Plotting Apples, Oranges, and Distributions

A New Taxonomy to Prevent Information Inequalities In Uncertainty Visualisations

Harriet Mason

6/6/23

Motivation

Think back to the last time you made some sort of data visualisation. What was the purpose of that visualisation? Was it to better understand your data? Was it to help you make a decision? Was it to to communicate that decision to someone else? There are many stages in our analysis that benefit from the power of data visualisation, however this does not mean it is always done with success. Visualization is an important step in exploratory data analysis and it is often utilised to **learn** what is important about a data set. The importance of data driven discovery is highlighted by data sets such as Anscombe's quartet (Anscombe 1973) or the Datasaurus Dozen (Locke and D'Agostino McGowan 2018). Each of the pairwise plots in these data sets have the same summary statistics but strikingly different information when visualised. The anscombe quartet is shown in Figure **??**. Instead of having to repeatedly check endless hypothesis to find interesting numerical features, visualisations **tell** us what is important about the data set. This powerful aspect of data visualization is poorly or seldom used in later stages when we are communicating our findings, specifically with respect to uncertainty.

Utilising visualisation can give people a more complete understanding of risks. Studies asking participants to sketch a distribution allowed them to better compute statistics about that distribution and improve predictions (Hullman et al. 2018; Goldstein and Rothschild 2014). While there is some evidence that confidence provided in text form only are less likely to be misinterpretated than grahics (Savelli and Joslyn 2013), text is insufficient to express more complicated aspects of a distribution, such as mass. The confusion caused by visualisation could also be due to a lack of expose, since Kay et al. (2016) found that people exposed to the same uncertainty visualisation get better at making judgements the more they are exposed to them. Additionally, visualisation allows for interactive graphics that provide a more in depth understanding of probability (Potter et al. 2009a; Ancker, Chan, and Kukafka 2009) and infographics that make uncertainty more accessible for people with poor numeracy skills (Ancker, Chan, and Kukafka 2009).

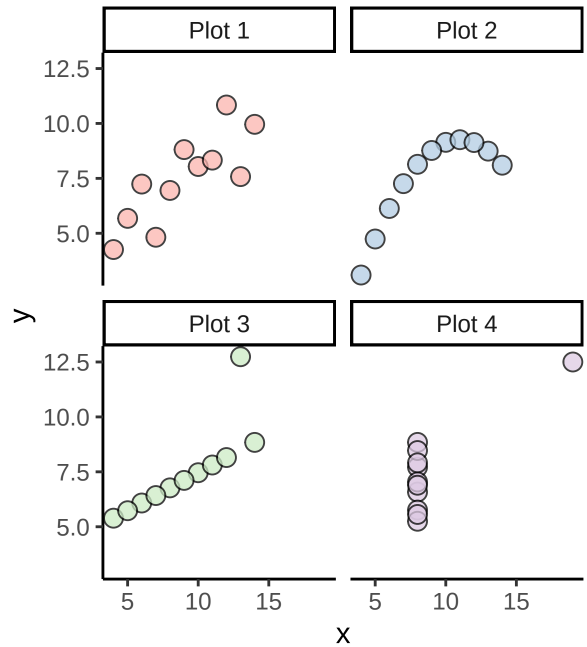


Figure 1: The four scatter plots that make up the anscombe quartet. Each x and y variable has the same mean, standard deviation, and correlation

Despite these benefits, there is a reasonable amount of annecdotal and survey evidence that we don't visualise uncertainty as often as we should. Some economists suggest that visualisation authors are responding to incentives that make it tempting to avoid visualising uncertainty, even if those incentives are based more in perception than reality (Manski 2020). A survey conducted by Hullman (2020) found that majority of visualisation authors agreed that expressing uncertainty is important and should be done more often than it currently is, some even agreed that failing to do so is tantamount to fraud. Despite this, only a quarter of respondents included uncertainty in 50% or more of their visualisations (Hullman 2020). Meaning participants were convinced that visualising uncertainty is morally important but were able to provide self sufficient reasoning that allows them to avoid doing it. The study by Hullman (2020) found that the most common reasons authors don't visualise uncertainty in practice despite knowing it's moral importance are: not wanting to overwhelm the audience; an inability to calculate the uncertainty; a lack of access to the uncertainty information; and not wanting to make their data seem questionable (Hullman 2020).

If decision markers are not presented with the uncertainty about an estimate the data analysts have, for all intents and purposes, made the decision for the decision maker. Upon further interviews Hullman (2020) found that authors believed uncertainty would overwhelm the audience and make their data seem questionable because decision makers are unable to understand uncertainty. This belief, while pervasive, is not true. While some research suggests that laypeople cannot understand complicated concepts in statistical thinking (such as trick questions on hypothesis tests or the difference between Frequentist and Baysian thinking) (Hoekstra et al. 2014; Bella et al. 2005) there is a large amount of research suggesting that presenting uncertainty information improves decision making, both experimentally (Joslyn and LeClerc 2012; Savelli and Joslyn 2013; Kay et al. 2016; Fernandes et al. 2018) and in practice (Al-Kassab et al. 2014). As a matter of fact, doing what many authors currently do (providing only a deterministic outcome with no uncertainty) causes decision makers to be *less* decisive and have completely unbounded expectations on an outcome (Savelli and Joslyn 2013). This reality cannot be avoided by providing secondary or non-specific information such as explaining calculations (Joslyn and LeClerc 2012), explaining the advantages of a recommendation (Joslyn and LeClerc 2012), or expressing uncertainty in vague terms (Erev and Cohen 1990; Olson and Budescu 1997), all of which are undesirable for decision makers and lead to measurably worse decisions (Joslyn and LeClerc 2012; Erev and Cohen 1990; Olson and Budescu 1997). Expressing uncertainty verbally additionally decreases the percieved reliability and trustworthiness of the source (Bles et al. 2020). One of the most popular depictions of uncertainty for decision making is a quantile dotplot, shown in Figure **??**.

