

Deception in Data Visualization

by

Kyle Kenneth James Adams

A thesis

presented to the University of Waterloo

in fulfillment of the

thesis requirement for the degree of

Master of Arts

in

Philosophy

Waterloo, Ontario, Canada, 2022

© Kyle Kenneth James Adams 2022

Author's Declaration

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Abstract

In recent years, there has been a great deal of public discussion of misleading graphs and statistics, but this phenomenon seems to have fallen outside the scope of philosophical analysis. The philosophy of language has produced many well-developed accounts of lying with words, but so far, has not given much attention to deception with data visualizations. This thesis is a first step towards filling that gap. I consider several existing accounts which might be expected to handle cases of lying and misleading in data visualizations, and I argue that none of them do an adequate job of analyzing these kinds of lies.

The first account that I consider is from Edward R. Tufte's *The Visual Display of Quantitative Information*, wherein he describes principles for integrity in data visualizations. I argue that these principles do not constitute an adequate account of deception in data visualization, because they do not reflect the widely recognized lying-misleading distinction, and also his prescriptions are too vague to constitute an effective definition.

I also consider accounts of lying and deception from philosophy of language, from Thomas Carson, Jennifer Saul, and Andreas Stokke. Each of these three accounts also falls short of adequately addressing lying and misleading in data visualization. In analyzing why this is, the question of what is said by a graph becomes crucial, and ultimately, turns out to be an extremely difficult question to answer. There are several confounding factors in trying to understand what is said by a graph, and therefore when it lies. The most important, I suggest, is that what is said by a graph is massively pluralistic—each contains many utterances that it can be thought of as saying. Although Stokke's account allows for flexible contextual supplementation in response to the structure of the discourse in which the utterances take place, and therefore seems to stand the best chance of addressing this issue, it does not allow for the sort of true pluralism that is necessary here. Ultimately, I conclude that none of the existing literature is suited to handling this problem.

Acknowledgements

I am grateful to have been a graduate student at University of Waterloo's Department of Philosophy. I have received extensive support, both academic and financial, from the department, and it is deeply appreciated. First and foremost, I would like to thank my supervisor, Prof. Jennifer Saul, for her invaluable feedback and guidance as this project transformed from a scattered collection of notes into a thesis. I would like to thank Prof. Patricia Marino and Prof. Carla Fehr for their generous and helpful revisions. This thesis would not be anywhere near its current state without the three of you, and your insights. Thanks is also due to Prof. Marino for setting me on the path towards graduate school in philosophy in the first place.

I also am very grateful for support from my family and friends. This would never have been possible without the support of my parents, and I am very thankful. I am thankful to my dear friends Tyler, Maddy, Joe, and Jyler, for patiently humouring me while I slowly turned barely connected thoughts into a coherent argument, often out loud, and in their presence. You're the best friends a scatter-brained graduate student could ask for. I would also like to thank Morgan, Isabella, Stella, Josué, John, Julien, Alex, Nancy, Hannah, Susannah, and Andrea. Our adventures in D&D, and your friendship, have been invaluable during the process of writing this thesis.

Finally, I want to thank Mr. Andy Kerr, for teaching me the brief motto that went on to become the motivation for this project: "the only time someone will ever tell you that 'the numbers never lie' is when they are about to use numbers to lie to you." I did not, back in grade twelve, imagine that the introduction to the Data Management class would turn out to motivate my graduate research, but it seems to have done so.

To all these people, and many more, I am very thankful.

Table of Contents

Author’s Declaration	ii
Abstract	iii
Acknowledgements	iv
List of Figures	vii
Chapter 1 Introduction.....	1
1.1 Data Fabrication	3
1.2 Adjusting Axes.....	4
1.2.1 False Starts	4
1.2.2 Scale Manipulation	5
1.2.3 Unconventional Design Choices.....	7
1.2.4 Multiple Axes and Axis Hiding.....	8
1.2.5 The Audience Question.....	9
1.3 Norms of Data Visualization—Edward R. Tufte	11
Chapter 2 Lying, Misleading, and Deception	19
2.1 Thomas L. Carson	19
2.1.1 Carson’s Account of Lying	20
2.1.2 Carson’s Accounts of Deception and Misleading.....	24
2.2 Jennifer Saul.....	26
2.2.1 What is Said.....	30
2.3 Andreas Stokke	33
2.3.1 The Stalnakerian Model of Discourse and Questions Under Discussion.....	34
2.3.2 What is Said.....	35
2.3.3 The Common Ground	37

2.4 The Promised Complications	38
2.5 Bullshit.....	39
2.6 Conclusions.....	40
Chapter 3 What Does the Graph Say?.....	41
3.1 An Indeterminate Utterance	47
3.2 Plurality.....	48
3.3 Does it Say Anything at All?.....	50
3.4 Damned Liars, All of Them	50
3.5 By Whom and To Whom.....	51
Chapter 4 Analysis of Examples	53
4.1 FOX News	54
4.1.1 Carson	55
4.1.2 Saul	56
4.1.3 Stokke	58
4.2 Planned Parenthood Services	61
4.2.1 Graphical Bullshit.....	63
4.2.2 The Most Basic Lie.....	65
Chapter 5 Conclusion.....	67
References	69

List of Figures

Figure 1: A graph I made up.....	3
Figure 2: A FOX News graph of tax rates (Hickey 2012).....	4
Figure 3: A tweet from the White House with a graph of economic growth (White House 2022).....	6
Figure 4: A graph of gun deaths in Florida (Lallanilla 2014).....	7
Figure 5: A graph of some Planned Parenthood services (Qiu 2015)	8
Figure 6: A cumulative sales graph presented by Tim Cook (Yanofsky 2013).....	10
Figure 7: A graph without enough context	16
Figure 8: Possible contexts for the graph in Figure 7.....	16
Figure 9: Data from Statistics Canada. Table 17-10-0009-01. Population estimates, quarterly	41
Figure 10: A graph that is bullshit (Statistics How To).....	64
Figure 11: Fabricated Data	65

Chapter 1

Introduction

Data visualization offers powerful tools of analysis and demonstration, but these tools can be very easily put to nefarious use. Despite the prevalence of deceptive graphs in public and private discussion, attempts to work towards a systematic understanding of these graphs are startlingly rare. Every first-year student in a statistics course has some basic familiarity with what might make a graph misleading or otherwise deceptive, but this information has most often been presented as lists of red flags: things to watch out for when making or reading a graph.

Although anti-misleading checklists of this sort are undoubtedly valuable in inoculating people against graphical misinformation, they do not tend to go far enough in identifying what it is to be misleading, or lying, or deceptive, for a graph. In the first chapter of this thesis, I will introduce a variety of nefarious graphs, and discuss the best existing attempt to address the problems of deception in graphics: *The Visual Display of Quantitative Information*, by Edward R. Tufte, a statistician and professor emeritus of political science, statistics, and computer science. Tufte lays out a series of principles of graphical integrity, which outline some of the hazards that await the creator of a graph.

I will argue that even Tufte's account does not suffice to deal with the problem of deception in data visualization. So, since even the best of the statistics-inspired accounts does not adequately address the problem, another angle should be taken. A great deal has been made of the particulars of lying, and of the distinctions between lying and misleading, in the literature of the philosophy of language. In the second chapter of this thesis, I will discuss several recent accounts of lying and deception. The goal of this analysis will be to see how these accounts operate, and which of them might be suitable for applying to graphs. Because these accounts all originate within the philosophy of language, they are specifically focused on utterances: usually spoken, and sometimes written. Thus, it is not always immediately clear how to apply these definitions to graphs, which are largely

visual displays. Most of my focus in this chapter will be on three accounts: from Thomas Carson, Jennifer Saul, and Andreas Stokke. It will also be necessary to introduce some supporting literature, as well. This includes Harry Frankfurt's essay *On Bullshit*, linguistics work from Craige Roberts that features prominently in Stokke's account of lying, and some complications about recorded utterances from Alan Sidelle and Stefano Predelli. The goal of this chapter will be, by necessity, largely expository, and will lead to the conclusion that in order to make sense of whether or not a graph lies or misleads, it will be necessary to develop a robust conception of what is said by a graph.

Determining what is said by a graph is the focus of my third chapter. Here, I will apply the content of several ideas of what is said, drawn from the accounts of deception in Chapter Two, and attempt to use them to make sense of what is said by, or with, a graph. As I will argue, there are a great many complications in attempting to determine what is said by a graph. I will show that every graph communicates a *staggering* plurality of propositions, and that it may be necessary to embrace this plurality in order to make some kind of sense of what is said. This plurality, as it turns out, poses a serious challenge to making sense of what is said by a graph. I will also explore direct applications of the various theories from Chapter Two, and identify some particular shortcomings from them.

In Chapter Four, I take the work I have done from the preceding chapters and apply it to several instances of misleading and lying graphs. These graphs allow me to demonstrate more precisely the problems that arise in an effort to directly apply the philosophy of language definitions to the analysis of graphs. Ultimately, I argue that existing accounts of lying, misleading, and other forms of deception do not offer sufficient tools for understanding deception that takes place in data visualizations.

In order to motivate the content of the later chapters of this thesis, I will first present a survey of some graphs which are, in some way or another, deceptive. The purpose of this section is simply to lay out intuitions, about what sorts of graphs might be misleading, and what might be an outright lie,

as well as to give a sketch of some of the types of ‘red flag’ intuitions that are often used in assessing whether or not a graph is deceptive.

1.1 Data Fabrication

The most direct way to deceive with a graph is to simply fabricate a bunch of data and stick it in a graph. Want to ‘prove’ that women are taking over the world, driving men into abject poverty? All-too-easily done!

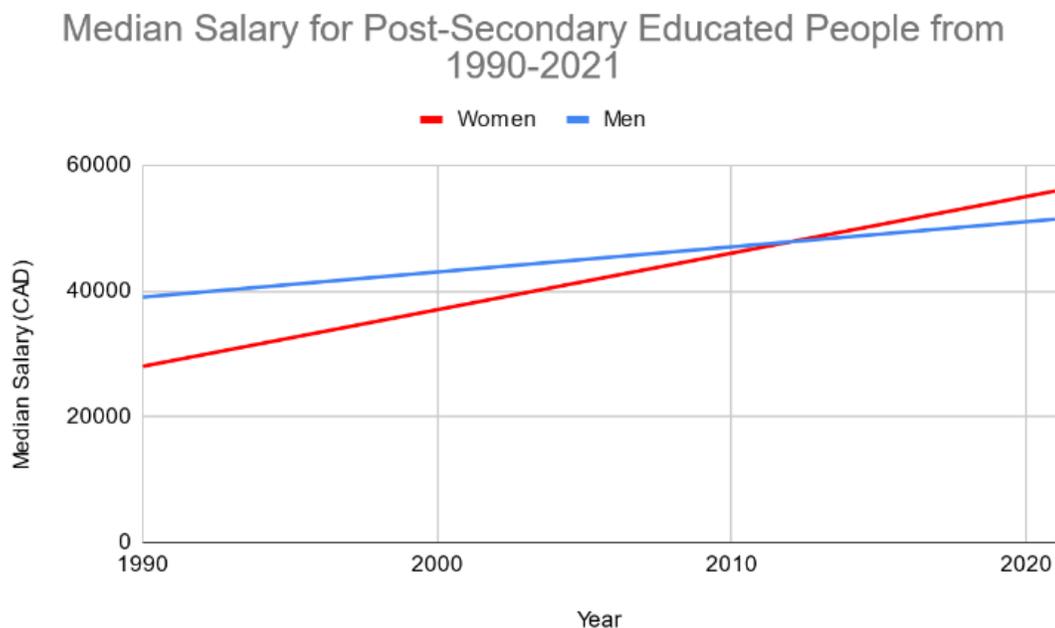


Figure 1: A graph I made up

Now, lo and behold, a graph assures the truth of the claim. That the data are not accurate need not concern the author of the graph, because a large part of the audience is highly unlikely to verify those figures. It is often said that the numbers never lie. There is, I believe, no expression that is more indicative of impending falsehood. This sort of behaviour seems to be a very close analogue to a direct lie. The ‘author’ of the graph presents a claim that they do not believe to be true (those

believing that a lie must in fact be false would tack on that criterion as well—as the author of this graph I assure you that the data is in fact false).

1.2 Adjusting Axes

Adjusting the axes of a graph is a classic means of presenting true information in a misleading way. This sort of graphical deception can be done in a wide variety of ways, which merit their own brief discussions.

1.2.1 False Starts

One of the most direct is a non-zero start. Consider the following example from FOX News:

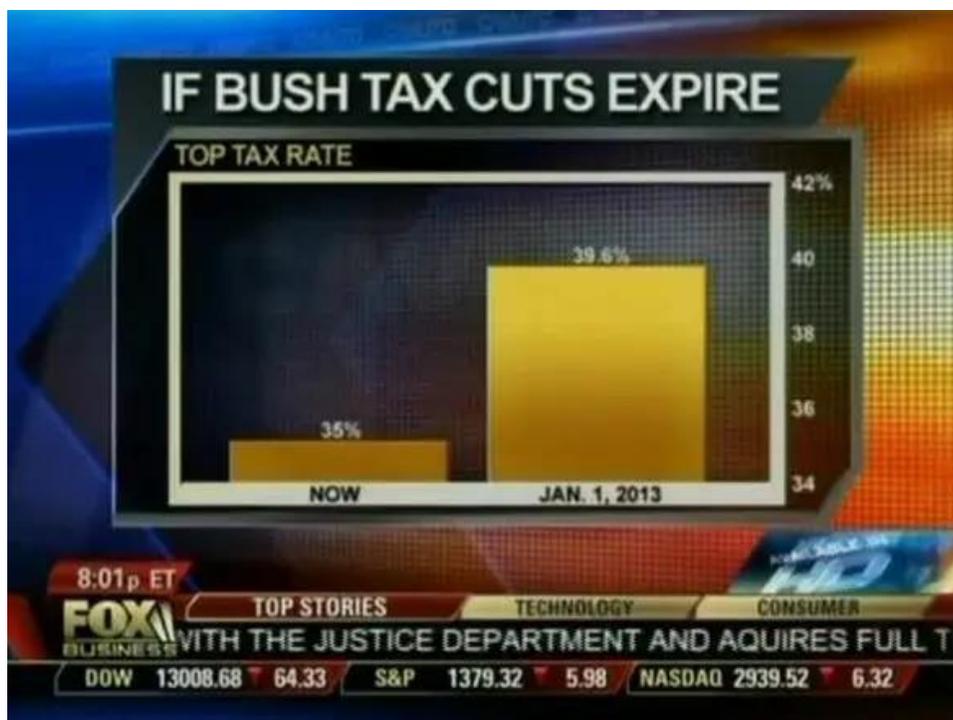


Figure 2: A FOX News graph of tax rates (Hickey 2012)

At a glance, it seems that allowing the Bush tax cuts to expire will result in a huge spike in taxes, increasing roughly by a factor of six (a 500% increase over their 'current' rate). Closer

inspection, however, reveals that this is not the case. The ‘NOW’ bar corresponds to a top tax rate of 35%, and the ‘JAN. 1, 2013’ bar corresponds to a top tax rate of 39.6%. Far from what it seems, this is an increase of only 4.6 percentage points: a roughly 13.1% increase. How can this ever be accurate? Well, examination of the vertical axis reveals that it has its origin not at 0, as would be customary, but at 34. By omitting the bottom portion of both bars, and including only the smallest amount necessary to indicate that are really two bars here, the designer responsible for this graph presents a hugely exaggerated picture of the impact of the tax cuts expiring (this is compounded by the fact that what they are showing is only the *top* tax rate, and therefore this rate is not even going to apply to folks who are not already very rich, but a breakdown of tax code is rather outside the scope of this paper).

