids på [Instagram](#)
- Korta informationsfilmer om covid-19: CRUSH Covids kanal på [YouTube](#)

Undantag och detaljer

Mätarna visar information om testning för befolkningen 15 år och äldre i respektive område..

- ! Befolkningsantalet mätarna bygger på är invånare 15 år eller äldre.
- ! Baserat på befolkningens mängd 31 mars 2021.

https://crush-covid.shinyapps.io/crush_covid/

Processing

A Processing implementation of Game of Life

by Joan Soler-Adillon

Press SPACE BAR to pause and change the cell's values with the mouse. On pause, click to activate/deactivate cells. Press 'R' to randomly reset the cells' grid. Press 'C' to clear the cells' grid. The original Game of Life was created by John Conway in 1970.

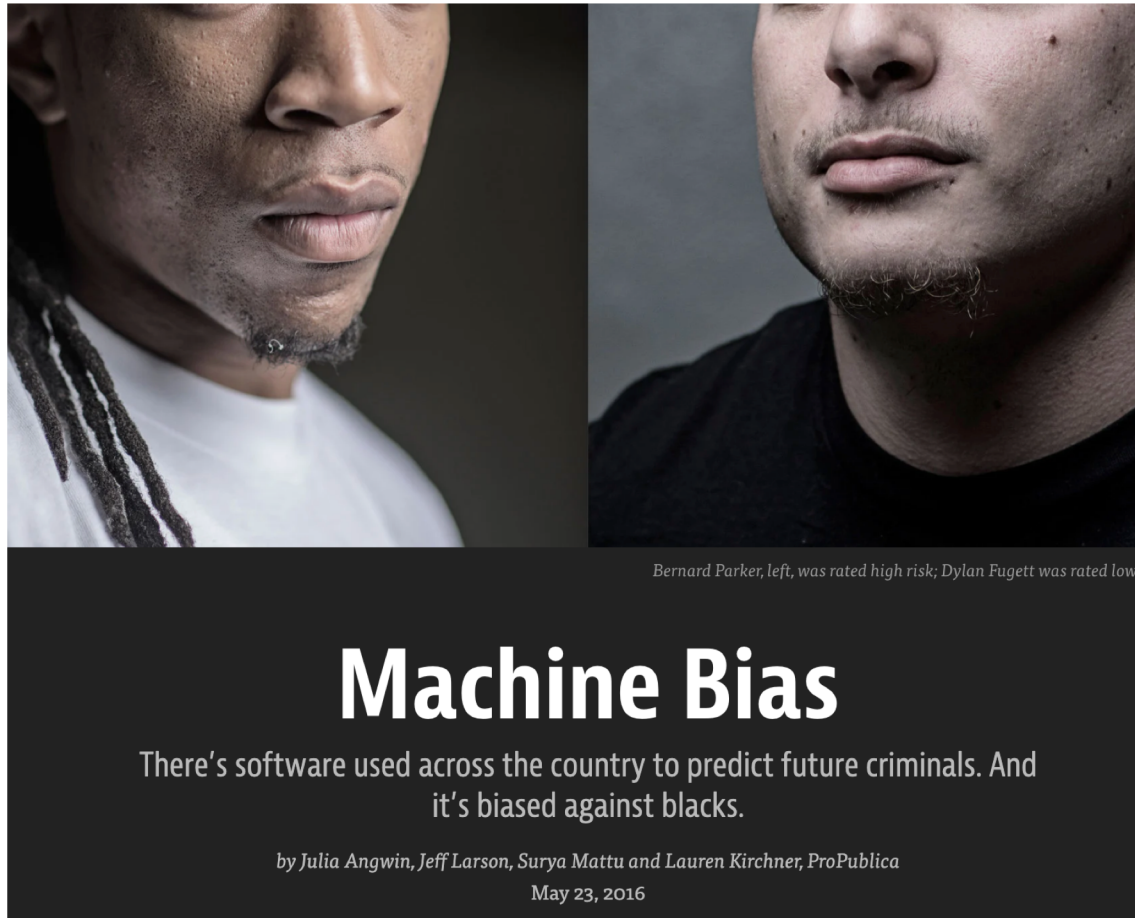
Featured functions

Array
random()
map()
constrain()
keyPressed()



