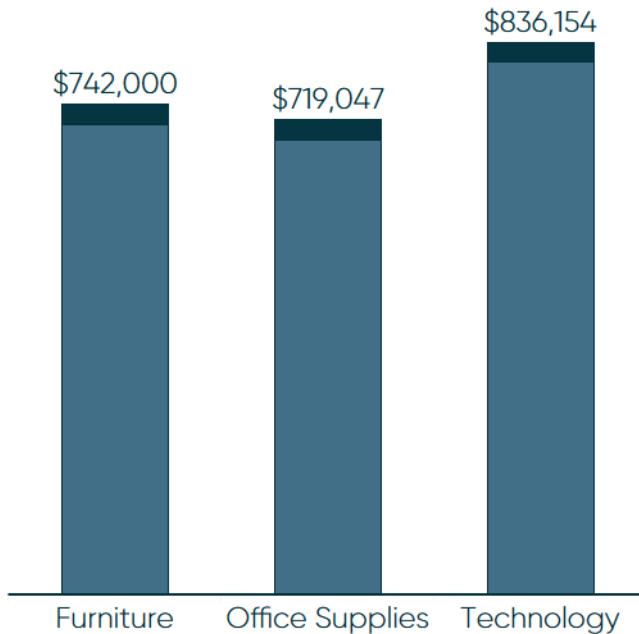
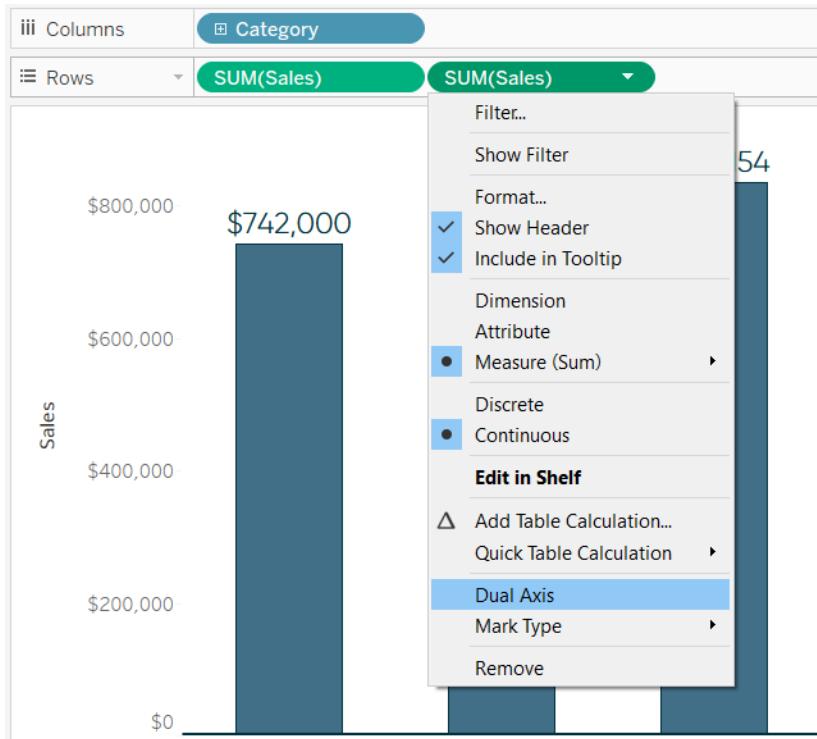


## Approach #3: Add Caps to Bars

As I explained in my blog post, “[Tablueprint 2: My Tableau Public Viz Views](#)”, capped bar charts have two benefits: (a) they enhance the design of a bar chart, and (b) add value by potentially doubling as a secondary comparison point. For example, the size of the cap can represent the value of a goal or be colored by a different measure (i.e., length of the bars represent sales; color of caps represent profitability). While I like the style of capped bars and the subtle way they can provide additional information, beware they can easily be confused with stacked bar charts, and should be used with caution. Here’s what a capped bar chart looks like when added to our previous example:



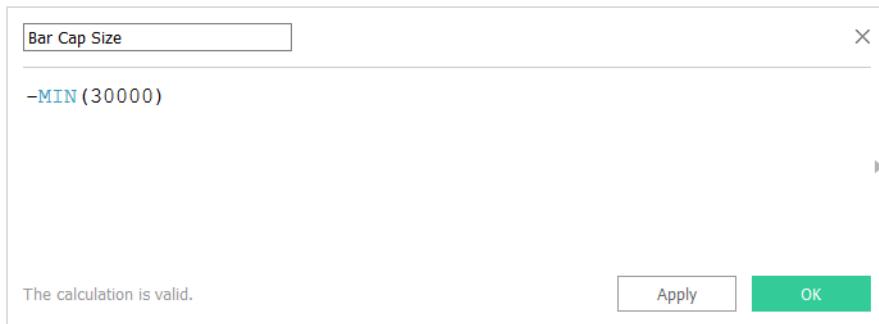
To create a capped bar chart in Tableau, start by adding a dual axis to the original bar chart. This can be accomplished by dragging the measure onto the Rows Shelf a second time, clicking the second occurrence of the pill, and choosing Dual Axis:



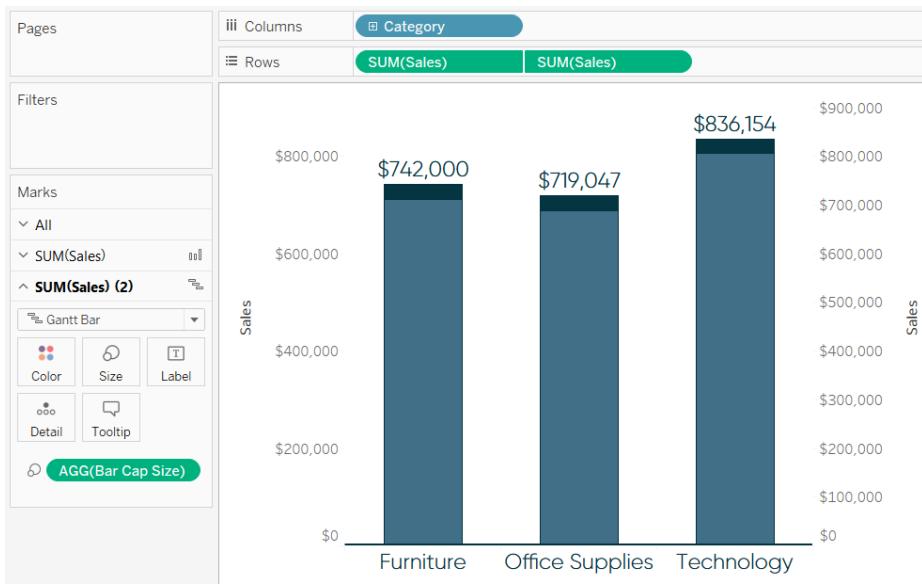
Once you've created the dual-axis chart, change the mark type of the first measure back to Bar and the mark type for the second measure to Gantt Bar. Also ensure the axes are synchronized by right-clicking the right axis and choosing Synchronize Axis. At this point, my view looks like this:



The next step to creating capped bar charts in Tableau is to create a calculated field for the size of the caps. The calculation is simply `-MIN([insert size])`. It is critical to add the negative sign before the bar size to ensure the caps go the correct direction. Here's how my calculated field looks when setting the cap size at 30,000:



Lastly, drag this newly created Bar Cap Size calculated field to the Size Marks Card on the Marks Shelf for the measure you set as a Gantt Bar:



To finalize my capped bar chart, I colored the caps to match the bar borders and hid both axes.



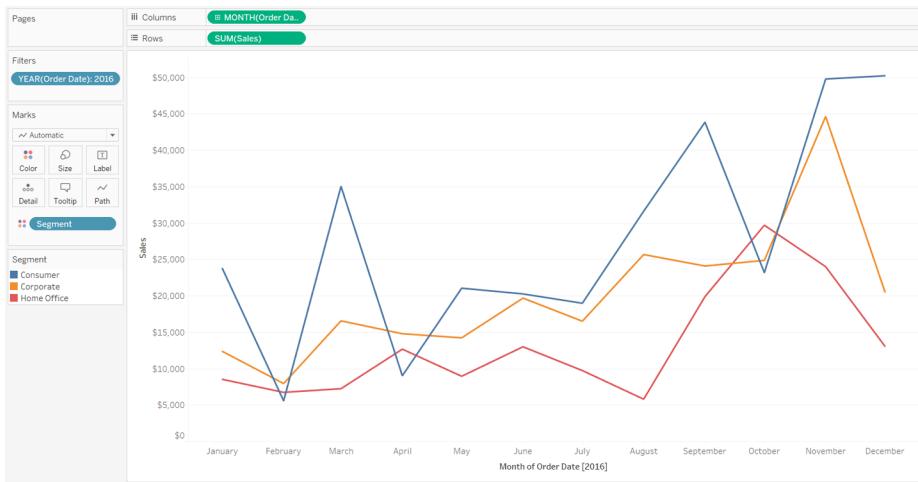
---

# **Three Ways to Make Lovely Line Graphs in Tableau**

Line graphs are a close second to bar charts as my favorite fundamental visualization type and are the obvious choice for evaluating trends over time.. Like bar charts, the invention of line graphs is generally credited to William Playfair at the end of the 18<sup>th</sup> century. Also like bar charts, I blame their age and people's familiarity with line graphs as the reason some data visualization enthusiasts look for "more engaging" choices. Line graphs have stood the test of time and their effectiveness cannot be denied. I'm hoping that these three approaches help cement line graphs as a top choice for your Tableau data visualizations.

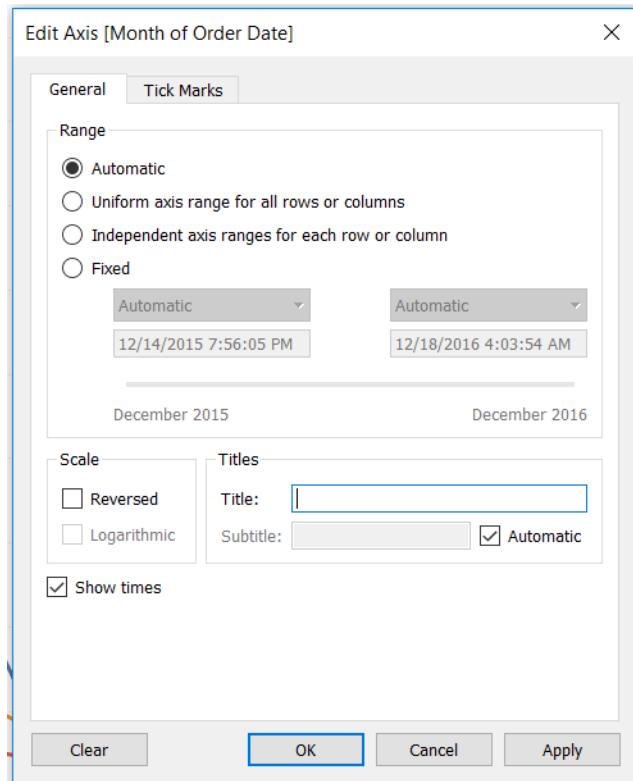
## **Approach #1: Use Formatting Available in Tableau**

My first tip for making line graphs more engaging is to use the formatting options available to you in Tableau. Consider the following sales by segment line graph with all of the default Tableau format settings. Note that I've used the Month([Order Date]) field as continuous and filtered the data to the year 2016:

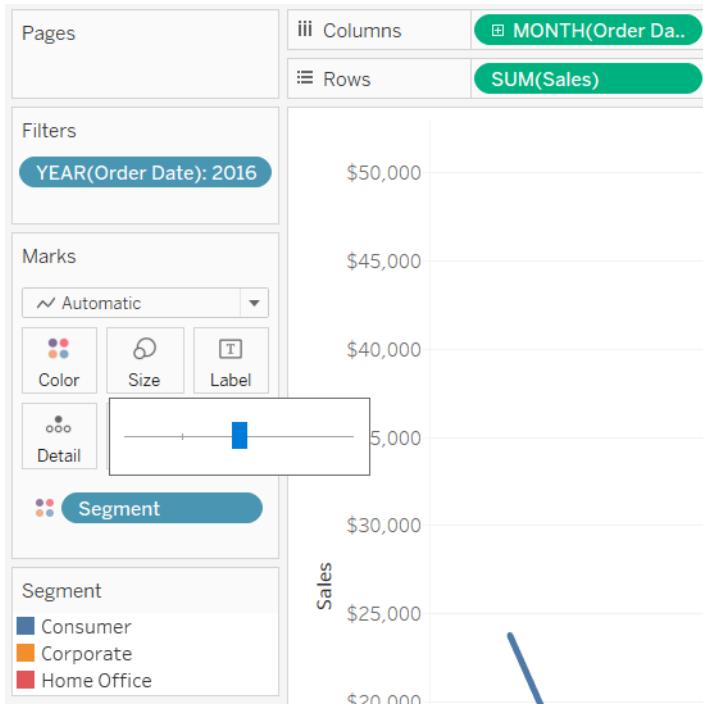


A great deal of thought went into Tableau's default formatting including the fonts, colors, and mark sizing. So as can be expected from my data visualization tool of choice, this graph is already getting the job done and providing good insights. That being said, if one of your objectives is to make the graph as engaging as possible, there are several opportunities to add value to this view.

First, I've never personally liked the axis title that is used to designate the date part when you are using a date field continuously. While I appreciate the clarity the axis title provides, this information is typically already implied in the graph itself and usually stated in surrounding context (i.e., "This is a graph about Sales by Month last year"). I almost always get rid of this axis title. Unfortunately, if you right-click the axis and deselect Show Header, both the axis title and month names disappear. To get rid of only the axis title but keep the month names, right-click the axis, and choose Edit Axis. This will open a dialog box where you can change the axis title. The trick is to simply delete all of the text on this line; then click the OK button:

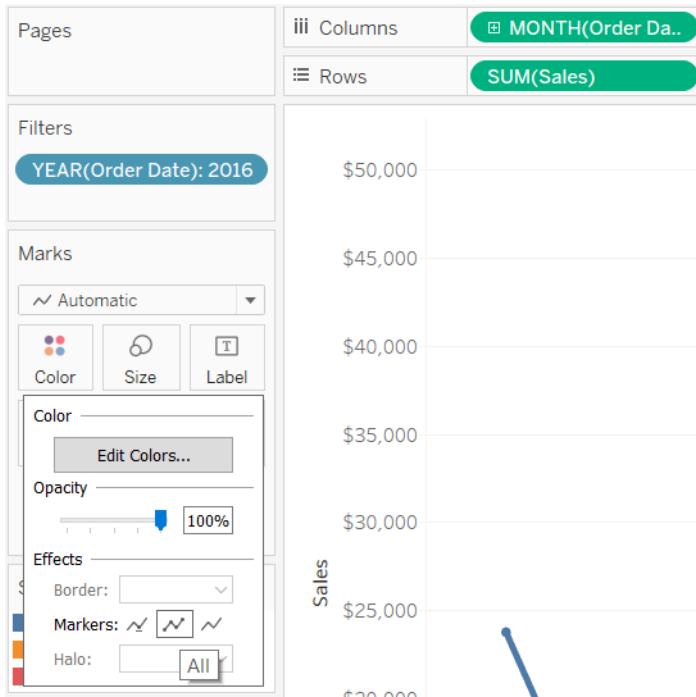


Next, if this is a standalone visualization, the weight of the lines is too light. To change the weight of the lines, click the Size Marks Card and drag the slider left or right until the lines are weighted as desired. I normally find that the second “notch” on the Size Marks Card is a good option for lines:

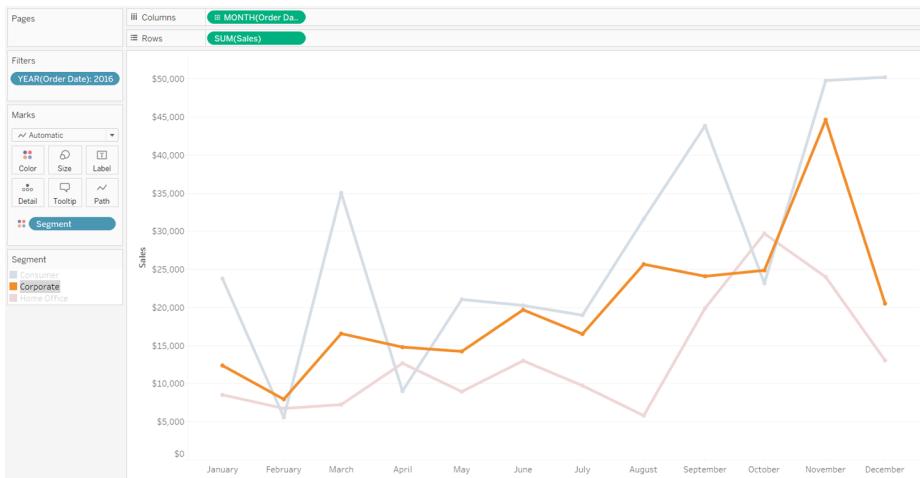


Take this on a case-by-case basis because sometimes heavier lines won't look good when used in a tighter space (such as one of several objects on a dashboard).

