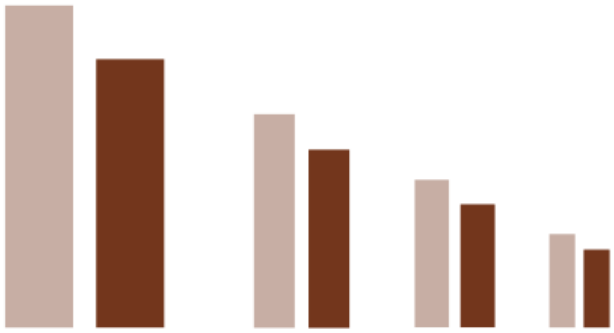
COMMUNICATING WITH **YOUR REPORT**



CONTENTS

INTRODUCTION: START

Great white sharks are one of the most feared creatures in the sea, and rightly so. As one of the apex predators of the world's oceans, great white sharks have historically been responsible for most of the reported and fatal attacks on humans from sharks.

Another reason that white sharks are feared more than other shark species is that they swim with their mouths open, making them appear even more menacing. Unlike other shark species, great whites breathe through "ram ventilation," where water forced into the mouth while swimming enables water to pass through the gills. If great white sharks stop moving, they stop breathing.

Most people believe that the ability to do creative work, like data visualization, is a gift that you're born with; some people have it, while others don't. This is a lie. Creative work, like data visualization, isn't about being inspired. It's about starting and persistently moving forward.

Most of the time, data visualization projects start with a vague idea of what needs to be communicated, or maybe no idea at all. If the idea needs to be clarified, getting started is what provides clarification. You don't need to convince yourself that you are a data visualization engineer before building a data visualization. Like a great white shark that has to move forward to stay alive, an effective data visualization project is never finished on the first draft. It starts and keeps pushing forward, improving with persistent forward movement.

This guide is a crash course in data communication, with a focus on using Microsoft Power BI[1,2]. The guide makes several opinionated claims, but the most important one is that data visualization is a skill that can be developed. Communicating with data is difficult, and even expert practitioners can struggle with doing it effectively. Fortunately, the goal isn't perfection – the goal is consistent, forward motion and improvement each step of the way.

PART ONE: CLARIFY

For decades, Microsoft Excel has been the tool of choice for data analysis for hundreds of thousands of organizations. Excel is a remarkable tool – using Excel, data can be stored, manipulated, processed, and analyzed as a table or as a visualization. But the default way to analyze data in Excel is as a table. There are

---

[1] Broadly a data visualization can be categorized as visualization for analysis or visualization for communication, and there is a difference in the two. Data visualization for analysis *precedes* visualization for communication. This guide assumes that you have analyzed your dataset and wish to present that information to someone so that they can understand it and use it.

[2] The ideas in this guide are heavily influenced by Nathan Yau, a statistician who runs FlowingData and Donald Miller's, "HowtoTellAStory," which focuses on telling stories with clarity. It's important to note both of their influences here: Donald Miller for his ideas on telling a good story and Nathan Yau for his guidance around Visualization for Clarity. They are both incredible resources, and again, have shaped the ideas here.

some professionals that work in Excel every day for entire careers but never actually take the time to learn how to visualize data using the software.

The problem with this is that data is challenging to understand from tables alone. Data analysis is hard – there's a reason that organizations are always trying to hire professionals with data analysis skills. But it's even harder when findings and insights are reported primarily as tables. Although tables can provide quick summary estimates efficiently, once you have more than a few data points to understand, it's hard to make sense of it all. With a table, how does anyone know what needs to be tackled first? Or what really matters? Or what to look for?

In contrast to a table, a well-designed data visualization *communicates*. With a table, the inference has to be massaged out. With a data visualization, the patterns, comparisons, and trends are apparent.

The first lesson, then, is that more than anything, a well-designed data visualization should communicate insights clearly[3].

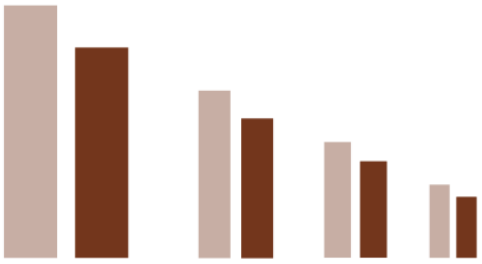
PART TWO: THE FORMULA

What makes a visualization clear?

Like a well-written story, there's a formula to creating a data visualization that communicates insights clearly. Here are the steps, listed and plotted:

1-Define the question(s)
2-Sketch it out
3-Add context (labels, hierarchy, and highlights)
4-Get feedback
5-Iterate

(Plot of Steps)



[3] When a data visualization is created for analysis, the level of visual clarity is up to the researcher. Visualization for communication is inherently candid; its primary job is to convey a message from an analysis. The more that message is ambiguous, the less likely end-users are to engage with the insights. So it makes sense to ensure that message is crystal clear. It's a nice idea that data speak for itself. Anyone whose done analysis knows, though, that data doesn't speak for itself as much as it has to be interpreted.

