

Group I

Vincent Wahid Ajoubi	viaj@itu.dk	BSc Data Science
Edi Begovic	edbe@itu.dk	BSc Data Science
Høgni Jacobsen	hoja@itu.dk	BSc Data Science
Mourad Baptiste Karib	mouk@itu.dk	BSc Data Science

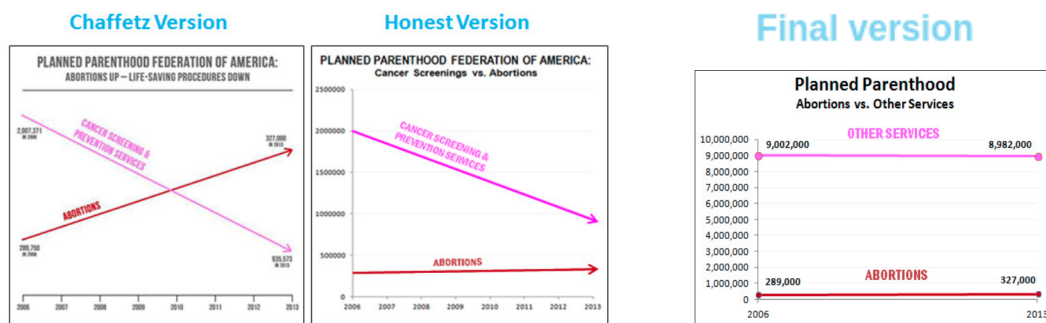
Reflections on Data Science 2021 Journal

Mourad Baptiste Karib: Planned parenthood is not trying to save lives	1
Mourad Baptiste Karib: Michigan motorcyclists deadly accident's numbers decline since helmet mandate-repeal	2
Høgni Jacobsen: COVID-19 measures vs Flu measures	3
Høgni Jacobsen: The Philippines doing well with its vaccination rollout?	4
Vincent Wahid Ajoubi: Television Viewing and Aggressive Behavior During Adolescence and Adulthood	5
Vincent Wahid Ajoubi: Women remain married if they marry after the age of 26 with stable finances.	6
Edi Begovic: Turn off video on Zoom and help the climate	7
Edi Begovic: Kids got fatter due to inactivity during lockdown	8

Mourad Baptiste Karib:

Planned parenthood is not trying to save lives

During a [hearing](#) [1] at the parliament on the results of planned parenthood, Jason Chaffetz, a representant of the parliament, used a [chart](#) [2] [3] to show that the planned parenthood unit increased the number of abortions at the expense of life saving procedures (graph on the left). Here we are going to show the problems with his claim and propose other visualizations that are more accurates.



<https://www.motherjones.com/kevin-drum/2015/09/lying-charts-anti-abortion-edition/>

Firstly, the y-axis is quite problematic or actually missing. The slopes of the two lines are not proportional. One line goes from 289'750 to 327'000 (1.13) whereas the other goes from 2'007'371 to 935'573 (0.46).

Secondly, they sum up "life saving procedures" to only cancer screening, when planned parenthood actually has a broad spectrum of life saving services for their patients, like for example STD testing and contraceptive services.

The chart Jason Chaffetz uses comes from "America United for Life" (AUL), an anti-abortion organisation, which explains some of the reasons for these biases.

In this example, there is an obvious example of lying with visualization. It was quite easy to spot since the y-axis is missing and the numbers are really different. But that's not all, we also have the problem of unfair comparisons. The intent is to compare abortions with life-saving procedures, but they disregard many other procedures that can save people lives.

The results of these false claims are that a national organisation that provides essential services to many people could lose funding just because some people have a negative opinion on abortion and have a biased perspective on the services offered by planned parenthood.