The next formatting tip addresses one of the most common questions I am asked while [speaking about Tableau](#): Markers. Markers are the small circles that are sometimes seen on data points of a line graph. Markers serve the practical purpose of telling the end user where there is data in a subtle way, and they also are an easy way to enhance a line graph. To add markers, click the Color Marks Card and select the second Markers option:

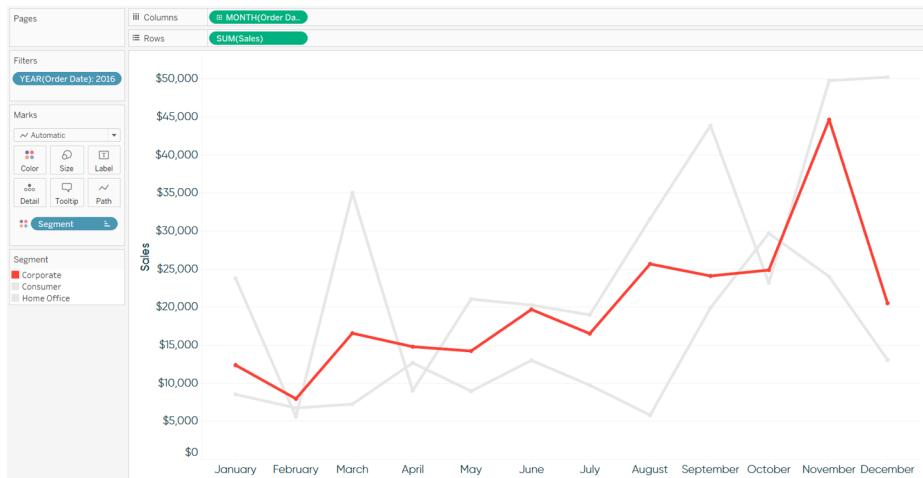


The next tip is optional, but many times I like to make the line or insight that I care about “pop” on the view. This can be accomplished in several ways; the easiest of which is to click a dimension member on the color legend:



This approach also allows the end users to make the selection that they care about. Another way to highlight a dimension member would be to show a highlighter for the Segment dimension. This is accomplished by right-clicking the Segment dimension (currently on the Color Marks Card) and choosing Show Highlighter.

There are other times that you want to make the highlight more permanent. Maybe I am the manager of the corporate segment, so I always want it highlighted throughout all views. Or maybe I'm not allowed to see the performance of the other dimension members, but my boss wanted me to see my segment's performance in context of the others. There is a way to do this with parameters that is beyond the scope of this chapter, but it can also be accomplished easily by changing the colors of the dimension members. Here's how my line graph looks after "graying out" two out of three dimension members by clicking the color legend and remapping the colors. To finalize the view, I've also put the colors and fonts in brand:

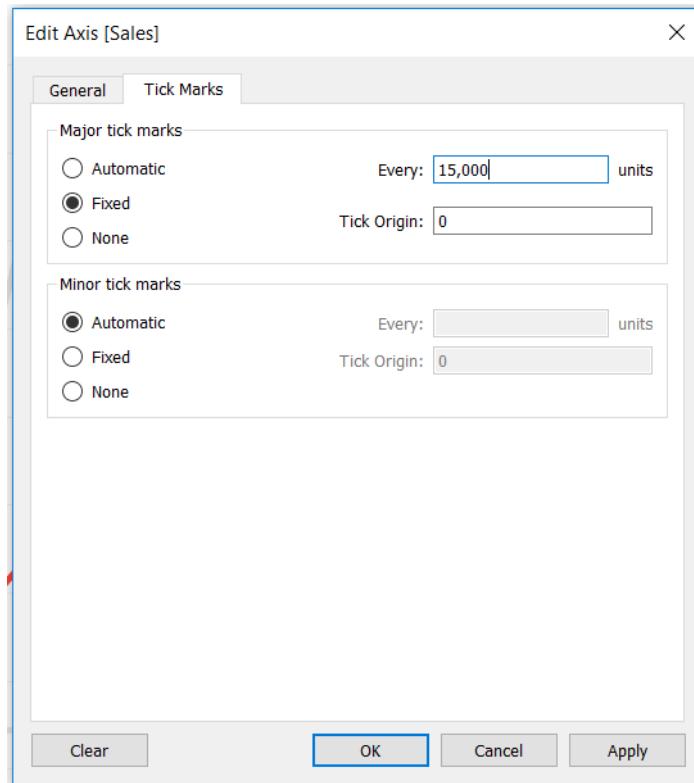


## Approach #2: Maximize the Data-Ink Ratio

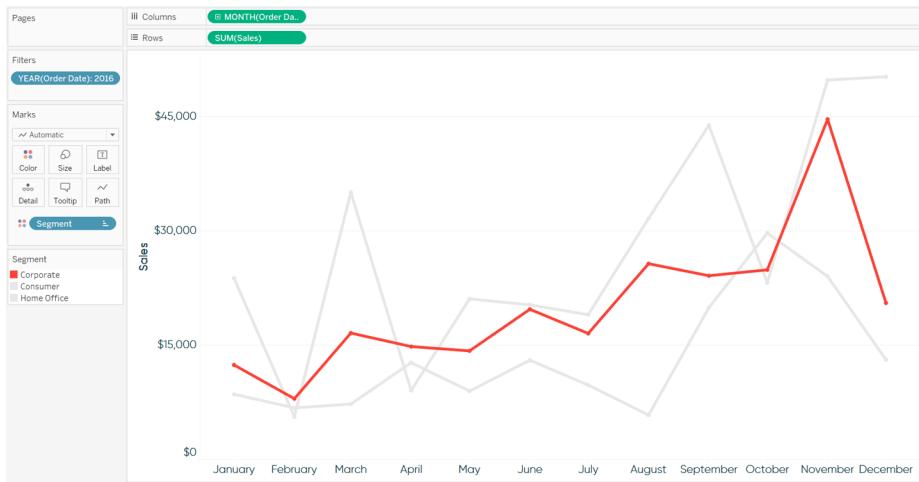
As mentioned in the last chapter, the next tip should be considered for any visualization you create: Maximize the Data-Ink Ratio. The Data-Ink Ratio is a concept introduced by Edward Tufte, who says you should dedicate as much "ink" on a view to the data as possible. This means getting rid of unnecessary lines, effects, and anything else that detracts from the data itself.

There are a couple of opportunities to make our line graph even more minimalist than it already is. First, the axis ticks on the Sales axis can be reduced by at least half. The axis ticks for Sales on this line graph are the multiples of \$5,000 going up the y-axis. By default, these are often too granular for my taste, and cause too many extra lines and too much ink. To fix the axis ticks, right-click the y-axis, click Edit Axis, and

navigate to the Tick Marks tab. You should beware that this option truly does fix the axis tick marks, so if you filter the view later, this setting will stick:



Here's how my line graph looks after fixing the tick marks at 15,000 units:



This same approach can be taken on the x-axis with months if you are using a continuous axis. This is not always an appropriate choice, but often the range of the axis is enough to communicate what the visualization is about. I often show only the starting point and end point of the x-axis of a line graph, especially when using sparklines. Here's how my line graph looks after fixing the axis ticks at eleven months with an origin of January 1<sup>st</sup>:



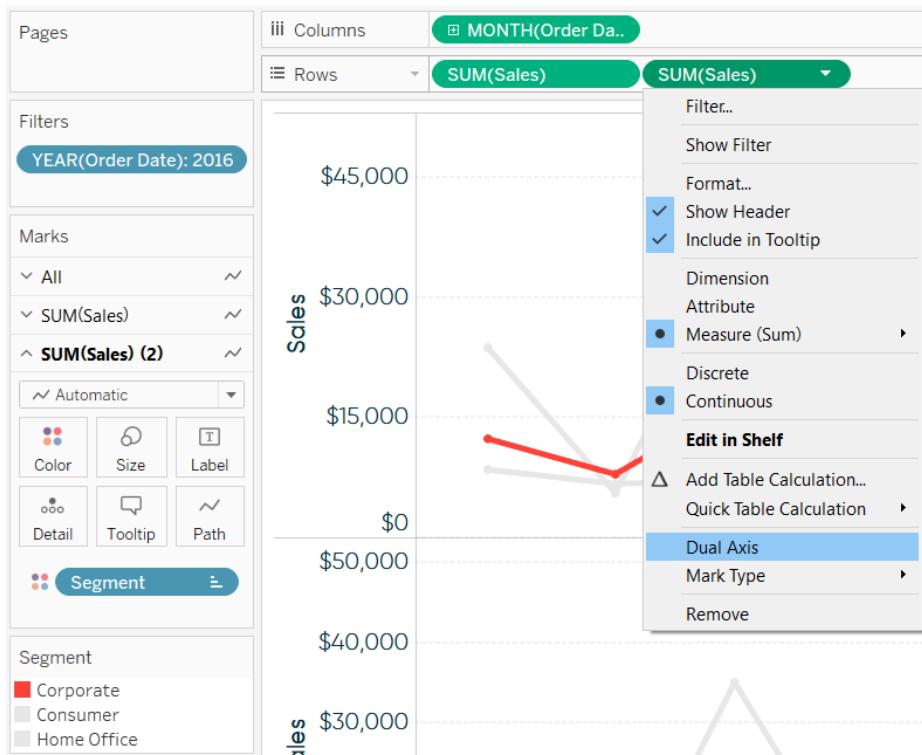
To finalize the view, I will hide the axis rulers and change the gridlines to dotted. Both of these options are found in the formatting pane, which you can reach by right-clicking the graph and choosing Format:



Compare this to the default line graph at the beginning of the chapter!

## Approach #3: Leverage the Dual-Axis

There are at least two ways the second axis on a line graph can be used to enhance the visualization. Let's say that we really like the markers from the formatting tip, but would like to make them more pronounced. To do this, start by putting the Sales measure on the Rows Shelf a second time, right-click it, and choose Dual Axis:



Ensure the axes line up by right-clicking the right axis and choosing Synchronize Axis.

Now that you've got two measures on the Rows Shelf, you've also got two sets of Marks Cards for the SUM(Sales) measure. These can be edited independently, which means you can keep the mark type for the first occurrence of Sales as Line, but change the mark type for the second occurrence of Sales to Circle. After sizing the circles, which are your new markers, hide the right-axis by clicking it and choosing Show Header:



The last tip for using the dual-axis to enhance a line graph is a design inspired by Google Analytics. When showing only one measure and dimension member at a time, Google Analytics shows a dual-axis line graph/area graph combination. To show you this approach, I am going to filter the view to only the segment that I care about—Corporate:



Now that we've isolated a single dimension member, change the mark type on the second occurrence of Sales to Area. Lastly, reduce the opacity of the area to something very light, such as 10%. Here is how my final view looks:



# Three Ways Psychological Schemas Can Improve Your Data Visualization

Whether you are aware of it or not, you are constantly recognizing and processing patterns in your everyday life. Think about when you go to dinner at a sit-down restaurant that you've never been to before. Even though you're trying the restaurant for the first time, you will have some expectation about the order of events, which usually looks like this:

1. You walk in and are greeted by a host or hostess who shows you to a table.
2. The waiter comes by to introduce himself and takes your drink order.
3. After returning with your drinks, the waiter takes your dinner order.
4. Throughout the meal, the waiter may stop by every once in a while to make sure you're doing OK and refill your drinks.
5. After the meal, the waiter will make a bad joke about "saving room for dessert."
6. After declining, he will either produce your bill from his pocket or run over to the register to get it for you.
7. You pay and go on your way.

This is one of many examples of a *psychological schema* (not to be confused with a *database schema*) in your everyday life. These patterns help society align diverse audiences and help us process varying situations very efficiently. These schemas are so powerful and ingrained that a disruption to the pattern can be confusing and challenging to overcome. As one extreme example, imagine showing up to a restaurant and having the waiter bring you a check before you've sat down.

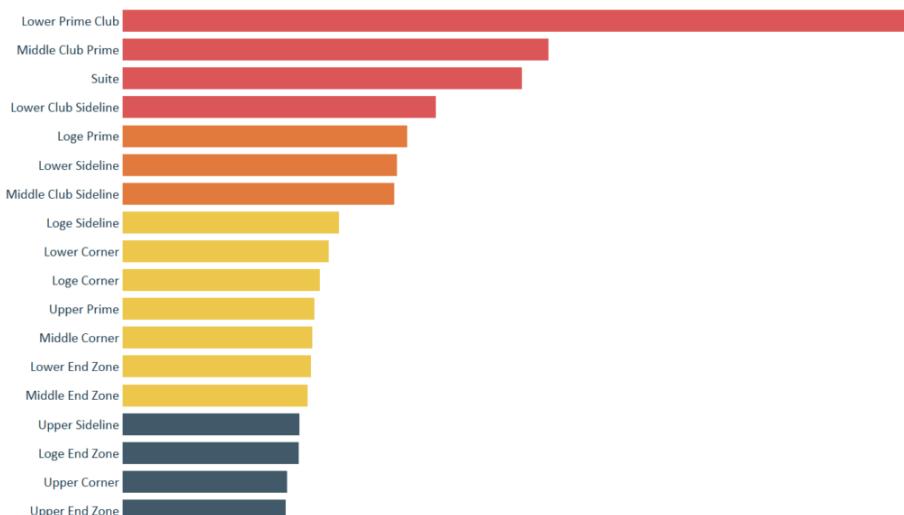
The restaurant pattern goes in an intuitive order so it is unlikely to vary much, but you also create schemas that are personal to you based on your own life experiences and worldview. These expectations help you avoid “reinventing the wheel” because you’ve experienced the same or similar situation before and know how to handle it.

Schemas play an important role in data visualization because they have the ability to make or break the two biggest benefits of visualizing data: reducing time to insight and improving the accuracy of insights. Tap into your audience’s schemas and you improve their experience; disrupt their schemas and you run the risk of leading your audience in the wrong direction.

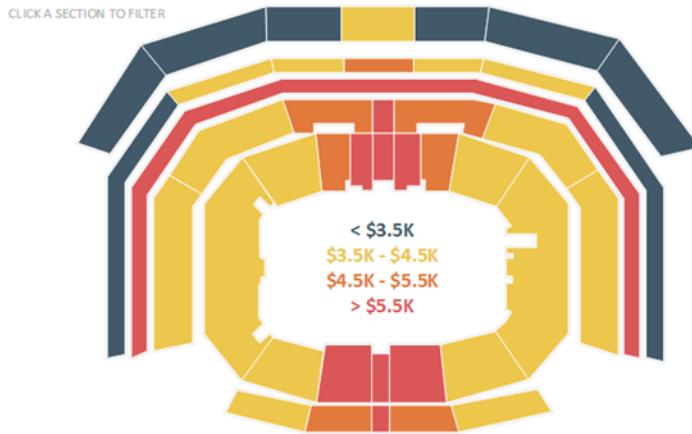
This chapter shares three ways to leverage schemas to improve your data visualization.

## Schema #1: Spatial Context

Maps help us process data because in addition to the data point, they provide spatial context that help our analyses. Consider the following bar chart showing the lowest cost per section to attend the 50th American Football Championship:



This a good data visualization within best practices, and there are definitely insights to be found in this chart. However, adding spatial context immediately helps the analysis make sense, even if you are not familiar with the stadium where this game was played:



I can use the schema I've constructed in my mind from my experiences buying tickets to many sporting events over my lifetime to know that the lower and closer to mid-field you are, the more expensive the ticket will be. This will help reduce my time to insight because it's much faster for me to determine if the numbers on the bar chart make sense to me intuitively or if there is a disruption to my schema, which in this case would also lead to insight (i.e., if lower bowl tickets are going for less money than the upper deck). Depending on what story you are trying to tell, it may make sense to display a map like this in addition to, or even instead of, the bar chart. For my *Cost of Attending the 50th American Football Championship* visualization, I chose to show the stadium map and a line graph.