<https://processing.org/examples/gameoflife.html>

Article, Analysis, Github and Colab



<https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing>

Article, Analysis, Github and Colab

Risk of Violent Recidivism Logistic Model

Dependent variable:

Score (Low vs Medium and High)

Female	-0.729*** (0.127)
Age: Greater than 45	-1.742*** (0.184)
Age: Less than 25	3.146*** (0.115)
Black	0.659*** (0.108)
Asian	-0.985 (0.705)
Hispanic	-0.064 (0.191)
Native American	0.448 (1.035)
Other	-0.205 (0.225)
Number of Priors	0.138*** (0.012)
Misdemeanor	-0.164* (0.098)
Two Year Recidivism	0.934*** (0.115)
Constant	-2.243*** (0.113)
Observations	4,020
Akaike Inf. Crit.	3,022.779

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

The COMPAS software also has a score for risk of violent recidivism. We analyzed 4,020 people who were scored for violent recidivism over a period of two years (not including time spent incarcerated). We ran a similar regression model for these scores.

Age was an even stronger predictor of a higher score for violent recidivism. Our regression showed that young defendants were 6.4 times more likely to get a higher score than middle age defendants, when correcting for criminal history, gender, race and future violent recidivism.

Race was also predictive of a higher score for violent recidivism. Black defendants were 77.3 percent more likely than white defendants to receive a higher score, correcting for criminal history and future violent recidivism.

<https://www.propublica.org/article/how-we-analyzed-the-compas-recidivism-algorithm>

Article, Analysis, Github and Colab

propublica / compas-analysis Public

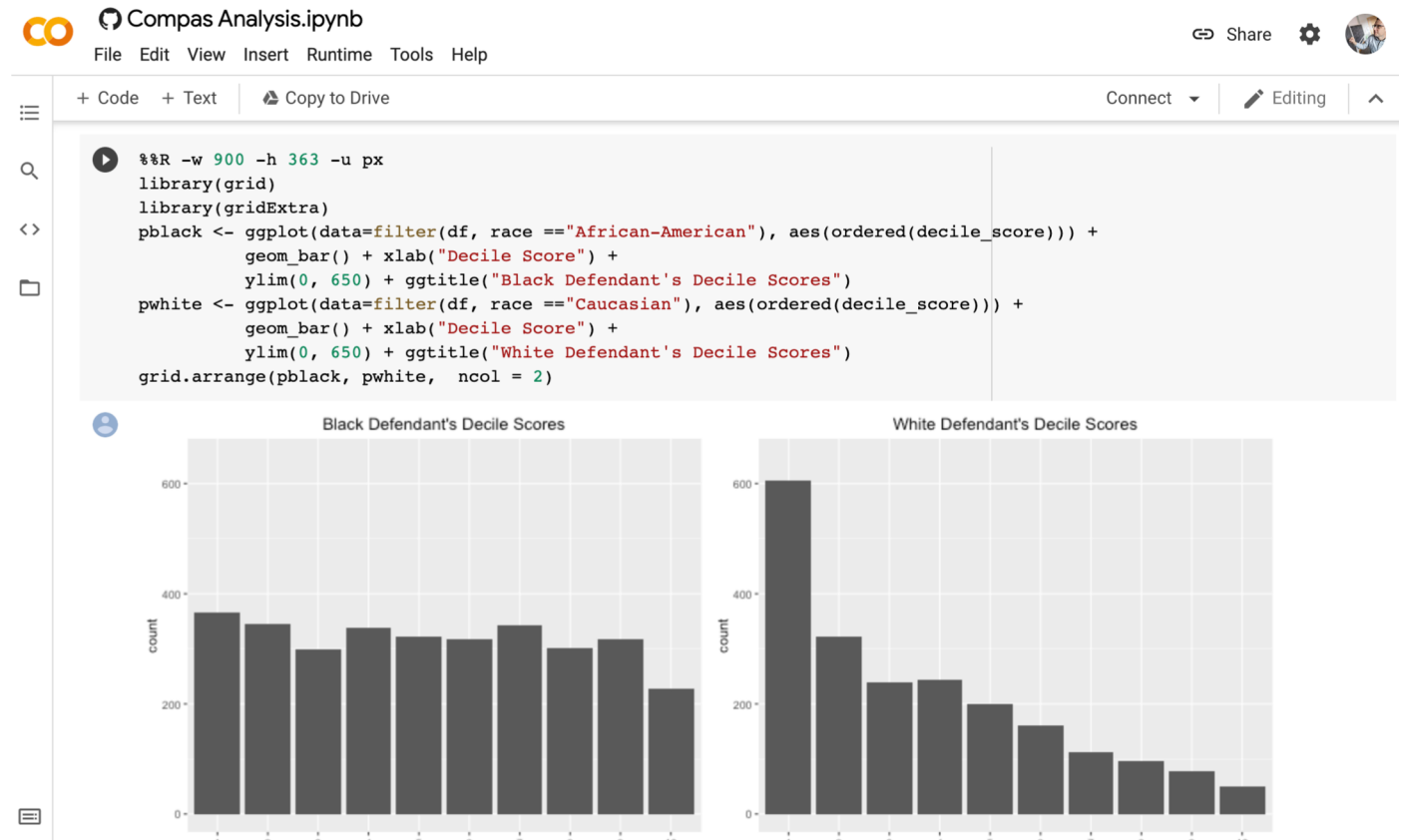
Code Pull requests Actions Projects Security Insights