The best rock and roll bands give the impression that they're improvising their live performances, but most people recognize that they're backed by hours of structured practice. Each show is a little bit different, but most of what's performed is rehearsed and delivered according to a particular form. Building a clear visualization is similar; there is a form to it – define the question, sketch it out, add context, get feedback, and iterate – but there's also room to deviate from form if needed. There's room for creativity within the structure, but the structure is necessary for clarity. And clarity is what matters above all else.

Now, let's look at each step in more detail.

STEP ONE: DEFINE THE QUESTION(S)

It seems obvious, but any data visualization project should begin not by coding, but by thinking. Once you have a dataset and insights in hand[4], it's tempting to just get into it and see where you end up, like a drunk stumbling his way home after a night out on the town. For a visualization to communicate, though, this should be avoided.

Instead, it's worth taking a few minutes to attempt defining what it is exactly that you're wanting to show[5]. In the real-world, datasets are messy and insights are rarely apparent so it's difficult to pin down a flow chart of questions that will always put you on the right track. But there are a few that are close to universal.

First, *who is your end-user?* This should be crystal clear, so that you're sure to meet their needs. The last thing that you want is to spend your valuable time developing a visualization or report that is never used, and that's exactly happens more often than you would think. The best way to avoid this is to make sure you're designing for the right audience.

Which leads us to the next question, *how will the insights be used?* Answering this question helps when you're adding context to your report (step 3 of the formula). For example, if your end-users are wanting to see performance over time, you will want to make sure that the visual hierarchy (what stands out the most in your visual) illustrates that trend. If your end-users need to pinpoint outliers (over or under-performing employees or business units, for example), you'll want to highlight those outliers with a label or coloring.

Lastly, you should pause and make sure that you're able to answer: *What needs to be communicated so that your end-users can take the right action?*

In the end, you ought to be able to write out something like this with each visual that you create:

> *With this visual, I want to show [my end-user] [some interesting insight from your analysis] so that they can [take the right action].*

---

[4] Remember, if you're visualizing data for communication, you've already completed your analysis. You are trying to make a particular point in the clearest way possible.
[5] Same thing – you know the point and you're trying to tell it.

If you can answer these questions, at least you've oriented yourself towards your end-user in a way that primes your visualizations and overall report for clarity[6,7].

STEP TWO: SKETCH IT OUT

Once you've defined your end-user and what you'd like to communicate to them, the next step is to sketch out a wireframe of the visuals that you'd like to use. A wireframe is a simple, sketched out version of your report or visual. If you were building a home or another structure, you (or your architect) would start by creating a sketch of the end result. Building a report is no different – you should wireframe a simple outline of the visuals that you will be providing to your end users[8].

For those with less experience, it's a good idea to do what human beings do better than any other species in the history of planet earth – leverage the expertise of other people. First, take a quick second to make sure that you understand the data types for the insight you're wanting to communicate (Is it a time series? Multiple continuous data points? Categorical data?). Then, reference one of the many chart selection tools on the world wide web. Here are a few good ones from people and organizations who do this work each and every day:

Data to Viz
Financial Times/chart-doctor
experCeption
The Extreme Presentation™ Method

Once you've decided on a type(s) of visualization(s), add it(them) to your wireframe.

After you've chose visualizations and added them to your wireframe (and before moving on to the next step!), pause to ask yourself whether the visuals that you've chosen will be able to communicate the insights from your analysis to your end user, in a way that will enable them to take the right action. If this is true, then move on to the next step.

---

[6] There are different questions to ask for your preceding, exploratory analysis. They're always specific to the project that you're working on and can lead in a thousand different directions.

[7] What if the insights aren't clear? Then you want to communicate that fact to your end user.

[8] How do you build a wireframe? Perhaps the easiest approach is to use an Excel document. Using excel, you can build your wireframe by using cell outlines and images. For example, you can select a range of cells to represent the layout of your dashboard. Within that range of cells, you can outline where your data visualizations will live with shapes and borders. If you would like a line chart, you can copy and paste an image of a line chart, or you could create a mock version using simulated data on another tab of your spreadsheet. There are other methods you could use as well – PowerPoint, a simple Paint document, or even a picture of an actual hand sketch would work. The important thing is to build a wireframe so that you know what type of information you will need to present to end-users in order for your project to be a success.

STEP THREE: ADD CONTEXT

When you're confident that your wireframe will be able to communicate the insights you've uncovered, the next step in building a clear visualization is to add context. Context is the collection of visual cues that help to enable a full understanding. Again, the focus should be on clearly communicating your insights to your end-user, and for this context is key.

In general, there are some context defaults that you'll want to work through, and a set of other options that you'll want to consider.

The Defaults

1-*Good titles* are short and tell your end-users what they're seeing in plain language.

2-*Use subtitles* if and when you need to further explain your title or your insights. As a rule of thumb, *you should almost always use a subtitle*. Even if your visual is something generic like "Annual Sales," you can use a subtitle to add details like, "This chart shows annual sales beginning in 2019. Sales are calculated by summing the total sales volume in each region each year." Again, you're working to provide clarity, and subtitles are an overlooked method for providing additional detail.

