**UNDERSTAND THE POWER OF DATA**

**DATA TRIALS AND TRIUMPHS**

**Data** is a collection of facts.

**Data analysis** reveals important patterns and insights about that data and it can help us make more informed decisions.

An **algorithm** is a process or set of rules to be followed for a specific task.

Businesses and other organizations use data to make better decisions all the time. There's two ways they can do this:

**1-** **Data-driven** decision-making**(DDDM)** means using facts to guide business strategy.The phrase “data-driven decisions” means exactly that: Data is used to arrive at a decision.**This approach is limited by the quantity and quality of readily-available data**. If the quality and quantity of the data is sufficient, this approach can far improve decision-making. But if the data is insufficient or biased, this can create problems for decision-makers.

**Potential dangers of relying entirely on data-driven decision-making** can include overreliance on historical data, a tendency to ignore qualitative insights, and potential biases in data collection and analysis

**Example:** 2 widgets websites, collected data from both and had the final say based on their performance.

**2- Data-inspired** decision-making**(DIDM)** (it explores different data sources to find out what they have in common) includes the same considerations as data-driven decisions **while adding another layer of complexity.** They create space for people using data to consider a broader range of ideas: **drawing on comparisons to related concepts, giving weight to feelings and experiences, and considering other qualities that may be more difficult to measure**.

**Data-inspired decision-making can avoid some of the pitfalls that data-driven decisions might be prone to**.

**Example:** Customer support center makes a survey with scale from 1-10 with additional comments and interview people working in the customer support center. Then the manager formulates a strategy to improve customer satisfaction based on the data from the survey.

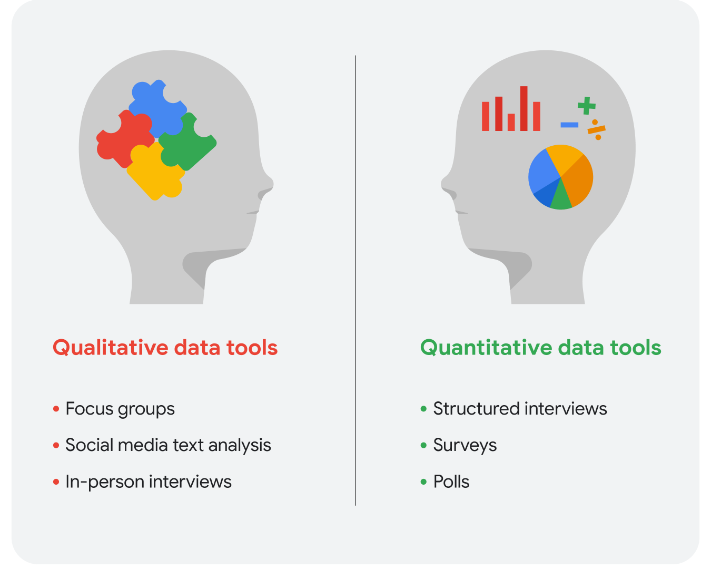
## **Key takeaways**

As a data analyst, you’ll rarely need to consider, “Am I being data-driven or data-inspired?” It’s helpful to have some context for these two approaches, though your own skills and knowledge will be the most important parts of any analysis project. So, keep a data-driven mindset and ask lots of questions. Experiment with many different possibilities. And use both logic and creativity along the way. Using this approach, you’ll be prepared to interpret your data with the highest levels of care and accuracy.

**-----------------------------------------------------------------------------------**

**QUALITATIVE AND QUANTITATIVE DATA**

**Quantitative data** is all about the specific and objective measures of numerical facts.This can often be the **what, how many,** and **how often** about a problem. In other words, things you can measure, like how many commuters take the train to work every week. We can see numbers visualized as charts and graphs.

**Qualitative data** describes subjective or explanatory measures of qualities and characteristics or **things that can't be measured with numerical data**, like your hair color. Qualitative data can give us a more high-level understanding of why the numbers are the way they are, it is important because it helps us to **add context** to a problem.

## 

## **Qualitative data for all three trends plus ticket pricing example:** Since you know that the theater is planning to raise ticket prices for evening showtimes in a few months, you will also include a question in the survey to get an idea of customers’ price sensitivity.