master 1 branch 0 tags

Go to file Code

thejefflarson Merge pull request #4 from mroswell/patch-1 ... bafff5d on 13 Jun 2017 8 commits

.gitignore	this->sky	5 years ago
Compas Analysis.ipynb	Typo fix	4 years ago
Cox with interac	Compas Analysis.ipynb	5 years ago
README	add in links to main story and methodology.	5 years ago
compas-scores-raw.csv	raw foia data	5 years ago
compas-scores-two-years-v...	this->sky	5 years ago



Professional looking Webapp



blog · [faq/contact](#) · [toss monies at me](#) · [thanks!](#)

Hi, I'm Nicky! I make shtuff for curious & playful folks.

Wanna know when there's new shtuff? Well, the algorithms would rather show you democracy-eroding clickbait, so let's get around 'em with...



[my low-volume newsletter!](#) or even better, [let's do RSS!](#)

max 1 update per month · [see full archive](#)



SHTUFF YOU CAN PLAY



**EXPLORABLE
EXPLANATIONS**

<https://ncase.me/>

<https://www.11ty.dev/>

Data sources...

Statistikdatabasen

Uppdatering av Statistikdatabasen sker måndag-fredag kl. 9.30.

Kom igång med Statistikdatabasen

1

Välj tabell

2

Välj variabel

3

Visa tabell

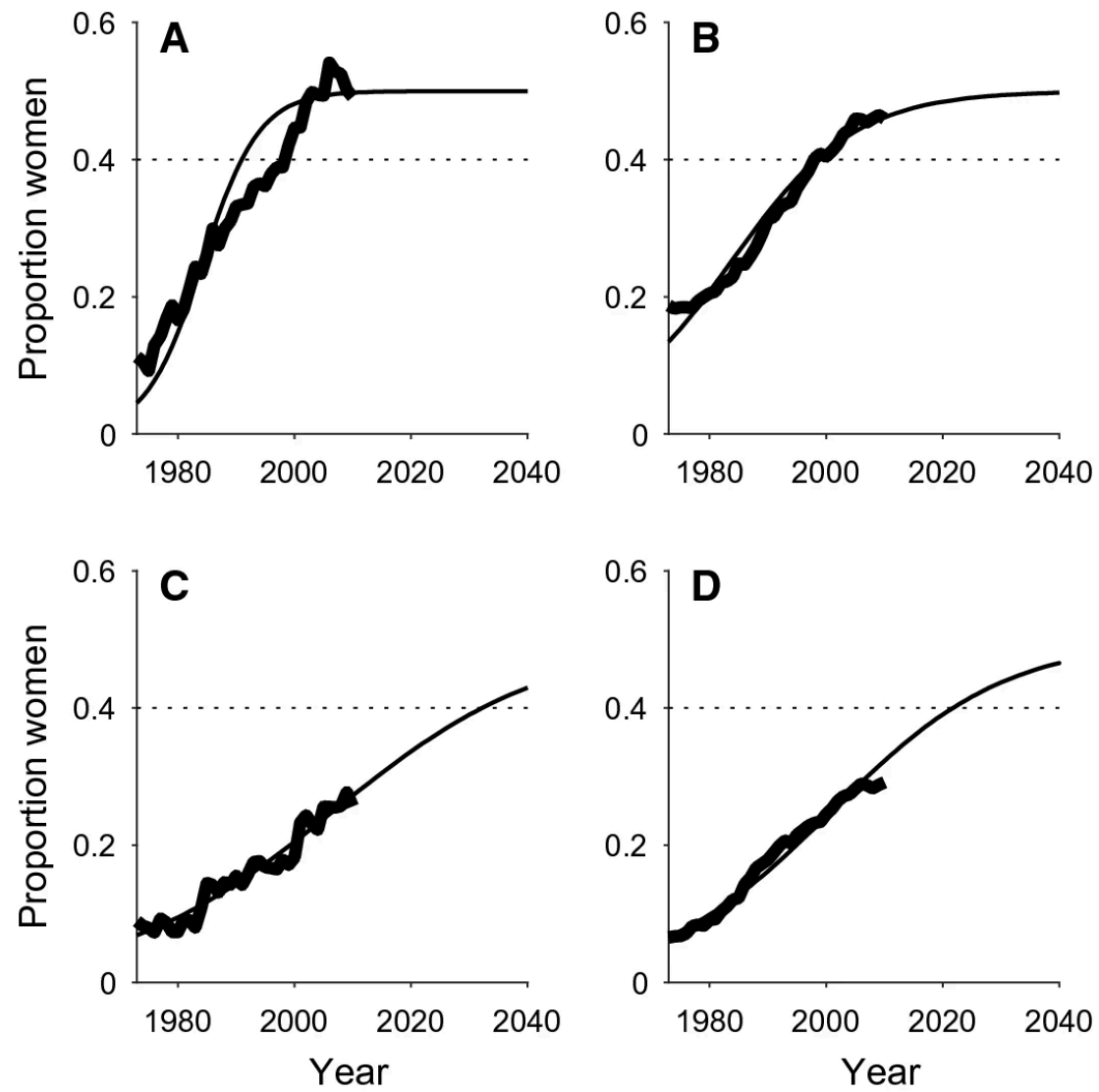
⊕ Arbetsmarknad

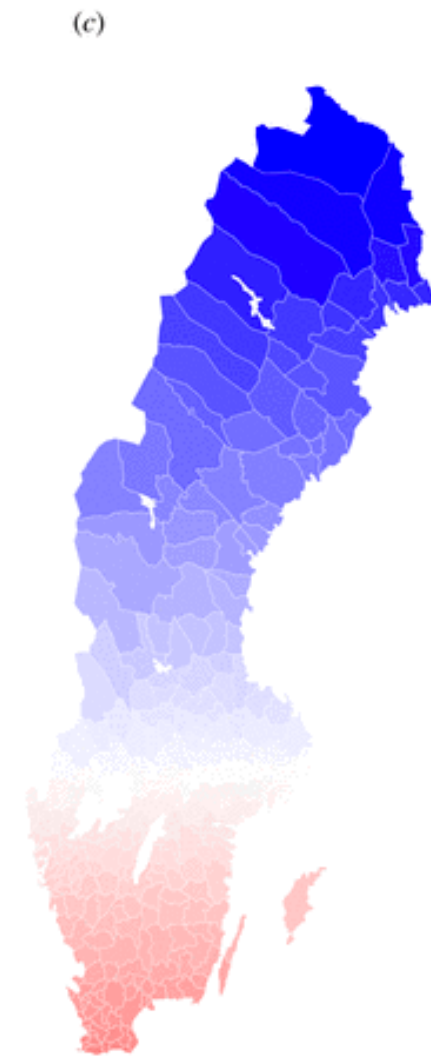
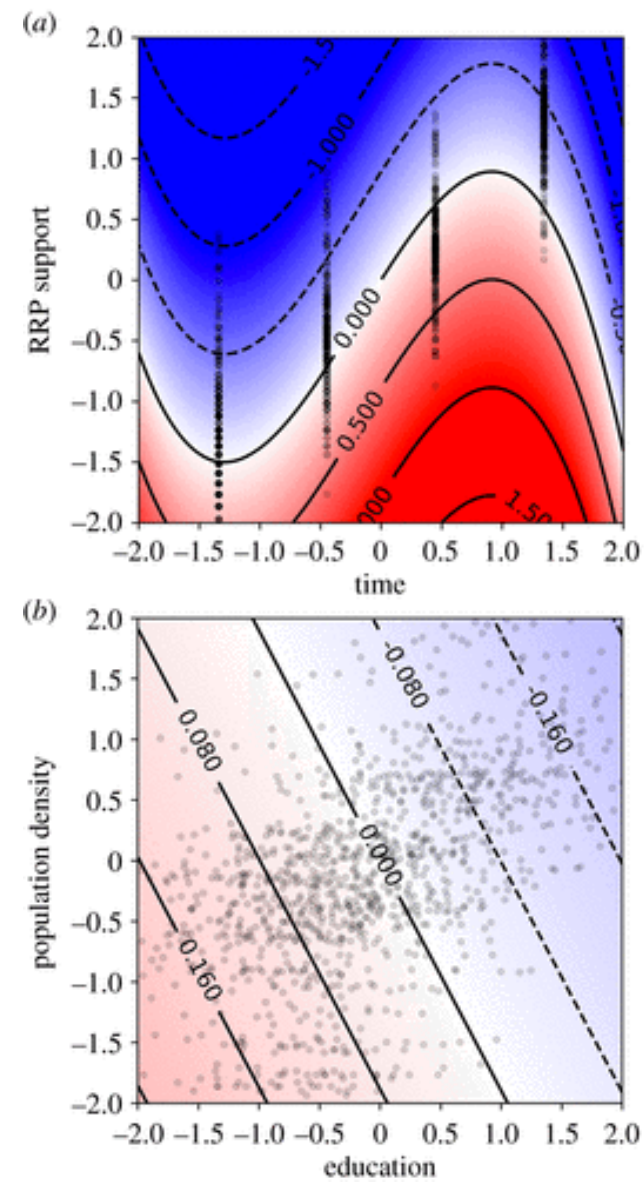
⊕ Befolkning

⊕ Boende, byggande och bebyggelse

⊕ Demokrati

<https://www.statistikdatabasen.scb.se/pxweb/sv/ssd/>





<https://royalsocietypublishing.org/doi/full/10.1098/rsta.2019.0145>

Individual level data.



A service provided by the Swedish Research Council

Registerforskning.se


Registers in Sweden

In Sweden, there are many reliable data sources for register-based research. One reason is because Sweden maintains population-based registers with personal data.

Sweden has a system of unique personal identity numbers that allows researchers to link data from different registers to a specific individual. Registers of interest from a research point of view can be divided into national public authority registers, quality records in healthcare, biobanks and research-generated data.