3-Label your axes. Without the units of measurement, it's hard to understand any visual.

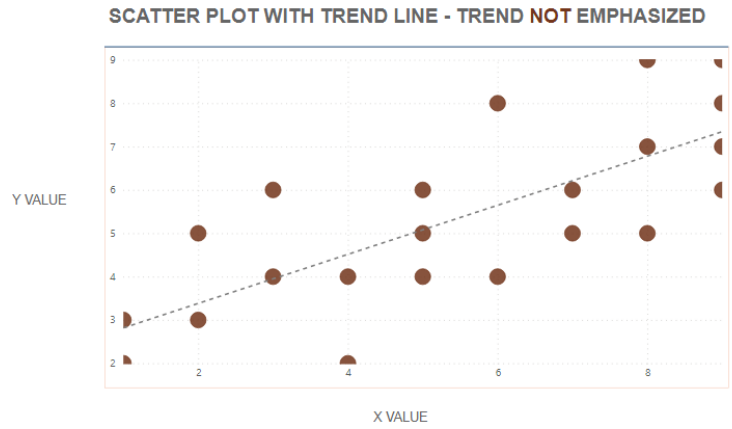
4-*Use a legend* when there are multiple colors that need to be registered by your end user.

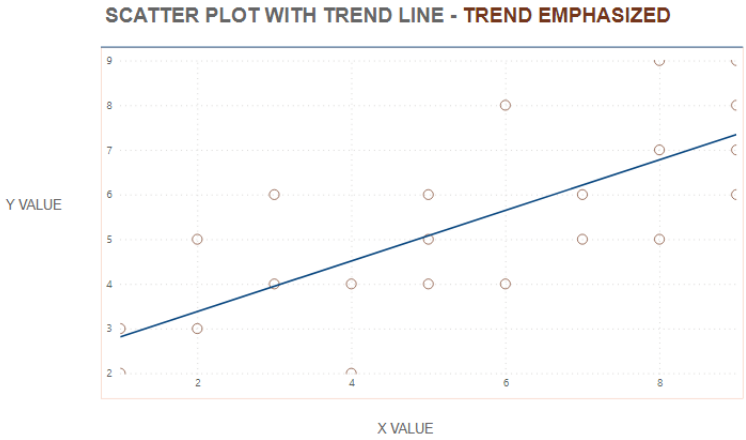
5-*Sequence any ordered data* to match your end-user's expectations. For example, make sure that any visualizations with a time trend are oriented so that the latest time(s) are on the right and the earlier periods are on the left.

Other Options

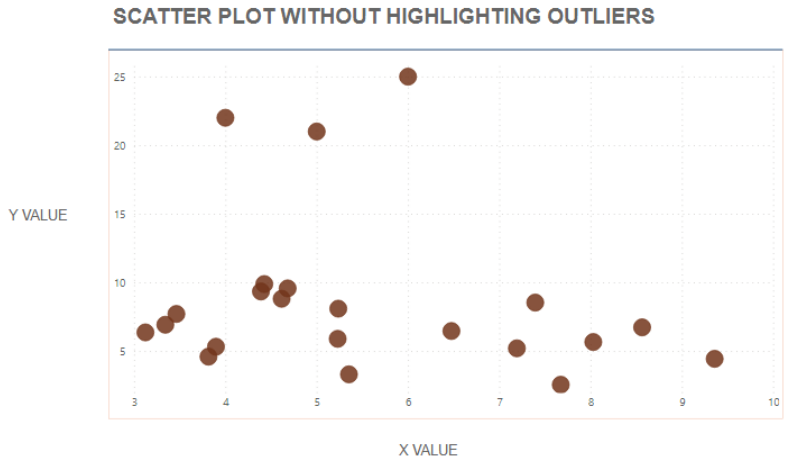
1-Heirarchy refers to the layering of your visuals. In many ways, hierarchy is what your end-users will notice first about the actual visual, beyond things like title and other highlighted elements. It's what the end-user registers in the brain without knowing that they've even noticed anything. *The hierarchy of your visualizations should foreground the message you want to emphasize*.

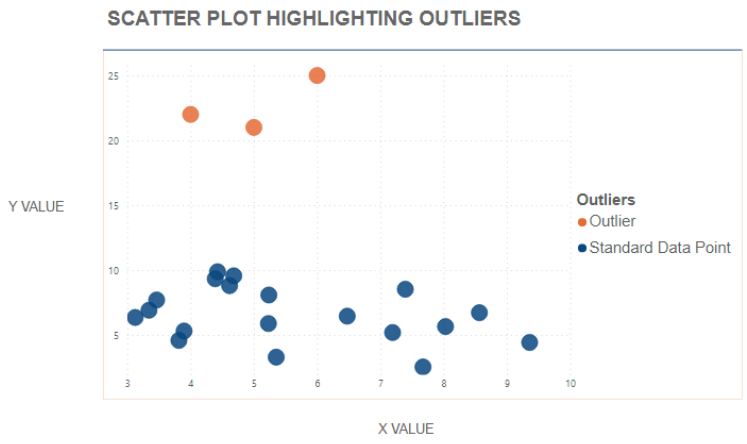
For example, consider a scatter plot with a trend line. Suppose you want to show the relationship between your x and y values, but you want stress the overall trend to your end-users. One way to layer the elements of your visual would be to place the trend on the same level of the visual hierarchy (the first visual below). But it would be better to place the trend line closer to the foreground of the visual hierarchy (the second visual below). Overall, the difference(s) is(are) subtle. But several subtle changes to clarify your visualization add up to a more explicit message.