1.2.2 Scale Manipulation

When looking at a graph, especially on a small phone screen, especially when simply scrolling around on Twitter or Instagram, how many of us really take the time to dig carefully into the details of what that graph is communicating? I submit that it’s not very many. Most folks take a look at the picture, get a feel for the rough trends, and move on. That is what makes the following example such an effective piece of trickery. These sorts of graphs are relatively common in all manner of politically-charged contexts: consider this recent graph from the official Twitter account of the Biden administration:

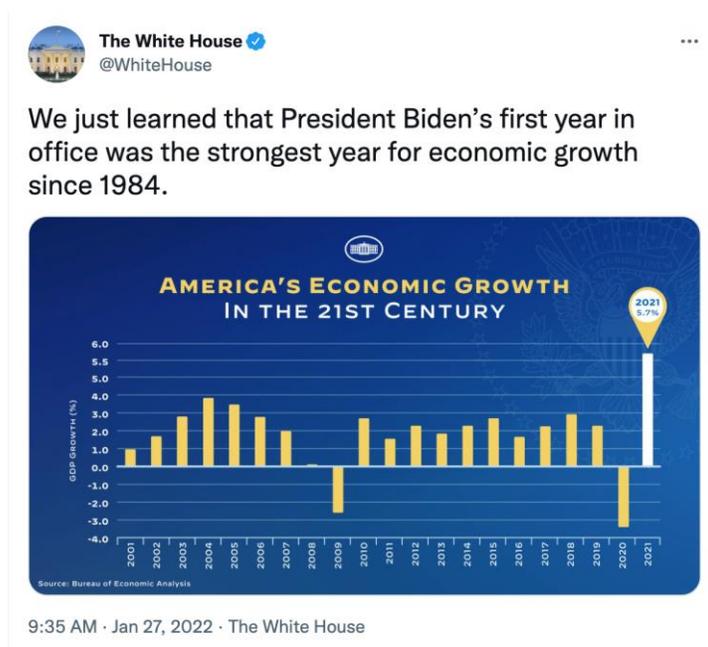


Figure 3: A tweet from the White House with a graph of economic growth (White House 2022)

Blown up to this size, it may be easier to notice the flaw. There is a clearly marked origin at zero, and the data are correct. But looking closely at the vertical axis, there is some subtle trickery going on. Starting from 0.0, it counts up in unit increments until 5.0, at which point its next increment is 5.5. This switch of incrementation partway up the graph serves to exaggerate the extent to which 2021 was year of biggest economic growth. I also have some reservations here about choice of statistic: presenting the percentage growth really goes a long way to making Biden look better, because it is inflated by the preceding year's decrease, which pretty clearly has more to with the pandemic than with anything else. It also does not really seem to acknowledge the impact of the pandemic in any specific way. By making the only relevant consideration about Biden as president, important background information has been left out.

1.2.3 Unconventional Design Choices

In order to discuss the type of examples that I wish to discuss, it will be helpful to begin with a motivating instance, because the actual practice is very strange. Consider the following graph:

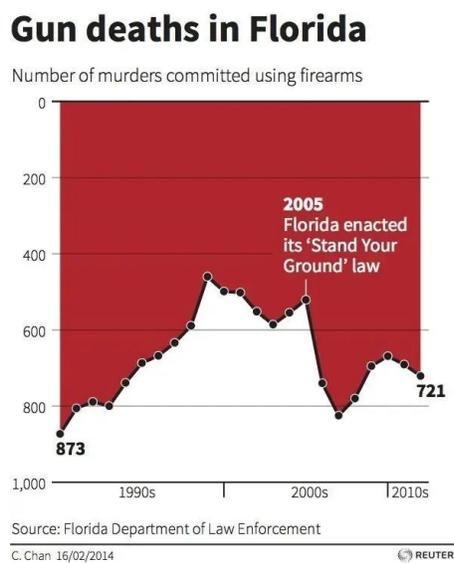


Figure 4: A graph of gun deaths in Florida (Lallanilla 2014)

The casual observer will look at the dramatically coloured graph, see the flagged and labelled year 2005, in which Florida enacted its ‘Stand Your Ground’ law, and then note a remarkable drop in the number of firearms murders the very next year. Only, that’s not what is happening.

A closer look, taking care to notice the scale on the left-hand side of the graph, reveals that the *axis is inverted*. The lower the data point in 2D space, the higher it is in the natural numbers. This choice to flip the axis, running strongly in the face of conventions for graphical design, seems like it is almost certainly an act with deceptive intent, trying to send a dishonest message about the effects of the Stand Your Ground law.

Once again, though, the analysis gets a little bit murkier, because Cynthia Chan, the creator of the graph above, was apparently not trying to deceive anyone. She was, instead, motivated by an

infographic she had seen, which portrayed the death toll of war using visualization that resembled blood, dripping down the display (Lallanilla 2014). In an attempt to create a similar effect, the graph above had its vertical axis flipped, obscuring its accurate (and intended) message. This phenomenon strikes me as something interesting about graphs of this sort: they are misleading against the wishes of their creators.

1.2.4 Multiple Axes and Axis Hiding

The following graph is, I believe, a lie.

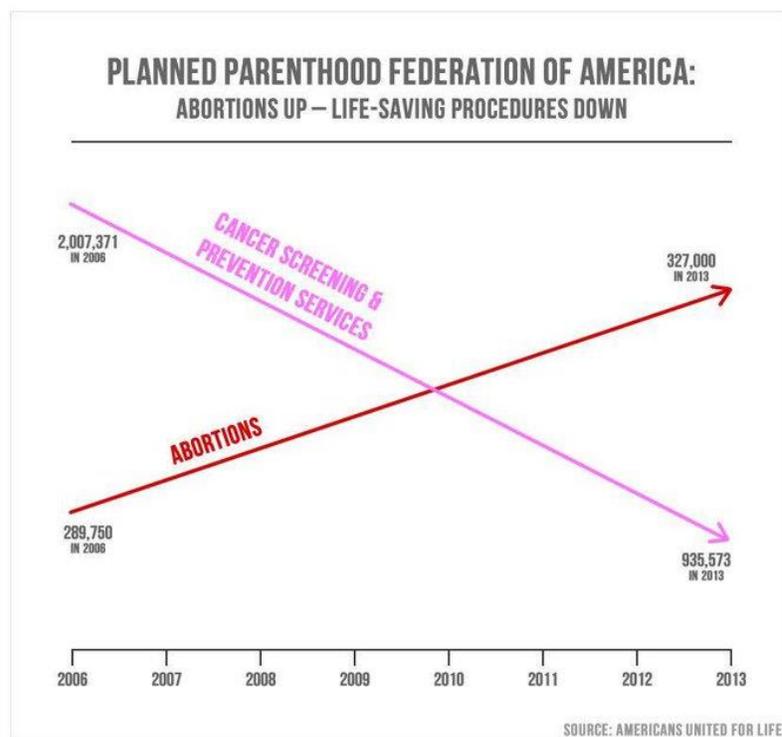


Figure 5: A graph of some Planned Parenthood services (Qiu 2015)

The graph in Figure 5 is one of the most nefarious graphics that I have found in my research so far. Here we see a horizontal axis ranging from 2006 to 2013. Two lines, one pink and one red, and displayed. The pink line is labelled “Cancer Screening and Prevention Services” and runs from

2,007,371 in 2006 to 935,573 in 2013. The red line is labelled “Abortions” and runs from 289,750 in 2006 to 327,000 in 2013. The graph angrily proclaims that life saving procedures are down, and abortions are up. Now, setting aside for a moment that they are placing abortions *in contrast* with life-saving procedures—a deeply horrible move that merits some examination as a deceptive act of speech in its own right—I turn to focus on the construction of the graph itself.

There is something immediately visible about this graph: namely that its two lines intersect, suggesting that the number of abortions has not only increased, but in fact *surpassed* the number of cancer screening and prevention surfaces. That is clearly false, from the very labels on the graph. 935 573 is, after all, considerably greater than 327 000. So, although it is true that the raw number of abortions has increased, and the raw number of cancer screening and prevention services has gone down, the graph is nevertheless at *best* deeply misleading, and indeed I would say perniciously so, given the extent to which this is a violation of graphical conventions. My strong intuition is that in at least many cases, this graph is in fact a lie.

This graph obscures the truth with a two-step approach. First, it uses two distinct vertical axes: one to count abortions, and one to count cancer screening and prevention services. There are sometimes good reasons to use two vertical axes, such as when comparing correlations of quantities that operate on vastly different scales. This is a legitimate move at best when a graph is meant to compare *correlation*, not raw number. Since the graph in Figure 5 is already not about correlation, but really is comparing raw numbers, it has already failed to make the use of dual axes appropriate. Moreover, the graph in Figure 5 goes even further by *omitting* the axes.

1.2.5 The Audience Question

One additional complication that arises when considering questions of lying and misleadingness as they pertain to graphs is the fact that graphs, unlike utterances, are persistent.

Therefore, the audience and context of utterance of a graph is not necessarily clear-cut. Many of them have audiences that change with time, and are differently equipped to make sense of the possibly misleading implications of the graph presentation.

A graph that I take to be a striking example of this is the following, from a presentation by Tim Cook of Apple:



Figure 6: A cumulative sales graph presented by Tim Cook (Yanofsky 2013)

Here, Cook presents one visualization of iPhone sales data (it seems, as far as I can tell, to be broken down by sales quarters, although only the year is marked on the horizontal axis. This graph also does not include a vertical scale, but from context it is reasonably safe to assume it starts at 0 and counts up in some increment, probably of millions of units). At a glance, this graph makes it look like business is booming. Apple is thriving, just look at those sales!

Only, wait a minute. Those are *cumulative* sales, and that doesn't quite mean sales. To somebody not in the know about business mathematics, or with a little bit of calculus background, this graph is hiding a secret. We get the actual per-quarter sales from this graph not just by reading it off, but by taking the first derivative of the cumulative sales curve. To some people, this graph is clear

as day: last quarter was a noticeable fourth-quarter drop in sales. To others, it misleads them into believing that sales are constantly on the rise (and of course, cumulatively they are, because you can't ever have a decrease in cumulative sales).

1.3 Norms of Data Visualization—Edward R. Tufte

What the preceding sections have shown is that in many cases of misleadingness and deception in graphs, there is some sort of norm that has been violated. This might be the origin values of an axis, the orientation of that axis, or any number of other things. Thinking of deception in terms of norm violation, or a list of 'red flags' for graphs, is an approach that is commonly taken among mathematicians, data analysts, and statisticians. To give this type of account the best possible chance of success, I turn to *The Visual Display of Quantitative Information* by Edward R. Tufte. Tufte's work in this book is generally quite excellent, and provides a lot of guidelines for the design of data display graphics. He tends to use the terminology of 'lying' more or less exclusively, which I intend to move beyond, but the book provides some terrific insights. In his second chapter, titled *Graphical Integrity*, he lays out six principles. I present them numbered, but otherwise they are as written by Tufte himself.

- (1) The representation of numbers, as physically measured on the surface of the graphic itself, should be directly proportional to the numerical quantities represented.
- (2) Clear, detailed, and thorough labeling should be used to defeat graphical distortion and ambiguity. Write out explanations of the data on the graphic itself. Label important events in the data.
- (3) Show data variation, not design variation.
- (4) In time-series display of money, deflated and standardized units of monetary measurement are nearly always better than nominal units.

(5) The number of information-carrying (variable) dimensions depicted should not exceed the number of dimensions in the data.

(6) Graphics must not quote data out of context. (Tufte, 77).

I will now go into some detail about each of these principles, in order to make them as clear as possible.

I begin with Tufte's (1) and (2). These two principles are intended to offer some sort of standardization in the presentation of graphics, in order to limit the variation in perception to factors coming from the perceiver, rather than the designer as well (Tufte, 56).

He uses the first principle to introduce another core concept in his analysis of graphical integrity, which he terms the "Lie Factor" (Tufte, 57). Tufte defines the Lie Factor as the quotient of the size of the effect shown in the graphic with the size of the effect in the data (Tufte, 57).

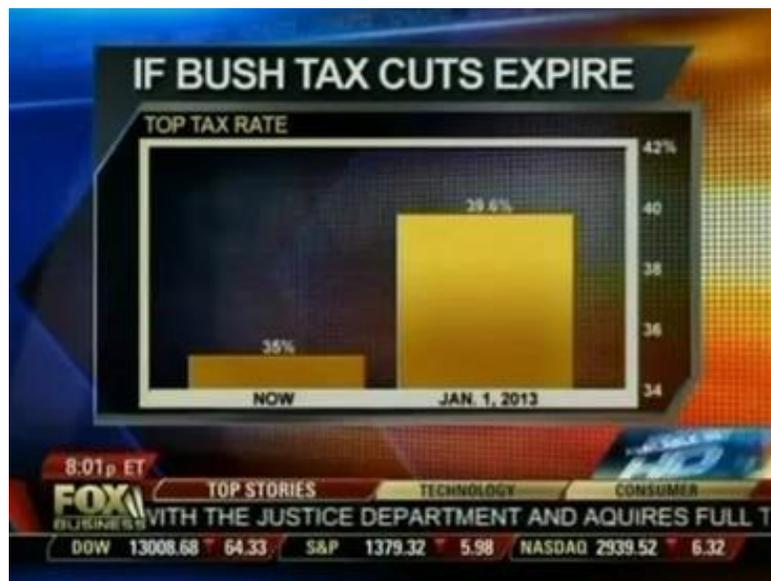


Figure 2

For example, considering the graph from FOX News from Figure 2, the change in actual top rates is from 35% to 39.6%, which represents a percentage increase of roughly 13%¹. The bars of the graph go from 6mm to 32mm in height, with constant width, which represents a percentage increase of roughly 433%. This, then, results in a Lie Factor of approximately 33.3. The Lie Factor returns throughout the chapter as a regular feature of analyses by Tufte.

Tufte’s (2) requires somewhat less explanation. Clear and descriptive labels are a perfectly reasonable thing to expect from graphics.

Principle (3) is a bit of a difficult one to fully explain, because it is not entirely clear what precisely Tufte takes to be a design variation. He provides many examples, but no system or precise definition. He writes that “the eye may mix up changes in design with changes in the data,” which gives some idea of the sorts of things he may mean. The two examples that he uses to first introduce this principle are time-series graphics that change their time intervals partway through.

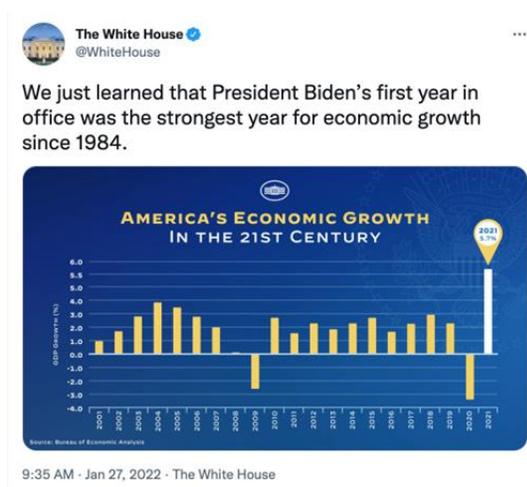


Figure 3

¹ Percentage increase here must not be confused with the increase in percentage points, which is 4.6%. The 13% figure is the result of the calculation $\frac{39.6\% - 35\%}{35\%} \times 100\%$, taken to two significant figures.

My Figure 3, the graph showing the strong economic growth under Joe Biden's first term as President, displays a similar fault. Partway up its vertical axis, it switches from unit intervals to half-unit intervals, which serves to exaggerate the size of the Biden Administration's economic growth. The examples that Tufte gives seem to all suffer from internal design variation, and so this principle seems like it might not address cases in which the graph's whole design is somehow deviant from the norms of graph design, like my Figure 4, which has a wholly inverted vertical axis.