Sources

[1] [Videos of the hearing](#)

[2] [Previous blog post with the graph \[down\]](#)

[3] <https://twitter.com/Bound4LIFE/status/648873097454579712>

Mourad Baptiste Karib:

Michigan motorcyclists deadly accident's numbers decline since helmet mandate-repeal

In this [article](#) [1] written in the Michigan Capitol Confidential in April 2016, the author claims that the repealing of the mandate making helmets mandatory in Michigan, led to a decrease of deadly motorbike accidents.

They selected some specific statistics to support their claim.

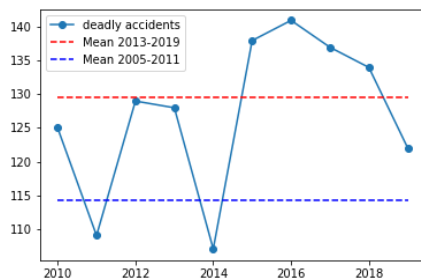
"Between 2005 to 2011, while the helmet mandate was still in effect, an average of 114.2 people were killed each year in Michigan motorcycle accidents and 2,757.6 injured."

The state repealed the mandatory helmet law in 2012.

In 2012, 129 people were killed in Michigan motorcycle accidents and 2,870 injured. Those numbers dropped to 128 and 2,497 respectively in 2013. In 2014, the most recent year for which data is available, 107 people were killed and 2,309 injured."

Based on these numbers, they support the repealing of the mandate making helmets mandatory for the motorbike riders.

The main problem with this claim is that they compare the average over 10 years with data from a two year period. Their goal is to find an argument to motivate their free-market agenda, as they call it in the description of their web-site, without using any methodology to back up their arguments. It's also interesting to note that to be allowed to not wear a helmet, you need to pay more than \$20,000 extra in insurance.



I got the data from their [sources](#) [2] and used the data they had access to, to calculate the average over a longer period. If we take the years from 2012 to 2019, the average of death is actually higher now. This case is a simple case of unfair comparison where people jumped on data. I also calculated the ratio of deadly accidents to see if the helmets were saving lives in case of accidents and we can see that the ratio of deadly accidents increased since 2012.

This claim has fatal consequences since it is trying to legitimize people deciding to take the risk to ride motorbikes without helmets, when all the professionals of the field repeat the importance of wearing helmets to save life. It is mis-informing the public about health related subjects.

Sources [1][Link to Article](#) [2][Link to police data](#)

Høgni Jacobsen:

COVID-19 measures vs Flu measures

In this [article](#) [1], Dr. Cyril Wecht is claiming that COVID-19 measures are draconian and an overreaction. He makes a comparison between COVID-19 and the seasonal flu. The goal being to bring into question COVID-19 measures. The specific paragraph being “According to the website for the Centers for Disease Control and Prevention, there were 61,000 deaths attributed to influenza in the 2018-19 season. Somehow we dealt with that epidemic without closing down society.”

This is an unfair and biased (confirmation) comparison for a few reasons. Not to mention that the 61,000 deaths figure is from the year [2017-18](#) [2] and not 2018-19. Also, it is an estimate.

It is unfair as COVID-19 is far more infectious and deadly than the seasonal flu. How easily a virus spreads is the R0 score (basic reproduction number). R1 would mean an infected person will infect 1 other person. According to the Virginia Department of Health ([VDH](#) [3]) the seasonal flu has an R0 between 1 and 2 and the COVID-19 virus has an R0 of between 2 and 3, meaning COVID-19 is far more infectious than the flu. COVID-19 is also more deadly, “Both COVID-19 and flu can cause severe illness and even death, but a larger proportion of COVID-19 cases result in hospitalization or death.” - [VDH](#) [3].

The article is biased on the fact that the CDC has a comprehensive comparison between the seasonal flu and COVID-19 ([CDC](#) [4]), which also corroborates the data from VDH. The CDC is the same resource that was used to make his claim, therefore he selectively chose his data.

The consequences of this BS is that it has the potential to cause resistance to necessary COVID-19 measures. Beyond this resistance it contributes to the erosion of trust in public institutions.