2-Similarly, *use highlighting to stress inferences that can't be missed*. If you need to direct your end-user's attention to a particular point, one of the best ways is to use highlighting to make a particular point stand out. For example, it may be important to call out any outliers. With the first visual below, you can see that there are a few points that are distanced further from the cluster, but it isn't obvious that they're outliers - even when you're looking for it. In contrast, the second visual uses highlighting to draw attention to the same set of outliers. In the second visual, the outliers are the very first thing that most people's brains will register. Some other criteria would determine which observations were outliers (business rules, some arbitrary statistical procedure, etc.), but the visualization makes certain to highlight the point that there are a few observations that are very, very different from the others.





Overall, the approach is to focus each element of your visual on the insights that you want to communicate, both the defaults and other options. Start with the defaults; when you're happy that your defaults communicate the insights you'd like to share, use the other options to further call attention to those insights.

STEP FOUR: GET FEEDBACK

Once you've clarified what you're trying to communicate to your end-users, wireframed your report/visualizations, and added context to each visual, the next step in the process is to get feedback. There is a tendency to skip this step, but you shouldn't – *it's perhaps the most important of all for communicating with clarity*.

For most people, the thought of getting direct feedback on something creative is unimaginable. It's unnerving enough to be doing data visualization in the first place, much less getting direct feedback on it. Unfortunately, getting feedback is an uncomfortable but necessary part of effective data communication.

Usually, the individual who builds the report to communicate is also the person who has analyzed the data. The individual who is doing the analysis, then, *already knows the point they're trying to make*. Most people have heard of the curse of knowledge – when an analyst does the analysis, and then attempts to communicate those insights, her/his brain will automatically sense the insights even if they aren't apparent. If you've ever written an article or lengthy email, you know this all too well – even when you proofread, it's nearly impossible to catch all errors because your own brain will read what *you are trying to communicate instead of what's being communicated.*

In academia, getting feedback on your work is called peer review. In software engineering, there are code reviews. In human resources, it's called a performance review. In all three areas, inference is strengthened through feedback - the papers improve, the software is less buggy, and the employee knows what she/he is doing well. Designing a report or data visualization that communicates an intended message is no different – you have to get feedback to ensure that the message you are working to craft is being received.

Who should you ask for feedback? Preferably, this will be someone who you've worked with in the past, someone who you trust, and someone who isn't afraid to provide an honest and fair critique. In short, the best person is someone you can trust to contribute to the development without bogging down progress; a partner who understands how to provide feedback without derailing forward momentum.

And what should you ask them? Here, it's a good idea to avoid priming the person who is giving you feedback; simply ask them what the visualization says and see if they are able to easily decode the intended message. If you first explain to them what you are wanting to say, you can bias the feedback that you're going to receive. In the end, you want to know whether you have clearly communicated the point(s) that you are attempting to make. Start by simply asking them to tell you what they're seeing; from there, you can work through how you can clarify the inference even further.

STEP FIVE: ITERATE

Your first time through the steps will most likely leave you with several improvements to make, so after you've gotten feedback you should iterate and build improvements into your report/visualization(s). Developing a report that communicates is inherently iterative, and the message you're wanting to send should clarify with each build.

One last time, the steps:

1-Define the question(s)
2-Sketch it out
3-Add context (labels, highlights, and hierarchy)
4-Get feedback
5-Iterate

As new data and requests are made, or new insights are uncovered, the process begins again, incrementally driving towards a report or visualization that clearly communicates the insights needed for decision making.