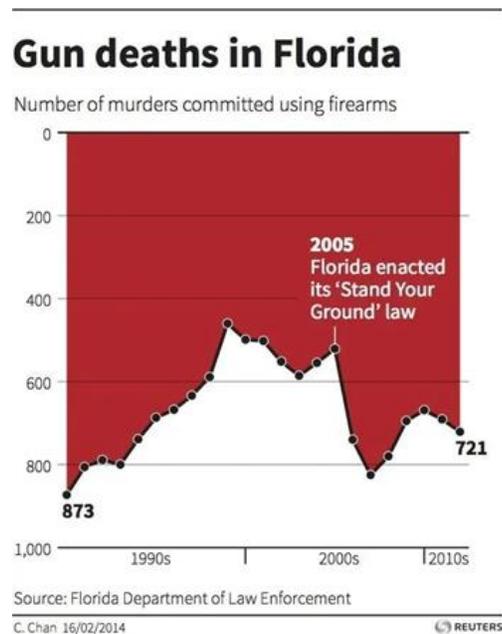


Figure 4

Principle (4) is to say that a standardized unit of currency, correcting for inflation and possible deflation, is preferred to using fluctuating values of currency. I find this just about entirely uncontroversial. There are perhaps some rare cases in which the nominal (i.e. pre-adjustment for inflation) costs are more interesting for analysis, but in general this is not the case. Common baselines of measurement are hugely important; we are all familiar with the old adage about the folly of comparing apples with oranges.

Tufte's principle (5), that the number of variable dimensions should not exceed the dimensionality of the represented data, is perhaps his most mathematically dense. It is closely related to his concerns about the Lie Factor. Consider a graphic which represents the GDP of nations using circular marks, where the radius of the circle is given by the GDP in trillions. Now suppose that Anland has a GDP of \$1 trillion, and Belmere has a GDP of \$2 trillion. In this graph, Anland will be represented by a circle with area π , but Belmere will be represented by a circle with area 4π . The area of the second circle is four times larger, instead of twice as large.

This may sound like a somewhat esoteric construction, but is in fact fairly common, especially when it comes to money. Often graphics that depict inflation, or the purchasing power of a currency over time, do so using representations of bills in that currency, scaled with side length proportional to their purchasing power. This results in visualizations that overrepresent periods where currencies are highly inflated, sometimes by quite significant ratios.

Finally, Tufte's principle (6), that graphics must not present data out of context, seems to me to be all but impossible to actually adhere to. On a charitable reading, his principle is only against excluding crucially relevant context, but even this strikes me as being a very lofty bar. There is, by sheer virtue of necessity, always context that will be omitted from any graphic, since any attempt to include all possible context would result in a graphic so mind-boggling complicated as to be entirely useless. Tufte takes the core question behind all quantitative analysis to be "compared to what?" and he emphasizes the importance of context for this question to be answerable (Tufte, 74). Although this principle is, in itself, dreadfully vague, examples might help to illuminate what Tufte is describing as the problem. In what follows, I have recreated a slight variation on the graphs that Tufte uses to develop his own account. Consider the following graph:

Yearly Profits of Green Team, Inc. Before (2005) and After (2006) New Regulation

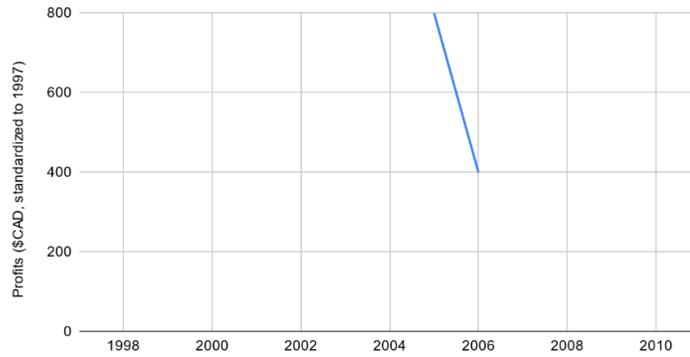


Figure 7: A graph without enough context

This graph offers very little context for anything. There is what looks like a sharp decline in profits, from \$800 to \$400. But depending on the added context, this sharp decline could tell any of a set of very different stories. Consider the following graphs, where more context has been provided around the one-year change in regulations on Green Team Inc.

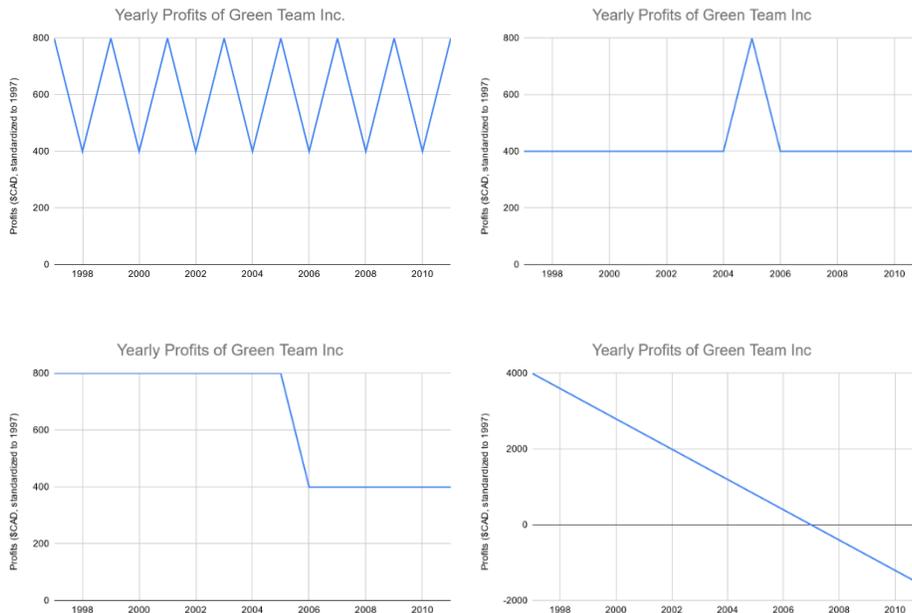


Figure 8: Possible contexts for the graph in Figure 7

Although each of the graphs has the exact same drop, from \$800 to \$400 between 2005 and 2006, the different contexts around that drop lead a reader to wildly different interpretations of what that drop represents.

Tufte's approach is typical of those generally taken by statisticians and mathematicians. Every student who has received some basic education in statistics, data management, or numerical reasoning will be familiar with certain red flags for graphics that indicate a potential area of misleadingness.

Ultimately, I do not believe that Tufte's principles do an adequate job of capturing the details that I hope to. That is, admittedly, not their intent. Tufte himself says only that non-deceptive graphs are "more likely to result if these six principles are followed," and does not make any sort of claim to have created a set of necessary and sufficient conditions for a graphic to be deceptive (Tufte, 77).

Moreover, Tufte's analysis, not coming from a language angle, is somewhat simplistic in its categories. There are a few components that would be somewhat unusual to include in an account of lying and misleading. Firstly, he does not seem to allow for the possibility of misleading: every time he refers to a graph that my intuition would suggest is misleading, he says it is lying. There is a widely intuitive distinction between lying and misleading, and it is one that I think bears quite strongly upon deception with graphs. It is often taken to be a point of moral distinction whether someone has lied or merely misled, and many interesting cases of graphical deception seem to be better captured by misleading than by lying. I will discuss technical accounts of the distinction between lying and misleading in greater detail in Chapter 2. Any graph which is misrepresenting any part of the actual facts of the matter is, for Tufte, a graph that lies, which seems to be at odds with the natural understanding of lying and misleading. This is captured nicely by his ready use of the Lie Factor.

Another shortcoming is that certain conditions are somewhat vague in their prescriptions. Consider principle (3), that a graphic must show variation in data, not in design. I have already

mentioned that this seems to fall short of capturing problems with graphs that wholly defy convention, as it focuses on those which contain internal deviations, but there is another problem with the principle. It does not strike me as obvious that any and all design variation within a graph should be in some way deceptive. Consider for instance, a scatterplot style graph, but where all of the points on the graph are assigned colours at random. This would certainly show design variation that does not reflect any sensible change in the data. It would certainly be jarring, but I don't think it would speak to any lack of integrity on the part of the graph.

The challenge of vagueness is at its most evident with respect to principle (6), that graphics must not present information out of context. Although context, as I have explained, is essential to making informed inferences from graphics, and therefore also to the creation of responsible graphs, this principle does not give any guidance on how much context should be taken to be enough.

Chapter 2

Lying, Misleading, and Deception

So, if Tufte's approach will not suffice for understanding graphical deception, it may be sensible to turn elsewhere. Deception, lying, misleading, and insincerity have all featured prominently as topics of discussion in the philosophy of language. For the purposes of my efforts to understand what it means for a graph to be lying or misleading, I will consider some of these resources from philosophy of language, in order to see how they might be made to serve the purpose of making this distinction for graphs, as they make it now for spoken and written utterances. I will survey several accounts of lying, misleading, and other forms of deception. Since these accounts originate with the philosophy of language, they presuppose linguistic (and for the sake largely of convenience, oral) communication. This premise introduces a sharp disanalogy with the domain to which I seek to apply these models, but one must begin somewhere, and there is no sense in reinventing the wheel from the ground up.

The purpose of this chapter is to discuss the existing literature about lying and misleading. What I will show is that in order to have any hope of translating these results to the domain of analyzing graphs, it will be necessary to discuss the problems that arise in trying to determine what is said by a graph. The distinction between lying and other forms of deception, like misleading, hangs on what is said by an utterance, and this chapter will show the importance of that condition for the rest of the thesis.

2.1 Thomas L. Carson

Thomas L. Carson presents an account of lying in his 2010 book *Lying and Deception*. He divides this book into three parts. The first is a conceptual roadmap, which contains the bulk of his

definitional work, explaining his positions about lying, deception, and several other related concepts (Carson 2010, 1). His Part II is focused on moral concerns about when, if ever, lying and other deceptions might be permissible (Carson 2010, 1-2). Part III is a survey of applications, taking the moral framework developed in Part II, and applying it to a range of cases, covering sales, advertising, and political and military cases (Carson 2010, 2). For my purposes, Part I is the most important, and I will dedicate most of my space here to summarizing his conclusions from this portion of the book. In this thesis, I am not particularly concerned with normative moral questions about when lying or misleading might be permitted, but simply with addressing shortcomings of the descriptive accounts of lying and deception when they are extended in an attempt to cover non-linguistic practices that are intuitively thought to be misleading. Personally, I find it largely intuitive that some lies or deceptions might be morally permissible, but I am not going to attempt to draw that distinction here.

2.1.1 Carson's Account of Lying

Carson's Part I encompasses Chapters 1 and 2 of his book. Chapter 1 focuses on lying in particular, whereas Chapter 2 covers other forms of deception. Over the course of the first chapter, Carson refines his definition of lying several times, but his original position is that a lie is, roughly, "a deliberate false statement that the speaker warrants to be true" (Carson 2010, 15). For Carson, notably, making a false statement is a necessary condition for telling a lie. He claims that "showing that a statement is true is always sufficient to counter the accusation that one has told a lie" (Carson 2010, 16). This position seems to be motivated by cases in which a speaker chooses their words to carefully avoid saying anything false, but that nevertheless result in their listeners arriving at false beliefs. He acknowledges the more challenging case, which arises when a speaker professes a true statement that they merely *believe* to be false. Carson's intuitions are that such cases are not lies, as all lies must be false statements, but he recognizes that this is a case in which intuitions are varied,

and leaves the question of if a lie must be false open, to the extent to which those intuitions vary (Carson 2010, 16). Here I will add that my own intuition is that a lie need not be a false statement, but will now continue in Carson's footsteps.

Carson begins his work towards a rigorous definition by starting with two dictionary definitions of lying. They both can be summarized as saying that a lie is a false statement made with the intent to deceive (Carson 2010, 17). From here, he distinguishes between lies and other falsehoods, illustrating that not all false statements are lies. Even in cases where one speaks a falsehood with the intent to deceive, it is possible that the falsehood might not be a lie if the speaker *believes* it to be true. The example that Carson presents is of a speaker telling a listener that Joe is away from home, hoping to deceive the listener into thinking it will be easy to burgle Joe while he is away. The deception here arises because the speaker knows that Joe has a state-of-the-art anti-burglary system, and it will not be at all easy to burgle his home. However, as it turns out, Joe is at home. The speaker here, on Carson's account, has not lied, even though they spoke a falsehood with the intent to deceive, because they believed the falsehood to be true when they spoke it (Carson 2010, 17)².

Carson takes this example to show that a definition of lying must somehow exclude cases where the liar believes what they say to be true, and he offers two possible formulations. A strong condition, that the speaker believes what they say is false (or probably false), and a weak condition, that the speaker does not believe what they say is true (or probably true) (Carson 2010, 17). The distinction between these conditions is relevant in cases of Frankfurtian bullshit, in which the speaker is indifferent to and unaware of the truth values of the statement they make (Frankfurt 2005, 33-34).

² I find it at least modestly interesting that this class of examples seem to entail that the intended deception must be about something other than the content of the utterance, but that line of questioning is not terribly relevant to my purposes here.

Frankfurtian bullshit is something of an elusive phenomenon in reality, but at a conceptual level it is indeed relevant as a distinction between the two conditions. I will return to the subject of bullshit later in this chapter. Carson ultimately concludes that the best course is to simply hold that there could be stronger and weaker conceptions of lying, rather than try to argue for one or the other of the two formulations of the belief condition (Carson 2010, 17-18).

Carson highlights that there are some gaps yet to be filled in this definition. In particular, condition 1., that S makes a false statement, requires more explanation, and as we have already discussed, Carson will ultimately conclude that the intent to deceive is not necessary for a person to tell a lie. He nevertheless comments that he is “fairly close to the mark” (Carson 2010, 20).

At this point Carson digresses from his own refinements to consider a definition which considers a statement a lie only if it violates a person’s right to know the truth. He does so by considering the following condition: “the person to whom he makes the statement has the right to know the truth about the matter in question” (Carson 2010, 19). This is a definition motivated by absolutist ideas that lying should always be impermissible, and it will count many things that would seem, intuitively, to be lies as not being lies at all, in order to make them morally permissible. I agree with Carson that such a definition is likely to result in confusion, since it is at odds with a great deal of common usage (Carson 2010, 19). It also does not properly succeed in ruling out all cases of lying that I—and Carson—consider to be morally permissible. Carson describes one such case in which a man, immediately after a heart surgery, is at risk of death if exposed to any emotional shock. The man asks where his son is. Carson, knowing that the son has been killed in a car accident quite recently, says that the son has been delayed. The proposed criterion, Carson says, “does not exclude this a case of lying—surely, the father has a right to know the truth about his son’s death (Carson 2010, 20). Nevertheless, this seems to be a morally permissible case of lying, in order to save the life of the man. So, even for the moral absolutist about lying, the condition may not suffice.

As we have already heard, Carson will ultimately conclude that the intent to deceive is not necessary to lie. He introduces cases of such lies with several examples, many of which have since become standard examples in the discipline of so-called bald-faced lies—precisely those lies which do not have an intent to deceive. Consider Carson’s first example, that of a witness to crime who is threatened by the culprit, and then, when asked in court if they saw the culprit commit the crime, denies it out of fear of reprisal. They make this denial despite clear video evidence of them witnessing the crime. In such a case, the witness clearly does not intend to deceive the jury—in fact, they likely hope that their testimony will be disbelieved, and the culprit convicted. What matters is that they tell the lie, in order to avoid reprisals (Carson 2010, 20).