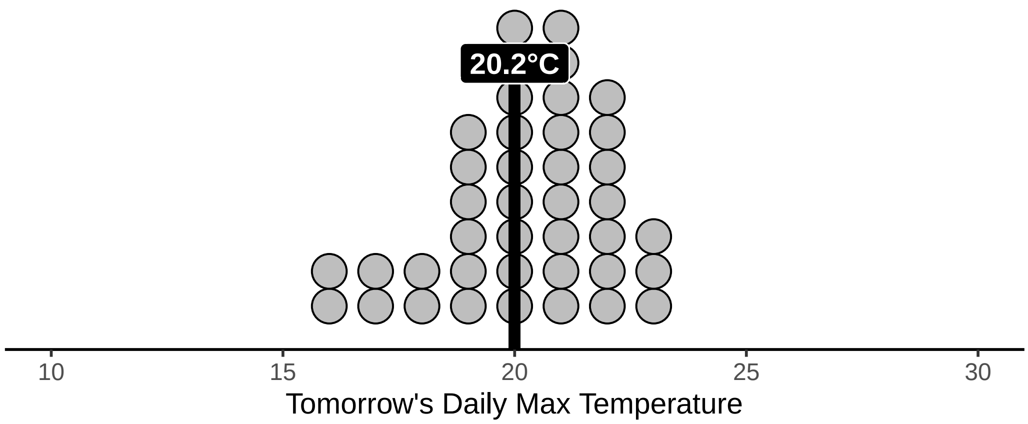


Figure 2: A quantile dotplot is an uncertainty visualisation that provides a discrete display of uncertainty to aid decision making. This example plot provides an estimate for a daily maximum temperature.

Not only does communicating uncertainty improve decisions but the mistrust created by communicating certainty in uncertain situations can be exploited. A 6-month survey of anti-mask groups on Facebook during to COVID-19 pandemic showed that the anti-maskers thought carefully about their grammer of graphics and made pursuasive visualisations using the same data as pro-mask groups by exploiting information ignored by the pro-maskers (Lee et al. 2021). It is understood that deceptive plots can lead viewers to come to incorrect conclusions or significantly overstate effects or risks (Pandey et al. 2015; L. Padilla et al. 2022) but these incorrect takeaways cannot be mitigated with instructions in how to correctly understand the plot (Boone, Gunalp, and Hegarty 2018). This evidence indicates we are more likely than not to hurt our message when we ignore uncertainty information and trying to raise the general publics plot literacy is an insufficient strategy to curb conspiracy theories and misguided scientific communication. In direct contrast to this, displaying numerical estimates of uncertainty information has shown to lead to greater trust in predicitions (Joslyn and LeClerc 2012; Bles et al. 2020). While Han et al. (2009) found people have more worry when presented with uncertainty reguarding health outcomes, this worry is not a bad thing if the concern is warranted given the ambiguous situation.

The disconnect between the research supporting uncertainty and the consensus aganinst may not be entirely driven by a lack of understanding of the literature. For example, at least one interviewee from the study by Hullman (2020) claimed that expertise implies that the signal being conveyed is significant, but also said they would omit uncertainty if it obfuscated the message they were trying to convey (Hullman 2020). Other authors who were capable of calculating and and representing uncertainty well did not do it, and were unable to provide a self-satisfying reason why (Hullman 2020). These conflicting motivations are acknowledged in the paper itself where Hullman (2020) says:

> "It is worth noting that many authors seemed confident in stating rationales, as though they perceived them to be truths that do not require examples to demonstrate. It is possible that rationales for omission represent ingrained beliefs more than conclusions authors have drawn from concrete experiences attempting to convey uncertainty".

An overwhelming consensus among visualisation authors seems to be that uncertainty is secondary to estimations. There is a belief help by those that work with data that the uncertainty associated with an estimate (the noise) only exists to hide the estimate itself (the signal). From this view, uncertainty is only seen as additional or hindering information, therefore despite its alleged importance, when simplifying a plot uncertainty the first thing to go. This belief is also reflected in the development of new uncertainty visualisations. Often when trying to visualise a high problem, uncertainty is relegated to unimportant aesthetics in the plot, often of lower importance than the estimate (Correll, Moritz, and Heer 2018; Lydia R. Lucchesi and Wikle 2017). This is not uncommon with high dimensional data considering spatial temporal data often Cases where uncertainty is not relegated to an undesirable aesthetic instead incorporate interactivity to allow users to explore the complicated space themselves (Potter et al. 2009b, 2009a). Even the literature about uncertainty communication expresses an implicit belief that it is of secondary importance to the estimates or context of the data.

An adjacent issue is *how much* uncertainty we should include when trying to quantify our associated distribution. Obviously the correct answer is somewhere between "every possible outcome" which would result in an unbounded uncertainty and "a simple confidence interval based on a set of strict and unrealistic assumptions" that would result in an interval that is far too narrow. Unfortunately between those two extremes the solution is largely a judgement decision which can sometimes be overwhelming. This is why software that provides *some* uncertainty visualisation as a default, such as forecasts in the `fable` package are useful (O'Hara-Wild, Hyndman, and Wang 2023). It prevents authors omitting uncertainty through inaction.

There is the possibility, however, that default uncertainty visualisations facilitate the poor understanding of which conclusions are relevant to the uncertainty visualised.

This issue is not helped by the fact that the term “uncertainty” lacks a commonly accepted definition in the literature. Lipshitz and Strauss (1997) even commented that “there are almost as many definitions of uncertainty as there are treatments of the subject” . This mishmash of terminology leads to a large body of work, all claiming to finding the best visualisation or expression of of “uncertainty” but most don’t even seem to agree on what uncertainty is. The most encompassing definition of uncertainty I have seen comes from Walker et al. (2003) who define uncertainty as **“any deviation from the unachievable ideal of completely deterministic knowledge of the relevant system”**. This definition does not completely align with the distribution conceptualisation I discussed earlier in this chapter, but the cases where they diverge are rare and can still be handled by focusing on the relevant information. More commonly, uncertainty is defined using a taxonomy rather than a strict definition. There are a few of taxonomies for uncertainty, but, just like the definition, most of them are a subset of the one laid out by Walker et al. (2003).