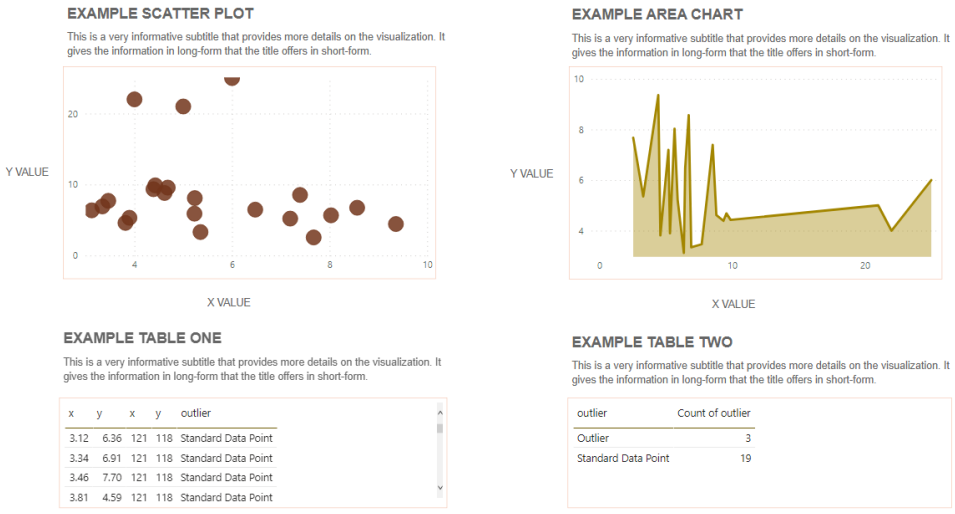
PART THREE: POWER BI TIPS

This section offers a collection of tips to give your reports and visualizations a better overall look and feel. Unlike the steps for clear communication, the suggestions made here are non-essentials so they may not apply universally. But in general, these suggestions will improve the overall look and feel of your report and give your end-user a better experience.

Leave Space

When designing to communicate, be careful not to overfill your page with visuals. The human brain has evolved to monotask, so packing in multiple visualizations in your report works against communicating with clarity. Most reports require multiple visualizations, however, so make sure to leave adequate white space between each element. *As a rule of thumb, limit visualizations to a maximum of five per page; as a second rule, use fewer than five whenever possible.*

For example, here's a snapshot of a report with four visuals fully-labeled for clarity. Here, even four visuals make the appearance a bit crowded; it would be better to leave more space – perhaps removing a visual or two - just to ensure that each visual communicates the intended message without clutter.

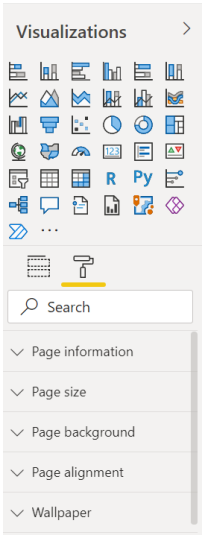


Make the Page Larger

In Power BI, the default display view for each tab is "Fit to Page" and the default display size is "16:9," but most of the time this will feel like rather limited in canvas space (like the visual above). Luckily, you can manually configure the width and height of the page to fit the visualizations you need.

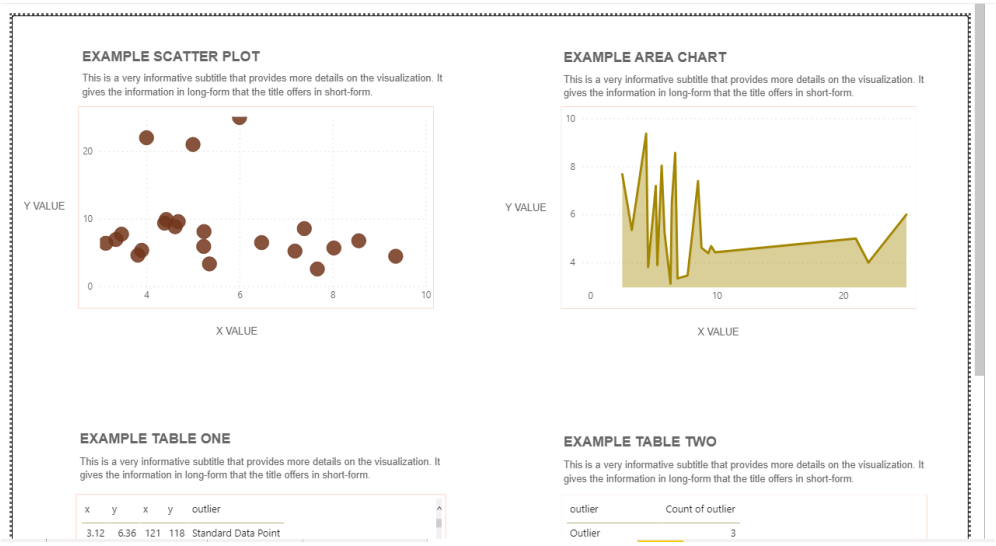
The typical approach in Power BI is to vary the page height but not the width. This gives a feel for your end-user that is closer to an actual web application. To increase the display height:

1-Select View → Page View → Actual Size.
2-Within the visualizations Pane (below),