The example of the intimidated witness, together with the other examples that Carson includes, conclusively demonstrate that intention to deceive is not necessary for lying. This motivates a search for a new condition to replace the intention to deceive, which for Carson will involve the warranting of the truth of what is said. Carson holds that saying X typically involves warranting that X is true (Carson 2010, 24).

Carson’s discussion of warranting the truth of statements includes the observation that whether or not a speaker warrants the truth of what they say is not entirely up to them, and that it may not even always be clear to the audience whether or not the speaker is in fact warranting the truth of any given statement (Carson 2010, 27-29). Carson takes this to pose a problem for his definition, in the case of a speaker who warrants the truth of what they say without realizing that they have done so. A comedian, for instance, who has found his way into the wrong theatre, and presents a false story to an audience expecting a serious presentation, might be taken to be warranting the truth of his statements, when he does not intend to do so. Carson observes that it seems unfair to say this comedian has told a lie, and so proposes another condition, “that it is not the case that the speaker takes himself to be not warranting the truth of what he says” (Carson 2010, 29).

At this point, Carson adds an audience-relative piece to his definition of lying, in order to account for the possibility that different members of the audience of a lie might have different levels of competence. A statement that could be clearly non-warranting to one part of the audience might be taken as warranting to another. In such a case, a definition that does not have a relativistic component requires some expansion. Carson arrives at:

A person S tells a lie to another person S1 iff: 1. S makes a false statement X to S1, 2. S believes that X is false or probably false (or, alternatively, S does not believe that X is true), 3. S states X in a context in which S thereby warrants the truth of X to S1, and 4. S does not take herself to be not warranting the truth of what she says to S1 (Carson 2010, 30).

The definition has external features that he endorses: the falsity of the statement, and the success of the warranting of the statement (Carson 2010, 39). He concedes that those who are more attracted to an internal conception of lying will prefer other versions of the definition, and he presents several variations. I do not share Carson's intuition that a lie must necessarily be false, and I am also somewhat suspicious of external conditions that hang the question of lying on whether or not the audience is sufficiently receptive to a warrant that the liar is attempting to put into force.

2.1.2 Carson's Accounts of Deception and Misleading

The second chapter of Carson's book is dedicated to explaining other kinds of deceptive speech. As in the first chapter, Carson begins with a dictionary definition: that to deceive someone is to cause them to believe what is false (Carson 2010, 47). Carson rejects this immediately as too broad, with good reason. A listener might conclude something wildly implausible from a reasonable and true description—Carson gives the example of a customer wrongly concluding that a car will be safe in high-speed collisions with large vehicles, based on a salesperson's clear and accurate description of the car's safety measures (Carson 2010, 47). Carson, therefore, reasons that deception

must be intentional. Moreover, if A causes B to have a false belief X, it is not deception if A herself believes that X.

Carson ultimately presents two different definitions of deception more broadly. He also makes modifications to accommodate the following sorts of cases. A tries to cause B to believe the false proposition X, but unknown to A, B already believes X. Later, B discovers information that would have caused them to disbelieve X, except that their belief in X has been strengthened by A's remarks. Carson, following Mahon (2007), agrees that this is a case of deception on A's part, and amends his definitions as follows:

D1a. A person S deceives another person S1 if, and only if, S intentionally causes S1 to believe X (or persist in believing X), where X is false and S knows or believes that X is false.

D2a. A person S deceives another person S1 if, and only if, S intentionally causes S1 to believe X (or persist in believing X), where X is false and S does not believe that X is true.

(Carson 2010, 50)

As in the analogous cases of lying, these differ in circumstances where the speaker has not given any thought to the truth or falsity of X. By way of further analogy, Carson again does not have a strong reason to prefer one over the other, and suggests that we think of there being stronger and weaker notions of deception. With his treatment of deception settled, Carson turns to other, related concerns. The first of these is the question of misleading. Carson takes misleading to be more general than deceiving: simply the causing of others to have false beliefs (Carson 2010, 49). Misleading, then can be intentional or unwitting, which is in contrast to deceiving, which for Carson is always intended. These three concepts—lying, deception, and misleading—are the three that I need from Carson, and so I turn now to other accounts of these notions.

2.2 Jennifer Saul

Jennifer Saul's 2012 book *Lying, Misleading, & What is Said* focuses specifically on the notion of lying that is taken to contrast, in the popular consciousness, with misleading (Saul 2012, 1). This focus, being largely shared by my project, motivates the hope that this account might be more useful for my purposes in attempting to make sense of what makes a graph lying or misleading. The goal of Saul's Chapter 1 is to arrive at a definition of lying that is designed to contrast it with merely misleading. Ultimately, she arrives at the following definition:

If the speaker is not the victim of linguistic error/malapropism or using metaphor, hyperbole, or irony, then they lie iff (1) they say that P; (2) they believe p to be false; (3) they take themselves to be in a warranting context. (Saul 2012, 3).

I will say a few words in summary about how that definition is uncovered.

Saul takes it that the focus on what is *said* (or asserted, or whichever other notion plays that role in a particular formulation) is what allows for the intuitive distinction between lying and misleading (Saul 2012, 4). Any definition of lying that did not include some understanding that something was being said would run together all the forms of deliberate deceptive activity (Saul 2012, 4). Saul begins the construction of her definition of lying with a dictionary definition familiar to us through earlier discussion by Thomas Carson: that a lie is a false statement made with the intent to deceive. Saul also rightly moves past this first brush quickly. A person who says something false that they *believed* to be true is not telling a lie, they are simply mistaken.

Here, Saul and Carson will begin to diverge somewhat. Saul and I share the intuition that a statement need not be false to be a lie. Suppose that Alice and Bob, two high-achieving students in a philosophy course, are speaking about an upcoming assignment. Alice, believing it to be false, tells Bob that it will be fine if he just hands in an essay late without asking for an extension, because Alice

wants Bob to do poorly, in order to improve her own relative standing in the class. Bob takes this advice, and it turns out that the professor is in fact extremely indulgent, and does not penalize the late submission. So, what Alice told Bob was true, but I find it far more intuitive to say that she lied, rather than not. Carson and Saul both include provisos to accommodate the opposing intuition, but this point of agreement is where I begin to prefer Saul's account. This new adjustment allows for the lie itself to be a true statement.

At this point, Saul's account includes a requirement for an intention to deceive. She observes that this helps to work around counting statements that are jokes, or include metaphors, as being lies, but that as it stands the condition needs some refinement. Saul uses the example of Beau, a cantankerous and unpleasant fellow who wants to convince the people of Sheffield that he is nice, and so remarks "it's freezing here" in an attempt to bond over the weather, when it is quite cold, but clearly not below 0\deg C (Saul 2012, 8). As it stands, this definition would count Beau as lying, which seems to be inaccurate. Here, Saul proposes the following amendment: "A person lies iff (1) They say that P; (2) They believe P to be false; and (3) They intend their audience to believe that P" (Saul 2012, 8)³. This revision addresses the problem with the Beau example by narrowing the scope of deceptions that would count as lying. The deception must be about the statement that is said.

Earlier, I spent some time discussing objections to the inclusion of intent to deceive, and so does Saul—in fact, the very same courtroom testimony example from Carson. This type of example, now generally referred to as bald-faced lies, following the terminology of Sorenson, poses a serious problem for the inclusion of intent to deceive in the definition of lying (Sorenson 2007). But, as Saul points out, that condition was doing important work in keeping metaphors, jokes, and other figures of speech from being counted as lies. Here, Saul adopts the idea of warranting from Carson, and arrives

³ Saul begins numbering her conditions at this point for ease of reading.

at a new definition: “a person lies iff (1) They say that P; (2) They believe P to be false; (3) They say that P in a warranting context; (4) They do not take themselves to be in a non-warranting context” (Saul 2012, 11). This is, for now, very close to Carson’s definition. Saul’s account, however, engages with objections to this formulation of warranting from Don Fallis and James Mahon, and thus is refined again in short order (Fallis 2009, Mahon 2008). Saul introduces the example of a crooked politician who has been invited to give a speech on a serious topic, but finds himself in the wrong room, speaking to a crowd expecting a fictional monologue. Having been bought off by lobbyists, the politician tells the crowd many falsehoods, hoping that they will be believed, but they are not, because the audience is expecting fiction. The current form of the definition counts this as not being a lie, since the speaker is not in a warranting context, but this result seems unpalatable (Saul 2012, 12). So, the definition is refined to its penultimate version: “a person lies iff (1) They say that P; (2) They believe P to be false; (3) They take themselves to be in a warranting context” (Saul 2012, 12). This reformulation addresses the case of a speaker who wrongly takes themselves to be in a warranting context, as does the crooked politician above.

In working towards the final definition of lying, Saul addresses a series of lingering complications. The first of these is metaphor. I present a slight variation on the example used by Saul to make this point⁴. Imagine Alice saying to a friend that “Bob is a lamb,” here intending to suggest that Bob is gentle, meek, and innocent. So long as Bob is in fact a human being, this is untrue. Alice certainly knows that it is literally untrue, and supposing she takes herself to be in a warranting context—I see no reason why this could not be so—she has, according to the current version of the definition, lied, which is not an acceptable interpretation of this speech. Saul points out that irony and hyperbole will lead to very similar issues, and so these concerns are grouped together.

⁴ In order to make it a metaphor with which I am more familiar.

The next group of problematic speech that Saul addresses is the accidental falsehood. There are many ways to get it wrong when one is speaking. One sort is when a speaker says something that they believe to be true, but are mistaken, and say something false. For instance, if I hold the mistaken belief that all prime numbers are odd, having forgotten about 2, and say as much, I have spoken a falsehood but did not mean to do so.

Alternatively, one might speak the wrong words and thereby exonerate themselves from speaking a falsehood: suppose a racist politician, who is aware of the systemic racism embedded in his government, attempts to deny this by saying “there is no systemic racism in our government.” But now suppose also that he misspeaks, saying instead a true statement that “there is no *systematic* racism in our government.” Has he lied? It is certainly clear that he *meant* to lie, but in bungling his speech, has he managed to fail to do so? It is not entirely clear how these sorts of cases should be organized, and intuitions are unclear surrounding them.

Saul first introduces two clauses to act as patches to fix these sets of problems, giving us the definition: “a person lies iff (1) They say that P; (2) They believe P to be false; (3) They take themselves to be in a warranting context; (4) They are not speaking metaphorically, hyperbolically, or ironically; (5) They are not a victim of linguistic error or malapropism” (Saul 2012, 15). This, Saul explains, is actually a problematic definition as well, since it rules out the possibility of lying in any context where metaphor, hyperbole, irony, linguistic error, or malapropism are taking place. Saul presents several examples of these features being used in utterances that might be lies. One of the best is a competitive gardener with a poor crop of tomatoes this year declaring “We’ve got tomatoes coming out of our ears,” which is metaphorical, and thus cannot be a lie under the current definition (Saul 2012, 16). I share Saul’s intuition, though, that this might very well be considered a lie. Saul ultimately concludes that it is best for the purposes of the book, which is to zero in on the distinction between lying and misleading, to avoid passing judgement on these convoluted edge cases (Saul

2012, 16). This brings us to the definition used for the remainder of the book: “If the speaker is not the victim of linguistic error/malapropism or using metaphor, hyperbole, or irony, then they lie iff (1) they say that P; (2) they believe p to be false; (3) they take themselves to be in a warranting context” (Saul 2012, 18). This definition of lying has been constructed with the lying-misleading distinction in mind, which stands a very good chance of making it useful for me.

2.2.1 What is Said

Because the lying-misleading distinction in ordinary speech hangs so heavily upon the distinction between what is said as opposed to merely implied, Saul’s account lays out an idea of what is said to help capture that difference. Saul’s Chapter 2 traces the course of contemporary analysis of what is said, and her Chapter 3 proposes a new solution that better captures the intuitive lying-misleading distinction.

2.2.1.1 Problems with What is Said

Saul’s analysis of the issues facing existing concepts of what is said begins with discussion of H. P. Grice. Grice does not give a particularly clear definition of what is said, but highlights the distinction between what is said, and what is suggested by a given utterance that might go beyond what is said. Grice calls these unsaid suggestions “conversational implicatures” and takes them to be what is necessarily understood from an utterance such that the utterer can be thought of as being cooperative (Grice 1989; Saul 2012, 22). Saul addresses a number of important classes of examples that are not discussed by Grice. Among these are indexicals, and demonstratives, as well as cases requiring completion and expansion. Indexicals are “terms whose reference varies with context, but for which that reference is completely determined in each context by a linguistic rule” (Saul 2012, 23). Paradigmatic examples are words like ‘I’, ‘here’, and ‘now’, which refer to the speaker, the place

of the utterance, and the time of utterance.⁵ Demonstratives, like ‘this’, ‘she’, or ‘that’ also require supplementation from the context of the utterance, but that supplementation is not nearly so well-governed as in the cases of the indexicals. There are also problems posed by completions and expansions of an utterance. The classic examples of completion are sentences like ‘Alice has had enough’ and ‘Bob is ready’, where the meaning of the sentence varies significantly with context. Has Alice had enough heart medicine, or whiskey? Is Bob ready to go out on a date, or for an intense surgery?

Working towards an understanding of what is said that will function well for the purposes of the lying-misleading distinction, Saul roughly divides theories of what is said into three rough groups, which she emphasizes are based on theories from the literature, rather than being direct adoptions, which she calls unconstrained, constrained, and austere (Saul 2012, 27). These three sorts proceed in decreasing order of acceptable contextual addition to what is said, beyond the literal and explicit content of the uttered words (Saul 2012, 27). As she argues over the course of her Chapter 2, none of these three types of accounts will adequately suit the purposes of the lying-misleading distinction.

Unconstrained conceptions count many disbelieved statements as said, and therefore as lies, that are intuitively only misleading. An example given by Saul is the infamous statement from Bill Clinton that “there is no improper relationship” (Saul 2012, 30). Some people would (and did) interpret this statement as saying that had never been an improper relationship, and so on an unconstrained conception of what is said, the statement would say that. This is strongly against the intuition that Clinton spoke misleadingly, and very carefully, to avoid telling an explicit lie. The

⁵ Ignoring, for the moment, complications about recorded or otherwise asynchronous utterances, such as discussed by Predelli (1998) and Sidelle (1991). These complications will return later.

lying-misleading distinction hangs quite closely on the literal meaning of what is said, and so an unconstrained account will not deal with it very well (Saul 2012, 30).

Constrained conceptions fail to properly address cases of expansion. Saul gives the following example, where we are supposing that Amanda and Beau had children out of wedlock and then later married: “Amanda and Beau got married and had children” (Saul 2012, 37). On a constrained theory, this utterance will say what is expressed by “Amanda and Beau got married and then had children” (Saul 2012, 37). This raises problems for the lying-misleading distinction, because it seems that if I were to say the former utterance in order to deceive someone concerned about the propriety of Amanda and Beau’s relationship, then I have merely misled them, whereas the second utterance would be a lie. However, on a constrained account, the former says the same thing as the latter, which contradicts what we would hope for. So, a constrained account does not suffice (Saul 2012, 37-8).

Although austere conceptions draw the right conclusions about the previous sorts of cases, they still fall short of Saul’s goals for the lying-misleading distinction. Saul discusses three variations of austere conceptions⁶, in more detail than I have space to discuss here. Various, these accounts fail to properly address cases that involve demonstratives or completion (Saul 2012, 40-5). In any case, ultimately austere conceptions of what is said do not quite fit the bill for the lying-misleading distinction, either, and it is therefore necessary to look somewhere new (Saul 2012, 50).