Sources

[1]<https://www.post-gazette.com/opinion/Op-Ed/2020/05/14/Cyril-H-Wecht-Time-to-end-the-COVID-19-hysteria/stories/202005140031>

[2]<https://www.cdc.gov/flu/about/burden/index.html>

[3]<https://www.vdh.virginia.gov/coronavirus/2020/12/07/covid-19-and-influenza-surveillance/>

[4]<https://www.cdc.gov/flu/symptoms/flu-vs-covid19.htm>

Høgni Jacobsen:

The Philippines doing well with its vaccination rollout?

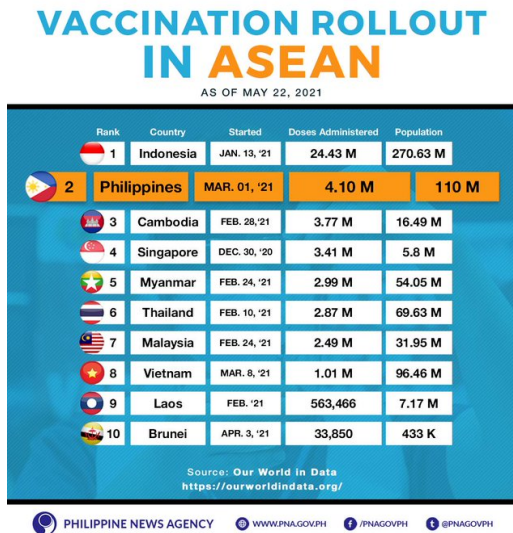


Figure 1

% of population who have received a vaccine dose in ASEAN			
	Doses Administered	Population	% of population
Singapore	3410000	5800000	58.79%
Cambodia	3770000	16490000	22.86%
Indonesia	24430000	270630000	9.03%
Laos	563,466	7170000	7.86%
Brunei	33,850	433000	7.82%
Malaysia	2490000	31950000	7.79%
Myanmar	2990000	54050000	5.53%
Thailand	2870000	69630000	4.12%
Philippines	4100000	110000000	3.73%
Vietnam	1010000	96460000	1.05%

Figure 2

With reference to this Twitter [post](#) [1] by the Philippines News Agency. The underlying claim in Figure 1 is that the Philippines is doing well with its vaccine rollout and is in 2nd place in the Association of Southeast Asian Nations (ASEAN).

While the data in the visualization is correct, I would argue that the amount of doses administered is an irrelevant metric, as in lecture 1 “Lying with Visualizations”, slide 33, graph titled “Number of Drivers in Fatal Crashes, 1988”.

A better way to represent the standing of the Philippines would be to represent the doses administered as a percentage of the population, as in Figure 2, however this shows an undesirable picture. This visualization would place the Philippines at 9th place with 3,73% of the population receiving a dose (Assuming 1 dose per person).

Making a visualization like this hides the true state of the vaccine rollout in the Philippines and is an example of lying with visualizations.

The consequences could be that it instills a false sense of security in the public which might lead to undesirable behaviour in times of a pandemic. In some ways this could also act as propaganda for questionable governments, as in the same way it instills a false sense of security during the pandemic it can also lend credibility where none is due.

Sources

[1]<https://twitter.com/pnagovph/status/1396805779848679424?s=20>

Vincent Wahid Ajoubi:

Television Viewing and Aggressive Behavior During Adolescence and Adulthood

Kids who watch TV for more than 1 hour a day are more than twice as likely to commit an act of violence in adulthood, compared to those who watched less than 1 hour a day, claims this [post](#) [1] based on a 17 years long [study](#) [2] on 707 families.

Previously similar studies claimed that watching violent televised contents contributes strongly to a child's behavior, however the study claims that the amount of time spent is a big factor. Risk factors such as childhood neglect, residence area, low family income, low parental education and psychiatric disorder were taken into account but the content shown on the TV was not recorded. 88 kids, 14 to 22 years old, randomly chosen to watch less than an hour and 619 people watched over 1 hour. Given the data they claim that there's a significant association between the amount of time spent watching television during adolescence and early adulthood and the likelihood of subsequent aggressive acts against others.