Figure **??** is an illustration of the taxonomy presented by Walker et al. (2003). In this taxonomy, there are three things we need to consider for each “uncertainty” we encounter through the modelling process. First, consider the source of the uncertainty. Is this uncertainty coming from inaccurate measurements or a poorly defined model? This is the *location* of the uncertainty. Second, consider how well you can quantify this uncertainty. Do you know exactly how much measurement error there is in each observation or are you not even aware if there is or isn’t measurement error? This is the *level* of your uncertainty, and it ranges from discrete to total ignorance. Finally, consider how this uncertainty came into existence. Is it a result of a naturally random process (epistemic) or is it due to imperfect information and could be improved (aleatory). This is the *nature* of your uncertainty. Walker et al. (2003) then go on to describe mapping our uncertainty in a 3D space that is defined by its location, level, and nature, but I think the taxonomy is more easily understood as a series of questions we need to consider when asking how much unertainty we should express to our audience.

While information about the sources of our uncertainty and the type of uncertainty may seem like an unimportant secondary step in uncertainty visualisation, communicating these features of uncertainty helps decision makers make more informed choices. L. M. K. Padilla et al. (2021) found that low forecaster confidence or high model uncertainty both contribute to more conservative judgements by decision makers. Failling to communicate the nature of your uncertainty can result in underestimation or overestimation of failure probabilities (Kiureghian and Ditlevsen 2009). Additionally Gustafson and Rice (2019) found that the framing of our uncertainty, (i.e. informing the reading if the uncertainty came from a lack of knowledge, approximations, unknown unknowns, or disagreement among parties) does not have a significant effect on the belief in the estimates, perceived credibility, or behavioral intentions of the decision makers. This means communicating secondary information about your uncertainty is unlikely to negatively affect your communication and is important in understanding the scope of your uncertainty.

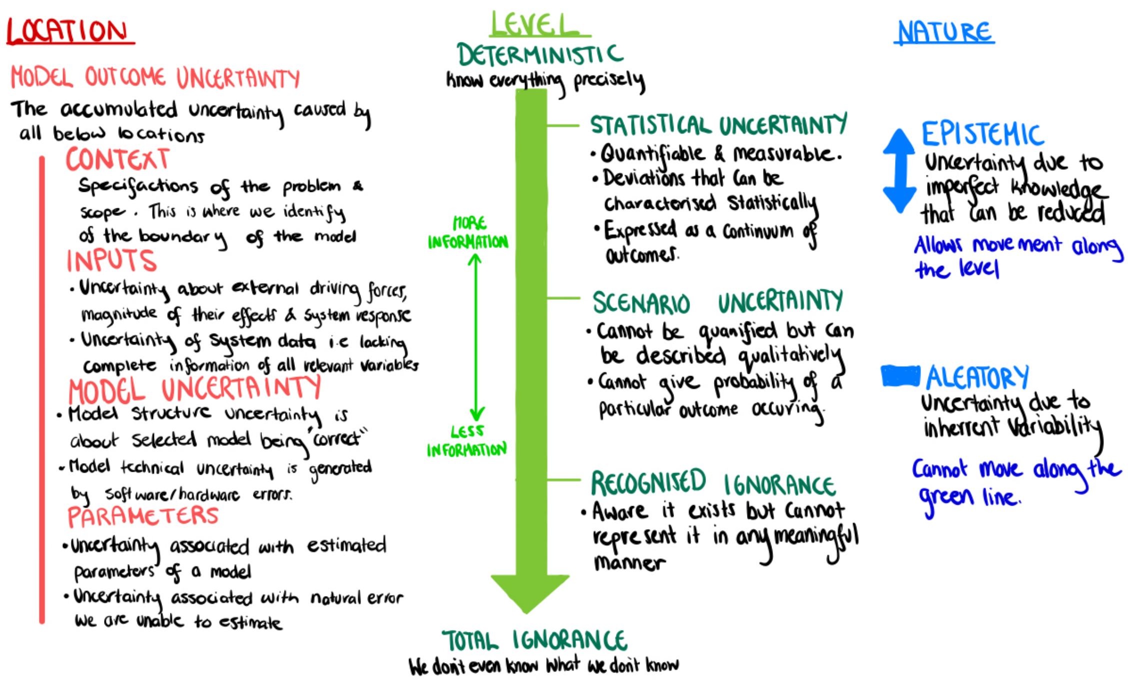


Figure 3: Illustration of the taxonomy described in Walker et al. (2003)

The use of uncertainty in high dimensional environments is especially important in energy data. Large models that incorporate spatial-temporal data from many sources and systems are used to predict energy uses in the short and long term. Understanding how to improve and make better decisions in these models is imperative in both the daily operation of the energy sector as well as the transition from fossile fuels to clean energy. Therefore the energy sector is an incredibly relevant application of research in uncertainty visualisation techniques.

Thesis Overview

The overarching theme of my thesis is a change in the way we understand uncertainty visualisations, specifically in the case of communication. This work will be divided into three chapters.

