

1.



Overview

Now that you have been introduced to decision trees, you can pause for a moment and think about what you are learning. In this self-reflection, you will review a scenario, consider your thoughts about choosing a data visualization, and respond to brief questions.

This self-reflection will help you develop insights into your own learning and prepare you to make a decision tree for your next visualization. As you answer questions—and come up with questions of your own—you will consider concepts, practices, and principles to help refine your understanding and reinforce your learning. You've done the hard work, so make sure to get the most out of it: This reflection will help your knowledge stick!

Use a decision tree to choose a visualization

Choosing a data visualization to communicate your message may seem more like an art than a science. Fortunately, a decision tree can help you select the best visualization. To construct a decision tree, you simply need a basic understanding of design principles and the necessary knowledge of your data, message, and audience. This process will help add some structure to your decision making process, making it easier to pick the right data visualization.

First, take a moment to review some of the design concepts and best practices you've learned so far.

Create successful data visualizations

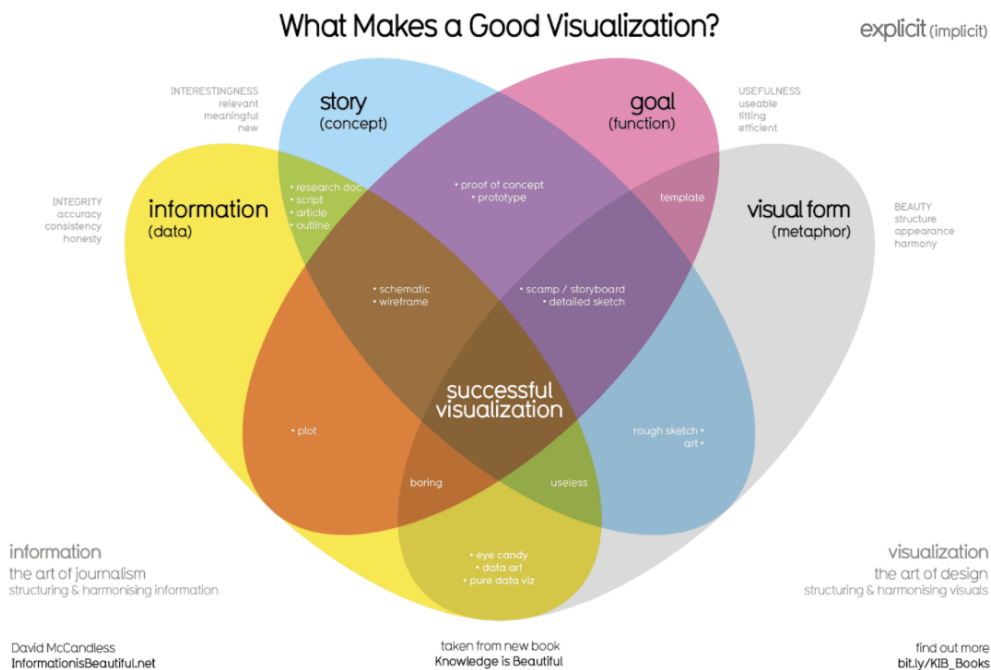
Recall that there are four elements of successful data visualization:

Information: reflects the *conclusion* you've drawn from the data, which you will communicate with visualization

Story: adds *meaning* to the data and makes it *interesting*

Goals: makes the *data* usable and *useful*

Visual form: creates both *beauty* and *structure*



Keep these elements in mind as you review the scenario below. This will help you make better data visualizations by helping you connect the information you want to communicate with your audience and your goals.

Here are additional best practices that are helpful to keep in mind:

Your audience should know what they are observing within five seconds of being shown a data visualization. Visuals should be clear and easy to follow.

In the five seconds after that, your audience should understand the conclusion your visualization is making—even if they aren't familiar with your research.

As long as it's not misleading, you should visually represent only the data that your audience needs in order to understand your findings. Including irrelevant data may confuse, distract, or overwhelm your audience.

These rules will guide you as you create your visualizations and you can apply them as you consider the following scenario.

The scenario

Imagine you are a junior data analyst at a local company. In your current data analysis project, you've been exploring sales data for all of your company's products, including the top products and sales trends for the last year. You need to present the results of this analysis to the company executives. Your goal for this presentation is to demonstrate how sales of the company's products have changed over the last 12 months. Your findings about the two top-selling products include:

Product name	Units sold	Time period with highest sales
Product A	2.5 million	70% of total annual sales occur in October, November, and December
Product B	1.9 million	Mostly consistent sales year-round, with a slight increase in November, December, and January

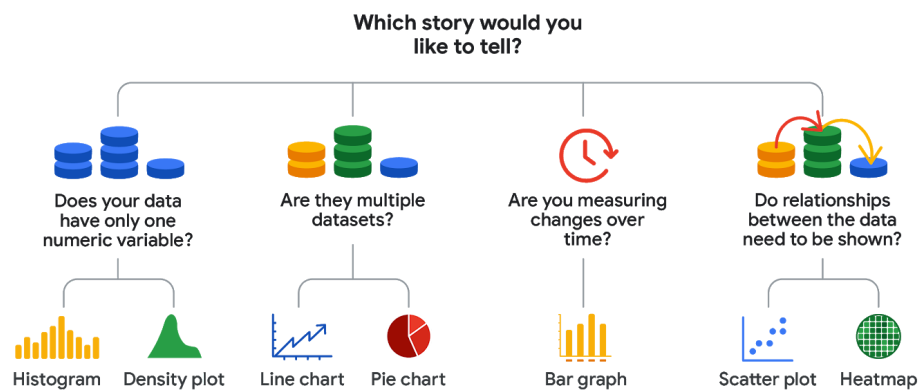
Your audience is made up of the chief marketing officer (CMO) and marketing department vice presidents (VPs), not other data analysts or engineers. The audience will use the information you present to make decisions on how to allocate the advertising budget for each product for the coming year.

Use a decision tree

A decision tree is a flowchart that you can use to help frame larger decisions as a series of smaller yes/no decisions. These are very useful when trying to choose the best data visualization to communicate a given message to your audience. Since different visualizations have different strengths and weaknesses, a decision tree can help you pick the best visualization for your given information and audience.

Observe the following decision tree. Consider what you know about the type of data you have and the kind of relationships you are trying to communicate. Use this to navigate the flowchart and determine what kind of visualizations are most appropriate.

Decision tree



Based on how you want to represent the sales data, there are several good choices here. You could frame this by using a comparison of different categories using a bar graph, or by showing how the composition of total sales changes month-to-month as a line chart with lines for each product.

Either of these would be a good choice. Combined with effective use of design principles, either could accurately communicate the message you need to convey to company leadership. This message being: sales for Product A are extremely seasonal, so they may want to consider reducing their advertising spend for this product when it's out of season.

Reflection

Consider the scenario and the decision tree you used in this reflection:

What kind of graph did you consider using before consulting the decision tree? What was your reasoning?

How did a decision tree simplify your decision-making process? Did the result of the decision tree differ from your first instinct?