Your final online survey might include these questions for qualitative data:

1. What went into your decision to see a movie in our theater today? (movie attendance)
2. What do you think about the quality and value of your purchases at the concession stand? (concession stand profitability)
3. Which showtime do you prefer, 8:00 PM or 8:30 PM, and why do you prefer that time? (evening movie-goer preferences)
4. Under what circumstances would you choose a matinee over a nighttime showing? (ticket price increase)

## **Key takeaways**

Data analysts will generally use both types of data in their work. Usually, qualitative data can help analysts better understand their quantitative data by providing a reason or more thorough explanation. In other words, **quantitative data generally gives you the what, and qualitative data generally gives you the why**. By using both quantitative and qualitative data, you can learn when people like to go to the movies and why they chose the theater. Maybe they really like the reclining chairs, so your manager can purchase more recliners. Maybe the theater is the only one that serves root beer. Maybe a later show time gives them more time to drive to the theater from where popular restaurants are located. Maybe they go to matinees because they have kids and want to save money. You wouldn’t have discovered this information by analyzing only the quantitative data for attendance, profit, and showtimes.

**-----------------------------------------------------------------------------------**

**FOLLOW THE EVIDENCE**

**THE BIG REVEAL: SHARING YOUR FINDINGS**

Data is great, but if we can't communicate the story data is telling, it isn't useful to anyone. **We** **need ways to organize data that help us turn it into information**. There are all kinds of tools out there to help you visualize and share your data analysis with stakeholders.

PRESENTATION TOOLS

**REPORTS**

**PROS:**

a **static collection of data** (because of that they reflect data that's already been cleaned and sorted) given to stakeholders periodically. Reports are **great for giving snapshots of high level historical data for an organization**. They're **quick to design and easy to use** as long as you continually maintain them.

**CONS:**

They need regular maintenance and aren't very visually appealing. Because they aren't automatic or dynamic, reports don't show live, evolving data.

**DASHBOARD**

**PROS:**

For a **live reflection of incoming data**, you'll want to design a dashboard, they **give your team more access to information being recorded**, you can **interact through data** by playing with filters, and because **they're dynamic**, **they have long-term value**. If stakeholders need to continually access information, a dashboard can be more efficient than having to pull reports over and over, which is a big time saver for you.

**CONS:**

They **take a lot of time to design and can actually be less efficient than reports**, **if they're not used very often**. If the base table breaks at any point, they **need a lot of maintenance to get back up and running again**. Dashboards can sometimes **overwhelm people with information** too.

As a data analyst, you need to decide the best way to communicate information to your stakeholders.

**PIVOT TABLE**

A pivot table is a data summarization tool that is used in data processing.Pivot tables are used to summarize, sort, re-organize, group, count, total, or average data stored in a database. It allows its users to transform columns into rows and rows into columns.

**-----------------------------------------------------------------------------------**

**DATA VERSUS METRICS**

That specific measurement of data is done using metrics. Now, I want to tell you a little bit more about the **difference between data and metrics**. And how metrics can be used to turn data into useful information.

**METRIC**

is a **single, quantifiable type of data that can be used for measurement**. Think of it this way. Data starts as a collection of raw facts, until we organize them into individual metrics that represent a single type of data. Choosing the right metric is key.

**-----------------------------------------------------------------------------------**

# 

# **Tools for visualizing data**

In this course, you’ll work with Tableau and spreadsheets. Both of these tools have advantages and disadvantages Often, data analysts will discover they need to use multiple tools, even on a single project. What you use will largely be determined by the work you’re doing and your goals. This reading explores two of the tools you might use to visualize and present data: spreadsheets and Tableau.

## 

## 

## 

## 

## 

## 

## 

## **SPREADSHEETS**

****

Google Workspace and Microsoft Office Suite both offer spreadsheet applications. You’ve worked with Google Sheets in this course, and it’s very similar in function to Microsoft Excel. If you want to compare some of the features of Sheets to Excel, check out the Microsoft video [Create a chart from start to finish](https://support.microsoft.com/en-us/office/create-a-chart-from-start-to-finish-0baf399e-dd61-4e18-8a73-b3fd5d5680c2).

Both Sheets and Excel are go-to choices for creating static charts and graphs. They offer basic data visualization capabilities that are often enough for simple visualizations. In addition, you can use them to clean, sort, and filter data. And both offer a range of chart types, graphing tools, and pivot tables for creating effective data visualizations. These charts are easy to manage; they update when the source data is updated, so they don’t require much manual intervention once implemented.

Sheets and Excel are connected to other apps in their product suites. Google Docs and Slides are very similar to Microsoft Word and Powerpoint, for example. You can incorporate data visualizations from Sheets or Excel into reports and documents in Docs and Word. Presentation programs such as Slides and Powerpoint allow you to create engaging presentations that include data visualizations so you can share insights in a presentation format. Learn more about the power of this interconnectivity among Google tools in the article [Link a chart, table, or slides to Google Docs or Slides](https://support.google.com/docs/answer/7009814?hl=en&co=GENIE.Platform%3DDesktop).