Chapter 1: A new theoretical framework

Chapter 1 discuss the current state of the research surrounding uncertainty visualisation and describe two fundamental mistakes in the way researchers conceptualise uncertainty visualisations. The first mistake is in the role of distributions in the visualisation framework, where irrelevant distributions are used to answer questions and relevant distributions are ignored. These distribution issues are mistakenly reported as visualisation issues in the literature, ignoring a fundamental statistical problem in the way we visualise uncertainty. The second mistake is a belief that uncertainty visualisation will improve while uncertainty is seen as of low importance. I highlight how a lot of research in imporoving uncertainty visualisation reflects the belief that uncertainty is inherrently unimportant. I suggest that there is no overarching best uncertainty visualisation, but rather uncertainty visualisations should depend on a motivating question. Finally I provide a framework for visualising uncertainty that mitigates these issues by ensuring the visualisation author decides on the motivations and of their graphic, can correctly identify the relevant distribution and the aspects of the distribution that are needed for their motivation; and correctly assigns priority to these apects by using the correct aesthetics.

Chapter 2: Applications of the framework

Chapter 2 applies this framework and investigates its practical usefullness through experiemnts and discussions with AEMO. The purpose of data visualisation is insight, but due to time limits or other constraints, most visualisation studies use multiple benchmark tests as a substitute for measuring the complicated phenomena of insight (North 2006). Unfortunately the validity of these results hinge on the large insights we gain from graphics being the sum total of these small incremental insights which may not always be true (North 2006). Specifically in uncertainty visualisation, there is a focus on performance and accuracy based measures that assume more predictabe behaviour from people than what research on human decision making suggests. The work in Chapter 2 should avoid these common pitfalls that arrise from experimental plot evaluations by working closely with people at AEMO. Our partners at AEMO will be able to provide detailled and open ended descriptions of the beenfits and struggles of our uncertainty visualisation techniques. This allow us to directly see the improvements (or lack thereof) in insight due to the suggested framework.

Chapter 3: Translation

Chapter 3 will be a translation of chapter 1 and 2 into an R package. This will make this research more accessible and allow others to easily implement this visualisation framework in their own work.

Chapter 1: A new theortical framework

The current landscape of visualisation experiments

Current research in visualising uncertainty seems to have a focus on designing new options for visualising uncertainty for specific types of data. Typical papers comparing visualisation methods seem to fit into one or both of the follow categories:

1) a paper that suggests or compares uncertainty visualisations that depict a single distribution.
2) a paper that suggests or compares uncertainty visualisations for a specific type of data

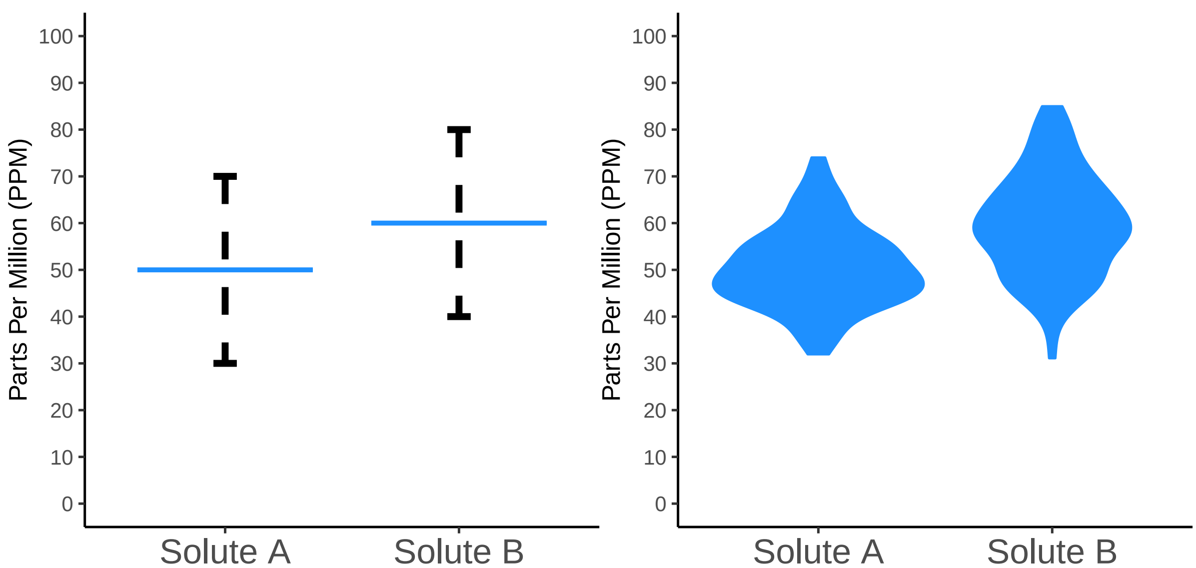
Generally the goal seems to be to increase the number of options we have in our "visualisation bag" so that we always have a plot at the ready if we want to visualise uncertainty. These new plots also tend to focus on a catch all plot that is "best for making decision" or "best for visualising spatial uncertainty". Data visualisation is commonly utilised as a tool in data exploration, so it is not uncommon for a data analyst to make a plot with only a vague goal and use it to pull out a large number of adjacent observations. For example, a data analyst might make a scatter plot to find out the relationship between two variables, but in doing do also finds out that one variable has discrete outcomes over a range of values. It is common knowledge that any single plot cannot reveal all information about a data set, all plot types obfuscate and uncover different types of information, however this aspect of plotting becomes more apparent in uncertainty visualisation. Often when discussing "uncertainty" information, we expect readers to be able to draw information from a plot that was not estimated, prioritised, or visualised. Communicating uncertainty can be boilled down into two simple steps, first we need to quantify the uncertainty, and second, we need to communicate it (Webster 2003). The two issues I have noticed in uncertainty visualisation literature each come from one of these steps. The first issue is failing to identify the correct distribution (quantifying) and the second is failling to select a visualisation that highlights the important aspects of the distribution (communicating). These issues run deep in the literature, but they are easiest to understand with an example that I will repeatedly return to as we develop this idea.

Part 1: Distribution

Example: Comparing HOPs, error bars, and violin plots

The issues that are rampant in the uncertainty visualisation literature are easily seen if we zoom in on one example. The study done by Hullman, Resnick, and Adar (2015) is a great illustration in the importance of having a clear motivation when you design a graphic. It is important to keep in mind that while I am primarily discussing one paper to illustrate a point, the issues I am brining to light are the standard in the uncertainty visualisation literature. This paper is no outlier.