To analyse the result of this study on these two groups, we conducted a null hypothesis examination ([Journal.ipynb](#) [3]). We receive a z-score of -4.04 which is far out in the left tail. Given that we get a p-value of $(2 \times (0.00003))$ which is significantly lower than 0.05, considering it as our significance level. We conclude that there's extremely strong evidence against the null hypothesis and therefore we reject it and accept the alternative hypothesis.

This was an observational study. In practice, it would be impossible to conduct an experimental study by randomly assigning teenagers to watch little or much TV over several years. That itself might be a reason to provoke a kid over several years of repetition. Also, spending more time on TV does not necessarily mean watching more violence. A very strong factor was ignored in this paper.

First of all, these types of posts exaggerate strongly the content of such a study which should not be around, or at least be ignored. Second, these types of papers and posts might trigger parents to follow the pattern. They unintentionally might put a grudge on their kids for not letting them watch TV. Besides, there's a reason that all televised contents are rated upon age. Many programs are broadcasted on TV just for the sake of good for kids. Of course anything more than a threshold would be bad for anybody.

Sources

[1]<https://www.newscientist.com/article/dn2109-tv-viewing-linked-to-adult-violence/>

[2]https://www.researchgate.net/publication/11440839_Television_Viewing_and_Aggressive_Behavior_During_Adolescence_and_Adulthood

[3]<https://github.com/edibegovic/ReflectionsDS/blob/main/Journal.ipynb>

Vincent Wahid Ajoubi:

Women remain married if they marry after the age of 26 with stable finances

According to [Time](#) [1] "women who get married for the first time while they're very young or over the age of 45 are more likely to see their union end in a divorce than others, statistics show" and women who get married after 26 are more likely to have a successful marriage.

The most important issue is that the study has taken just women into account., while a divorce is the result of two people not satisfied with their current life. Men should have been included and studied as well. Personality factors such as mental health, upbringing or class are all correlated with marital quality and could be a crucial factor. If we divide graph [3] [2] into three sections, we have *teenage marriages* at the top (red), middle-age marriages (green), and *older marriages*(red) at the bottom of the graph.

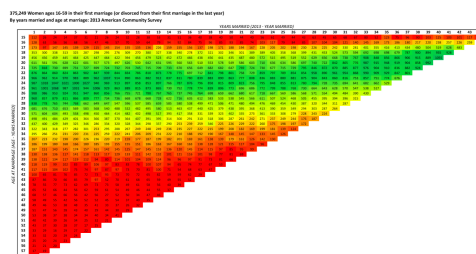
1 - Economy is a vital factor for a successful marriage. There's a high chance that women in their late 20's or early 30's have a high income job in comparison to women in their late teenage years or even early 20's.

2 - Kids and family affairs from previous marriages could contribute to a divorce. These previous issues such as step families, mental health, occupation and x-partners could be a contributing factor to divorce, marriages without these issues might be more successful.

3 - There is too little data. They selected a subset of 375,249 women in the 2013 American Community Survey (ACS) public use file, ages 15-59, which is very selective. How about other years, or even older women!

They have conducted an unfair comparison because the way this study has been carried out includes selecting data that fits the intended message (filtering men out). Furthermore, there is no way we can check that the different age groups have the same representation. Some data-sets might have different sizes and therefore bias the study.

This claim is misinforming since it is not following a clear methodology to base their claim. It might put pressure on women who want to get married when actually it's not their responsibility, just because of misinformation like this article.