2.2.1.2 What is Said

And so, Saul develops her own account of what is said. The goal is to arrive at a sort of middle ground. Unconstrained and constrained conceptions both allow for too much contextual flexibility in what is said, and austere conceptions do not allow for enough (Saul 2012, 54). Saul

⁶ From Kent Bach, Emma Borg, and Herman Cappelen and Ernie Lepore.

arrives at a necessary criterion for evaluating potential contextual contributions to what is said, which she calls Needed for Truth Evaluability (NTE): “A putative contextual contribution to what is said is a part of what is said only if without this contextually-supplied material, S would not have a truth-evaluable semantic content in C,” where S is a sentence and C is a context (Saul 2012, 57). This is a fairly strict criterion, but allows for more flexibility than the austere conceptions encountered thus far.

2.3 Andreas Stokke

And so, I turn to *Lying & Insincerity*, the 2018 book by Andreas Stokke. Here, Stokke presents a rigorous and technical account of lying, which draws on traditions from linguistics. These linguistic theories of information structure in a conversation allow for a great degree of flexibility in Stokke’s analysis of what is said, which is core to his definition of lying. Stokke’s account of lying is as follows:

A lies to B if and only if there is a proposition p such that:

(L1) A says that p to B, and

(L2) A proposes to make it common ground that p, and

(L3) A believes that p is false. (Stokke, 31)

For Stokke, (L1) and (L2) above together capture a specific notion of the general idea that A asserts that p to B. This general idea, of asserting something that one believes to be false, is a through-line in many definitions of lying. We see it, for instance, in both Saul and Carson, though Carson requires that the proposition is in fact false.

In order to make use of Stokke’s definition of lying as a launching point, it is necessary to be clear about what he means by ‘saying’ and by ‘common ground’. Owing to its more explicit definition in Stokke’s own writing, I would like to begin with ‘saying’. However, owing to the

linguistic ground on which Stokke's account is built, I must first take a brief detour to explain the notion of a question under discussion, or QUD.

2.3.1 The Stalnakerian Model of Discourse and Questions Under Discussion

The Question Under Discussion model is developed by Craige Roberts, and builds upon the understanding of discourse due to Robert Stalnaker. So, in order to properly explain QUDs, I will first need to briefly summarize the understanding that Stalnaker develops for discourse. Stalnaker takes the primary goal of all discourse to be inquiry: the process of coming to understand the way that things really are (Stalnaker 1978).

Stokke uses the account of questions under discussion (QUDs) developed by Roberts. What is said by an utterance depends on the current QUD, which is taken to be a sort of topic of the conversation. This framework follows Stalnaker in presenting all conversation as a sort of mutual inquiry, which is a presentation that I am often skeptical of in general, but for the purposes of what I am intending to do in later sections of this paper, it will suit me rather nicely.

The core thrust of Roberts's account is that information in a conversation can be hierarchically understood in terms of the structure of the QUDs. The goal of conversation is answering the current QUD, which is often—perhaps always—a sub-QUD for some higher-level question. Roberts reinterprets Stalnaker's formulation of the goal of all conversation, taking the overarching goal of discourse to be “the Big Question: what is the way that things are?” (Roberts, 5). On this view, all QUDs that are tied to any particular instance of conversation are always, in some sense, sub-QUDs for the Big Question.

An example of this hierarchy would be the Big Question, followed by ‘who likes to eat what?’, followed by ‘what does Alice like to eat?’. An answer to each sub-question informs a partial answer to the super-questions. Knowing what Alice likes to eat says something about who likes to eat

what, and something about the way things are. In this way, QUDs inform the structure and strategy of a conversation. Which sub-QUDs are asked says something about how the interlocutors intend to answer the super-QUDs, and indeed the Big Question. The hierarchy of QUDs, coupled with their answers, form a conversational strategy for achieving the goal of conversation, which is again, taken to be answering the Big Question.

2.3.2 What is Said

Now that the idea of a QUD has been appropriately introduced, Stokke gives a precise definition of what is said. Using S to denote a sentence, c to denote a context, and $\mu_c(S)$ to denote a minimal proposition expressed by S in c , he defines what is said as:

What is said by S in c relative to a QUD q is the weakest proposition p such that (i) p is an answer (partial or complete) to q and (ii) either $p \subseteq \mu_c(S)$ or $\mu_c(S) \subseteq p$. (Stokke, 99)

This is quite a technical definition, and merits some further explanation. Crucially, p is not simply a weakest answer to the QUD q , but it also must stand in close relation to the minimal proposition expressed by the sentence S . Stokke encapsulated this close relationship through entailment: either p entails $\mu_c(S)$, the minimal proposition expressed by S in c , or else $\mu_c(S)$ entails p . The original formulation of the definition only allows for $\mu_c(S)$ to entail p , but the converse option is added in order to deal with objections about utterances like “You aren’t going to die” made to someone suffering from an injury (Schoubye and Stokke 2016, 783). In this case, it seems that what is said is really “You aren’t going to die *from this injury*,” which is an instance of the reverse entailment that is incorporated into the definition. Note that here, “you aren’t going to die” does entail “you aren’t going to die from this injury” but the converse is not true.

The core insight in this definition of what is said is the relationship between the statement S and the QUD q . The same sentence, even when it is a fully truth-evaluable proposition on its own, can

still say different things in response to different questions. Stokke illustrates this with a terrific example, which I recreate here. Suppose that Elizabeth and William are talking about a co-worker, Doris. Doris has been recently fired due to downsizing in the company. A little while prior, Doris had insulted Sean, the company's CEO, at a party, but Sean did not take this to heart, and they are still on good terms. William knows all of this, but wants to give Elizabeth the impression that there is bad blood between Doris and Sean. Consider the two following brief conversations:

(1) **Elizabeth:** Why did Doris lose her job?

William: She insulted Sean at a party.

(2) **Elizabeth:** How is Doris's relationship with Sean?

William: She insulted Sean at a party.

As Stokke points out, an account which is not sensitive to the structure of the discourse, and the current QUD, cannot explain why William's statement in (1) is a lie, whereas in (2) the same statement is merely misleading. Stokke's account does make that distinction, because what is *said* changes between (1) and (2), in response to the changing QUD. In both cases, William's utterances lead Elizabeth to wrong answers to the question that she has asked. However, in (1), this answer is given explicitly: she lost her job because she insulted Sean at a party. In (2), that Doris and Sean have a bad relationship is communicated only implicitly, and therefore is not said (Stokke 2018, 82).

Stokke lays out the details of his definition of what is said in a 2016 paper with Anders Schoubye called *What is Said?* (Schoubye and Stokke 2016, 783). The meanings of the constituents of any particular sentence constrain what that sentence could possibly say, but some contextual supplementation is also clearly required in some cases. Stokke's $\mu_c(S)$ is the minimal content of the sentence *S* in the context *c*, which is derived from the constituents, their order in the sentence, and necessary contextual supplementations (Schoubye and Stokke 2016, 773).

2.3.3 The Common Ground

For Stokke's account to consider an utterance of a statement S to be a lie, that utterance must also propose that the relevant proposition p (that is, the one that is said) be added to the common ground. The common ground is a collection of propositions that are treated as if believed by the participants in a conversation (Stalnaker 2002). Although it is most simply modeled by straightforward common belief among interlocutors, this does not encompass the full range of possible contents of the common ground. Stalnaker specifically refers to assumptions, presumptions, and pretense as possible additions to the common ground (Stalnaker 2002, 704).

If interlocutors begin a line of conversation with an assumption, perhaps with the classic phrase "for the sake of argument", then that assumption may very well enter into the common ground, even if only temporarily. The conversation can then proceed as if all parties do believe the assumption, even though one or more of them may not. For instance, suppose that Alice and Bob are discussing issues of refugee policy as follows, and that Bob believes refugees bring crime with them, while Alice does not believe that refugees bring crime with them.

Alice: We should accept more refugees to this country.

Bob: No way, refugees bring crime with them.

Alice: Even if that's true, we still have a duty to accept them.

Here, Bob proposes to update the common ground with the believed proposition that refugees bring crime with them. Alice rejects the move to add that proposition as common belief, but nevertheless accepts it to the common ground as something that they can both *make as if to believe*. This is a core distinction between common ground and common belief.

2.4 The Promised Complications

As I suggested in my discussion of Saul’s account, there are going to be problems about the temporally-disjointed nature of recorded utterances that will cause difficulties for the graphs I am focused on. I give a brief overview of some literature that deals with those complications. The concern about temporally-disjointed remarks starts with Alan Sidelle’s 1991 paper *The Answering Machine Paradox*. Sidelle is motivated into the study of indexicals in this paper by the seemingly-paradoxical utterance of “I am not here now” as delivered through the recording of an answering machine (Sidelle 1991, 526). Sidelle briefly makes the case that it seems clear that any utterance of ‘I am here now’ should be true, and thus ‘I am not here now’ should always be false, because roughly, the utterer is always located at the place of utterance at the time of utterance (Sidelle 1991, 526-8). Sidelle’s focus is on the examples and issues that arise out of this case of the answering machine case, but it motivates some discussion of problems related to asynchronous utterances—like with graphs. Sidelle’s purpose in his own paper is to open the door to a new line of inquiry, but he does offer a proposed solution to the problem: thinking of these sorts of utterances as being deferred: the relevant indexicals are to be contextually supplemented based on the context of their delivery, not the context of their initial inscription (Sidelle 1991). He uses, among others, the example of the reading of a dead man’s will, which contains the phrase “If my nephew is by now no longer an alcoholic,” to make this point (Sidelle 1991, 538). As he rightly points out, it does not make any sense at all to take that ‘now’ as referring to the time that the author wrote it down; clearly at that time the author would have known the answer to the hypothetical, and the hypothetical would therefore not have been necessary. The writing of a note, or recording of a message, is “arranging to make an utterance at a later time” (Sidelle 1991, 535). Sidelle does not make the case that this will be the best possible interpretation, but it seems a promising start.

Sidelle's account is refined and expanded upon by Stefano Predelli, who makes the move away from the deferred utterance having contextual parameters set at the time of its writing, towards the context of interpretation being that which is intended or expected by the person inscribing the message. This change avoids some of the problems that arise in cases where a written message finds its way into an unintended location, or is read at an unintended time. Following Predelli, we would not want to call the writer of a note that reads "I am not here now" a liar simply because by the time the note was read, they had in fact returned to the place where the note was left (Predelli 1998). The intended context of delivery for the utterance strikes me as a satisfying way to address these cases, for the purposes of my own analysis.

2.5 Bullshit

One area of study within the broader study of deception that I had not originally expected to be relevant to my purposes here is the study of bullshit. Bullshit, philosophically speaking, was made popular by Harry Frankfurt's *On Bullshit*, and has been a fascinating area of discussion ever since. Frankfurtian bullshit is characterized by the speaker's indifference to the truth (Frankfurt 2005, 33-4). On Frankfurt's account, this is a remarkably pervasive phenomenon, that is prevalent in many aspects of our society. Frankfurt also takes bullshit to be a greater enemy of the truth than lying, as on his account, bullshit demonstrates a total disregard for the relevance of truth as an important facet of communication (Frankfurt 2005, 61).

The "essence of bullshit," says Frankfurt, "is not that it is *false* but that it is *phony*" (Frankfurt 2005, 47). The bullshitter, although they are in some sense faking their way through something, is not necessarily wrong. They are not trying to deceive the listener about any particular aspect of the state of affairs that they are discussing, but rather about what they are up to. The bullshitter is not concerned with the truth or falsity of their utterance, but rather with some ulterior motive. I will later

introduce an example of a graph that I take to be, in fact, best described by bullshit, as opposed to either lying or misleading.

2.6 Conclusions

There are important throughlines common to all of these accounts of lying and deception. Chief among these is that in order for any of them to be applied, it is important to have a notion of what is said, and some sense of what it means for something to be said *seriously*. This sense of ‘seriousness’ is what is captured variously by asserting, warranting contexts, or proposals for addition to the common ground, depending on the theory currently under discussion. So, the next step must be to try to make some sense of the notion of what is said by a graph. The problem of what is said by a graph, under various accounts, will occupy the following chapter.

Chapter 3

What Does the Graph Say?

The question of what is said by a particular utterance is thoroughly covered by a robust body of literature in philosophy of language and in linguistics. What we have seen so far, specifically tailored to the question of lying, is only barely scratching the surface. Much of this discussion is centred on the model of an in-person, synchronous conversation, usually in which two interlocutors are discussing something. There are very often, as we have seen in Stokke’s account, underlying assumptions about the purposes of inquiry in general. A great deal of this literature, though, begins to be strained when a broader notion of ‘conversation’ is considered. If my goal is to conceive of graphs as ‘lying’ or ‘misleading’—and this is my goal—then it will be important to understand what is *said* by a particular graph. Consider the following graph, which I have created for the purposes of this discussion:

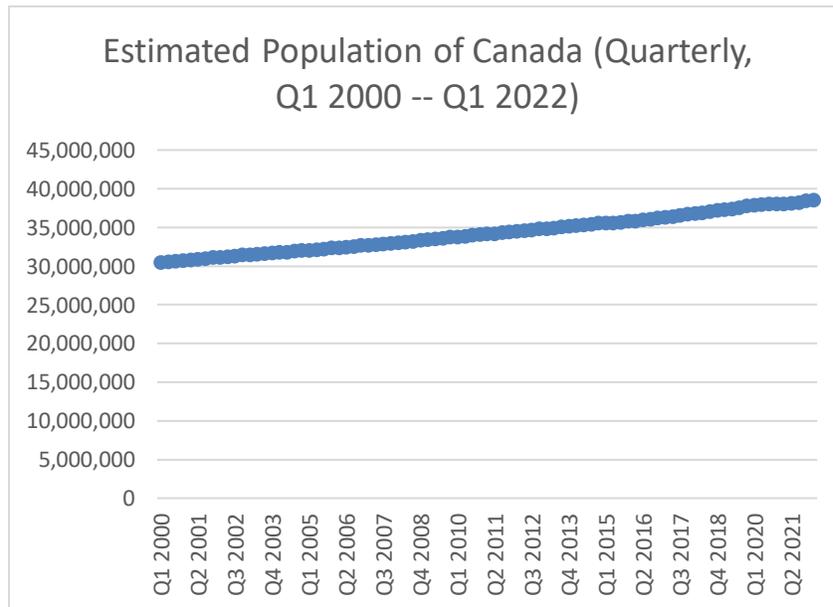


Figure 9: Data from Statistics Canada. Table 17-10-0009-01. Population estimates, quarterly

The graph above displays population estimates for Canada on a quarterly basis, from Q1 in 2000, to Q1 in 2022. The estimates are from Statistics Canada.

It is reasonable to ask what this graph says, but it will turn out to be very difficult to arrive at a satisfying answer. I start with a very small-scale question. The first data point chronologically, for Q1 2000, appears just slightly above the 30,000,000-line on the vertical axis. I begin by asking: what does this graph *say* about the estimated population of Canada in Q1 2000? There are, I think, some obvious contenders, some of which will entail others. Some things that the graph *might* say are, in decreasing order of strength:

- (a) The estimated population of Canada in Q1 2000 is 30,525,872⁷
- (b) The estimated population of Canada in Q1 2000 is approximately 30,500,000⁸
- (c) The estimated population of Canada in Q1 2000 is between 30,000,000 and 35,000,000, and is closer to 30,000,000 than to 35,000,000
- (d) The estimated population of Canada in Q1 2000 is between 30,000,000 and 35,000,000

I will argue that the graph *does* say at least (d), and also that it does *not* say (a). The data point is distinctly between 30,000,000 and 35,000,000, so it definitely says at least (d).