The study by Hullman, Resnick, and Adar (2015) asked participants to provide some numerical properties of a distribution using a hypothetical outcome plot (HOPs), an error bar plot or violin plot. Participants were given questions relating to individual and multiple distributions. When shown a single distribution participants were asked about the mean of the distribution, the probability of an outcome being above some threshold (indicated on the plot with a red dot), or the probability of an outcome being between two given values. When shown two distributions the participants were asked "How often is measurement of solute B larger than the measurement of solute A?", and when shown three distributions "How often is measurement of solute B larger than the measurement of solute A and solute C?". Figure **??** shows an example error bar plot and violin plot for the two distribution case. There was also a high and low variance case for every plot and question combination. The study decided which technique was better using absolute error between the subjects response and the true value. They specified that the error bars are not confidence intervals as they represent the true underlying distribution, not the sampling distribution of the mean. They found that the HOPs reliably outperformed the violin and error bar plots for all the two and three distribution questions, but were no better than the violin or error bar plot in the other univariate cases (and in some cases worse).



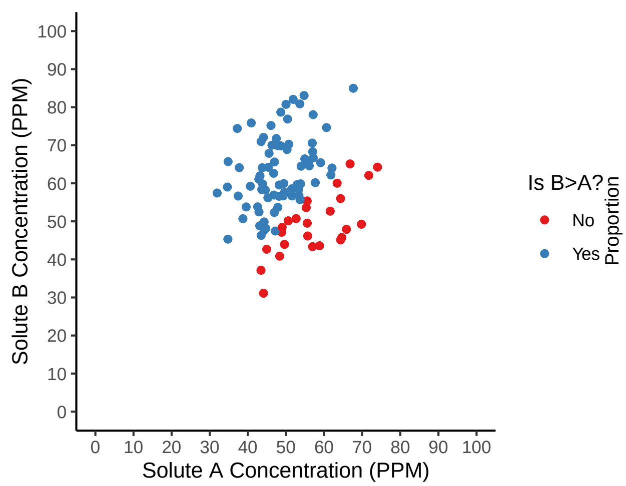
(a) Error bar plot

(b) Violin plot

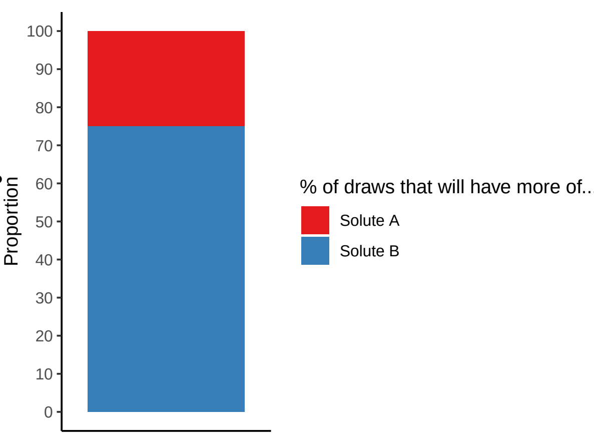
Figure 4: An example of the plots shown in the two variable tasks given in (Hullman, Resnick, and Adar 2015). The example question provided with this plot was 'In what percentage of vials is there more of solute B than A (Probability(B > A)?'

My first issue with the plots in Hullman, Resnick, and Adar (2015) study is that the violin plot and error bars are visualising a different distribution to the HOPs plot. The error bar plot and the violin plot in Figure **??** visualise the marginal distributions of solutions A and B and provides no information on the joint distribution. The HOPs plot is similar to a slope graph except instead of connecting the outcomes from A and B with a line, they are connected through frames of an animation. The HOPs, just like a slope graph, depicts a relationship between two varibles, i.e. the *joint* distribution. This key feature can be used to improve upon the HOPs and direct our attention to the distribution that is even more useful for answering this question.

Two alternative graphics that could be used to answer the question “In what percentage of vials is there more of solute B than A (Probability(B > A)?” are provided in Figure **??**. The scatter plot depicts the joint distribution of the two solutions and uses colour to highlight the Bernoulli distribution that is more closely aligned with the question. The stacked bar exclusively visualises the Bernouli distribution that describes the event $B > A$ and ignores the joint distribution highlighted by the scatter plot. Through this process of moving from the marginal distributions, to the joint distribution to the bernouli distributionm we moved the information needed to answer the question “What is $P(B > A)$?” from something you needed to calculate in your head (when looking at the error bar plots) to something you can **see** in the bar chart. While this process is illuminating, it is important to avoid whittling down the problem **too** much. Providing a categorical decision alone is somewhat useless, it is important to ballance advice with uncertainty estimates as a ballance of the two results in the most accurate decisions (Joslyn and LeClerc 2012).



(a) Scatter Plot



(b) Stacked Bar Chart

Figure 5: Two plots that could be used as alternatives to answer to question ‘In what percentage of vials is there more of solute B than A (Probability(B > A)?’. The scatter plot in (a) focuses on the relationship between the concentration of solute A and solute B, while the stacked bar chart in (b) highlights the frequency with which each solute is greater than the other.

Theory: Selecting the correct distribution

In order to fairly compare two uncertainty visualisations, they need to provide the same information. This is rarely done in plots comparing distribution visualisations. The important distinction between the distribution displayed and the distribution required to answer a question is often ignored in our discussions good or bad uncertainty visualizations. This means that studies identifying some graphics as better than others may only do so because the two plots displayed different distributions. Figure **??** illustrates how we think about distributions when performing statsitical tests compared to when we create visualisations. When we compute statistics or perform a hypothesis test, we typically put a lot of thought into what the correct distribution is for our question, however, when we perform visualisation we typically plot a collection of normal marginal distributions and ignore the actual distribution required. This extends to our perception of visualisation in general as most of the visualisation we consider to be for "distributions" are actually just tools for visualising marginal distributions.