## **TABLEAU**

Tableau is used to create powerful and interactive visualizations, making it an excellent choice for data visualizations such as live dashboards. Tableau also makes it easy to create charts, graphs, and dashboards in a drag-and-drop interface. The application supports a wide range of data sources and provides advanced analytics capabilities. These features allow for in-depth exploration of data trends and patterns.

Tableau is particularly useful for creating visualizations using huge datasets, like in this [World Happiness Report](https://www.kaggle.com/datasets/unsdsn/world-happiness) by Sustainable Development Solutions which uses global reporting data on different countries' happiness ratings. Likewise, this visualization of [Population and Housing State Data](https://www.census.gov/library/visualizations/interactive/2020-population-and-housing-state-data.html) from 2020 United States Census Data compares population rates in the United States and available housing.

Tableau is widely known and used for its versatility and power, but it can take quite a bit of time to learn to use Tableau effectively. Soon, you’ll begin practicing with Tableau. But if you’d like to check it out now, there is a free environment you can access at [Tableau Public](https://www.tableau.com/products/public).

## **Key takeaways**

There are many visualization tools you will have the opportunity to use as a data professional. Different tools have different advantages and disadvantages. Although Tableau ultimately has more power than a basic spreadsheet application, it’s most often used for specific cases and to work with large datasets. Don’t underestimate how much you can do with spreadsheets or how powerful interconnectivity between apps can be!

Most of the time, especially for something like a quick report, you’re more likely to reach into your toolkit for your spreadsheet app of choice. But your data career will definitely benefit from Tableau, so as you progress take advantage of opportunities to work with it. With so many different data analysis situations, familiarity with all of these tools will help you know which is the best for each situation.

**-----------------------------------------------------------------------------------**

# 

# 

# 

# 

# 

# 

# 

# 

# 

# 

# **DESIGN COMPELLING DASHBOARDS**

Dashboards are powerful visual tools that help you tell your data story. A dashboard is a tool that monitors live, incoming data. It organizes information from multiple datasets into one central location, offering huge time savings. Data analysts use dashboards to track, analyze, and visualize data in order to answer questions and solve problems. For a basic idea of what dashboards look like, refer to this article: “[Real-world examples of business intelligence dashboards.](https://www.tableau.com/learn/articles/business-intelligence-dashboards-examples)”

## 

## **The beauty of dashboards** The following table summarizes the benefits of using a dashboard for both data analysts and their stakeholders.

| **Benefits** | **For data analysts** | **For stakeholders** |
| --- | --- | --- |
| **Centralization** | Share a single source of data with all stakeholders | Work with a comprehensive view of data, initiatives, objectives, projects, processes, and more |
| **Visualization** | Show and update live, incoming data in real time\* | Spot changing trends and patterns more quickly |
| **Insightfulness** | Pull relevant information from different datasets | Understand the story behind the numbers to keep track of goals and make data-driven decisions |
| **Customization** | Create custom views dedicated to a specific person, project, or presentation of the data | Drill down to more specific areas of specialized interest or concern |

It’s important to remember that changed data is pulled into dashboards automatically only if the data structure is the same. If the data structure changes, you have to update the dashboard design before the data can update live.

## **TABLEAU**

There are many different visualization tools available. One of the most powerful is Tableau, which supports a range of data sources and has advanced analytics capabilities that allow for in-depth exploration of data trends and patterns. Tableau can handle more data and larger datasets than many other tools and offers real-time data availability.

It does take some time to learn to use Tableau, but your efforts can be well-rewarded, as Tableau visualizations are pleasantly interactive. For a dashboard to be successful, it needs to engage users and help them learn. Tableau has put in a lot of effort to ensure that its users have a great experience and the platform is accessible to everyone.

**CREATE A DASHBOARD**

Here’s a process you can follow to create a dashboard, whether in Tableau or another visualization tool:

### **1. Identify the stakeholders who need to see the data and how they will use it**

Begin by asking effective questions. Check out this [dashboard requirements gathering worksheet](https://s3.amazonaws.com/looker-elearning-resources/Requirements+Gathering+Worksheet.pdf) to explore a wide range of good questions you can use to identify relevant stakeholders and their data needs. This is a great resource to help guide you through this process again and again.