## Schema #2: Icons/Shapes/Symbols

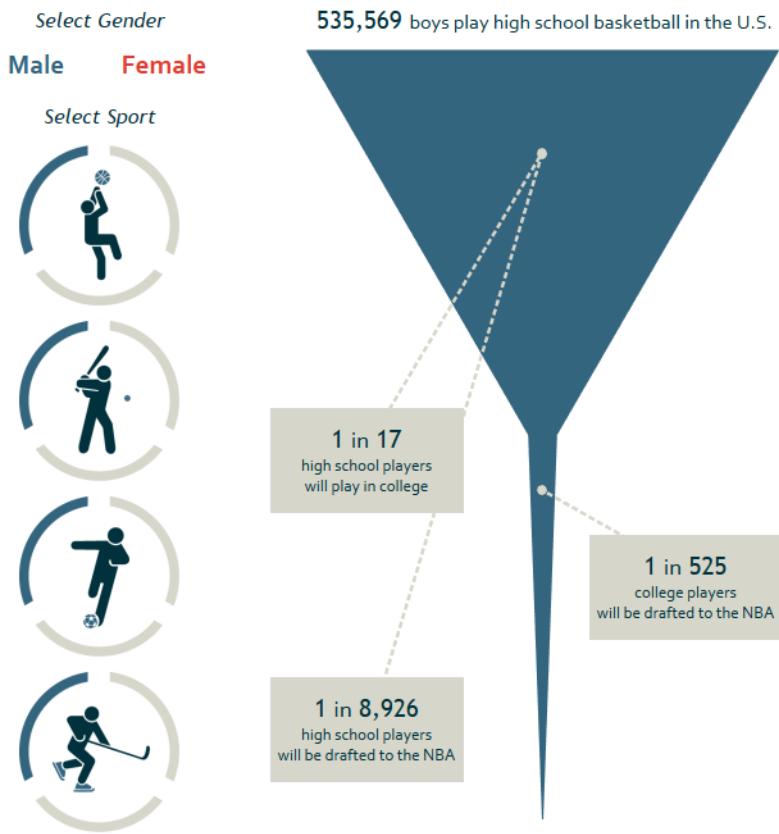
A picture tells a thousand words. The second schema that can improve your data visualization is the use of icons, shapes, and/or symbols. There's a thin line between graphics that enhance a data visualization and chartjunk (discussed in [Chapter 98](#)), but when done tastefully, graphics have the ability to provide much more information than words alone.

At the most basic level, think about how much value a “+” or “-” sign adds when it precedes a number on a dashboard. When used to show if there was a positive or negative change in a KPI, just one character reduces the time to insight and increases the accuracy of insights. Arrows or triangles pointing up or down work the same way.

That's a basic example, but many graphics work in this way because we have preconceived notions about what they mean. Take a look at the navigation I used on my visualization, *The Odds of Going Pro in Sports*:

## WHAT ARE THE ODDS OF GOING PRO IN SPORTS?

An analysis of high school, college, and pro sports in the United States by gender.

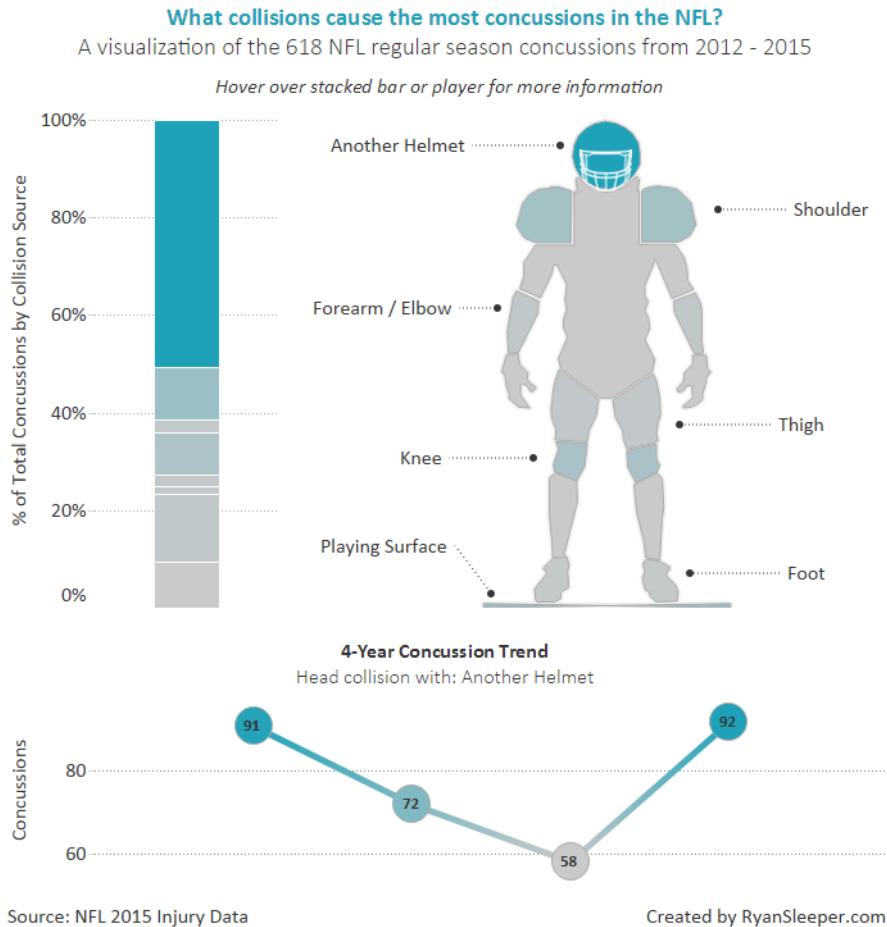


Data from | [www.scholarshipstats.com](http://www.scholarshipstats.com)

Data visualization by |

I used icons instead of words to display the different sport options. In my opinion, this makes the view more engaging, but icons also provide the advantage of reducing, and often eliminating, language barriers.

In another example, I used a polygon map of a football player to help illustrate what body parts cause the most collisions in the NFL:



I could have simply made a bar chart by collision source, which would be a fine approach, but by overlaying the heat map onto the player's shape, the viewers can tap into their existing schemas to process the view and discover insights.

## Schema #3: Color

For better or worse, I bet you associate green with good and red with bad. I'm not exactly sure where that preconception was born, but it seems to be with us to stay. It is a schema. You don't have to teach your audience what red and green mean; they know (or think they know, so be careful!).

I don't recommend this color combination for both scientific (color blindness) and personal reasons (I believe it's ugly), but this is an example of a schema you can leverage to help your audience make sense of your data visualizations.

Being aware of your audience's existing associations can help you help them decrease their time to insight and improve the accuracy of their insights. It works both ways though, so be careful not to completely disrupt their schemas. If you're making a visualization about fruit, don't color oranges purple and grapes orange.

If you're using color outside of the common green/red or blue/orange color palettes, be consistent so your audience becomes conditioned to understand your use of color.

Leveraging these three schemas in a thoughtful way can go a long way toward maximizing the two biggest benefits of data visualization: decreasing the time to insight and improving the accuracy of insights. At the very least, be aware that your audience has their own preconceptions, and disrupting them can make it more challenging for your audience to find value in your data visualization.

**PART IV**

---

**Framework**



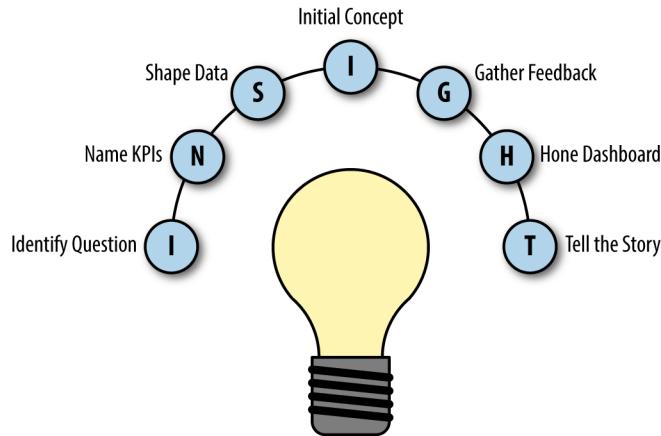
---

# Introducing the INSIGHT Framework for Data Visualization

To this point in *Practical Tableau*, we've introduced the basics, provided how-to tutorials for dozens of chart types, and covered some of my favorite tips and tricks. All of this will help you on your journey to becoming a Tableau expert, but it is not enough. It doesn't matter how good you are at the technical aspects of Tableau if you don't consider some intangibles that will help tie it all together. The information in [Part IV](#) is the missing link to help bring your data visualization from good to great.

Over the past eight years, I have personally designed and constructed hundreds of data visualizations. To this day, I have yet to build a dashboard that *every* viewer thought was perfect. Data visualization is an art, and its value is in the eye of the beholder. There are limitless possibilities and variables involved in creating a data visualization, too; no wonder that many people, when confronted with a data visualization issue, stick to an inoffensive template that does ... OK. But you can do better! With a structured approach to the task of data visualization, you can maximize the chances that your work will be well received—and communicate the valuable information inside the data.

The strategic framework I've developed—the INSIGHT framework—has saved me hundreds if not thousands of hours of iteration, increased value for stakeholders, and led to countless business insights—and now I'm sharing it with you!



### *Identify the business question*

One of the most common pitfalls in data visualization is trying to answer too many questions in one view. By prioritizing what the data visualization will answer, you are able to keep the view simple and make it clear to the end users what stories they should be looking for in the dashboard.

### *Name the KPIs*

These key performance indicators (KPIs) should all help answer the business question that the data visualization is trying to answer.

### *Shape the data*

In order to work with the data in a data visualization software, data should be prepared in tabular form in advance. For the most flexibility, each column header should represent a unique field, with subsequent rows representing every entry for each of those fields.

### *Initial concept*

Get an idea documented, even if it is simply a sketch of the direction you are planning to take. By not committing too much time to the initial concept, you will have the flexibility to pivot based on stakeholder input.

### *Gather feedback*

Ask end users what they think of your initial concept and if they believe it will meet their needs. This step helps instill some stakeholder ownership during the design process so they are more likely to support the final product.

### *Hone the dashboard*

This is your opportunity to incorporate stakeholder feedback and finalize the data visualization.

*Tell the story*

Finally, you are ready to distribute the dashboard and allow the stories in the data to begin making an impact.

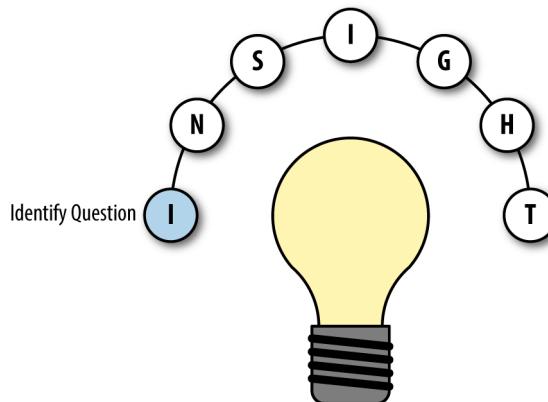
The following chapters will take you through each step of the process.



# Identify the Business Question

The first—and most critical step in the INSIGHT framework for data visualization—is to identify why you are making the data visualization to begin with! As a data visualization practitioner in a corporate environment, I call this step “Identify the Business Question,” but the concept can be applied to any question you are attempting to answer through visualizing data, whether that’s a business question or just something you’re curious about. For example:

- What is contributing to our business growth?
- Where is the best place to get a cup of coffee in Seattle?
- What are the odds of becoming a professional athlete?



It's incredible to see how often this fundamental step is overlooked in the real world. My theory as for why is that it is becoming increasingly easy to access data and it can

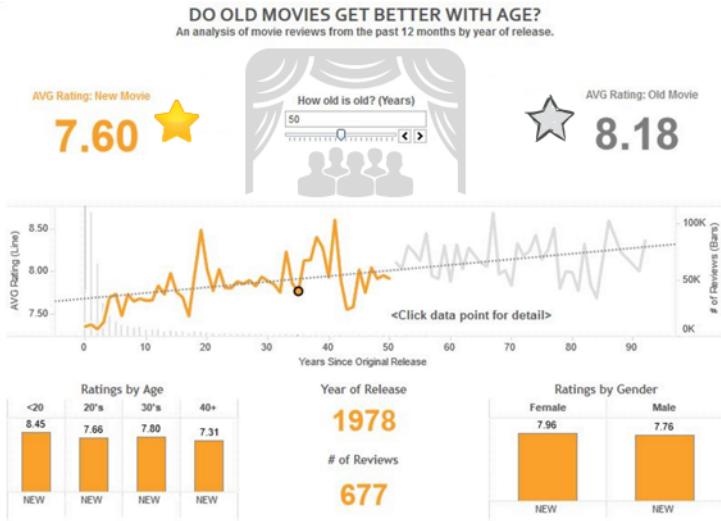
be tempting to look at too much at one time. While it may sound convenient to put every number you care about on a single dashboard, this approach typically does more harm than good.

By having a single business question in mind going into a data visualization design, you set yourself and your end users up for success by providing a clear purpose for the dashboard. Furthermore, most times it's appropriate to literally state the question to open a dashboard. This not only provides a purpose, but gives you an opportunity to frame what end users will be looking for in your dashboard even when you're not there to explain it to them.

Notice that this step is called Identify the Business *Question*, and I can tell you that the singularity of the word “question” is not coincidental. Some of the best data visualizations I’ve seen as well as some of my own most popular dashboards state a single question at the top. From there, the rest of the dashboard attempts to answer that single question. Sure, additional insights/context /“rabbit holes” are likely to emerge, but by going into a data visualization design with one question in mind, you improve the chances of your work providing value. This is because when you identify a single question, you’re almost guaranteed that your dashboard will serve at least one purpose: answering the question.

Throughout the chapters about the INSIGHT framework for data visualization, I will use my dashboard, *Do Old Movies Get Better with Age?*, to illustrate how the framework can be applied in a real-life scenario. This was the winning entry in Tableau’s 2013 Iron Viz Championship, a popular data visualization contest held live at the Tableau Conference each year. I’m using this as my example to show both that I practice what I preach, and that you can have success adopting this framework.

Here is a screenshot of the viz:



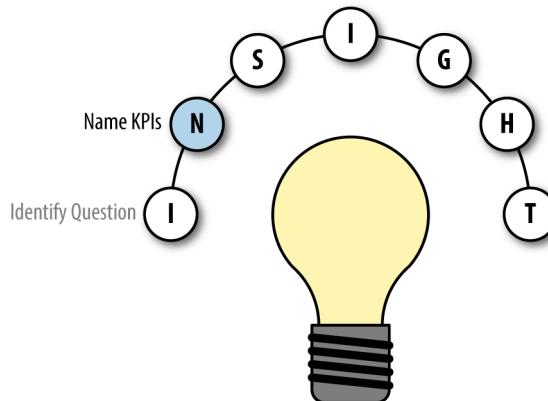
Notice that the dashboard asks a single question in the title, which makes it clear what the audience should be looking for throughout the rest of the dashboard. The next chapter will share how you help answer the business question.



# **Name KPIs**

The second step in the INSIGHT framework for data visualization is to name the key performance indicators, or KPIs, that will help answer the business question identified as the focus for the dashboard. More specifically, the KPIs should be documented, which serves three key purposes:

- Aligning stakeholders on the data they agree about will help answer the business question that is important to the organization. It is much easier to debate the merit of each KPI and make modifications to the list at this point rather than waiting until you have a data visualization developed.
- Informing the dashboard design. Once a consensus is reached on the KPIs that should be included in a data visualization, you basically have a blueprint in terms of what should be included in the final product.
- Providing a historical record so you can review how requirements change over time. It is common for a business environment to change to the point where new KPIs or ways of looking at the data emerge. You may discover a customer segment you want to focus on or realize that a KPI is no longer helping answer the business question. Allowing some room for change is OK, but it helps to maintain a record of the dashboard evolution:



The KPI-naming exercise is not a one-size-fits-all task. It's possible that a single KPI can answer the question at hand. Other times, several more KPIs will be required. I personally try to limit my dashboards to twelve total components, including KPIs. They are called *key* performance indicators for a reason. If you are finding you have more than five or six key metrics, you likely need to put some thought into what is really moving the needle for your business and differentiate between key and secondary indicators. The secondary indicators may be better served in a drilldown type of view, or you may run the risk of them muddying up the primary focus.

## So How Do You Name the KPIs?

I find the best KPIs come from one of two places, or a combination of both:

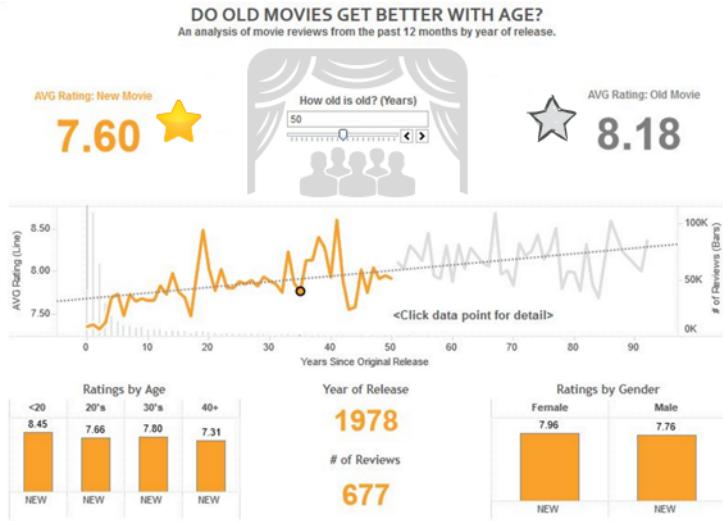
### *Measurement planning*

Putting some thought into how success should be measured and what information is required to answer a business question is half the battle when it comes to building valuable data visualizations for your organization. Many times, this step alone provides enough to inform a dashboard's requirements.