An example that might be more familiar to the average statistician is something I like to refer to as "the error bar problem". If you have ever looked at two overlapping error bars and said "oh these variables are not statistically significantly different" you have used error bars incorrectly. While it is true that error bars that do not overlap implies statistical significance, overlapping error bars do not imply the converse in true. The same is true for a lot of the other ad-hoc statistical tests we use error bars for. A study done by Bella et al. (2005) asked participants to adjust two error bars until the means were "just" statistically significantly different, and most people adjusted the error bars until they were just touching. In the case of independent confidence intervals that just touch, the p-value of the associated two-tailled t-test is about 0.006 (Schenker and Gentleman 2001). Bella et al. (2005) also found that few people could incorporate changing information about independence that arrises from repeated measure design and most participants were ignorant to the fact that error bars are used for both confidence intervals and standard error bars, two wildly different indicators of precision.

This is a classic example of expecting readers to draw conclusions about a distribution that was not visualised. Error bars typically represent the 95% confidence interval of a sampling distribution, most commonly the significant values of a t-distribution. This means that each error bar provides the range of significant values through the two end points, and a vague indicator of variance with the length of the bar, *of each variable independently*. What an error bar does **not** depict is the t-distribution associated with a difference of two means, the equivalent statistical test we utilise error bars for. The visualisation itself is not the problem, trying to draw conclusions that requires information that was not visualised is.

The papers that go on to cite these misunderstandings about error bars discuss the work as though the problems are caused by error bars themselves. They suggest that error bars should be avoided as a visualisation tool, ignoring the fact that these fundamental misunderstandings in uncertainty will likely follow other visual encodings of t-distributions. Maybe some visual aspect of error bars encourages these extrapolations, but the HOPs example I drew on

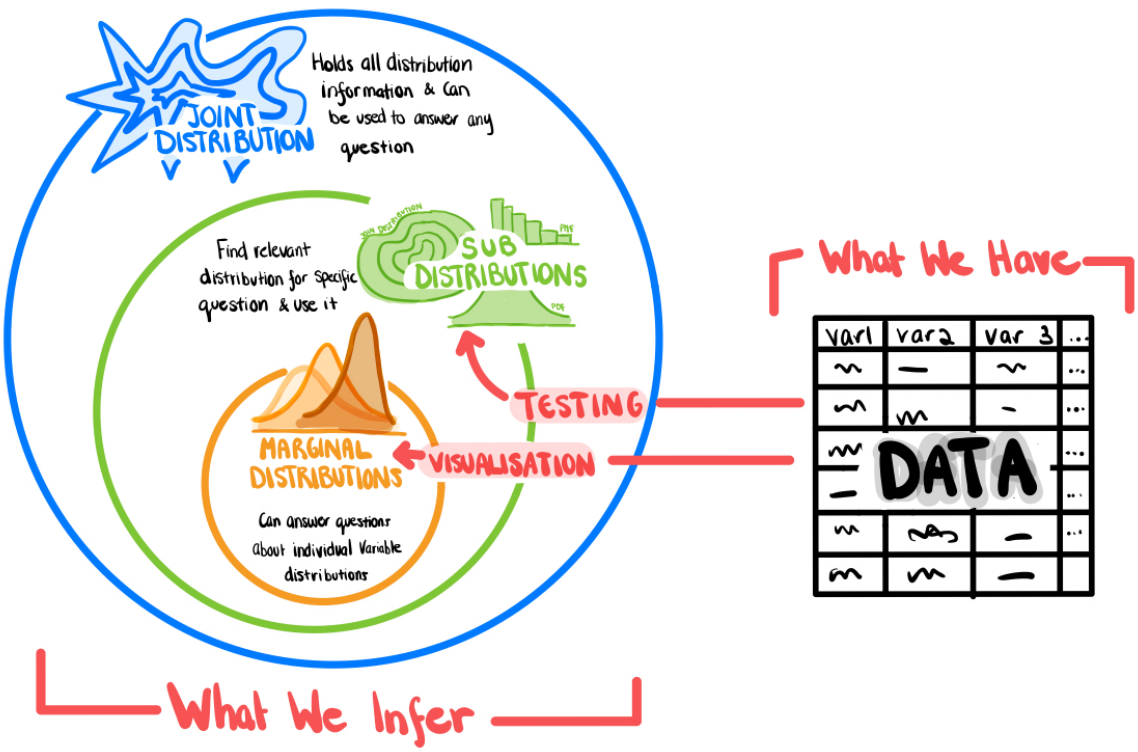


Figure 6: Illustration depicting the difference between the distributions we use for testing vs the visualisation we visualise

at the start of this chapter illustrates that this problem of expecting people to answer questions that are not directly related to the visualised distribution is larger than questions about significance.

This view of understanding uncertainty visualizations opens up a new gap in the literature. For example, if we accept that visualizing a collection of t-distributions is not a substitute for a collection of F-tests or paired t-tests, how *should* we visualise uncertainty if we want to draw those conclusions? Rather than shutting down the discussion on the usefulness of visualisation, I beleive it opens it and highlights areas of improvement in our current work.