### **2. Design the dashboard (what should be displayed)**

Use these tips to help make your dashboard design clear and easy to follow:

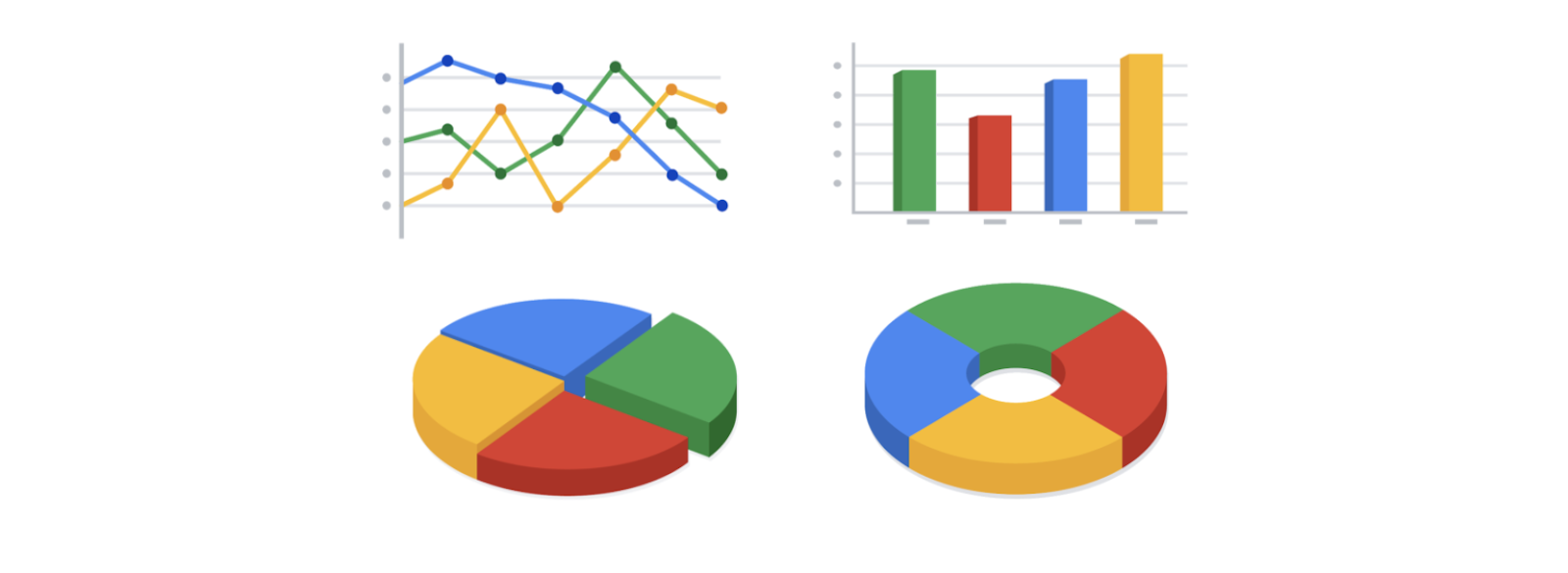
* Use a clear header to label the information.
* Add short text descriptions to each visualization.
* Show the most important information at the top.

### **3. Create mockups if desired**

A mockup is a simple draft of a visualization used for planning a dashboard and evaluating its progress. This is optional, but a lot of data analysts like to sketch out their dashboards before creating them.

### **4. Select the visualizations**

You have a lot of options here. Which visualizations you select depends on the data story you are telling. If you need to show a change in values over time, line charts or bar graphs might be the best choice. If your goal is to show how each part contributes to the whole amount being reported, a pie or donut chart is probably a better choice.



Two pie charts show an even distribution of 4 parts of a whole. The first pie chart is more traditional, appearing as a solid circle. The second pie chart is styled to show the same data in a doughnut shape.

To learn more about choosing the right visualizations, check out Tableau’s galleries:

* For more samples of area charts, column charts, and other visualizations, visit the [Tableau Dashboard Showcase](https://www.tableau.com/solutions/gallery). This gallery is full of great examples that were created using real data; explore this resource on your own to get some inspiration.
* Explore [Tableau’s Viz of the Day](https://public.tableau.com/en-us/gallery/?tab=viz-of-the-day&type=viz-of-the-day) to check out visualizations curated by the community. These are visualizations created by Tableau users and are a great way to learn more about how other data analysts are using data visualization tools.

### **5. Create filters as needed**

Filters show certain data while hiding the rest of the data in a dashboard. This can be a big help to identify patterns while keeping the original data intact. It’s common for data analysts to use and share the same dashboard, but manage their part of it with a filter. To dig deeper into filters and find an example of filters in action, visit Tableau’s page on [Filter Actions](https://help.tableau.com/current/pro/desktop/en-us/actions_filter.htm). This is a useful resource to save and come back to when you start practicing using filters in Tableau on your own.