It is, as a simple matter of fact, impossible for this graph to even indicate something as specific as (a). The image of this graph, saved as a PNG, is only 1558 pixels in height, which means that the horizontal band of space between 30,000,000 and 35,000,000 is at *most* 173 pixels high, and

⁷ This is the actual data point from Statistics Canada.

⁸ Based on a visual assessment of the data point in the graph.

is in fact notably less than that⁹. Thus, each possible displayed height in this band corresponds to at least 2890 different possible values¹⁰. So, even if a single pixel is correctly chosen as the height of the data point, it does not uniquely indicate a specific population estimate.

To me, this raises a natural alternative for what might be said. Perhaps what the graph says is instead:

(a*) The estimated population of Canada in Q1 2000 is between 30,522,982 and 30,528,762¹¹

I find the proposed solution of (a*) to still be unsatisfying. Although it addresses the immediate concern that is raised in this case, it seems unlikely to me that what is said by a graph should hang on the resolution of the particular instance of the graph in question. If we had a graph image that was suitably high in resolution, then maybe an individual pixel would uniquely determine a value. I do not wish to diverge too far from my topic, and I do not take it that the pixelation is the most relevant issue at play, so with that said, I return to discuss (a*), setting those concerns aside. Whether or not the graph says any of (a*), (b), or (c) is less immediately obvious than (a) or (d). To determine this, I now turn to some particular accounts of what is said.

Recall that, using S to denote a sentence, c to denote a context, and $\mu_c(S)$ to denote a minimal proposition expressed by S in c , Andreas Stokke defines what is said as:

What is said by S in c relative to a QUD q is the weakest proposition p such that (i) p is an answer (partial or complete) to q and (ii) either $p \subseteq \mu_c(S)$ or $\mu_c(S) \subseteq p$ (Stokke, 99).

⁹ This is a very coarse upper bound: there are nine equal sized horizontal bands on the graph, and 1558 pixels divided by 9 is roughly 173.1. Thus, any one of these bands must have height less than 174 pixels. Since there are titles and labels above and below the graph, this bound is an overestimate.

¹⁰ 500,000 divided by 173 is roughly 2890.2.

¹¹ The actual data point plus/minus the minimal necessary error required by the pixelation.

Recall from Chapter Two that Stokke is using entailment here to encapsulate the close relationship between what is said and the minimal content of the sentence that is uttered. So, for Stokke, we are interested in the weakest proposition p such that p at least partially answers the QUD, and it is entailed by, or entails, the minimal proposition $\mu_c(S)$. On Stokke's minimal account of what is said, we need to understand which QUDs this graph is meant to answer. This is already a source of possible disagreement, but I contend that one plausible choice arises from the question I asked to motivate the propositions in the first place:

Q₁: What was the quarterly estimate of the Canadian population in Q1 2000?

If what Stokke believes to be said can be applied to this data visualization, and if the QUD I suggest above is acceptable, then what is said by the graph should be (d), because what is said is the *weakest* proposition p that at least partially answers the QUD, and it is entailed by, or entails, the minimal proposition $\mu_c(S)$ (Stokke 2018, 99). Since any of those proposed statements do serve as at least partial answers to the QUD, the minimal among them, which is (d), is what is said. This strikes me as a frustrating consequence of Stokke's account. I would like to be able to claim that the graph says more than just the vague range denoted by its major gridlines. I think that the intuition we have of a graph is that it says something more than just that a data point falls in some broad range. Any account of what is said that cannot reflect that seems to be failing to capture an important part of the transition to understanding what is said by a graph. It should be possible for that graph to say more than just (d), but the minimality constraint on the answer to the QUD makes that impossible.

One significant disanalogy between the graphical depiction of x_1 having the value 30,525,872 and the spoken words " x_1 , the estimated population of Canada in Q1 2000, is 30,525,872" is that the graphical depiction, as we have seen, cannot be as precise as the written/spoken form, at least not at this scale. Another is that there are weaker propositions that are *necessarily* explicit in the visual display that are not explicitly present in the spoken words. For instance, that x_1 is between 30 and 35

million follows immediately from the spoken form, but it is not explicitly stated. Displaying the data value on a set of axes necessitates that the ordering is explicitly part of the representation. Once the point is marked, it is explicitly given as falling between the nearest bounding limits. Because giving the interval containing the point is logically weaker than giving the explicit value of the point, Stokke's definition of what is said puts us in a position where we have to consider the interval the best we can say. The image displays a weaker proposition that is not directly included in a spoken utterance of the data value. The minimality constraint is used by Stokke's account to closely tie what is said to the words that are uttered in a sentence (Schoubye and Stokke 2016, 15). Unfortunately, it causes problems for the purposes of determining what is said by a graph.

It seems unintuitive to give up on the idea that a graph says something about where an individual data point falls. The fabricated data point is, I think, the clearest possible case of a lie in a graph. Suppose that the graph from Figure 9 is taken, and the first data point, for Q1 2000, is moved up to the point of 34 million. My strong intuition is that this would make this graph a lie; it would be fabricated data being represented. But, as we have seen, on at least Stokke's account, the graph doesn't *say* anything different. It still only says that the Q1 2000 estimate was between 30 and 35 million. Now *this* is a problem. In order to call a single fabricated point a lie, we need to be able to say that the graph in Figure 9 says more about the value of the Q1 2000 population estimate than simply that it falls between 30 and 35 million people. The positioning of the point of data on the axis needs to matter. So, the minimality condition of Stokke's account of what is said seems to make what I think is the clearest possible example of graphical lying out to not be a lie. So, the emphasis on minimal content is definitely not the way to go for something like a graph.

At this point, it has begun to seem to me that my efforts to determine what is said by the graph are going in the wrong direction. Motivated by advice from Tufte, who says that the core question of all quantitative analysis is "compared to what?" maybe fixating on one point of data is

going in the wrong direction. Perhaps it would be better to conclude that what a graph *really* says is not best understood in terms of any one of the individual data points, but rather arises from a more general view. Perhaps the most relevant thing that is said by a graph is about relationships and comparisons. It might not make sense to say that a graph says things about the individual data points that compose it, but instead to take what it says as a higher-level statement about overall trends, or comparisons between those data points.

Indeed, consider what might be the most natural choice of QUD for the graph, which can be inferred from the title:

Q: How have quarterly estimates of the Canadian population changed from Q1 2000 to Q1 2022?

It is clear that this relational understanding of the data being presented gives a better account of answers to Q, and therefore, perhaps a better account of what is said. For instance, recall (d) above: The estimated population of Canada in Q1 2000 is between 30,000,000 and 35,000,000. Although this is certainly the minimal among the propositions that I originally proposed, it does not seem to be a very good answer to the QUD given by Q: How have quarterly estimates of the Canadian population changed from Q1 2000 to Q1 2022?

It does, however, seem like at least a plausible answer to the QUD Q_1 : What was the quarterly estimate of the Canadian population in Q1 2000? Moreover, Q_1 certainly could be taken to be a sub-QUD for Q, in that an answer to Q_1 does constitute a partial answer to Q. Here, I think, the issue with figuring out what a graph *says* starts to become clearer. The simple fact of the matter is that, compared to any normal sentence in English, a graph says a *tremendous* amount, in response to a wide range of QUDs. For although each well-presented graph has an obvious QUD built into its title, there are equally obvious QUDs that we might rightly ask about each of the individual data points that are represented, and the graph will *say* different things about each of them, and for some of these

possible QUDs, an account that requires minimality—like Stokke’s—is going to lead to some awkward results, as we have seen. So, I will now discuss some possible ways of dealing with this tremendous quantity of propositions.

3.1 An Indeterminate Utterance

One place to look for a way to interpret the example above is in Jennifer Saul’s discussion of indeterminacy in completion (Saul 2012). Although the issue facing us is not precisely one of completion, there seems to be some analogy that merits discussion. Saul claims that in cases of completion, it may be prudent to think that the completion is in fact indeterminate across the range of possible completions (Saul 2012, 64). If, for instance, I utter the sentence “Tyler is not ready” in response to Tyler’s partner Hannah asking if he is ready for their dinner date yet, it seems plausible that I could be taken to be saying any of the following:

1. Tyler is not ready for dinner.
2. Tyler is not ready for your date.
3. Tyler is not ready to meet you.

This is, incidentally, not at all an exhaustive list of possible completions that would be more or less equally reasonable. Each of these is appropriate, but it’s also equally clear that the completion “Tyler is not ready to swim across the Pacific Ocean” is not at all appropriate. So, what Saul calls the “Indeterminate Completion view” holds that none of the individual appropriate completions “has any claim to be the right one, although the range of acceptable completions does have a claim to correctness” (Saul 2012, 64). This indeterminate completion view results in an addition to Saul’s definition of lying to account for cases of this sort.

At first blush, it might seem that a similar line of reasoning would be helpful in dealing with cases of competing possible ideas of what is said, as with (a) through (d) from Figure X. They do all seem to be driving at similar ideas, after all. Perhaps it would be best to take the graph to be saying something indeterminate across those propositions. Taking this angle would involve the use of a more complete definition of ‘lying’ from Saul, where importantly, a lie told using indeterminate statements takes place when “for each complete proposition in the range CP1, ..., CPn, they believe that proposition to be false” (Saul 2012, 65). This complete definition of lying will introduce problems if the notion of indeterminate statements is taken seriously for graphs. For instance, if the graph from Figure X is taken, and the first data point is modified to sit just below the 35 million line, it seems to me that this, being an instance of a fabricated data point, would be one of the most straightforward possible ways to lie with a graph. However, if the indeterminate statements are taken seriously, then this data point still says (d), and (d) is still both true and believed to be true. So, it seems that the indeterminate statement approach does not solve the problem of the plurality of statements coming out of a graph.

3.2 Plurality

Ultimately, perhaps the best way to conceive of what is being said by a graph is really to take it as a straightforward plurality of statements. It is possible in ordinary speech, after all, to utter a conjunction of propositions: “Bob is tall and Bob is smart” is perfectly coherent. Perhaps a graph is in some sense an utterance of many propositions rolled into one: all of a)—or perhaps a*)—through d) are said, in addition to a great many more propositions relating to other data points, to overall trends, and to statements made by the legends and titles of the graph.

Plurality of this magnitude is somewhat troubling for any attempt at analysis of what is said by a graph, but it is not in itself deeply problematic. Setting aside the limitations arising from the finitude

of my life expectancy, and the possibility of infinite statements being made by a graph, I could utter an arbitrarily long conjunction of unrelated statements in plain English that would serve to say the same things, it would just be an incomprehensibly massive sentence. One of the points of data visualization as a practice is to make large data sets accessible to visual inspection (Tufte 2002, 13). It may be quite natural to think of the graph as amalgamating a huge range of utterances, on this understanding. Although this will not be my focus here, the plurality of propositions being communicated gives a basis for saying that a particular reader has understood a graph only partially: they have received some, but not all, of the propositions that it communicates.

The issue of plurality also helps address the concern with indeterminacy. If the first data point in Figure 9 is moved to just below 35 million, it now says many statements, including its nearest visual approximation, instead of saying something indeterminate across the many options. On this new understanding, the graph *is* now telling a lie: it lies about the analogues of (a) through (c). These new analogues might be put as:

- (a) The estimated population of Canada in Q1 2000 is 34,925,872
- (b) The estimated population of Canada in Q1 2000 is approximately 34,500,000
- (c) The estimated population of Canada in Q1 2000 is between 30,000,000 and 35,000,000, and is closer to 35,000,000 than to 30,000,000

Since the data point, in truth, has not changed (it remains 30,525,872), and I as the author know this, those utterances would be lies. The ability to handle this sort of change in a graph strikes me as a significant point in favour of the pluralistic notion of what is said by a graph.

3.3 Does it Say Anything at All?

A somewhat common opinion among statisticians is that data does not tell us—or say—anything at all. The data set simply is what it is, and any information beyond that which is reported is not a matter of the data saying anything, but rather of the analyst inferring or interpreting something (cf. Wijono 2020, Cianfrocco et al 2021). This is quite a minimalist account of what graphs can say. On an account such as this, the graph says nothing, and therefore, on every account of the lying-misleading distinction that I am aware of, cannot be lie. This strikes me as too simplistic, although I am deeply sympathetic to the idea that the data do not talk. Consider the graph I fabricated for Figure 1. All of the data points used in that graph are fictional. It seems very clear to me that if presented in a suitable context (one in which it is false, I know it to be false, and so on), such a graph would constitute a lie. This conclusion seems unavailable to us if a graph does not say anything at all. It may very well not constitute a lie in every possible circumstance—I do not believe that my including it here should be taken to be a lie—but the fact that it *could* be a lie under *any* circumstance at all means that it must say something. So, the aphorism of the data analysts, that the data doesn't say anything, will not help me to resolve my concerns with graphical deception.

3.4 Damned Liars, All of Them

Nearly the opposite approach is taken by Tufte, for whom it seems that any graphic which misrepresents anything is a lie. Recall the ubiquity with which he applies his metric of the Lie Factor; it applies, for Tufte, to many graphs which on my strong intuition are misleading without containing any lies. A good example of this is the FOX News graph from Figure 2. As we saw earlier, the so-called Lie Factor of this graph was roughly 33.3. In the language that Tufte would use, this graph is lying thirty-threefold. (Tufte 2001, 76).

Tufte's approach strikes me as reflecting a concern with deception, rather than with the notion of lying that is most typically taken to hang in some careful sense on what is said. It is a very broad notion of lying, one that encompasses any attempt to mislead or otherwise deceive, as well. Although this is undoubtedly helpful in improving the principles of data visualization, it does not do me much good for the purposes of investigating lying and misleading in graphs.

3.5 By Whom and To Whom

Another source of potential problems for any understanding of what is said by a graph is the temporal disjointedness of communication by way of graph. The definitions of lying that we have seen all make reference to at least one person—they all refer to a speaker—and both Carson and Stokke make explicit reference to a hearer or other sort of audience member in the definition as well (Carson 2010, Saul 2012, Stokke 2018). In figuring out whether a graph is lying, and if so, who is responsible, there are concerns with identifying both of these parties. Since graphics can be repeated over and over, in many instances of their use it is not clear who to hold accountable for their content. It will be necessary to preserve some sense of audience relativity, especially when it comes to misleadingness.

It seems to me that in general, the most sensible choice for the role of the speaker is taken by the author or designer of the graph, but this is not the only option we are confronted with. If a graph is re-used later, do we still hold the original author accountable? There may very well be graphs which are put together with the best intentions, that mislead only accidentally, and are then used to malicious ends by the nefarious. In my reuse of the graphs that I present as case studies, am I now responsible for their misleading contents? I am, after all, propagating them, and surely their authors could not possibly have imagined that their work should find its way into my thesis.

The most important takeaway from this chapter is the understanding that figuring out what exactly a graph says is made hugely difficult by the vast number of propositions that each one contains. Both specific and indeterminate understandings of what is said by a graph failed to reach the appropriate conclusions for several sorts of deception. With these caveats in mind, I will now turn to discuss the application of the theories more directly to a selection of graph examples that will allow me to highlight the precise ways in which existing accounts do not adequately capture the intuitions about misleadingness in graphs.

Chapter 4

Analysis of Examples

In order to make a proper judgement of the various accounts that I have analyzed up until this point, it will be helpful to see how they treat a range of examples of nefarious graphs. I will consider the applications of the theories of deception that I have so far discussed to several of the graphs that I introduced earlier. Ultimately, I will show that none of the theories of deception derived from an utterance-focused context do an adequate job of addressing the deception present in these graphs. This will involve building on the discussion of what is said by a graph from Chapter Three. I will also briefly revisit the problems with accounts like Tufte's, in order to keep my conclusions gathered together in one place.