### *Discovery analytics*

If you're attempting to answer a new business question, it is not always known what KPIs are most equipped to help answer that question. By using data visualization tools such as Tableau, there is an opportunity to quickly pivot through many metrics to get a better feel for which may be most helpful.

Let's take another look at my data visualization, *Do Old Movies Get Better with Age?*; this time to see how the KPIs ladder up to answering the identified question:



In the case of this question, “Do old movies get better with age?”, there is at least one metric that should obviously be required to help answer the question: age. From here, there was a choice on how success should be measured, or what constitutes a movie being “better.” Should the measurement of success be the amount of money the film grossed? Oscar nominations? Times it made me laugh? All potential choices (assuming the data is available), but I chose to focus on the average IMDb movie rating. I also included a third metric for number of reviews because I thought an important component of the analysis was that older movies tended to have a smaller sample size of reviews. But that’s it—three “KPIs.”

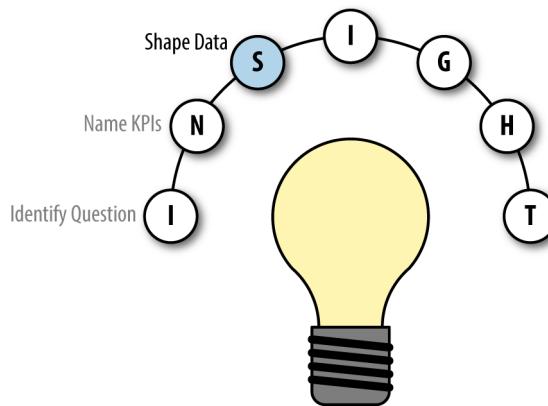
The rest of the dashboard looks at those three KPIs in different ways in an attempt to answer the question. The first and second parts of the dashboard are descriptive in nature, providing the answer and a trend right off the bat. The final third is more prescriptive in that it slices and dices average rating by demographics, which helps answer the “why” behind the results.

The next chapter will share how to prepare the KPI dataset so that it’s usable in your data visualization.



# Shape the Data

The first two steps in the INSIGHT framework for data visualization—identify the Business Question and name KPIs—will inform what should be included in the third step: Shape the Data. In [Chapter 4](#), I discussed the ideal way to shape data for use with Tableau, but this step is also about accessing the required data, doing any preparation such as joins or aggregations, and making sure the dataset is laid out for your data visualization requirements:



Mastering this step is truly half the battle when it comes to data visualization, and this chapter will touch on different tactics for ensuring your analyses are set up for success.

# Shaping Data for Use with Tableau

As mentioned in the introduction to this chapter, we have already covered the ideal way to shape data for use with Tableau, but it's worth repeating some of this critical concept. In general, it is easier to work with "vertical" tables than with "horizontal" tables. In most cases, each column should represent one field with the field name in the column header, and each subsequent row should be an entry under that column header.

One common type of data that usually requires reshaping is survey data. In survey data, each column corresponds to a different question, such as "on a scale of 1 to 5 indicate the degree to which you agree or disagree with the following statement," and each row represents the answers to every question from a single respondent. This is a recipe for disaster as the data is wide. It may be easy for analysts to look at but Tableau won't like it and you will be fighting to glean any insights. Now, if you reshape, or what Tableau calls "pivot," the data so that there are only a few columns with lots of rows, you may not like looking at the data in a spreadsheet but Tableau will be much happier.

## Joining and Aggregating Data

As your analyses become more sophisticated, it is not uncommon to require access to data across multiple sources—and when possible, it is often best to consolidate the data into one centralized data connection. Further, as the size of your data grows, you will need to start thinking strategically about how to aggregate your data so it continues to be manageable to work with.

One of the most common questions I receive is if it's better do this type of data preparation *in Tableau* or *before* you connect to the data with Tableau. Tableau comes with some useful data preparation functionality, including the ability to do joins, blending, interpretation, reshaping, and aggregation. My general recommendation is to prepare the data as much as possible before it gets to Tableau, but these capabilities are extremely handy for trying new things on the fly before making them more permanent in the underlying data source or stream.

One big exception to pre-Tableau preparation that I would be remiss if I didn't mention, is with calculated measures. One of the most powerful features of Tableau is its ability to do quantitative calculations on the fly across any dimension you throw on a view. If you are aggregating your dataset before it gets to Tableau to the point where every number is predetermined, you are basically working with an OLAP, or cube, data source. This takes away a lot of the exploratory value of Tableau. I suggest having the raw calculated measure inputs in the dataset before connecting with Tableau, but creating calculated measures themselves within Tableau.

# Laying Out Data for Specific Analyses

Certain chart types, such as Sankey Diagrams and Funnel Charts (pictured in [Chapter 40](#)), are much easier to create if there is some thought put into how the underlying dataset is laid out. Situations like this are a case-by-case basis, but I do want to point out again that the dataset itself can actually help you create data visualizations more efficiently. If you come across a situation like this where a certain chart type requires a unique data format, I advise keeping a master data connection for the rest of your workbook, but creating a second data connection for these unique requirements. The second dataset may have the exact same data, but it can be laid out in a way that makes it easier to accomplish the special use case.

## Shaping Data for the Iron Viz Example

Sticking with the *Do Old Movies Get Better with Age?* dashboard example, let's take a look at how the "Shape the Data" step was applied.

In the case of the Iron Viz competition, all three contestants were provided with a clean and easy-to-work-with dataset. After all, this contest is meant to be a data visualization contest, and not a data preparation contest. So I admit, this dataset was a lot easier to work with compared to many of the datasets I come across in my day job. That also means that I did not literally need to transpose or reshape the dataset in order to use it. That being said, there was one notable hole in the underlying data: age. As mentioned in the previous chapter, age was my number one KPI for my analysis, and I simply could not create the visualization I had envisioned without this field.

Rather than move on to a different concept, I evaluated the existing fields that I had to work with. Fortunately, I found that the titles of each movie included the year it was released in the text. If I could strip out the year of release from each title, I would have a numeric field. From there, I could subtract the year of release from the current year, creating my field for Age.

So I created a calculated field to isolate year of release that looked like this:

```
Year of Release = int(mid([Main Movie Title],find([Main Movie Title], "(")+1,4))
```

This calculated field tells Tableau to look at each dimension member in the Main Movie Title dimension and find the "(" character. The number four in the calculation tells Tableau to return the four characters following the "(" character.

After isolating the year of release, I created another calculated field for Age:

```
Age = Current Year - Year of Release
```

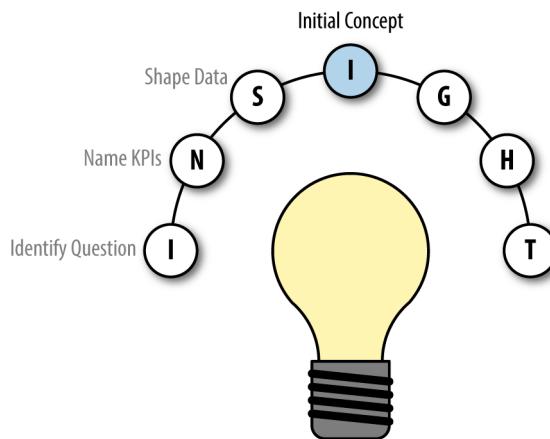
I was then able to use these fields as part of my analysis. This is a perfect example of creating something on the fly in Tableau by getting creative with the existing data. If

this were a situation in my day job, upon proving the value of this field, I would try to make Year of Release and Age permanent fields in my underlying dataset before it gets to Tableau.

In the next chapter, we will finally start to visualize data!

# Initial Concept

Are you ready to finally start visualizing data? I feel like Mr. Miyagi in *The Karate Kid*—making us spend months “waxing on” and “waxing off” before getting to the reason we came here. The truth of the matter is the Identify Business Question, Name KPIs, and Shape Data steps in the INSIGHT framework for data visualization are the foundation required to make your visualizations as effective as possible:



They also make the visualization component itself more efficient, and dare I say *fun*, because the preceding steps help reduce frustration and potential pitfalls. The strategic thinking put in during the I and N steps inform what should be included in the dashboard and the S step makes working with the data more seamless.

After this groundwork has been laid, you’re ready to create an initial concept for your data visualization.

# Creating an Initial Concept

There is no single best way to create a data visualization. There are seemingly endless factors that can influence a dashboard concept, which was actually my motivation behind creating this framework to begin with. If I had to choose the two largest factors though, I would say they are the data visualization's (a) audience and (b) purpose. These two factors should at least provide a starting point for what your data visualization will look like. When creating an initial concept, think about these questions:

*Who is my audience?*

- If the audience is fellow analysts, you may be able to provide more advanced chart types (i.e., box-and-whisker plots, Pareto charts, etc.).
- If the visualization is meant for a mass audience, consider simplifying the concept both in terms of layout and chart selections to make your story clearer.
- If the visualization is being created specifically for a C-level audience, more weight should probably be put on boiling explicit insights to the top before diving into any in-depth prescriptive analyses.

See [Chapter 86](#) for more on this topic.

*What is the purpose of the data visualization?*

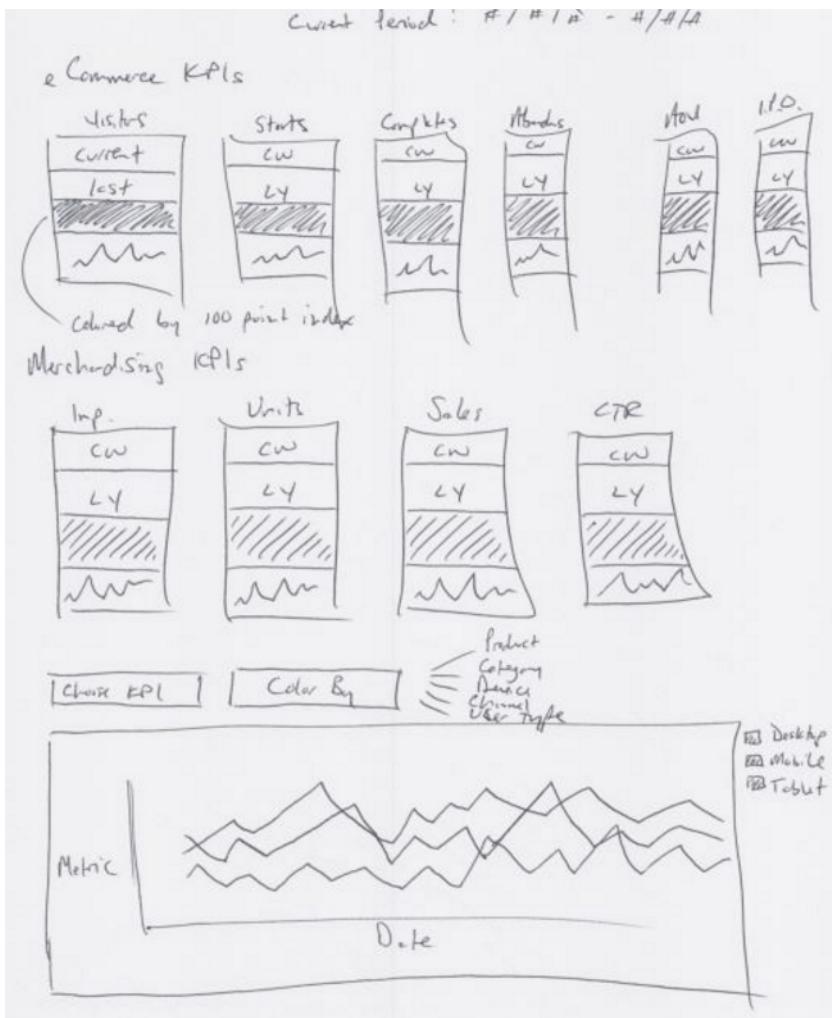
We have covered this in the Identify Business Question step, as well as put some thought into answering the question in the Name the KPIs step. During the initial concept phase, you have a chance to prioritize the KPIs in order to communicate the answer in a way that is intuitive to the end users.

With the answers to these two questions in mind, I create an initial concept in one of two ways:

*A hand-drawn sketch*

That's right—I will literally draw out a concept. In fact, this has become my preferred method of creating an initial concept because it is extremely efficient, and since I am not investing much time building an idea out, I do not get frustrated when I inevitably have to make a change.

I cannot draw, but that is OK! The outline, prioritization, and chart selections will come through and it will make it much easier to pivot if and when I receive feedback. Here's a real-life example from a project with one of my data visualization partners:



### A partially developed workbook

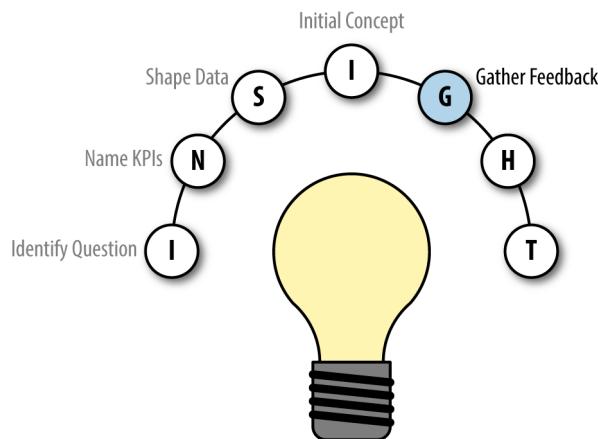
In some cases, there are certain dashboard elements that I know will be included in my final deliverable. If so, I don't mind spending some time building these out and will experiment with sizing, layout, and flow on a dashboard "canvas." When I take this approach, I usually at least have an idea of how the dashboard will be distributed, which will inform the appropriate canvas size (i.e., optimized for mobile, optimized for printing, etc.). Once I have a blank canvas in Tableau, I can begin placing elements to create an initial concept before requesting critiques.

No matter which approach to creating an initial concept you choose, the key is to be open to change at this point in the process. I recommend not investing a great deal of

your effort during this step because *your concept will change*—which is discussed in the next step—and you will be happy you saved your effort for later in the framework.

# Gather Feedback

As mentioned earlier, there is no one best way to create a data visualization and there are almost infinite factors that can influence a dashboard concept. I have always said that data visualization is a form of art, and with that, “beauty is in the eye of the beholder.” One anecdote that I enjoy sharing is that regardless of the Tableau contests I’ve won, and the popular publications that have picked up my work, and the *hundreds* of times I have created data visualizations...



I have yet to build a dashboard that every single viewer thought was perfect.

And that’s OK! I understand that it is impossible to please everyone all the time. Among a diverse audience, different people will have different ideas of what the business questions are, opinions on which KPIs should be used to answer those business questions, varying levels of knowledge on data visualization best practices, personal design preferences, and much more.

Gathering feedback should be an expected—and welcome—step in your data visualization process. Gathering feedback on your initial concept from the audience that will be using it provides at least three major benefits:

#### *Improved efficiency*

By taking time to gather feedback, you are able to get to the final product faster. When you dedicate a step to gathering feedback in an organized fashion, you are much less likely to receive sporadic, overly detailed feedback after you bring the concept to life.

#### *Reduced frustration*

In my personal experience, when I know that incorporating audience feedback is part of the design process, it makes it less frustrating when somebody requests that we move in a different direction.

I suggested in the previous chapter that you shouldn't invest a lot of time on the initial concept. It is easier to hear criticism or varying opinions when you haven't spent time actually building out a vision before gathering feedback.

#### *Ownership*

I have found that when you proactively gather feedback from end users, they take ownership in your final dashboard. Things tend to go smoother when the audience has some stake in your work. I don't have any science behind this, but I think subconsciously, people are less likely to criticize something if they had some say in its creation.

How you gather feedback is up to you and is largely dependent on your audience. For example, if you are building a dashboard for a small audience of one to three, you may simply send them an image of your initial concept and ask them to respond with their thoughts. If the audience is slightly larger, you may opt to set up an hour-long brainstorming meeting to think through different approaches as a group.

When designing for a mass audience, consider sending the initial concept to two or three individuals in your personal network to get a sample of how the overall population may look at your visualization. The key with your “preview” audience is that you choose individuals with diverse backgrounds that better reflect a larger, and diverse, public audience.

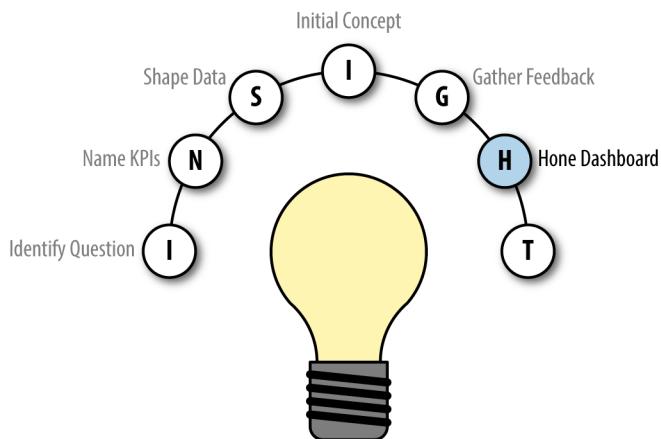
Even if you are creating a visualization for yourself, I encourage you to at least sleep on your initial concept. By stepping away from the design process, you can determine if you are still happy with the approach or have some ideas for improvement.