[3]

Sources

[1]<https://time.com/4358792/woman-age-married-how-long/>

[2]<https://familyinequality.wordpress.com/2015/07/22/how-we-really-can-study-divorce/>

Edi Begovic:

Turn off video on Zoom and help the climate

[Article by DR](#) [1] (in Danish) on the negative environmental impact of having video on during online conferencing. The article is based on a [recent study](#) [2] which generally covers the impact of increased global internet traffic.

There are many layers of issues in the article. Firstly, they reference the upper limits. While you can save up to 1kg/hour of CO₂, the estimate is probably closer to 1/8th of that, as stated by the study. But that doesn't matter, because they shouldn't use absolutes in the first place. Nobody knows what the impact 1kg of CO₂ has relative to their other daily activities. Lastly, if one knew the scale of the footprint, it would be clear that it is negligible. Much of the variance in a personal climate footprint can be attributed to other factors, such as using an oven or showering for an extra 30 seconds.

The issue is mostly a matter of the magnitude (Fermi estimation). The life cycle of a plastic bag emits ~10g of CO₂_(equivalent); a metal straw ~10² g; incandescent lightbulb (24h) ~10³ g; Oven (2h) ~10⁴ g; flight (4h) ~10⁵ g. Thus, one can quickly estimate the impact of a month's non-stop video-conferencing and compare that to, for example, running an oven for an hour. While everything stated is technically true — you do in fact reduce your climate footprint by turning video off — so does holding your breath for 10 seconds. The point is that the effect size is negligible.

While DR is nuanced in its coverage, including comments from professors at both ITU and DTU — both arguing it's not a big deal — this requires people to actually read the article. The narrative implied by the title and subtitle is clearly different. It places a kind of shame on the reader. The same study has also been cited and propelled using the same narrative by many other news outlets. The consequences of these kinds of articles is the spread of [greenwashing](#). Specifically the 5th sin of greenwashing, '*sin of irrelevance*'. While a claim may be truthful, it is unimportant or unhelpful for consumers seeking environmentally preferable products/actions. It contributes to the consumer skepticism of all green claims and diminishes the power of the consumer acting in truly impactful ways.

Sources

[1]<https://www.dr.dk/nyheder/viden/klima/hjaelp-klimaet-og-sluk-kameraet-uden-video-udleder-dine-onlinemoeder-96-procent>

[2]<https://www.sciencedirect.com/science/article/pii/S0921344920307072?via=ihub>

Edi Begovic:

Kids got fatter due to inactivity during lockdown

[TV2](#) [1] earlier this year reported that school children had gained significantly more fat over the previous year. They claimed this was due to the reduction in physical activity because of the COVID-19 lockdowns.

There are multiple reasons as to why this claim is shaky in its foundation (as also covered by [Detektor DR](#) [2]), including the sample size only consisting of eleven 8th grade children from a single school. While the study that the article was based on did conclude abnormalities in the development of some children, this was attributed to the higher-than-normal blood pressure and HR — not weight or body-fat percentage (BF%) as stated by the article. While there was a substantial difference in both weight and BF% between initial and last measurements, this could simply be explained by the fact that 8th grade children naturally undergo a lot of physical changes at their age.

The tool, if you will, would be a combination of multiple hypotheses and confirmation bias. A lot of people have speculated what the impact of the lockdown inactivity would have on their health. Thus it's easy to hypothesise (and get confirmed) that the change in weight and BF% of kids must be due to the inactivity, and not the more obvious reason, that kids grow.

At a greater scale, the societal impact of the above claim is arguably not that big. It can lead parents to “blame” children for natural processes assuming that it is behavioural. This might lead to physiological consequences for children such as eating disorders.

Sources

[1]<https://nyheder.tv2.dk/samfund/2021-02-19-skoleelever-toq-knap-8-kilo-paa-under-nedlukning-vi-risikerer-at-de-blive-syge>

[2]<https://www.dr.dk/nyheder/detektor/detektor-kun-11-boern-blev-maalt-og-vejet-da-tv2-fortalte-om-en-langt-fra-normal>