I have selected some of the graphs that I introduced in earlier chapters of this thesis, which will allow me to illustrate which parts of the theories of lying and deception fail to address certain types of graph. First, I will consider what I take to be the most archetypal example of a misleading graph—one in which an altered axis exaggerates a change. We will see that although the accounts of Stokke and Carson handle this case well enough, there is some ambiguity to do with a possible use of hyperbole that poses a challenge for Saul's account, and that Tufte, in taking nearly all forms of graphical deception to be lying, does not have the kind of nuance in his analysis that would permit the desired result.

The second graph that I will consider is a graph which is something of a borderline case: in some contexts it may be misleading, but in others, I will argue, it is a lie. Both the Saul and Carson accounts reflect that this is a lie in those contexts. Ultimately, this graph seems to be a lie in cases where the labels are not clearly legible, in which case it bears more than a passing resemblance to, of all things, the doctrine of mental reservation.

To my surprise, I have also come upon a graph that I take to be best understood in terms of bullshit. We will see a graph used in a commercial, which I will argue is entirely unconcerned with truth or falsity, and is thus better understood through the lens of Frankfurtian bullshit than through any of the various accounts of lying and misleading that I have otherwise been focused on.

Finally, I consider a constructed toy example of data fabrication. I take this to be one of the clearest possible cases of lying with a graph. However, it seems that the Stokke account, because of its minimality conditions, does not adequately capture the lie of this sort of graph. Moreover, this is the sort of example where Carson's insistence that a lie be false runs into some deeply counterintuitive results. Altogether, this results in the conclusion that none of the existing accounts of lying and misleading work perfectly for graphs. They all have various strengths and weaknesses, but ultimately some sort of new approach will be necessary.

4.1 FOX News

Let's look at the graph from Figure 2, which we discussed in Chapter 1, in a little bit more detail.

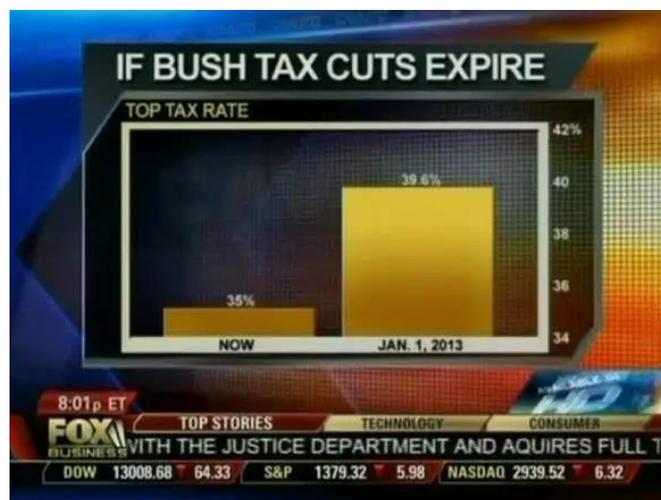


Figure 2

The core issue with this graph arises from the starting point of the vertical axis scale. Since the scale doesn't start at zero, it makes the two bars look far more different in size than the actual numbers that they are meant to reflect. As we have calculated earlier, the percentage increase in the height of the bars is slightly more than 33 times greater than the percentage increase in the top tax rate. Although Tufte would call this a Lie Factor of 33 times, it is, I think, the most classical sort of example of a misleading (rather than lying) graph, and so I will apply the three main definitions of lying that I have been using to it, in order to see how they handle it.

4.1.1 Carson

Recall that Carson's preferred final definition of a lie is "L5. A person S tells a lie to another person S1 iff: 1. S makes a false statement X to S1, 2. S believes that X is false or probably false (or, alternatively, S does not believe that X is true), 3. S states X in a context in which S thereby warrants the truth of X to S1, and 4. S does not take herself to be not warranting the truth of what she says to S1" (Carson 2010, 30). There are some technicalities that get us in to trouble right away: firstly, we are not clear on which persons S and S1 are meant to be. This is a complicated problem to resolve because graphs, once created, can be presented in a wide range of contexts. There is no uniquely-defined context of utterance, and no uniquely-defined audience, for most graphs. It is necessary to specify, in most cases, specific instances of graphs, and the audiences to which those graphs are being presented.

In the particular case of Carson's definition, it is not at all immediately clear who S and S1 should be. There are, following in Sidelle's example, two clear candidates for the utterer: either the graph's author, or the graph itself. Since my general practice will be to take the graph itself as being analogous to an utterance, that means that the graph author should be taken as the utterer. This instinct will ultimately raise problems in the case of repeated and propagated graphics.

For the purposes of continuing my own analysis of Carson's theory, I will simplify matters by focusing on an original presentation of this graph on the FOX News network. The speaker S, then, is the news anchor, and we can take an audience member for the news broadcast to fill the role of S1. The graph also cannot, on Carson's account, be a lie, in virtue of being true. No matter what the author of the graph might have believed, because the tax rates are true and accurate, it's not a lie. This is as it should be. The graph, I take it, is not a lie, but is misleading.

For the sake of completeness in the analysis, I will say that I do take it that the graph author believes the figures being reported by the graph, that the context of a news program is one in which they are warranting the truth of their graph, and the author takes that to be true as well.

4.1.2 Saul

Now, recall the definition of lying given by Saul, which is developed particularly for the purposes of understanding the lying-misleading distinction: "If the speaker is not the victim of linguistic error/malapropism or using metaphor, hyperbole, or irony, then they lie iff (1) they say that P; (2) they believe P to be false; (3) they take themselves to be in a warranting context" (Saul 2012, 18). Although the truth of the content of the graph does not let the author off the hook as immediately here, given that the speaker, we assume, believes what they are saying about those contents, if this definition does pass judgement on this graph, it will probably not be a lie. The author, as above, takes themselves to be in a warranting context.

What is said, on Saul's account, is the semantic content of the sentence, together with the contextual supplementation that is needed for truth evaluability. The latter is allowed in order to deal with cases of completion in sentences. In the case of this graph, which is—for all its faults, and setting aside the indexical 'now' labelling the first column—clearly and comprehensively labelled, it is not clear that anything resembling completion is going to apply, and so we can be contented with

the direct content of the graph as what is said. This results in an understanding of what is said, and a verdict on its status as a lie, similar to the verdict from Carson, with the notable exception that being true does not immediately get the graph off the hook. We need the extra piece that the author believes it to be true. That this is the case strikes me as very likely indeed.

The trickiest aspect of Saul's definition of lying is in the limitations on when it can be applied. Because I take it as quite clear that this graph is not lying, we must ensure that the graph is not somehow using some figure of speech, or the victim of some kind of linguistic error.

As far as most of the figures of speech that Saul mentions, this is fairly obvious. There are no linguistic errors here, and no malapropism. The graph does not seem to be ironic or metaphorical in any way. There are nevertheless two things that might be a small problem for this account. First, we should convince ourselves that the graph's exaggeration of the difference between the pre- and post-cut tax rates is not hyperbole, or something like it. Second, we should make sure that this is not some kind of error analogous to malapropism or linguistic error.

These two issues are somewhat troubling to resolve. It is not altogether clear whether or not there even *is* an analogue to linguistic error that would make sense in this context. I suspect that there probably is, but I don't think it is at play here. There is nothing in the presentation of the graph that seems like a slip-up of words, or a case of mere mistake. The question of hyperbole merits some consideration. The core of what we find misleading about this graph is that it exaggerates an effect, by a factor of roughly 33. We do not usually think of hyperbole taking place in a graph, but imagine someone who, after scoring a 30 and then a 90 on two successive exams, proudly declares to their friend that they "did a hundred times better on the second one". It seems intuitively clear to me that this is a case of hyperbole, but it is *precisely* a case in which the speaker has exaggerated a change by a factor of roughly 33. I do not yet see a strong argument to claim that the graphical version is or is not a case of hyperbole, and I find this troubling. There is a difference between the two, in that the

speaker about the exam may be taken as engaging in hyperbole in a way that the presenter of the graph might not be, but this seems to put too much weight on how the speaker is taken, when it comes to determining if an utterance is hyperbole or not. In any event, it remains a troubling possibility.

Saul's definition does not take the graph to be a *lie* in either case, but since my intuition is that this graph is one of the clearest possible cases of misleading that is not a lie, the risk that the definition might fail to deliver a verdict is concerning. So much of graphical misleading is tied up in exaggeration, and there is a blurry distinction between hyperbole and exaggeration at the best of times. This example might suggest that the categories that are excluded from Saul's definition are a particularly relevant class of examples for the purposes of graphical deception. In general, I find this to be a solid account of the lying-misleading distinction, but in the specific case of graphical misleadingness it seems that the caveats happen to render it unable to make a verdict on a particularly common phenomenon.

4.1.3 Stokke

Stokke's account of lying says that: "A lies to B if and only if there is a proposition p such that: (L1) A says that p to B, and (L2) A proposes to make it common ground that p, and (L3) A believes that p is false" (Stokke, 31). As with Carson's definition, the diachronic nature of communication via graph raises the now-familiar challenges of figuring out who A and B could possibly be in this case, but we will simply specify the author and an arbitrary reader for the purposes of having something concrete to analyze.

Firstly, I will note that there are a great many propositions contained in the graph. For instance, it certainly contains:

- i) The top tax rate is now 35%
- ii) If the Bush-era tax cuts are allowed to expire, the top tax rate in 2013 will be 39.6%

It might also contain the following:

- iii) 39.6% is greater than 35%
- iv) 39.6% is *considerably* more significant than 35%, on the order of four times as much
- v) The first figure is considerably more significant than the second figure

I take it that ii) and iii) are both reasonably clear. Proposition i) raises problems about the use of the indexical ‘now’—but if we take to ‘now’ to refer to the time of the creation of the graph, then there is no issue to be had with this proposition either. A proposition like iv) or v) is what is suggested by the heights of the bars in the graphic, and are the only two of the five that strike me as being problematic in their content. However, I do not think that the graph *says* that iv) or v). Instead, it seems to me that this is an inference that will be reached by those who are merely glancing at the graphic as it appears on the screen. Recall that for Stokke, what an utterance S in a context c , taking $\mu_c(S)$ to be the minimal proposition of S , says in response to a QUD q is: “the weakest proposition p such that (i) p is an answer (partial or complete) to q and (ii) either $p \subseteq \mu_c(S)$ or $\mu_c(S) \subseteq p$ ” (Stokke, 99). So, in order to make sense of what is *said* by the graphic, it is necessary to identify S , c , $\mu_c(S)$, and q . The formidable linguistic machinery that Stokke brings to bear in this definition makes it challenging to permute into an understanding of mostly-visual communication, but let us roll up our sleeves and say *bon courage!*

The statement, or sentence, S , is best taken to be the graph itself, and the context can stay as the context (i.e., that S is presented by FOX News, in 2012, *et cetera*). The determination of $\mu_c(S)$ is considerably more challenging, because the graphic presents a multitude of information. As we have discussed already, it is not sufficient to identify any one proposition as being the core of what is contained by a graph, since their purpose is in large part to present a considerable aggregation of information. Since there is no elegant way of choosing a single $\mu_c(S)$, I will table it for the moment.

The ability to identify a QUD is of huge benefit to the project of identifying what is lying as opposed to misleading. In this particular case, if we choose q to be something like Q_1: “How will the top tax rate change if Bush-era tax cuts expire?” then a minimal proposition that serves as an answer to Q_1 might be something like A_1: “The current top tax rate of 35% will increase to 39.6%.” This proposition is clearly both true and believed, and so, on Stokke’s account, is not a lie. This is just as I would hope. It also gives us the flexibility to understand what is said by a graph in a much more usefully pluralistic way. Earlier in this thesis, I introduced challenges to figuring out what is said by a graph, since it seems to make many different statements all at once. For instance, the graph seems to also say “the current top tax rate is 35%,” in response to the QUD “What is the current top tax rate?” and also to say “the top tax rate will increase” in response to a QUD focused only the relative magnitude of the tax rate over time.

This class of example seems to be unambiguously well-handled by the accounts of Carson and Stokke. Tufte’s principles, in blanketing all forms of deceptive graphic as being lying, are not well-suited to the purposes of my project, despite their valuable insights into shortcomings of integrity in graphics.

4.2 Planned Parenthood Services

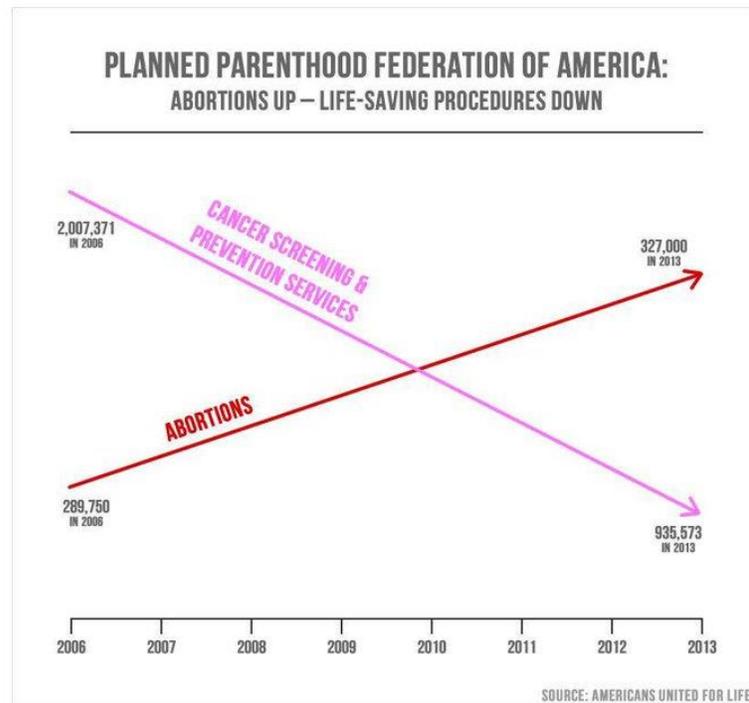


Figure 5

At this point, I return to the graph of Planned Parenthood services from Figure 5. The core deceptive act of this graph is not a truncation of its vertical axis (although that does happen here as well), but rather it is the use of double axes coupled with their omission from the graphic. This combination of tricks results in a graph that gives the impression that the number of abortions performed by Planned Parenthood has surpassed the number of cancer-screening and prevention services. Closer inspection reveals this to not be true; the points on the graph that are labelled reveal that despite a significant drop in cancer screening and prevention services, they still considerably outnumber abortions. There are other tricks involved here as well. The graph also makes it seem that the rates of change have been roughly equal in magnitude, which they absolutely have not, and also that the rates of change have been constant over time, which there is no real evidence here to support.

This graph is one which I ultimately take to be lying¹². Let us see which accounts of deception yield the desired result for this graph, as well.

For all three of the accounts of lying that we have seen, it will be paramount to figure out what it is that this graph is *saying*, and indeed it seems to me that there is one very important proposition which we must decide if the graph says or not. That proposition is (P): The number of abortions has surpassed the number of cancer screening and prevention services performed by Planned Parenthood.

It is, I think, uncontroversial that this graph strongly suggests that (P), by depicting the two data series lines intersecting. Whether or not the graph ultimately says this is somewhat less obvious. This proposition in particular stands out as important because it is so strongly implied by the visualization in place. It is the clearest false takeaway that we are invited to believe by looking at the graph.

On Carson's account, the relevant criteria are to verify that the graph contains a false statement that the author believes to be false, and that this statement is made in a warranting context that is not taken not to be a warranting context. This graph was created by Americans United for Life, as part of a report that purported to show that Planned Parenthood had become some sort of so-called abortion profiteer, but was presented by Congressman Jason Chaffetz during a public hearing (Roth 2015). Testimony during a congressional hearing should certainly be taken as a warranting context, and there is no reason for Rep. Chaffetz to not take it in this way. So, for Carson, it will hang on whether or not a statement like (P) has been made with this graph.