In the next chapter, I will share how to incorporate the feedback that you've gathered and hone the dashboard.

## CHAPTER 82

# Hone Dashboard

After gathering feedback from your stakeholders, you are ready to finalize your data visualization. I call this step “hone the dashboard” because you are refining your initial concept based on the audience input provided in the previous step. At this stage in the INSIGHT framework, the business question (or purpose of the data visualization), the KPIs that will be used to answer that business question, and the dataset that contains those KPIs, should all be in a solid—if not final—state:



The task that is left to complete is investing the time to actually build out the components of the dashboard and lay them out. Depending on how much time you invested in the Initial Concept step, you may have quite a bit of effort left to get to a working product. In fact, I recommended that in most situations, the initial concept should have very few working components or even none (in the case that your initial concept was a hand-drawn sketch). While this step leaves quite a bit of work to bring your

concept to life, it is much more efficient because you only have to build the dashboard elements once (versus building everything only to have the audience make you move in a completely different direction).

While there is no one best way to create a dashboard, here are a few general rules I have in the back of my mind when trying to finalize a view:

- Dashboard dimensions should be determined by the method of distribution. For example, if you know the dashboard will be printed out or attached to an email, consider making the dashboard the same size as an 8½- by 11-inch piece of paper. If you know the dashboard will be consumed on mobile devices, make the dashboard the same size as a tablet or mobile phone.
- Keep dashboard elements to 12 components or fewer. By capping the number of dashboard elements, you keep the design focused and help the story in your data emerge more efficiently and effectively. If you find yourself requiring more than 12 components, consider breaking up the dashboard into multiple views or even multiple business questions. Occasionally I will start with a descriptive overview, but use interior pages to provide more prescriptive analyses.
- *Keep It Simple* in terms of chart selections, user experience, and design. Being clear and concise is another best-practice method for telling your data-driven story, which we will discuss in the next and final step of the INSIGHT framework.

I will close this chapter by addressing one of the biggest questions in any design process: *What do you do if the audience feedback is conflicting or outside of data visualization best practices?* Remember, part of honing the dashboard is the opportunity to incorporate the stakeholder feedback, and there are several benefits to doing so which were outlined in [Chapter 81](#). So what should you do if the audience disagrees with you or each other? Or worse, they want to create a dashboard with 15 pie charts on it?

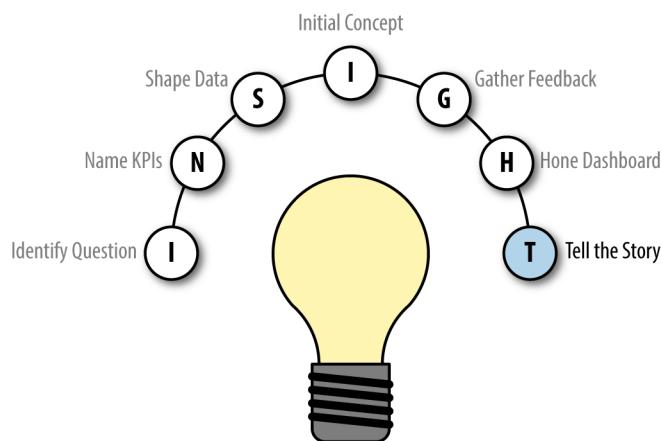
Ultimately, you are the final gatekeeper of the dashboard and need to prioritize which feedback makes sense and how it fits together. The designing outside of best practices question is one I have struggled with because I really do find it hard to put my name on views that are just a huge spreadsheet or include charts such as pies or packed bubbles. My solution is to provide the views in both ways: (a) the way that was requested, and (b) the way that I recommend. This way everybody gets what they want, and I at least have the opportunity to share the best-practice approach to a specific situation. Sometimes my advice is taken; sometimes it's not. But by providing both approaches, I don't lose any sleep because I know I tried to help.

Honing the dashboard can be an iterative step itself, so don't worry if you don't get it perfect on the very first try. Following the steps in the INSIGHT framework to this point will help make this step as efficient as possible, and you soon will be ready to distribute your work.

## CHAPTER 83

# Tell the Story

Once you have honed your dashboard and possibly iterated a couple of times to land on a final product, you are ready to distribute your data visualization. I call this step “Tell the Story,” because you have hopefully integrated some storytelling tactics into your dashboard and your data has a compelling story to share. Further, I believe that the entire *purpose* of data visualization is to find and share actionable stories that are based in quantitative evidence. If your dashboard does not provide insights that can help inspire action in your organization, it’s probably not worth sharing to begin with:



No, there's not always an earth-shattering realization that comes from every single weekly report you may create. That being said, if you have put some strategic thought into identifying the business question that your dashboard is answering and how the

KPIs you're measuring will answer that question, you drastically increase the chances of your dashboard providing actionable insight.

So assuming you've followed the INSIGHT framework to this point and your dashboard is sure to provide valuable insights, your method of distribution once again comes back to my largest factor when creating a data visualization: the audience. In my experience, the audiences for my dashboards are almost as diverse as the factors that go into creating the dashboard itself. They truly live on any and all points of a spectrum in terms of both their analytical sophistication and their technical infrastructure.

For example, I have created dashboards for audiences that have never heard of Tableau. With the analytics partner that I have in mind, I refreshed the dashboard data once per month and saved the dashboard as a PDF that I attached to an email. On the other end of that spectrum, I've worked with several partners that have a very robust Tableau Server implementation in place where I can post the dashboards and set them up to automatically update. This makes it a little easier to tell the story because users can subscribe for periodic email updates, and starting with Tableau 10, you can even subscribe other people to receive your dashboards with the most up-to-date data available.

Regardless of your audience, there are a few specific tactics you can use to help tell the story that emerges in your dashboard:

#### *Context*

You won't always be around to explain your dashboard to end users. Be sure to provide some context for your dashboard so your end users have an idea of what they should be looking for as they navigate your work. This context can be as simple as a title and subtitle on the dashboard itself, but can include additional information such as the source of the data or the date range being applied to the visualizations.

#### *Inline insights*

One of the most effective methods of providing findings and recommendations for next steps is to explicitly write them out within the dashboard. One way to do this is to design in a text box for "inline insights" that the developer or analyst can modify with their own thoughts. Some insights can even be automated through the use of calculated fields that will display a sentence based on the performance of the data.

#### *Annotations*

If a dip on a line graph, for example, is influenced by an event that is not readily apparent in the dashboard, add an annotation to the line graph to provide additional context. In just one hypothetical example, let's pretend that the line graph is looking at daily production. If you know the factory was closed for two days

due to unforeseen circumstances, but this is not data that is normally tracked or reported to the audience, help provide this context by annotating the graph.

I'll close the INSIGHT framework for data visualization part by using the *Do Old Movies Get Better with Age?* example—the winning visualization in the 2013 Tableau Iron Viz Championship—one more time. The three Iron Viz finalists each year are among the best of the best Tableau users in the world. While I pride myself on balancing data and design, I considered one of my competitors more skilled at design than me, and the other competitor more skilled at data. While I think my balance between the two helped, I think the primary reason I was able to win was *storytelling*.

At the end of the Iron Viz contest, each contestant gets to explain their data visualization. While the other two contestants used most of their time to walk through the technical aspects and user experience of their dashboards, I literally told the story behind what motivated my approach. I believe this helped me connect with the judges and gave purpose to my data visualization, providing the edge needed to pull off the win. In a corporate environment, the judges from the Iron Viz are your audience, and pulling off the win equates to finding and sharing a valuable insight that causes positive change for your business.

**Part V** will discuss more storytelling tactics you can integrate into your dashboard development process.



PART V

---

# Storytelling



## CHAPTER 84

# Introduction to Storytelling

So how exactly do you tell a data-driven “story”? I fear that storytelling in the context of data visualization has become a bit of a buzzword. What’s worse is the concept is often mentioned without practical examples of how you can apply storytelling techniques to data visualization. This section offers an introduction to storytelling and aims at providing several tangible tips for incorporating storytelling techniques into your data visualization. If the INSIGHT strategy described in **Part IV** is the key to taking your work from good to great; storytelling is the key to taking your data visualization from great to *actionable*: the “Holy Grail” of the profession.

Storytelling has been used in every culture as a means of engaging audiences and communicating everything from entertainment to moral values. Stories work because humans are wired to retain stories; not independent facts or statistics. In fact, after a presentation, 63% of attendees remember the stories presented, while just 5% of attendees remember the individual statistics.<sup>1</sup> This is particularly relevant to data, where business insights risk being lost amid a deluge of increasingly large sets of numbers. Think about what you remember after seeing an effective data visualization compared to a raw spreadsheet of numbers. The difference is inherent storytelling at work.

The parallels between data visualization and storytelling are undeniable. Much like stories, data storytelling includes three critical elements:

1. Characters
2. Plot or storylines

---

<sup>1</sup> Heath, C. and Heath, D. (2007). *Made to Stick: Why Some Ideas Survive and Others Die*. New York: Random House.

### 3. Narrative

With data, KPIs and contextual metrics can be considered the characters, storylines can be considered the insights, and the narrative is data visualization—or the style in which the insights are being communicated.

Data visualization practitioners have a lot of say with all three elements of data storytelling: your strategy informs the KPIs that are being used to answer the business questions at hand; your ability as an analyst helps unearth the insights from the data; and the style in which you communicate those insights goes a long way in determining if your visualization will be actionable and cause a change.

Interestingly, not to mention amusingly, data storylines tend to follow another parallel, *The Seven Basic Plots* from Christopher Booker:<sup>2</sup>

1. Overcoming the Monster; the protagonist (us), uses the power of data insight to make a recommendation that overcomes an obstacle to the business.
2. Rags to Riches; we leverage data visualization to add value and maximize our company's ROI.
3. The Quest; we have a business question or know something is impacting the business, but must use data visualization to locate the answer.
4. Voyage and Return; we travel to lengths only data visualization can bring us, returning heroically with business-changing insights.
5. Comedy; it really can be funny when data visualization reveals an insight that was seemingly under our nose for quite some time, but was disguised as a wall of numbers in our monthly Excel report.
6. Tragedy; unfortunately, this may be the most common storyline. At least data visualization helps us find the “tragedy” as efficiently as possible and, when done well, helps us know what to do about it.
7. Rebirth; this one is my favorite. This is when the main character realizes the error of his ways, and uses the power of data to become a better employee.

As our strategy generally informs the “characters” in our analyses and the plot is one of the seven just outlined, these tips will focus on narrative, or style of communication, to maximize the effectiveness of two types of stories:

- You, the visualization creator, know the outcome to the story. In this case, it is your responsibility to communicate the insight as effectively as possible so that action is taken.

---

<sup>2</sup> Booker, C. (2004). *The Seven Basic Plots: Why We Tell Stories*. London: Continuum.

- You make the end user part of the story. This happens when you include interactivity that allows users to find storylines about themselves or something they care about.



# A Data Visualization Competition— That's Also an Analogy for the Data Visualization Process

Before getting into the tips, I want to use the Iron Viz competition one more time to share a bit more about how I think about data visualization. If you have ever seen me present on the topic of visual analytics, you have heard me use Tableau's Iron Viz Championship as an analogy for good data visualization. The "Iron Viz" is a competition held annually at Tableau's conference at various locales throughout the United States. It is the culmination of a year's worth of Tableau users using the latest in Tableau product developments to create innovative Tableau Public dashboards. There is only one way to receive an invite to compete in the Iron Viz, and that is to win one of three Tableau Public feeder contests held throughout the year leading up to the annual conference. During the contest, the three "Iron Vizzers" compete live in front of several thousand conference attendees. Each contestant is given the same dataset, and they have exactly 20 minutes to create the best data visualization they can before being judged by the audience, the previous year's winner, and a panel of Tableau product experts.

It's quite a spectacle—and a lot of fun. As the contest begins, each competitor walks to the stage through dim lighting and heavy fog. Dramatic music is blaring. The competitors even don chef-like jackets—a nod to the contest's format borrowed from the *Iron Chef* television show. After some announcements and introductions from two emcees, the contest is played out on huge projection screens so the audience can watch every competitor's move in real time.

The reason I love this event as an analogy for data visualization is because it is a perfect balance of sizzle and steak—both at their highest levels. The presentation aspect is needed to get the largest audience possible interested in the event and make it an

entertaining experience from start to finish. It truly transforms the type of work I do every day into must-see theatre. It guarantees that the event will be (literally) remarkable, enticing the audience to share on social media, write blog posts about it when they return home, and of course, come back in the future for more. At the same time, the three contestants have been thoroughly vetted and are among the best in the world at what they do. The visualizations they are about to pull together promise to be top quality, and are the core of the event. As with data visualization, the data, analysis, and the designer's ability to communicate that analysis are the heart of the work. These key elements are strong enough to stand on their own, but the Iron Viz maximizes their effectiveness by helping make them memorable and shareable.

I argue that the best data visualizations are those that not only incorporate data visualization and analysis best practices, but those that also possess a sense of storytelling—whether that be through usability, graphic design, inline insights, or innovative visual approaches that help make the message of the dashboard stick.

My goal in **Part V** is to provide actionable tips that will help you round out your existing corporate dashboards in a way that maximizes their effectiveness in your workplace.

# Tip #1: Know Your Audience

There are not many tactics that are going to improve the chances of your data visualization making a difference around the workplace more than tip #1: *know your audience*. Before you create a single chart that you intend to share, putting some thought into who will see that chart will dramatically improve your data visualization.

My favorite analogy for this concept is the old Southwest Airlines *Wanna Get Away?* campaign. In one of the commercials, a rapper performing at a concert mistakenly and passionately thanks the audience by saying “Thank you, Detroit—we love you!” The crowd begins to boo and another performer informs the speaker that “Detroit was last night.” I have actually been at several concerts where the artist on stage thanks the home crowd by name. Even though the performer surely thanks every stop’s hometown in the same way, hearing your own city’s name provokes a great sense of pride and is usually an easy way to make a connection with the audience. Get it wrong, however, and the city shout-out has an equally, if not more, negative impact on the crowd.

This same principle can be applied to your own corporate dashboards. Knowing your audience goes a long way to making a connection and maximizes the chances that your end user will understand and happily adopt the reports that you have created. Get it wrong, and you risk permanently damaging the chances of getting your visualizations off the ground around the office. Think about this the next time you undertake a visualization project. Here are just a few examples:

- If your dashboard is intended for a C-level audience, keep your work simple and direct. Focus on KPIs and the progression toward goals. You may also consider creating dashboards that are optimal for being saved as PDFs, to improve the chances that your work is either attached to an email or printed out and handed to a C-level executive.

- If your work is intended for fellow analysts, build in interactivity that allows them to find their own stories in the data. In Tableau, filters, dashboard actions, and parameters are my go-to tools for achieving this.
- If you are using Tableau Public and attempting to make your data visualization connect with a mass audience, don't be afraid to use freeform dashboards, story points, and even incorporate some graphic design elements. Believe it or not, some people think of data as dry! Leveraging some of these tactics will make your visualizations much more shareable.

Data visualization is not a one-size-fits-all practice. Regardless of the storyline you have to share, knowing your audience will help you prioritize and make the most of the data-driven storytelling tips to follow.

## **Tip #2: Smooth the Excel Transition**

Tableau is not Excel. Excel is not Tableau. One of the most common barriers to Tableau adoption is the belief that similar data visualizations can be created more easily in Excel. This may be true for certain situations, but if your goal is to master the art of data-driven storytelling, you must leverage the strengths of each software individually. I do not use Tableau to create and store my datasets. While it's possible, the thought alone makes me cringe with frustration. That is not what Tableau is best at. On the other hand, I don't attempt to use Excel to make beautiful, interactive dashboards. I'm sure with some elbow grease, this could be achieved, but the one million row limit in Excel alone makes this solution impossible in the enterprise-level visual analytics projects I work on. That is not what Excel was designed to do.

And that is OK!

I love both of these programs and use them almost every day, but for different reasons. If Tableau's goal was to replace Excel, Excel would not be one of the primary connection types available in its software. Tableau Personal Desktop users, who are restricted to using flat data files, rely heavily on the ability to interact with their Excel files in Tableau. Tableau knows that it is not Excel.

Nevertheless, the Tableau versus Excel debate is one I encounter regularly. My recommendation for anybody experiencing the same pushback to leveraging the data visualization capability of Tableau, or moving from crosstab views to visual analytics, in general, is to start with highlight tables. The following exercise can be used to help illustrate the power of applying even a very simple visualization to your data. You'll find that even with the most basic visualization, you'll transition your forgettable spreadsheet to a remarkable data story.