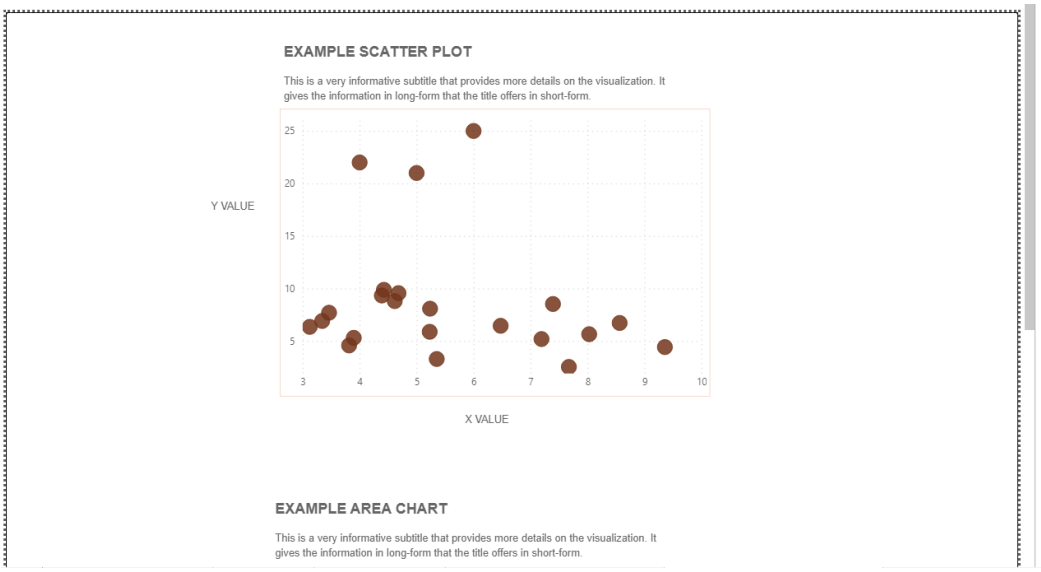


Select the Format  Icon → Page size → Type → Custom, and increase the page height.
3-Next, go back and select View → Page View → Fit to Width.

Now, you will have more room on your canvas for your visualizations. You should also see a scroll bar and have the ability to scroll down, depending on how much you increased the size of the page height.



With the additional canvas room, you can leave more room between visuals. If you really want to single-focus your end-user's attention, you can even do something like this:

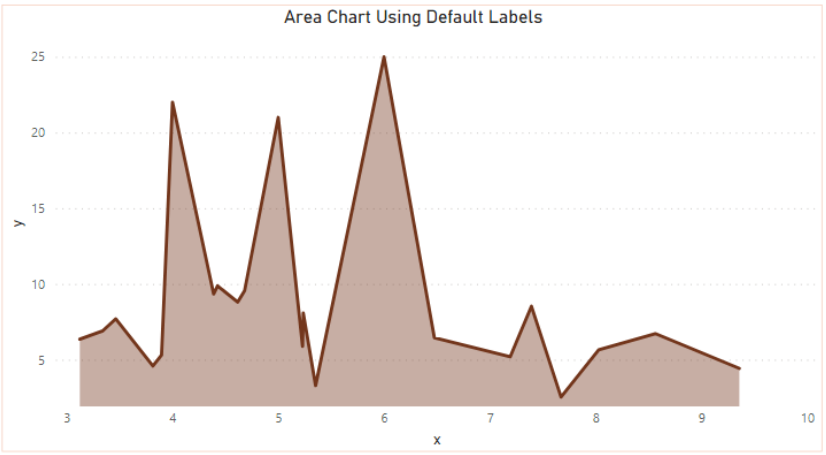


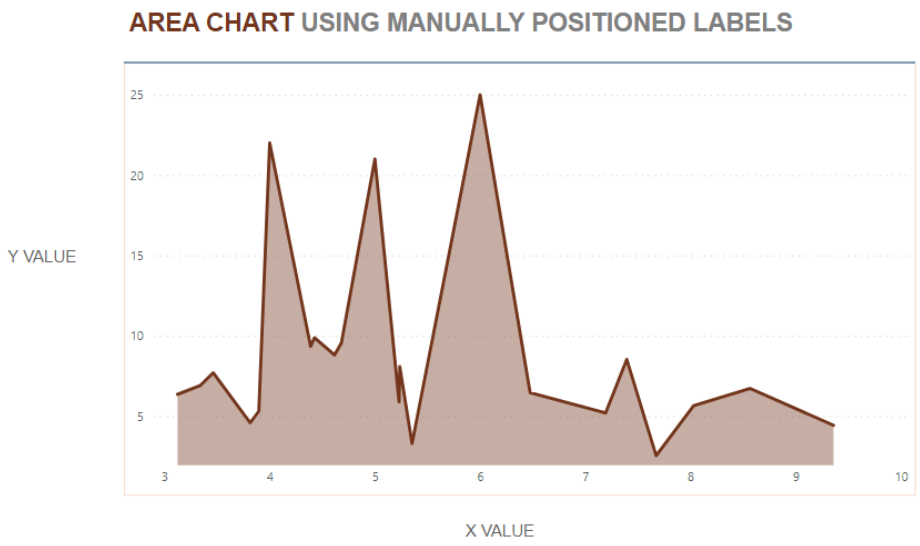
It's worth noting that this approach is only appropriate when your end-users are interacting with your report/data visualization from a desktop computer, where scrolling in a browser is intuitive. This approach could present problems if your end-users are primarily consuming your report on a mobile platform, or in an embedded environment. It's also problematic if you're looking to export your report or visuals to a .pdf file – Power BI only exports what's on the 16:9 canvas size.