Intuitions may differ on this matter, but I think that this graph does make the statement (P), especially in the context of Rep. Chaffetz's use of it, in which the labels are far too small to be

¹² I might, in fact, describe this graph as deeply evil, but that is a moral question for another paper.

comfortably read (Qiu 2015). In the absence of clarity and legibility of the labels that do show the accurate numbers, it seems like this graph must be a lie. I see this as being analogous to a case like the doctrine of mental reservation, in which a priest might say something like ‘I am not a priest...’ out loud, and then complete the sentence in their head: ‘... of Apollo’ and thereby exculpate themselves from the verbally uttered falsehood (Slater 1911). Despite what the doctrine of mental reservation might say, I take it that this is clearly a lie. The nearly-illegible labels of this graph fill a similar role to the mental reservation. This is perhaps a place where the importance of the audience who receives the statement is clear: it might be less obvious that this graph would be a lie when the numbers in the labels are clearly visible.

I think that Saul’s Needed for Truth Evaluability condition yields the same result here: in the absence of legible labels, interpreting the intersection of the lines as indicative of intersection in the data they represent becomes necessary to complete the utterance. Thus, this intersection becomes a part of what is said, and the graph becomes a lie. I am not, however, sure that Stokke’s account handles this particular case as well. The minimal content and the necessity of an appropriate QUD raise challenges, since between the two propositions “cancer screenings decreased in number and abortions increased in number” and “cancer screenings surpassed abortions in number” it is not clear which is the minimal. It seems more likely that these two propositions are responses to slightly different QUDs, but given the implicit nature of so many QUDs, it is tremendously difficult to pin down exactly which one the graph is serving to answer.

4.2.1 Graphical Bullshit

In Chapter 2, I briefly introduced Harry Frankfurt’s account of bullshit. When I began the process of researching these graphs, it never occurred to me that there might be a graph to which the Frankfurtian notion of bullshit could be helpfully applied. However, there is a type of graph for which

I think it is appropriate. One of the classic examples is this graph, from a commercial for Lanacane itch relief cream.

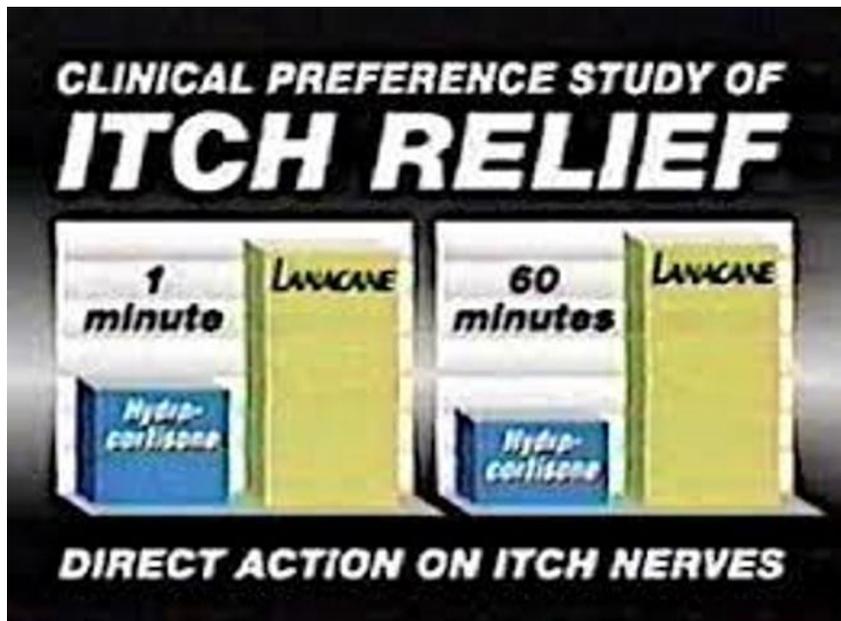


Figure 10: A graph that is bullshit (Statistics How To)

This graph is, I claim, a terrific analogue to the Frankfurtian notion of bullshit, as being entirely unconcerned with the truth. It would be, I think, very challenging to conclude anything about what a graph like this is meant to be communicate, or to say. It is clearly meant to be suggesting something about the superiority of Lanacane compared to other itch relief products, but I do not think that it makes much sense to talk about what this says as being meant to be true or false—there is simply not enough information in the utterance to make any sense of that whatsoever.

One of the classic examples of a bullshitter is the particularly underhanded salesperson who will say anything, true or false, without any concern for the truth value of the utterance, in order to secure a sale. Such a salesperson might make an utterance of “this car has no problem with overheating” without any idea of if it is true or false, in the hopes that it will make the prospective

buyer buy the car. This Lanacane graph is a very similar sort of utterance. There cannot possibly be concern for the truth of the utterance made with the graph, because the graph utterance is so incomprehensibly stupid as to not possibly have truth associated with it. If, as Frankfurt says, a lie is meant to “insert a particular falsehood at a specific point,” then the Lanacane graph certainly has not lied (Frankfurt 2005, 51). I would not wish to claim that this graph has inserted any specific falsehoods at all. Rather it seems to be deceptive in a way more analogous to bullshit: it is uttered in the service of some ulterior motivation, without the least regard for the truth of its contents.

4.2.2 The Most Basic Lie

Earlier, I suggested that the most straightforward possible way to lie with a graph is to simply make up data points and use them in the graph. I take it as uncontroversial that would be a lie. However, it turns out to be an example that poses a serious problem for Stokke’s account. As we have seen, the minimality constraint causes a lot of problems for us in attempting to apply Stokke’s account of what is said in the graphical context. Consider the following graphs, which I have fabricated, and suppose that the former graph depicts real data points, and the latter depicts a pure deceptive fabrication:

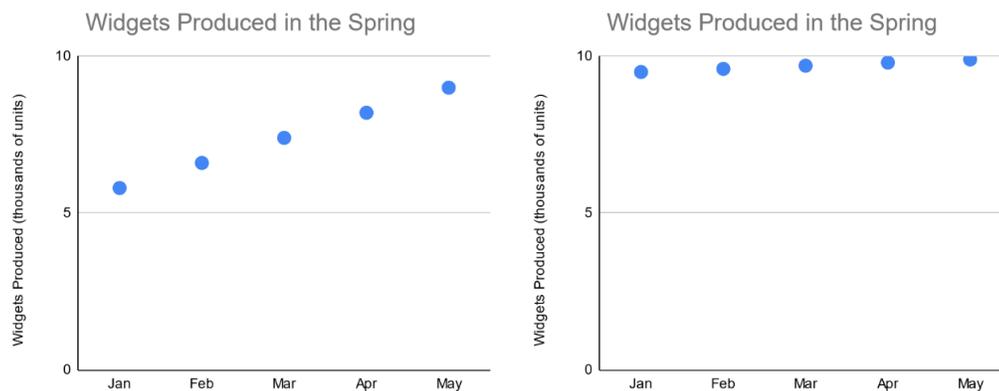


Figure 11: Fabricated Data

As I have argued in Chapter 3, Stokke's account, because of the minimality condition, does not let us claim that either graph says more about its individual points than that they fall between 5 and 10. Because "x is between 5 and 10" is a weaker proposition than "x is 6" or "x is 9," the former proposition is taken by Stokke's account to be what is said. Both series are strictly increasing, and so it seems like Stokke's account does not do a very effective job of distinguishing these two examples, despite the fact that one is the most intuitively clear case of lying. A completely fabricated graph also introduces a problem for Carson. Since his account of lying takes it as necessary that a falsehood has been told, if I turned out to have been right, despite believing that I am wrong, about the widget production in the graph above, then it would not be a lie. This is not unique to graphs, but it is the sort of case that arrives at deeply unintuitive results for the Carson account.

Chapter 5

Conclusion

In the course of this paper, I have examined several accounts of lying and deceptive speech, in order to see which of them are best able to shed light on the problem of misleadingness in data visualizations. I considered the principles of graphical integrity put forward by Edward Tufte, and three philosophical accounts of what it is to lie: from Thomas Carson, Jennifer Saul, and Andreas Stokke. Ultimately, none of these accounts are adequate for addressing the complexities of graphics as utterances.

Tufte's account, which is the best of the options from the mathematicians' side, is less rigorous and focused more on principles that can be employed to promote what he calls graphical integrity. Tufte would categorize many graphs which are merely misleading as being lies. The distinction between lying and misleading does not feature in his analysis, which is somewhat more mechanically focused on the creation of data visualizations. His account also does not offer a sufficiently systematized approach, focusing on principles that serve to pick out 'red flags' for graphs.

When considered from the angle of philosophy of language, understanding what it is that makes a graph lying, or misleading, hangs significantly on understanding what is said by a graph. This, unfortunately, is an area of major complication. Because graphics are intended, in large part, to communicate large aggregations of data in relatively little space, they communicate a vast spread of propositions: in most cases at least one for each point of data depicted on the graphic. Thus, any attempt to make sense of what is said requires either a pluralistic understanding of what is said, encapsulating each of these propositions, or else a reductive account that reduces what is said to a comment on relative standing of the data points contained in the graph. Neither option, as we have seen, is terribly satisfying.

It seemed that Stokke's account, which tied what is said to the current Question Under Discussion (QUD), would stand the best chance of capturing the pluralistic sense, since each possible proposition could be a possible response to a different QUD. However, it ultimately turned out that the minimality constraints on what is said pose a serious challenge to an understanding of what is said by a graph. This arose most notably in the case of scatterplot and line graphs, wherein the position of a data point not only gives a rough estimate—such as “the value of the data point is roughly x —but also always necessarily gives a coarse bound on the range that the point falls within—say that “the value of the data falls between a and b .” The minimality conditions mean that we are forced to take the graph as saying the weaker of the two choices: that the value is between its bounds. This lets some graphs that are definitely lying off the hook for their lies.

Ultimately, the problem of identifying lying and misleadingness in graphs is not something that any of the existing literature seems to be able to solve. In fact, the challenges that face any attempt to resolve the problem are so significant that I have been able to do nothing more than begin to identify them, in the limits of time and space that constrain this thesis. It seems clear that the best route for approaching these problems lies with a pluralistic account of what a graph can say, but none of the accounts of what is said are yet designed to address massively pluralistic utterances in this way.

References

- Carson, T. (2010). *Lying and Deception: Theory and Practice*. Oxford/New York: Oxford University Press.
- Chisholm, R. and T. Feehan (1977) 'The Intent to Deceive', *Journal of Philosophy* 74(3), 143-59.
- Cianoffro, A. et al. (2021). 'Data Does not Speak for Itself: You Have to Speak for It', *Industrial Safety and Hygiene News*, 6 May, at: <https://www.ishn.com/articles/112960-data-does-not-speak-for-itself-you-have-to-speak-for-it> (Accessed August 2022)
- Engel, P. (2014). 'This Chart Shows an Alarming Rise in Florida Gun Deaths After 'Stand Your Ground' was Enacted', *Business Insider*, 18 February, at: <https://www.businessinsider.com/gun-deaths-in-florida-increased-with-stand-your-ground-2014-2> (Accessed May 2022)
- Fallis, D. (2009). What is Lying? *Journal of Philosophy* 106, 29-56.
- Frankfurt, H. (2005 [1986]). *On Bullshit*. Princeton, NJ: Princeton University Press.
- Grice, H. P. (1989). *Studies in the Way of Words*. Cambridge, MA: Harvard University Press.
- Hickey, W. (2012). 'EXPOSED: Here are the Tricks that Fox News Uses to Manipulate Statistics on its Graphics', *Business Insider*, 28 November, at: <https://www.businessinsider.com/fox-news-charts-tricks-data-2012-11> (Accessed May 2022)
- Lallanilla, M. (2014). 'Misleading Gun-Death Chart Draws Fire', *Live Science*, 23 April, at: <https://www.livescience.com/45083-misleading-gun-death-chart.html> (Accessed May 2022)
- Levitin, D. J. (2017). *Weaponized Lies: How to Think Critically in the Post-Truth Era*. New York, New York: Dutton.
- Mahon, J. (2007). A Definition of Deceiving. *International Journal of Applied Philosophy* 21, 181-194.

- Mahon, J. (2008). Two Definitions of Lying. *International Journal of Applied Philosophy* 22(2), 211-30.
- Monmonier, M. S. (1991). *How to Lie with Maps*. Chicago: University of Chicago Press.
- O’Neil, C. (2017). *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy*. First paperback edition. New York: Broadway Books.
- Predelli, S. (1998) ‘I Am Not Here Now’, *Analysis*, 58: 107-12
- Qiu, L. (2015). ‘Chart Shown at Planned Parenthood Hearing is Misleading and ‘Ethically Wrong’’, *Politifact*, 1 October, at: <https://www.politifact.com/factchecks/2015/oct/01/jason-chaffetz/chart-shown-planned-parenthood-hearing-misleading-/> (Accessed May 2022)
- Roberts, C. (2012). Information Structure in Discourse: Towards an Integrated Formal Theory of Pragmatics. *Semantics and Pragmatics* 5, 1-69.
- Roth, Z. (2015). ‘Congressman Uses Misleading Graph to Smear Planned Parenthood’, *MSNBC*, 29 September, at: <https://www.msnbc.com/msnbc/congressman-chaffetz-misleading-graph-smear-planned-parenthood-msna692596> (Accessed May 2022)
- Saul, J. (2012). *Lying, Misleading, and What is Said: An Exploration in Philosophy of Language and in Ethics*. Oxford/New York: Oxford University Press.
- Schoubye, A. and A. Stokke (2016). What is Said? *Noûs* 50(4), 759-93.
- Sidelle, A. (1991). The Answering Machine Paradox. *Canadian Journal of Philosophy* 21(4), 525-39.
- Slater, T. (1911). ‘Mental Reservation’, *The Catholic Encyclopedia*. New York: Robert Appleton Company. Retrieved August 2022 from New Advent: <https://www.newadvent.org/cathen/10195b.htm>
- Sorenson, R. (2007). ‘Bald-faced Lies! Lying without the Intent to Deceive’, *Pacific Philosophical Quarterly* 88, 251-64.
- Stalnaker, R. (1978). ‘Assertion’, *Syntax and Semantics (New York Academic Press)* 9, 315-332.

- Stalnaker, R. (1984). *Inquiry*. Cambridge, MA: MIT Press.
- Stalnaker, R. (2002). Common Ground. *Linguistics and Philosophy* 25, 701-21.
- Statistics Canada. Table 17-10-0009-01 Population estimates, quarterly.
- Statistics How To. ‘Misleading Graphs: Real Life Examples’, *Statistics How To*, at:
<https://www.statisticshowto.com/probability-and-statistics/descriptive-statistics/misleading-graphs/> (Accessed July 2022)
- Stokke, A. (2014). Insincerity. *Noûs* 48(3), 496-520.
- Stokke, A. (2018). *Lying and Insincerity*. Oxford/New York: Oxford University Press.
- Tufte, E. (2001). *The Visual Display of Quantitative Information Second Edition*. Cheshire, CT: Graphics Press.
- White House (2022). [Twitter] 27 January. Available at:
<https://twitter.com/whitehouse/status/1486709480351952901?lang=en> (Accessed May 2022)
- Wijono, W. (2020). ‘Data Never Says Anything on its Own: Stop Putting Words in Data’s Mouth’, *Towards Data Science*, 4 August, at: <https://towardsdatascience.com/no-the-data-never-says-anything-510edbd9b43a> (Accessed August 2022)
- Yanofsky, D. (2013) ‘The Chart Tim Cook Doesn’t Want You to See’, *Quartz*, 10 September, at: <https://qz.com/122921/the-chart-tim-cook-doesnt-want-you-to-see/> (Accessed May 2022)