First, take a look at the following image, which is what an Excel spreadsheet looks like with no conditional formatting. See how many seconds (minutes, maybe?) it takes you to determine the top three values in the table:

State	Customer Segment			
	Consumer	Corporate	Home Office	Small Business
Alabama	\$45,552	\$20,843	\$52,121	\$8,191
Arizona	\$27,301	\$24,988	\$47,291	\$20,817
Arkansas	\$3,422	\$66,134	\$11,505	\$15,128
California	\$229,462	\$533,151	\$284,838	\$114,270
Colorado	\$10,172	\$38,715	\$40,599	\$42,725
Connecticut	\$10,495	\$15,899	\$6,469	\$9,439
Delaware	\$3,543			
District of Columbia	\$13,883	\$77,912	\$29,628	\$97,446
Florida	\$97,118	\$180,177	\$151,107	\$75,207
Georgia	\$23,659	\$65,199	\$68,561	\$38,920
Idaho	\$26,695	\$28,374	\$6,100	\$34,474
Illinois	\$94,632	\$316,880	\$152,589	\$103,696
Indiana	\$23,368	\$105,152	\$26,336	\$39,225
Iowa	\$8,548	\$33,265	\$42,301	\$4,587
Kansas	\$12,160	\$54,677	\$21,785	\$21,964
Kentucky	\$3,139	\$32,534	\$22,182	\$2,905
Louisiana	\$13,355	\$34,178	\$3,734	\$15,344
Maine	\$28,039	\$23,837	\$25,106	\$20,139
Maryland	\$49,992	\$19,570	\$44,077	\$11,266
Massachusetts	\$48,764	\$111,533	\$28,661	\$39,494
Michigan	\$93,473	\$88,627	\$79,004	\$63,490
Minnesota	\$25,044	\$82,637	\$60,135	\$22,673
Mississippi	\$9,916	\$2,488	\$17,883	\$11,631
Missouri	\$9,110	\$61,384	\$15,001	\$28,207
Montana	\$4,613	\$5,293	\$9,104	\$10,394
Nebraska	\$67	\$18,131	\$11,804	\$10,920
Nevada	\$11,978	\$4,937	\$3,114	

Next, take a look at the same data, encoded by color. In this case, the higher the sales, the darker the green. Now count how many seconds it takes you identify the three highest values:

State	Customer Segment			
	Consumer	Corporate	Home Office	Small Business
Alabama	\$45,552	\$20,843	\$52,121	\$8,191
Arizona	\$27,301	\$24,988	\$47,291	\$20,817
Arkansas	\$3,422	\$66,134	\$11,505	\$15,128
California	\$229,462	\$533,151	\$284,838	\$114,270
Colorado	\$10,172	\$38,715	\$40,599	\$42,725
Connecticut	\$10,495	\$15,899	\$6,469	\$9,439
Delaware	\$3,543			
District of Columbia	\$13,883	\$77,912	\$29,628	\$97,446
Florida	\$97,118	\$180,177	\$151,107	\$75,207
Georgia	\$23,659	\$65,199	\$68,561	\$38,920
Idaho	\$26,695	\$28,374	\$6,100	\$34,474
Illinois	\$94,632	\$316,880	\$152,589	\$103,696
Indiana	\$23,368	\$105,152	\$26,336	\$39,225
Iowa	\$8,548	\$33,265	\$42,301	\$4,587
Kansas	\$12,160	\$54,677	\$21,785	\$21,964
Kentucky	\$3,139	\$32,534	\$22,182	\$2,905
Louisiana	\$13,355	\$34,178	\$3,734	\$15,344
Maine	\$28,039	\$23,837	\$25,106	\$20,139
Maryland	\$49,992	\$19,570	\$44,077	\$11,266
Massachusetts	\$48,764	\$111,533	\$28,661	\$39,494
Michigan	\$93,473	\$88,627	\$79,004	\$63,490
Minnesota	\$25,044	\$82,637	\$60,135	\$22,673
Mississippi	\$9,916	\$2,488	\$17,883	\$11,631
Missouri	\$9,110	\$61,384	\$15,001	\$28,207
Montana	\$4,613	\$5,293	\$9,104	\$10,394
Nebraska	\$67	\$18,131	\$11,804	\$10,920
Nevada	\$11,978	\$4,937	\$3,114	

Much less time, right? This is the power of data visualization. This image is called a highlight table, and as you can see, even the simplest forms of data visualization can lead to much shorter time to insight. You can easily create highlight tables in Excel or Tableau, and I have found they are a great way to introduce the power of data visualization. In the tips to follow, I will introduce more complex data visualizations and data-driven storytelling techniques that can be integrated more efficiently in Tableau, but many times, you have to start by helping your audience understand why data visualization plays an important role in analytics.



## **Tip #3: Leverage Color**

I found I could say things with color and shapes that I couldn't say any other way—things I had no words for.

—Georgia O'Keeffe

During his keynote presentation at the 2014 Tableau Conference, Christian Chabot (Tableau co-founder and former CEO) talked about data analysis being a *creative* process. As an analyst, you may not think of yourself as an artist, but by its nature, data visualization is an art form. Visual analysts use data to express insights and provoke action. As Chabot put it, “Analysts and artists are both on a mission to reveal something new—to discover truth, to find meaning.” So if you are doing visual analytics, congratulations—you’re an artist! If that label makes you feel slightly uneasy, don’t worry, this chapter covers several tips on utilizing one of the most powerful forms of artistic expression: color.

While leveraging color is relatively easy to put into practice, it is also one of the most effective tools for discovering and sharing insights.

### **A Few Benefits of Leveraging Color in Your Data Visualization**

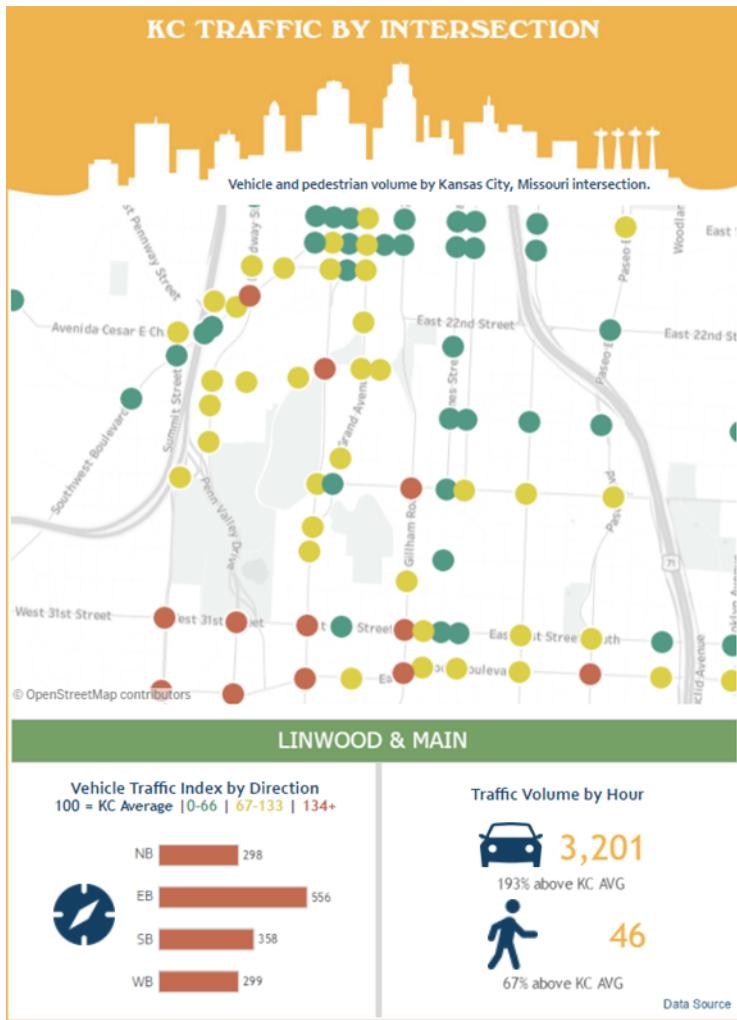
- Color makes the stories in your data pop. The primary use of color is a practical one: it helps the insights in your data emerge, both for you and your audience. Color helps accomplish a shorter time to insight by providing (a) a means to identify dimension members so you can quickly recognize strong and poor performers (i.e., each categorical value can have a different color such as West is colored orange, and East is colored blue), and (b) a scale to illustrate relative performance (i.e., a sequential palette that shows Sales turning a darker green as they

increase or a divergent palette that shows positive sales as blue and negative sales as orange).

- Color helps engage your audience. Careful color selections, or even the use of color at all, is an easy way to subliminally capture and keep the attention of your visualization users. I have found that the use of color is often a key ingredient in achieving remarkable in data visualization.
- Color provides an opportunity for you to show your narrative style. Perhaps your company's brand colors work well as primary or secondary palettes (this is not always the case, so consider with caution!). Maybe you want to use color to make an emotional alignment with the story the visualization is sharing (i.e., a darker look and feel when addressing a darker topic).

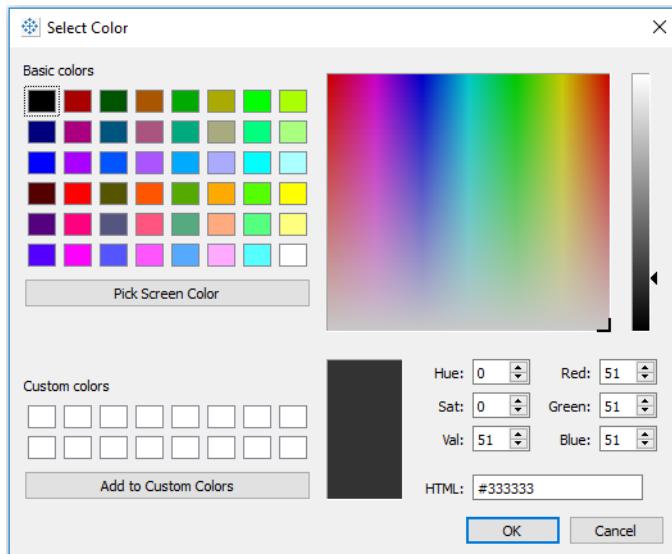
## Customizing Your Use of Color Is Easy with Tableau

The following is one colorful example from my Tableau Public portfolio. Remember, if you ever see something you like in a Tableau Public viz, you can download it to find out how it was created. The same is true with colors—if you see a color you like, you can download the workbook to find the values that generated the color. More on color values in just a moment:



This visualization uses the out-of-the-box “Temperature Diverging” color palette in Tableau. Even though it is typically not best practice to use red and green due to a common color blindness that impacts around 1 out of 10 men, I couldn’t resist using these colors for a viz specifically about traffic. Fortunately, the particular hues of red and green used here are actually color-blind friendly. The only downside was that using the color palette in Tableau was not as flexible as I would like because this certain color palette is only available for use when coloring continuous measures (i.e., lower values colored red; higher values colored green). The good news is Tableau allows you to customize colors and how you can use them if you know the values used to generate them.

There are three types of color values that you can use to customize colors in Tableau. The quickest way to customize colors in Tableau is to know either the (a) Hue-Saturation-Luminance, or (b) Red-Green-Blue values. Both HSL and RGB use a combination of three, three-digit numerical codes. You only need to know one of the combinations because when you use one, the other will automatically be generated. For example, if you type in the RGB values of a color, the corresponding HSL values will be generated. Any color can be added and used in Tableau by double-clicking a color in Tableau color legends until you see a color dialog box and enter the HSL or RGB values. The custom color dialog box looks like this on a PC:



Here are the RGB values for Tableau's temperature diverging color palette if you would like to use them in your own data visualization. The first value is the level of red, the second is green, and third is blue:

Green: 82.153.133

Yellow: 219.207.71

Red: 194.107.81

Orange: 239.180.78

The third type of color value that can be used with Tableau is called hex, which is a six-digit numerical code. Note that all colors have all three types of values (among others): HSL, RGB, and hex. Hex values can be used to create a variety of **custom color palettes in Tableau**. You can use a tool like [ColorSchemer](#) to find the color values you need.

For more tips on getting the most from color, review [Chapter 55](#).



# Tip #4: Keep It Simple

I didn't have time to write a short letter, so I wrote a long one instead.

—Mark Twain

I love this quote from Mark Twain because the author is putting a value on prioritizing content. It is said that Twain's "complete" bibliography remains incomplete due to the volume of his writings, and the fact that they were often completed for obscure publishers—not to mention under a variety of pen names. However, even as one of the most prolific writers of all time, this quote implies that Twain believed the most effective storytelling was done by being clear and concise.

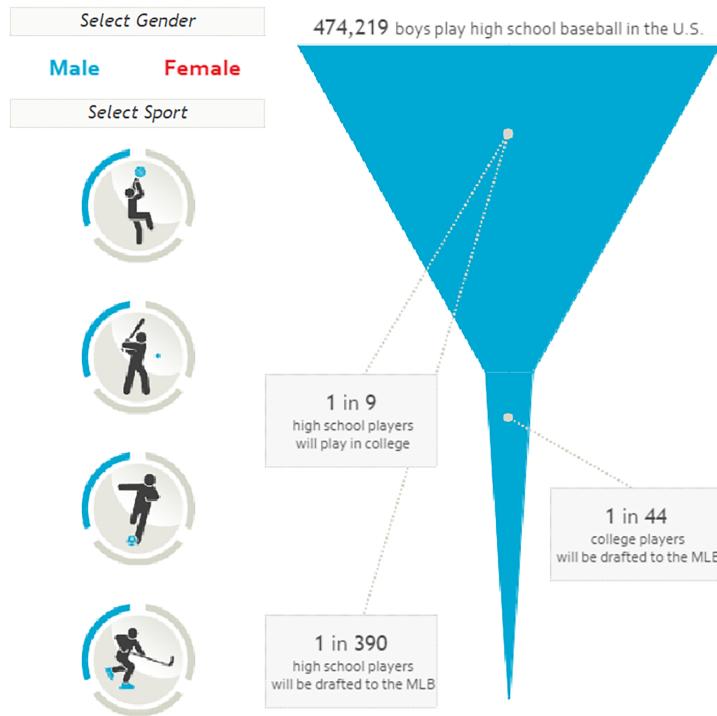
This same idea applies to data visualization: keep it simple. One of the most common mistakes I see in dashboard layout and design is attempting to create silver-bullet dashboards that provide every possible answer to the business question at hand—all in a single view. One of the things I find myself saying often is "just because it is possible in Tableau does not mean you should do it." Tableau makes it easy to add filters, charts, and objects to a dashboard, but there is a point when too many options for the end user actually detracts from your visualization, making it harder for the *story* in the data to emerge.

A concept that I often consider is Occam's razor, usually described as "the simplest answer is usually the correct one." William of Ockham did not have data visualization in mind when he devised this principle in the early 13th century, but I believe the concept fits quite nicely with the practice. Think about how bar charts have withstood the test of time, and despite their simplicity, continue to be one of the most effective ways to communicate the differences in data.

Less is almost always more when it comes to communicating your data-driven story. To help illustrate this tip, let's review one of my most-viewed Tableau Public vizs to date, which was built using four widgets: one title, two filters, and one chart:

## WHAT ARE THE ODDS OF GOING PRO IN SPORTS?

An analysis of high school, college, and pro sports in the United States by gender.



I primarily credit the success of this visualization with its simplicity. The visualization is simple in several different ways:

- It asks and answers a single question.
- It offers only two filters—one with two options; the other with four.
- The story is communicated using a single chart.

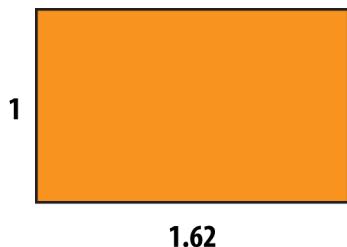
The story in this visualization is almost impossible not to understand—and understand very quickly—by analysts and nonanalysts alike. By keeping it simple, you maximize the effectiveness of your data-driven storytelling across the largest audience possible.

# Tip #5: Use the Golden Ratio

Without mathematics there is no art.

—Luca Pacioli

The Golden Ratio, or 1:1.618, is a number found in patterns that we see all around us every single day. In fact, the shape of this book and likely the screen you're reading this on (if you're an ebook reader) are probably pretty close to a Golden Rectangle, which is a rectangle where the long sides are 1.618 times longer than the short sides. The Golden Ratio is used in the design of the televisions we watch, architecture, paintings from the likes of Salvador Dali, and even occurs naturally in plants and our own DNA.



Shapes and patterns created with the Golden Ratio have been found aesthetically pleasing for centuries, and when it comes to data visualization, these patterns offer the added benefit of helping us prioritize the content we share in our dashboards. I've mentioned that I try to stick to twelve dashboard components or fewer, including titles, charts, and filters. The Golden Ratio is one of the tools I use to help guide how to select and lay out those twelve widgets.

The gist of the layout is that the most important data you are communicating should be toward the top and to the left, with lower-priority information appearing further

toward the bottom and right of your view. Just as you start reading a page in a book, you can expect your audience to approach your data visualization the same way.

Knowing your audience will help you choose which content earns the prime real estate on your dashboard. For example, if your audience is C-level executives, you should likely have your main measurements of success and your progress toward achieving that success near the top left of your dashboard. If you are creating self-serve reports that your audience is going to use to help answer their own needs, you may want to consider prioritizing the filters, which will be a natural starting point for this type of end user.