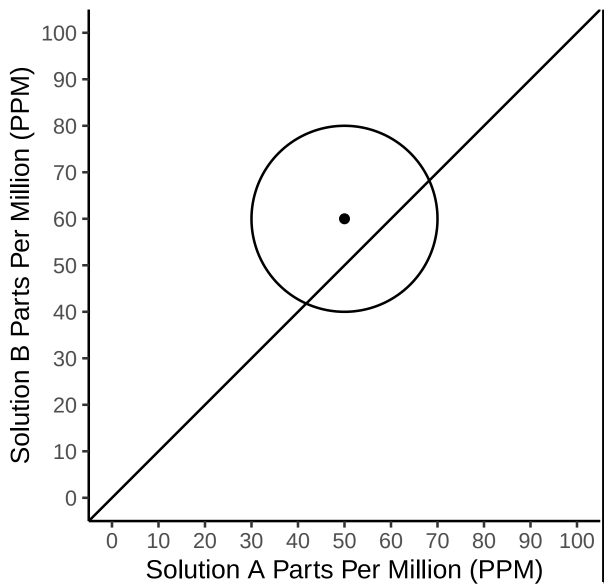
Part 2: Features

Example: more

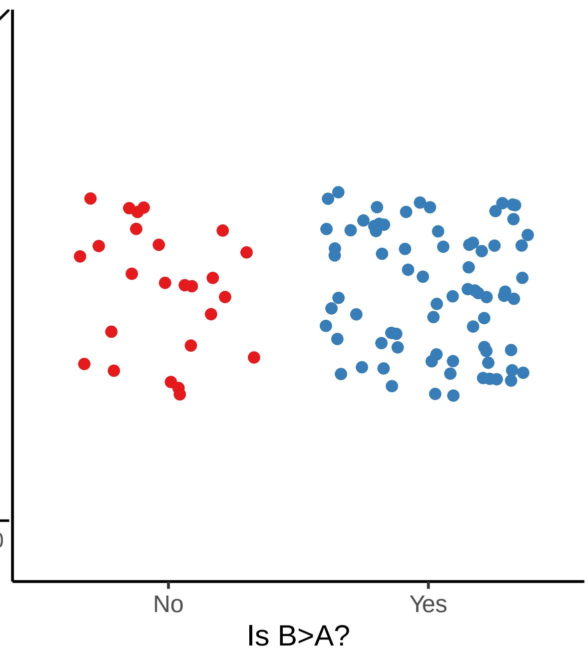
Let us return to the study done by Hullman, Resnick, and Adar (2015). Not only is the distribution depicted in the visualisation different, but the features of the distribution depicted are also different. In discussing the concept of a "best" visualisation, we rarely discuss which *features* of the distribution are being displayed in the graphic. The way we currently look at visualisation would classify the error bar plot and the violin plot as visualisations of a "distribution", the scatter plot is a visualisation of a "relationship" while the bar plot is a visualisation of "amounts" (Wilke (2019)), but this categorisation hides a lot of important details about drawing information from a graph. In this example the violin plot and the scatter plot both showed that each solution had an independent marginal normal distribution, the error bar (although technically a plot for visualising distributions) gives no concept of mass and would not give you the ability to identify even a simple distribution. Not only is it important to select a distribution that is appropriate for our question, it is also important to *show* the aspect of that distribution that holds the relevant information. The example asked about the *frequency* of a particular outcome which translates to visualising the *outcomes* of the joint distribution **or** the *parameter* of the Bernoulli distribution since these are the aspects of each distribution that hold the information we are looking for. If we visualised features of the distribution that don't hold the information we are looking for, the information is difficult to assertain and is more likely to be found through potentially faulty heuristics.

Figure **??** depicts two graphics that have the of the same distributions as those in Figure **??**, but each plot visualises an aspect of the distribution that is less relevant to the question. The 2D error bar plot (or circle?) highlights the mean and the values that are within the 95% confidence range. Since none of the parameters of the joint distribution are directly related to the question at hand, visualising the mean and significance instead of the outcomes made it harder to answer "What is $P(B > A)$?". Since the parameter of the Bernouli distribution *is* directly related to the question, visualising outcomes makes it harder to answer the question.

Not only does visualising the incorrect distribution that highlights incorrect features related to a question make it difficult to answer correctly, it might also completely disorientate the viewer.



(a) Error cirlce. The point in the middle represents the mean and the circle represents the area 95% of values will fall into.



(b) A scatter plot depicting the samples of the bernouli distribution

Figure 7: Two plots that are not useful in answering the question 'In what percentage of vials is there more of solute B than A (Probability(B > A)?'. While the distributions visualised are correct, we have visualised the incorrect feature of the repsective distributions. Plot (a) visualises mean and significance thresholds of for values (b) the.

In the discussion of the study Hullman, Resnick, and Adar (2015) notes: > "In light of the common usage of error bars for presenting multiple independent distributions, it is noteworthy how poorly subjects using these representations performed on tasks asking them to assess how reliably a draw from one variable was larger than the other…Many subjects reported values less than 50%, which are not plausible given that the mean of B was larger than the mean of A."

Theory: the four features of a distribution

Different features of a distribution have different questions they are mode adept adept at answering. The apsect of a distribution that are typically depicted in graphics can be organised into four key features. Figure **??** depicts these four features along with an example of how they are typically depicted in graphics. Each of these four features have a set of questions they are better equip to answer. The four features are:

1) **Mass** describes the PMF/PDF or CMF/CDF of the functions. Depictions of mass can inform us of the mode, likely or impossible values, and whether or not we have an identifiable distribution.

2) **Samples** are a set of actual or simulated outcomes of a distribution. Our data falls into this category as it can be seen as an outcome of some "underlying" distribution. Outcomes can also be simulated through techniques such as bootstrapping or random sampling. Outcomes are useful to answer questions about frequency, however since outcomes have a connection to mass (depending on the structure of the graphic, a mass plot may just be a smoothed plot of a visualisation of a sample) they can sometimes be used to answer the same questions.

3) **Parameters** are the statistics that are related to our distribution. They can be the sufficient statistics of the distribution, such as the mean, variance, minimum and maximum, or they can be other statistics such as the or correlation, median and mode. Visualising a specific parameter of our distribution typically gives us freedom because it allows us to express any aspect of a plot in terms of a single value. Unfortunately this flexibility means that the set of question questions a parameter plot can answer are limited. Questions about the mean, median, maximum, minimum, correlation, values of significance, etc would usually need multiple plots to answer.