Additionally, there are performance issues to consider; too many visuals on one page could impact performance. If you stick to the rule of thumb though – limiting visuals to a maximum of five per page – you shouldn't see any deficiencies in performance.

Use Manual Labels

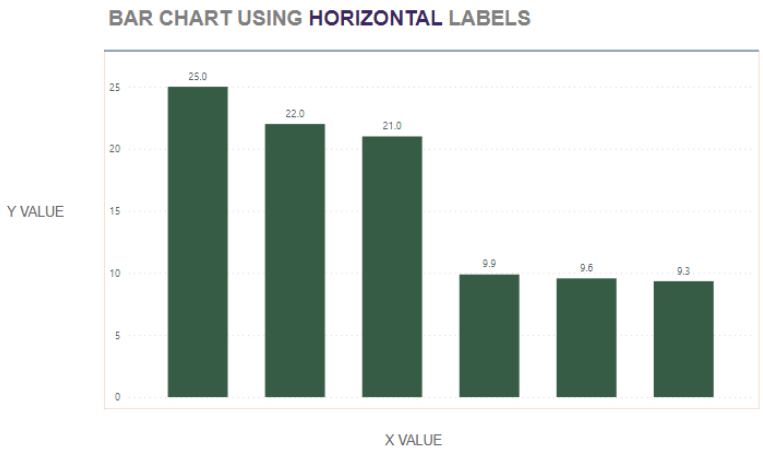
By default, Power BI applies labels to each visualization. To improve your labels, avoid using the Power BI defaults. Instead, create manual labels with text boxes. For example, in the first visual below, the default Power BI labels are applied. In the second visual, manual labels are created and positioned in place.

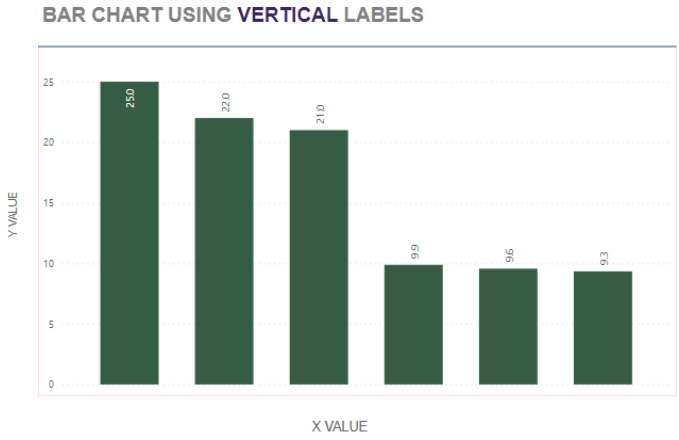




Keep Text Horizontal

Relatedly, you should try to keep text horizontal. For the English-speaking world, written language is read from left to right, upper to lower, so it makes sense to orient any text on your visuals in the same way. For example, here are two different visuals, one with horizontal labels and one with vertical labels. With horizontal labels, the information is consumed naturally by the end-user. The vertical labels, on the other hand, strain the brain's processing ability.





Be Consistent with Fonts (And Don't Use the Default Fonts)

A simple design upgrade is (arguably![9]) to avoid using the default font on your visuals (Segoe UI and DIN) and instead use something more familiar (Arial, Georgia, or my favorite, Consolas). Whatever font you choose, make sure that it's consistent across your report. The visuals in this guide use Arial, which is mostly frowned on by people who care about font decisions in the design world for its ubiquity. It is familiar, though, which probably helps with communication for clarity. And lucky for most of us, the typical end-user for an analytical report doesn't have well-formed opinions about font choices[10].

Choose Better Colors

---

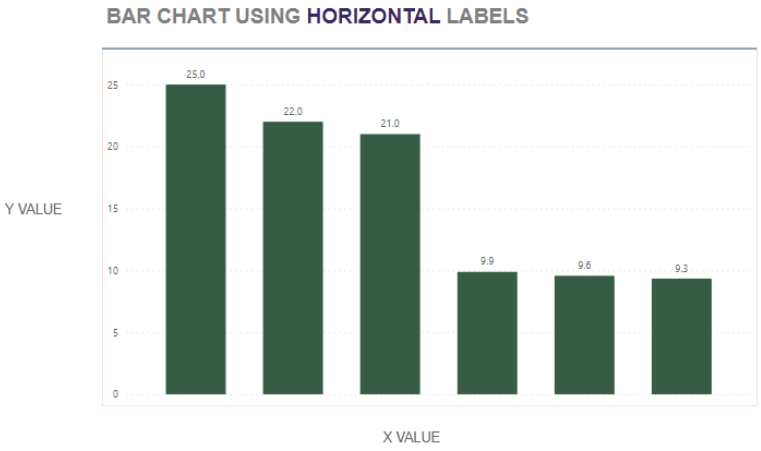
[9] If you do a simple Google search of, "Why Arial is a bad font" or "Why Arial is a good font" you'll get some strong and widely-varying opinions.