The Golden Ratio is effective in both a horizontal layout as well as a vertical layout. The choice of a vertical or horizontal dashboard should be based on how you intend to distribute your data visualization. For self-serve dashboards with many filters that other analysts are going to use to help find stories in the data, I typically opt for a horizontal layout. I prefer vertical layouts for dashboards that I know are going to be saved as PDFs and passed around, and for all of my Tableau Public dashboards, which are embedded in blogs that typically have vertical designs.

Using the Golden Ratio in your data visualization is an easy way to subconsciously provide a familiar and well-balanced layout to your audience. It also serves the practical purpose of helping you prioritize content and think about how your audience is going to consume your dashboard.

# Tip #6: Retell an Old Story

Our stories are not new; yet, in the retelling we are reborn as heroes.

—Harley King

I think we've all been there—a few months after introducing a new reporting format, you may start to get the sneaky suspicion that your weekly or monthly updates are not being opened by stakeholders. Or perhaps they are being opened, but that fresh new reporting format that was so well received upon release is no longer being used to gain as much insight as it is capable of providing. This may be due to rarely fluctuating KPIs, a lack of understanding on how to leverage the reporting, or maybe simply a shortage of time to dedicate to finding stories in the data.

The nature of many KPIs used to answer business questions is that they are steady and/or predictable. Think about a website's bounce rate and average time on site—for better or worse, two KPIs often used to measure a company's ability to engage website visitors. Without a site redesign, these two KPIs may never fluctuate more than 5%–10% over the entire life of the website. Once you have seen these numbers two or three times, they become stale, and end users become less and less motivated to use the reporting you're creating for them. In the same vein, you may have KPIs that follow a seasonal pattern, and your end users know exactly what to expect from week to week or month to month; this may make them indifferent to large spikes or dips spotted in your reporting.

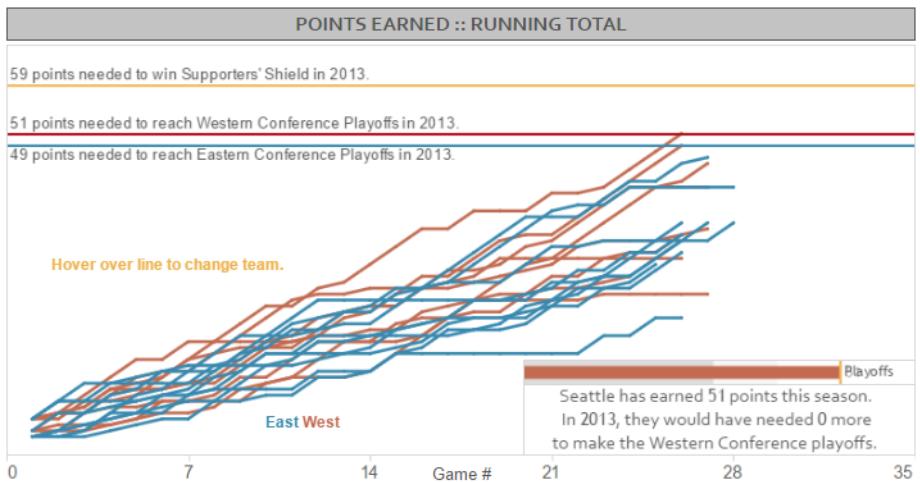
The good news is that data visualization provides a means for re-engaging your audience by retelling your data's stories in new ways. Data visualization tools like Tableau enable you to (a) make your dashboards more aesthetically pleasing, (b) add filters and functionality that allows end users to find their own stories in the data, and (c) makes your reporting more usable so that finding insights is more intuitive for the user.

As just one of infinite possible examples of retelling an old story, I took a shot at reimagining sports standings. As a data visualization evangelist, I view data tables as the least effective way of communicating data. As a sports fan, I am forced to view tables almost exclusively to see how my team is doing—and this is true for every league: NFL, NBA, MLB, NHL, MLS, among others. Tables have been used to track league standings in Europe’s “football” leagues for over a hundred years. They even affectionately call their standings, “The Table.” To show how much more value visualizing sports standings adds compared to the traditional table, I re-created the table in North America’s Major League Soccer as a Tableau Public data visualization.

First, take a look at the traditional standings from [mlssoccer.com](http://mlssoccer.com):

Eastern Conference	Points	Games Played	PPG	W	L	T
Kansas City	35	20	1.75	10	5	5
D.C.	34	19	1.79	10	5	4
Toronto	26	17	1.53	7	5	5
New York	24	20	1.2	5	6	9
New England	23	19	1.21	7	10	2
Philadelphia	23	21	1.1	5	8	8
Columbus	23	20	1.15	5	7	8
Chicago	20	19	1.05	3	5	11
Houston	19	20	0.95	5	11	4
Montreal	14	19	0.74	3	11	5
Western Conference	Points	Games Played	PPG	W	L	T
Seattle	38	18	2.11	12	4	2
Salt Lake	32	20	1.6	8	4	8
Colorado	30	20	1.5	8	6	6
Dallas	29	20	1.45	8	7	5
Los Angeles	27	17	1.59	7	4	6
Vancouver	27	19	1.42	6	4	9
Portland	24	20	1.2	5	6	9
Chivas	23	20	1.15	6	9	5
San Jose	20	18	1.11	5	8	5

Now, take a look at the data visualization created using the same data:



In addition to the table, you now have the ability to do the following in a single view:

- Filter between conference and league standings.
  - See how many games each team has played and how many games they have remaining.
  - Hover over every single game to see the teams that played and the final score.
  - Determine what it will take for each team to make the playoffs.

- Gain better understanding into how teams are performing relative to each other.

The same principles used to add value to this century-old story can be used to reinvigorate your reporting at work, engage your audience, and maximize the impact of your data visualization. I'm willing to bet that by retelling old stories, you will discover many new "Rags to Riches" storylines in your existing data!

## Tip #7: Don't Neglect the Setup

When you tell a story, you don't jump to the climax without first building up the characters and working your way through the plot. If you skipped the setup, the story would not be as effective as possible. Data visualization is the same way.

When you work closely with a dataset for a long time, the insights that emerge may begin to feel intuitive. It is natural to think that the audience for your data visualization will find the seemingly "obvious" stories in the data as easily as you. This may be true occasionally, but why not make it a guarantee by setting the user up for success by adding some context to your work? This will also allow you an opportunity to guide the way your data visualization is consumed and ensure that you and your audience will align on those aforementioned "obvious" stories in your data.

If you don't believe your data visualization will be interpreted differently by each stakeholder, you have likely never seen an Iron Viz Championship. I am always amazed by the diversity of stories that emerge from the same dataset each and every time the competition is held. The designers not only find unique stories in the data, but they communicate the stories using varying chart types, fonts, colors, and *context*.

Here are just a few tips on setting up your data visualization:

- *Always include a title.* This helps set an expectation for what your work is about and can provide valuable information to the end user such as data sources and the date range covered.
- *Ideally, you can ask a single question to open your data visualization.* Each item in your dashboard will then ladder up to answering that single question. Often-times, one dataset can be used to answer many different questions, but by stating up front what question you are focusing on answering, you guide how your audience will consume your data visualization.

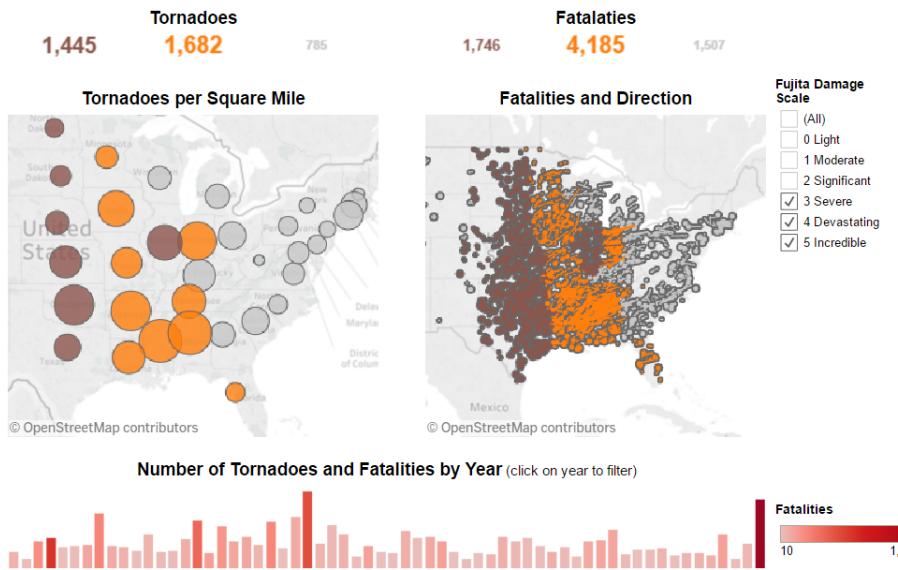
- Explain the features available in your dashboard up front. For example, if your visualization is interactive, explain the filtering options available.

Here is an example of a setup I like from Anya A'Hearn in her 2012 Iron Viz winner, *Does Tornado Alley Deserve Its Moniker?* Notice that the viz asks one question, and the rest of the viz attempts to answer the question. I also like how she saved real estate by building her color legend into the setup. This also helps the end user understand how to read the dashboard right from the start:

## Does "Tornado Alley" Deserve its Moniker?

**Severe Storms in Tornado Alley, Dixie Alley, and Other.**

First Dorothy was whisked away from Kansas, and then in 1952, the term "Tornado Alley" was coined by U.S. Air Force meteorologists as the title of a research project to study severe weather in parts of Texas and Oklahoma.



By not neglecting the setup, you help align your message with the audience's expectations by guiding the story that the data is telling.

# Tip #8: Don't Use Pie Charts

“The only worse design than a pie chart is several of them.”—Edward Tufte

Pie charts have become one of the most recognizable and widely adopted chart types in business and data journalism. Being that the message that pie charts convey is almost universally understood, it is tempting to incorporate pie charts into your own data visualization. However—data visualizer beware—while the message that pie charts convey is immediately recognized (parts of a whole), *the insights that pie charts provide are not*. Using pie charts works against you and your end users because when compared to other visualization choices, they make it more challenging to glean *accurate* stories from your visualization.

The pie chart’s primary limitation is that people are much better at comparing lengths and heights, as you would see in a bar chart or line graph, than they are at comparing areas within a pie. Further, the long-tail results, or the thinner pieces of a pie, tend to become unreadable. Pie charts are also a very poor vehicle for communicating changes over time.

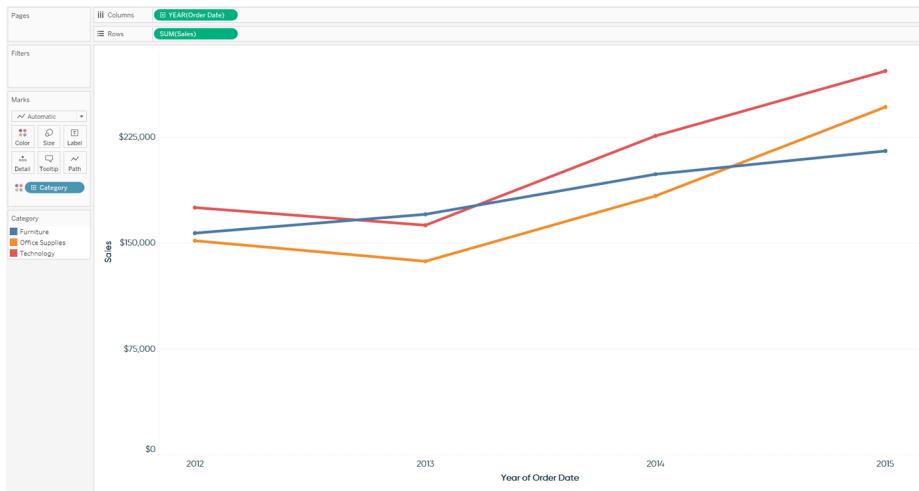
Despite the empirical evidence against pie charts, I continue to see them in some of the world’s most trusted news sources, stock reports, and corporate dashboards, among many other places. It is easy to make fun of people’s **use of pie charts**, but I want to be clear that when you hear people recommend avoiding pie charts, it is not some kind of elitist-data-visualization-specialist credence—it is a best practice based in evidence. I’ve always viewed the prolific use of pie charts as a huge opportunity to educate and improve the data visualization space. While graphical methods of statistics date back over 200 years, we still have a lot to improve—which is a great thing for you if you are reading this and are interested in evangelizing the benefits of data visualization.

To help, here is a simple exercise to illustrate the reduction in time to insight from a pie chart to more appropriate chart types.

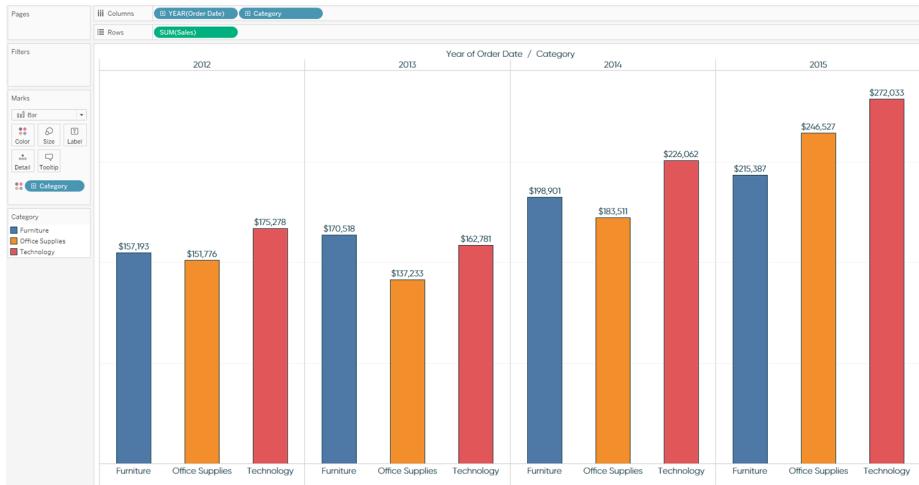
First, have a look at this series of pie charts, showing the sales by department over a four-year period. Take 15–30 seconds and think of a couple of insights that the charts provide:



You likely gathered that the entire pie was getting slightly larger over time (though it was hard to tell by how much and what that means), but it is pretty difficult to ascertain which pieces of the pie are causing the growth. With an easy tweak in Tableau, look at the exact same data as a line graph (take the same 15–30 seconds to understand what the data is telling you—though it likely won't take you nearly that long):

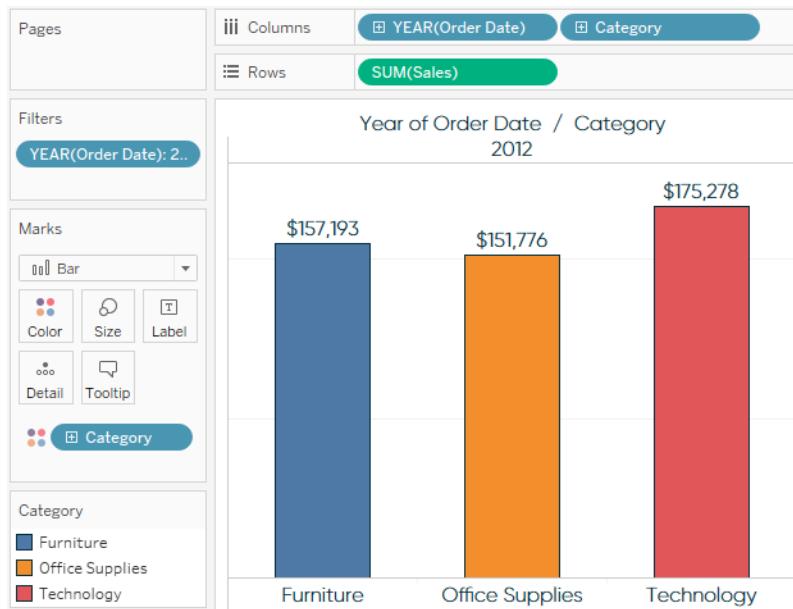


Much easier, right? By converting our pie chart to a line graph, we can easily see that all three categories are on the rise, and by how much. We were able to process the visual and get to the insight much faster. If you are not as interested in the performance over time, and want to see how each category performed compared to each other each year, use a bar chart:



Again, your brain processes heights of bars more easily than it processes areas in a pie. Using a bar chart also provides you more real estate to display the values, either as dollar amounts and/or percentages of the whole (as you were trying to convey by using a pie chart). To take this one final step further, if you only care about compar-

ing how each category performed against the others for one period in time, you can isolate the bar chart to show just the year in question:



I understand that you may have a stakeholder that has a hard time letting go of pie charts. If you or a boss can't quite go cold turkey, here are a couple of tips for using pie charts while you transition away from them:

- Stick to five slices or fewer, including the “other” category. As mentioned previously, the thinner, long-tail slices become unreadable. If you find yourself saying, “but I have to represent all fifteen of my categories,” that is another vote for you moving away from pie charts.
- Only use pie charts to show comparisons for one point in time. Avoid using them in a time-series analysis as shown in the first example.