4) **Exchangability** illustrates whether or not a elements in a sequence of random variables from this distribution can be swapped. Rather than answering questions about exchangability, it is something we typically want to highlight in our data. Exchangability is often expressed by having features touch in a graphic. A time series inechangability is expressed using a line plot, spatial data's inechangability is expressed using a map and the exchangability of randomly sampled data is often expressed using points. Exchangability is adjacent to continuity because visually connected features in a graphic are also used to express discretised outcomes of a continuous process (such as a ridgeline plot).

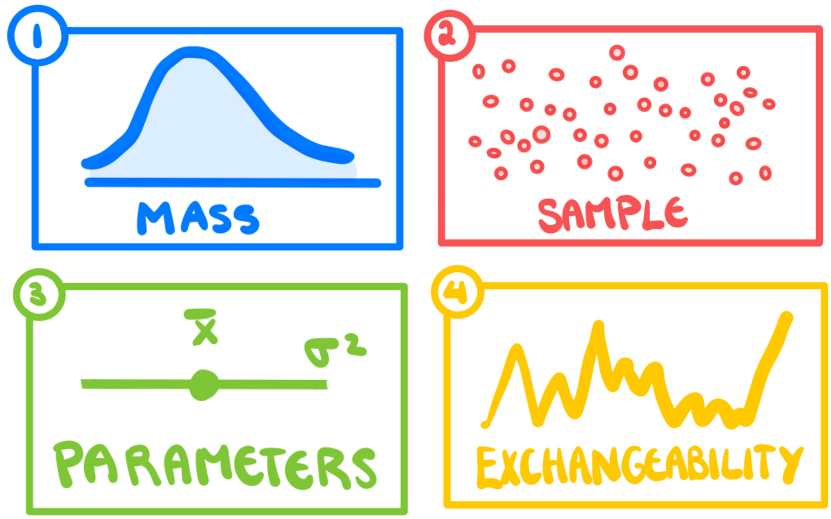


Figure 8: The four main features of a distribution graphics are typically used to represent

This is not an exhaustive list of every possible feature of a distribution, but rather the four most commonly visualised aspects of a distribution. Additionally, these features are not entirely distinct, particular visualsiations of parameters or samples may also double as a depiction of mass. Therefore, a graphic that is designed to depict one feature of a distribution may depict multiple features of multiple distributions. It also compares multiple ways of showing distribution to see which has more power.

Part 3: Hierarchy

Example: Looking to spatial uncertainty

This chapter so far has focused on the first way in which we talk about "visualising uncertainty" which is to depict some aspect of a distribution, however this does not cover the case of visualising uncertainty to express the impreciseness of estimates. Often our goal in visualisation is not to perform a statistical test or to make a decision based on a single estimate, but to identify some underlying structure in our data or make a decision based on a large number of connected systems. Visualising estimates and the uncertainty associated with those estimates separately can lead to the uncertainty being ignored (Moritz et al. 2017). This is typically the case when we try and visualise spatial uncertainty.

Trying to depict the uncertainty of an estimates with a spatial aspect is incredibly difficult. Even displaying an estimate in a spatial context can have problems. Improvements in spatial plots typically focus on ways to map a distribution that is dependent on something other than land size while still depicting the inexchangability and location that is relevant to spatial data. A common complaint about choropleth maps is that they depict the estimated parameter as a function of land size instead of location, even if land size is irrelevant to the estimate (such as in the case of voting). Visualisations that correct for this, such as hex maps, do so by colouring and plotting hexagonal tiles that each represent a portion of the dependent location (Kobakian and Cook, n.d.). This means an irrelevant feature, such as land size is not depicted as important in the map. Trying to highlight the uncertainty associated with these estimates makes the process even more difficult.

There are four proposed methods of visualising spatial uncertainty that can be made with the `Vizumap` R package (Lydia R. Lucchesi, Kuhnert, and Wikle 2021). These four plots each present uncertainty with varying degrees of success and failure.

The two visualiations that do a reasonably good job of expressing error are the bivariate map and the pixel map. Figure **??** depicts these two types of visualisation. Plot (a) depicts a bivariate map which uses a bivariate colour palette which is created by blending two single hue colour palettes. One colour represents the variable of interest while the other represents uncertainty. There are two immediate problems with this method. First of all, uncertainty is being mapped with hue and saturation. Figure **??** illustrates the differences between hue, saturation, and value when looking at a colour palette. Maceachren et al. (2012) found hue and saturation to be two of the worst aesthetics to map to uncertainty as they don't have an intuitive connection to uncertainty and so participants were worse. Value does have a natural connection to uncertainty (lighter values equate to higher uncertainty and darker values equate to more certainty) so it is a much more appropriate choice. While the `Vizumap` data does depict areas of light and darkenss, they are largely irrlevant to the uncertainty measure causing our heuristics to lead us to the incorrect conclusions. The Value-Suppressing Uncertainty Palettes (VSUP) shown in Figure **??** maps estimates to the hue and uncertainty to the value thereby creating a more intuitive plot (Correll, Moritz, and Heer 2018). Additionally, at high levels of uncertainty VSUP only has one output colour, which explicitly prevents any decoding of one output colour without enforcing a binary encoding of significance (Correll, Moritz, and Heer 2018). Unfortunately VSUP are not easy to combine with packages like `Vizumap` which leaves it still somewhat difficult to express this encoding in practice, however the combination of a bivariate map with VSUP has shown to improve decisions in the face of uncertainty (Correll, Moritz, and Heer 2018). Plot (b) from Figure **??** depicts a pixel map. Pixel maps are similar to HOPs plots since they present a sample of possible values for the estimate, rather than a single value and an uncertainty visualsiation. Some depictions of a pixel map are quite intuitive, such as those shown in Bauer and Rose (2015), however these cases link "uncertainty" to sparseness rather than a wide array of potential outcomes. Currently the effectivness of a pixel map is