[10] Anecdotally, at least. Again, though, the point is to focus your end-users on the insights, not the font. So just pick one you like and make it consistent.

Likewise, the default Power BI colors can be harsh. One way to soften the colors a bit is to avoid using Black in fonts and labels. Instead, use dark grey for fonts and labels (like the visuals in the guide). Additionally, dark blue, dark red/orange, and dark green are easier on the eyes compared to the default visualization colors. The "Data Colors" option within the Visualization pane can be used to alter colors as needed. When in doubt here, it's also good to get a second opinion from someone with a design background, or even a simple second opinion from a colleague. Again, getting feedback is part of what will improve the overall communication, and simple color-based improvements to your visuals helps.

Use Colors in the Title Instead of in the Legend[11]

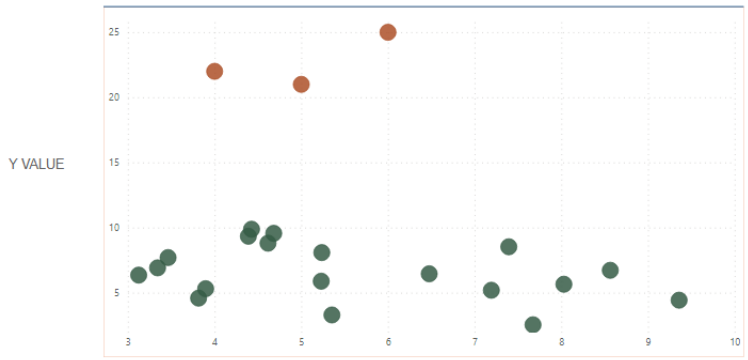
One simple way to add design flare is to use colors in titles to either stress particular elements or as a replacement for a legend. For example, most of the visuals in this guide use color to stress particular point. The visual below stresses that the visual uses horizontally oriented labels using a color highlight:



But you can also use color in the title *as the legend itself*, like the first visual below. If you need to use a legend, though, you should position it at the top or to the right center of the visual(s), since any text information will be read from left to right, upper to lower.

[11] For a lengthier discussion, and the individual that introduced me to this idea - https://flourish.studio/blog/legend-colors-in-header/

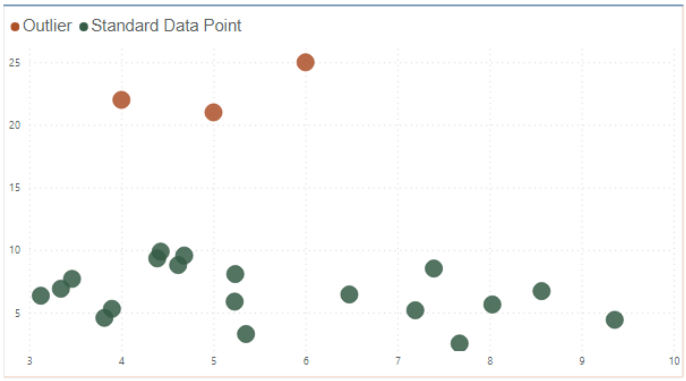
SCATTER PLOT WITH OUTLIERS AND NON OUTLIERS



Y VALUE

X VALUE

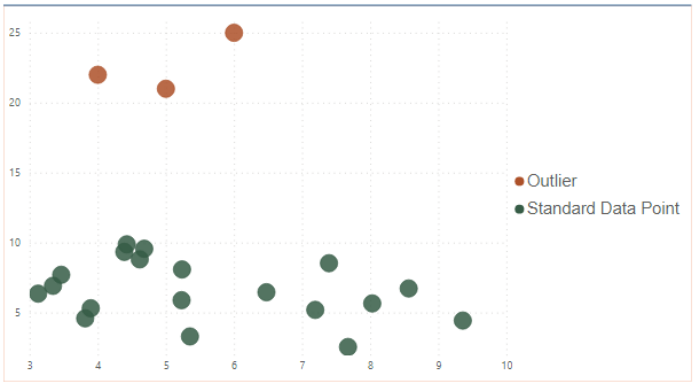
SCATTER PLOT WITH OUTLIERS AND NON OUTLIERS



Y VALUE

X VALUE

SCATTER PLOT WITH OUTLIERS AND NON OUTLIERS



Y VALUE

X VALUE

Use Navigation Bookmarks to Toggle Between Tabs

If you need to expand your report to multiple tabs, one way to give your end-users the experience that's closer to an actual web application is through the use of buttons and bookmarks. This is an advanced design pattern in Power BI, but it's a good way to drastically improve user experience. If you're interested in building your report this way, here's a link that will walk you through the steps. After setting up your bookmark navigation, you can then "Hide" each tab so that navigation depends solely on the bookmarked buttons in your report. Again, this is an advanced design pattern and you should only use it if you feel confident you're not going to introduce ambiguity or confusion for your end-user; but it's a great upgrade if you feel like your report communicates with clarity, and you want to add some extra flare.