For more thoughts on the topic, review [Chapter 52](#).

# Tip #9: Provide Visual Context

Content is king, but context is God.

—Gary Vaynerchuk

Context is useful in data visualization because it helps you avoid the dreaded question, “So what?” As powerful and simple as they are, fundamental chart types such as bar charts and line graphs are often used in a vacuum. Lack of context makes it difficult for an end user to understand, engage, and eventually *take action*. By building context into the visuals themselves, you help the stories in the data emerge, even if you aren’t there to explain them.

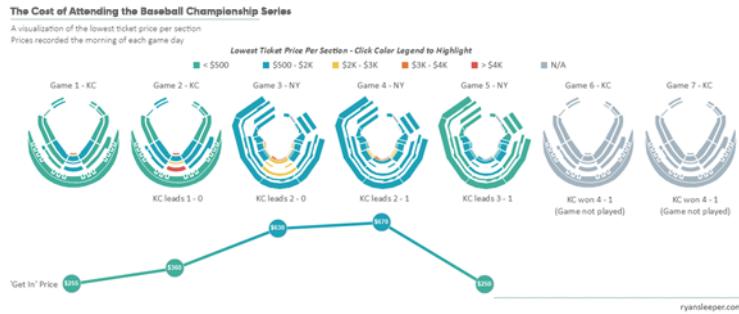
For this reason, context helps us tell both types of stories discussed in the introduction to this section. If you as the visualization creator know the endgame, you can use context to make your story stand out and “stick.” If you are building a self-service interactive data visualization, adding in context will help your end users analyze the data and find the storylines that are most relevant to them.

In [Chapter 92](#), I shared some recommendations for what to provide at the top of a dashboard. There are several ways to add context to the rest of the dashboard including:

- Index scores (covered in [Chapter 61](#)); not only do these scores provide comparisons, they help normalize the data.
- Sparklines (covered in [Chapter 24](#)); this chart type is essentially a small set of line graphs that allow you to view high-level trends across several metrics at the same time.
- Bullet graphs (covered in [Chapter 26](#)); this chart type is a variation on the bar chart that adds context in the form of comparison points such as prior period performance and/or goals.

- Small multiples (covered in [Chapter 25](#)); this chart type is similar to a table of visualizations instead of raw numbers, which allows the end user to compare results across several different dimensions and/or measures at the same time.

The Cost of Attending the Baseball Championship Series is an example of small multiples providing context that led to remarkable insight:



# Tip #10: Use Callout Numbers

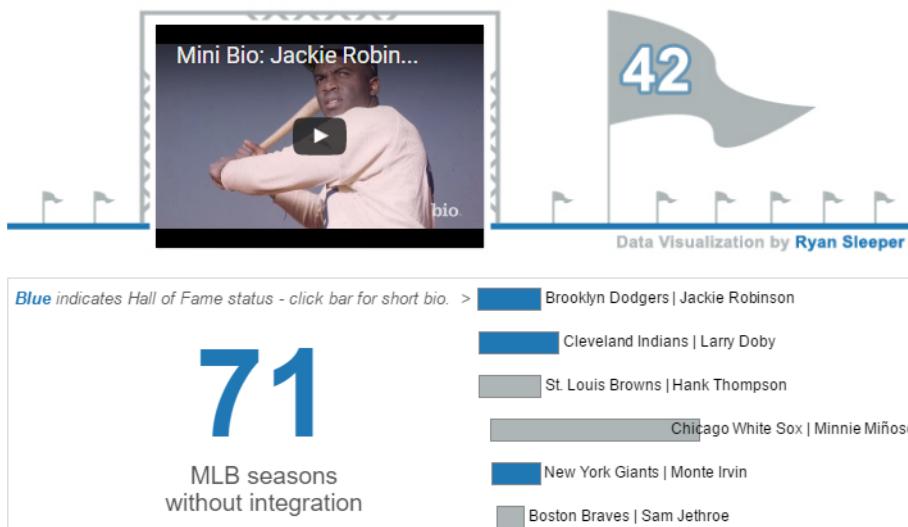
Callout numbers are one of the easiest ways to communicate what is most important to you in your data-driven story. A Google Images search for “callout” provides several examples of speech bubbles, thought bubbles, and Batman-style “BAM”/“POW” graphics. While I typically do not enclose my callout numbers in such an illustration, they share the purpose of explicitly telling the main story in your data visualization.

Callout numbers, as I’m calling them, are simply oversized numbers that should be in a legible font that is easy to consume. As these callout numbers will be communicating the most important numbers in your view, they should be prioritized near the top and left of your dashboard. For more on prioritizing and laying out content, review [Chapter 90](#).

In Tip #4 ([Chapter 89](#)), I discussed keeping your data visualizations simple. This is a critical concept if you want to cut through the noise and be as effective as possible at communicating the story in your data. Callout numbers provide a means for cutting through the noise, even within your own views.

Here is an example from my portfolio using callout numbers. This viz includes two numbers that stand out: 42 (Jackie Robinson’s number that is retired league-wide), and 71 (the number of seasons without integration in MLB). The end users can see instantly what the dashboard is about and why I think the content is important. From there, they can decide whether or not they want to explore the data:

## The Integration of Major League Baseball by Team



If you use callout numbers in your view, they will likely be the first thing that your end users are drawn to. This serves two important purposes: (a) it instantly tells the main point of your dashboard and (b) gives end users a natural place to decide whether or not they need to continue searching for the answers or context they are looking for. This could come into play with C-level executives, who may only want a top-level data point before leaving your visualization—and this is not necessarily bad.

I compare this behavior to a website's bounce rate. There is a common misconception that it is *always* bad to have a high bounce rate (or rate of people that view only one page on your website before leaving). However, a perfectly optimized web page that provides exactly what visitors are looking for without making them search for it on a second page will have a very high bounce rate. In the corporate world, you may have numbers that do not fluctuate very often, and callout numbers will be a valuable tool for either providing what users are looking for or communicating that it's important for them to investigate further.

# Tip #11: Allow Discovery

Tell me and I forget. Teach me and I remember. Involve me and I learn.

—Benjamin Franklin

I've always loved this quote from Benjamin Franklin, and especially in the context of data visualization, because I think it applies to the practice in a couple of ways. First, simply telling a stakeholder your results is the least effective way to get your data-driven story to stick. By its nature, visualizing data provides additional context beyond verbally sharing findings, so practitioners are already a step ahead in communicating actionable insights. Second, tools like Tableau allow data designers/artists/developers to build in interactivity that lets end users find stories in the data on their own. When end users are involved, whether it be in the iterative process of a dashboard design or in interacting with a dashboard, the shared sense of ownership goes a long way toward making your data visualization a success. When end users find an insight on their own, they are more likely to remember it, and what's better, *do something about it*.

Allowing discovery is a tip I take personally because, as a huge sports fan in a relatively small market (Kansas City), my teams do not get as much coverage as some of the teams in larger cities on the coasts. With tools like Tableau, there is no longer an excuse to not include relevant information for every fan. This is something I keep in mind every time I create a new data visualization. This same principle applies in a corporate setting, as you now have the capability to provide relevant information to a variety of stakeholders in the same amount of space. For example, you can provide filters that change the view based on product categories or regions. You may have data size considerations, but it is theoretically possible to allow end users to look at every order or every customer individually.

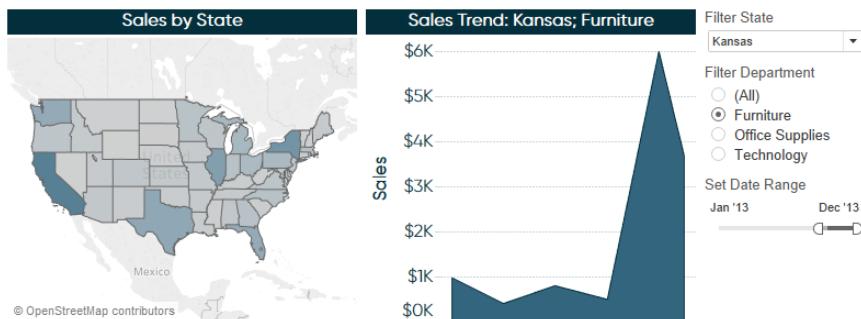
In Tableau, there are three easy-to-implement ways to allow discovery, which gives you the power to make your end user part of the story:

## *Filters*

We'll start with the most obvious tool for allowing discovery: filters. Any filter used in the making of a view can be added to a dashboard. Filters can be shown in two ways:

- From within a sheet view, right-click a filter on the Filters Shelf and select Show Filter. When the sheet is added to a dashboard, the filter will appear with the sheet.
- From within a dashboard view, (a) click the down arrow that appears when you hover over a sheet, (b) hover over Filters, and (c) make the appropriate selection.

What filters look like:

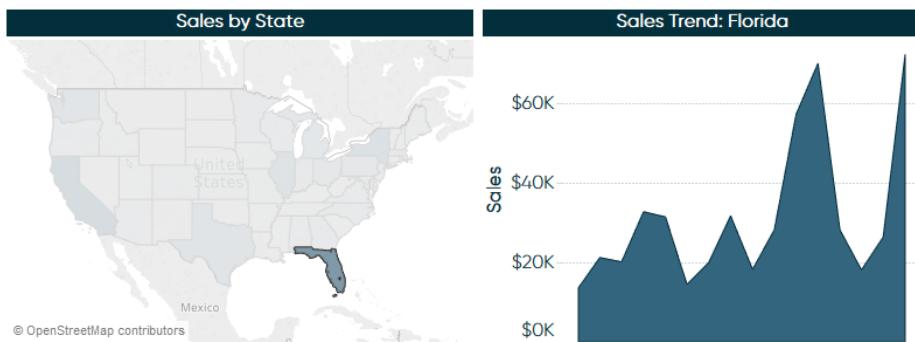


See [Chapter 11](#) for a review on filters.

## *Dashboard actions*

Dashboard actions are a more subtle way to add interactivity to a dashboard. They also have the added benefit of saving processing time if your Tableau workbooks are published to Tableau Server, Tableau Online, or Tableau Public. Dashboard actions are easy to create, but there are many options on how they can be utilized. The easiest way to implement dashboard actions is to click the down arrow that appears when you hover over a sheet and select "Use as Filter." This will create a simple dashboard action behind the scenes that will filter your entire dashboard based on the item clicked in the sheet that you opted to use as a filter.

What dashboard actions look like (clicking Florida filters the area chart on the right):



See [Chapter 56](#) for a review on dashboard actions.

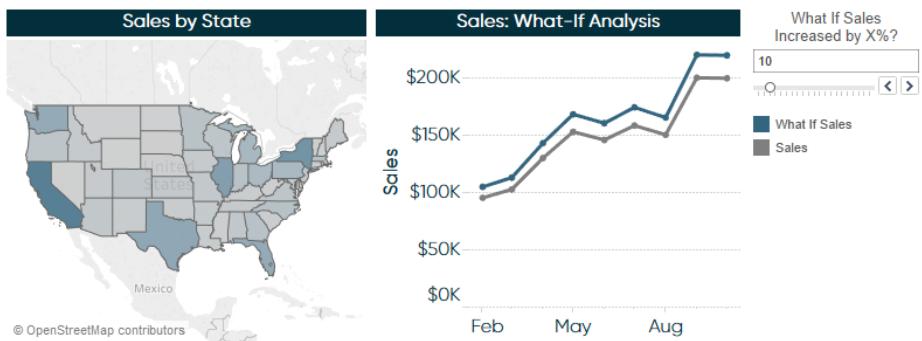
Many people do not realize that web-based Tableau workbooks are interactive. This will improve as more and more people are exposed to the functionality of the software and end users become comfortable exploring a dashboard on their own. In the meantime, you may want to include a brief instruction somewhere on your dashboard such as, “Click to filter by the element selected.”

You may even want to consider using hover actions so end users that are new to the software may stumble into the interaction capabilities that you’ve built in. For example, they may accidentally hover over the map in your dashboard, causing a change that makes them realize how to use your data visualization. This option is set when you build a dashboard action; instead of having the action execute on select (or click), choose to have the action execute on hover.

#### *Parameters*

Parameters are user-generated values that can be used in calculations. Another advantage of using parameters is that you can provide user-friendly names to the values, making it easier for your end users to understand what each selection will do. By showing the inputs for the parameters on your dashboards, end users can experiment with different scenarios on their own.

What parameters look like:



See [Chapter 14](#) for a review on parameters.

Parameters are a slightly more complex way of allowing discovery. To learn how to build the what-if analysis pictured, see [Chapter 48](#).

## Tip #12: Balance Data and Design

Data should always be the heart of any data visualization. Obviously, you cannot have a *data* visualization without data—and it is imperative that the data is accurate and honest. If I had to choose one element of a new data visualization to focus on between its data or design, I would make sure the data is right first every single time.

That being said, without making some design considerations, your data visualization is doomed to falling short of its full potential. Balancing data and design is another tip that I take personally because I often hear criticism from data purists that do not see the value in complementing the data with an appealing design. To those I ask:

*What is the purpose of data visualization?*

My answer is that the purpose of data visualization is to find and share actionable stories that are based in quantitative evidence. To make your insights actionable, this means sharing them with the most relevant, and many times, largest audience possible. If you don't provide your data stories in a well-packaged design, you drastically minimize the chances of your work spreading, and thus, it is consumed and acted on by fewer people.

See [Chapter 54](#) for some tactics for improving your balance between data and design—even if you don't have any design experience. This chapter also shows a before and after to help illustrate the value this balance can provide.

These tips will ensure your dashboard has its best chance at being shared and causing action.



# Tip #13: Eliminate Chartjunk (But Not Graphics)

Chartjunk is a term coined by Edward Tufte in his book, *The Visual Display of Quantitative Information*. In it, Tufte defines chartjunk as “non-data-ink or redundant data-ink.” My interpretation of chartjunk is that it is any design element that is counterproductive, actually detracting from a data visualization rather than adding value to it. Chartjunk can be as brash as charts that are made to fit within a graphic, with the data almost a complementary element to an illustration (see the [work of Nigel Holmes](#)). It can also be as subtle as a three-dimensional bar or extra gridlines. Calling nonessential elements “chartjunk” is a fair criticism, but the same person who invented the term also said the following:

“Only a picture can carry such a volume of data in such a small space.”—Edward Tufte

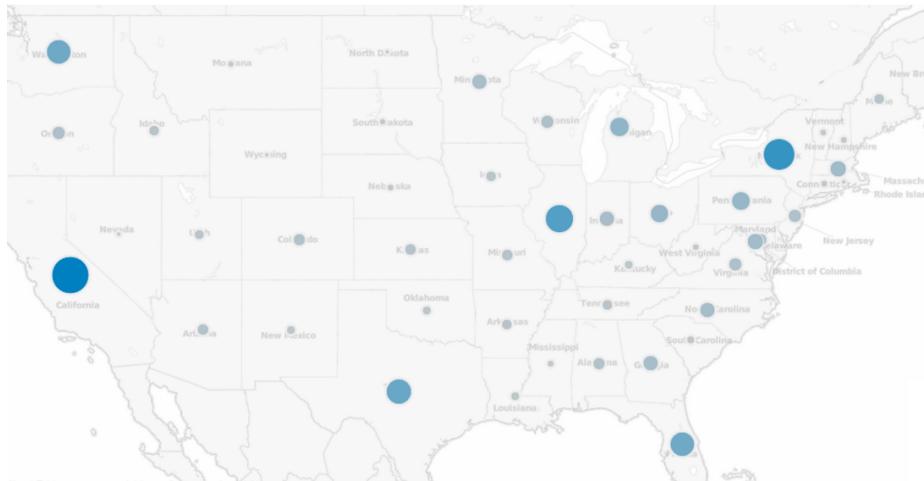
As with most debates, the truth probably lies somewhere in the middle. When the design becomes the primary purpose of a data visualization, it can become distracting and even misleading. Neglect design completely and a data visualization can be unengaging. In the last tip, I explained the importance of balancing data and design in data visualization. The tasteful use of graphics is one way to do that.

Chartjunk is not always composed of graphics. There are several examples of “non-data-ink” outside of graphics, including extra gridlines, shading, and three-dimensional marks. In the same way, graphics do not always create chartjunk. In fact, graphics are one of the most powerful ways to communicate information in a small space, and they also help engage an audience.

Consider the following “symbol map” showing sales by US state. This is technically a scatter plot with a circle on each combination of latitude and longitude. The circle is sized and colored by the sales amount. Here is the scatter plot with no image:



Now see what happens when we add a background image to the exact same scatter plot:



The map graphic helps you decode dozens of latitude/longitude pairs almost instantly, leveraging schemas for state locations that you have been building up in your mind over many years. The map also provides geospatial context, which can lead to additional insight that you may not have uncovered looking at a crosstab or