<https://www.registerforskning.se/en/registers-in-sweden/>

Most countries have data of this sort


 Office for
National Statistics

English (EN) | [Cymraeg \(CY\)](#)

[Release calendar](#) | [Methodology](#) | [Media](#) | [About](#) | [Blog](#)

[Home](#) | [Business, industry and trade](#) | [Economy](#) | [Employment and labour market](#) | [People, population and community](#) | [Taking part in a survey?](#)

Search for a keyword(s) or time series ID



Coronavirus (COVID-19)

[Get the latest data and analysis on coronavirus \(COVID-19\) in the UK.](#)

Main figures - [From our time series explorer](#)

Employment		Inflation	GDP
Employment rate	Unemployment rate	CPIH 12-month rate	Quarter on Quarter
Aged 16 to 64 seasonally adjusted (May - Jul 2021)	Aged 16+ seasonally adjusted (May - Jul 2021)	Aug 2021	Apr - Jun 2021
75.0%	25.0%	3.0%	4.8%

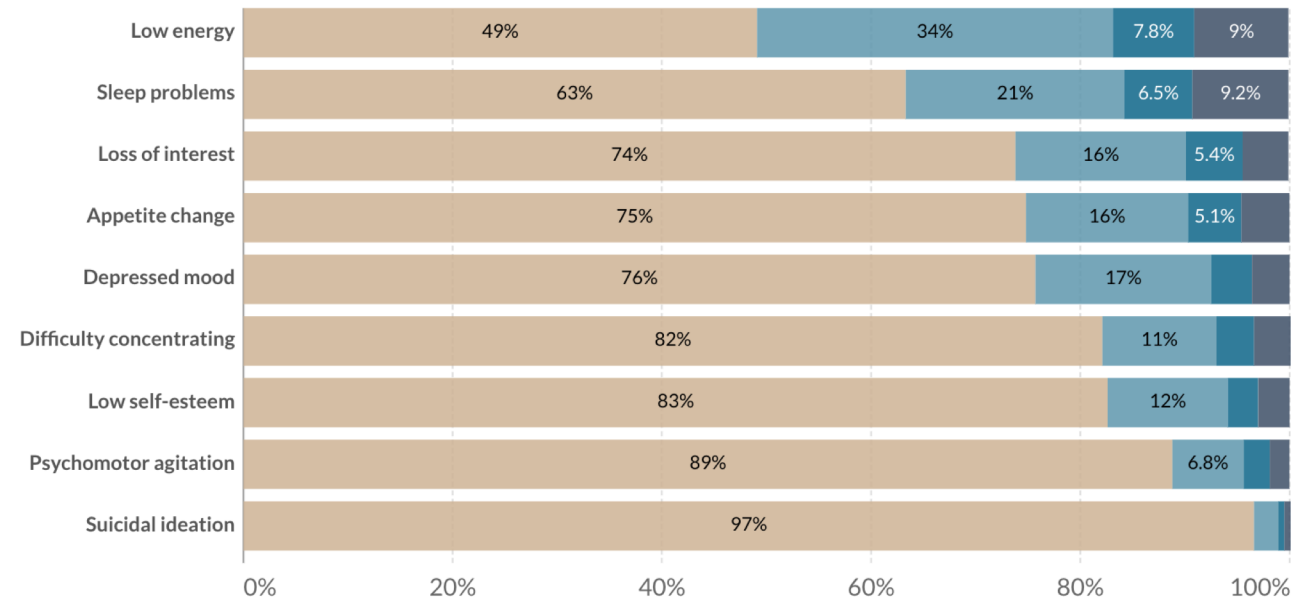
Our World In Data

Depressive symptoms across the total US population

Frequency of reported depressive symptoms from the 2013-14 National Health and Nutrition Examination Survey in the United States (NHANES). Respondents were asked about how frequently they had these symptoms in the previous two weeks.

Our World
in Data

Not at all Several days More than half the days Nearly every day



Source: Tomitaka et al. (2018). Distributional patterns of item responses and total scores on the PHQ-9 in the general population: data from the National Health and Nutrition Examination Survey.
[OurWorldInData.org/mental-health](https://ourworldindata.org/mental-health) • CC BY

<https://ourworldindata.org/>

And many more.....

Find the person with the data and ask them...

Need to consider data handling. Personal data etc.

Collect the data yourself...

Need to consider **(even more carefully)** data handling. Personal data etc.

A few thoughts on ambition
level...