## **Key takeaways**

Just like how the dashboard on an airplane shows the pilot their flight path, your dashboard does the same for your stakeholders. It helps them navigate the path of a project inside the data. If you add clear markers and highlight important points on your dashboard, users will understand where your data story is headed. Then, you can work together to make sure the business gets where it needs to go.

For a refresher, consider the different types of dashboards a business may use. Often, businesses will tailor a dashboard for a specific purpose. The three most common categories are:

* **Strategic:** focuses on long term goals and strategies at the highest level of metrics These dashboards provide information over the longest time frame—from a single financial quarter to years. They typically contain information that is useful for enterprise-wide decision-making.
* **Operational:** short-term performance tracking and intermediate goals, **the most common type of dashboard**. Because these dashboards contain information on a time scale of days, weeks, or months, they can provide performance insight almost in real-time. This allows businesses to track and maintain their immediate operational processes in light of their strategic goals.
* **Analytical:** consists of the datasets and the mathematics used in these sets, contain a vast amount of data used by data analysts. These dashboards contain the details involved in the usage, analysis, and predictions made by data scientists. Certainly **the most technical category**, analytic dashboards are usually created and maintained by data science teams and rarely shared with upper management as they can be very difficult to understand.

**-----------------------------------------------------------------------------------**

**CONNECT THE DATA DOTS  
MATHEMATICAL THINKING**

**Mathematical thinking** is a powerful skill you can use to help you solve problems and see new solutions. It’s looking at a problem and logically breaking it down step-by-step, so you can see the relationship of patterns in your data, and use that to analyze your problem. This kind of thinking can also help you figure out the best tools for analysis because it lets us see the different aspects of a problem and choose the best logical approach. There are a lot of factors to consider when choosing the most helpful tool for your analysis. One way you could **decide which tool to use is by the size of your dataset.** When working with data, you'll find that there's big and small data. Small data can be really small. These kinds of data tend to be made up of datasets concerned with specific metrics over a short, well defined period of time. Like how much water you drink in a day. Small data can be useful for making day-to-day decisions, like deciding to drink more water. But it doesn't have a huge impact on bigger frameworks like business operations. You might use spreadsheets to organize and analyze smaller datasets when you first start out. Big data on the other hand has larger, less specific datasets covering a longer period of time. They usually have to be broken down to be analyzed. Big data is useful for looking at large- scale questions and problems, and they help companies make big decisions.When you're working with data on this larger scale, you might switch to SQL.

**Practice Questions:**

* **Question 1:** What is the difference between small data and big data? Small data can be useful for making day-to-day decisions, it doesn’t have a huge impact on bigger frameworks like business operations. Big data has larger, less specific datasets covering a longer period of time, you usually will switch to SQL to deal with big data.
* **Question 2:** How can mathematical thinking help you choose the best tool for data analysis? Because it lets us see the different aspects of a problem and choose the best logical approach(One way you could decide which tool to use is by the size of your dataset).
* **Question 3:** How can SQL be used to analyze large datasets in a hospital setting? A tool that's capable of handling big datasets is a must, that’s when SQL shines.

**---------------------------------------------------------------------------------------------**

# 

# 

# 

# 

# **BIG AND SMALL DATA**

As a data analyst, you will work with data both big and small. Both kinds of data are valuable, but they play very different roles. 

Whether you work with big or small data, you can use it to help stakeholders improve business processes, answer questions, create new products, and much more. But there are certain challenges and benefits that come with big data and the following table explores the differences between big and small data.

| **Small data** | **Big data** |
| --- | --- |
| Describes a dataset made up of specific metrics over a short, well-defined time period | Describes large, less-specific datasets that cover a long time period |
| Usually organized and analyzed in spreadsheets | Usually kept in a database and queried |
| Likely to be used by small and midsize businesses | Likely to be used by large organizations |
| Simple to collect, store, manage, sort, and visually represent | Takes a lot of effort to collect, store, manage, sort, and visually represent |
| Usually already a manageable size for analysis | Usually needs to be broken into smaller pieces in order to be organized and analyzed effectively for decision-making |

## 

## **Challenges and benefits**

Here are some **challenges** you might face when working with big data:

* A lot of organizations deal with data overload and way too much unimportant or irrelevant information.
* Important data can be hidden deep down with all of the non-important data, which makes it harder to find and use. This can lead to slower and more inefficient decision-making time frames.
* The data you need isn’t always easily accessible.
* Current technology tools and solutions still struggle to provide measurable and reportable data. This can lead to unfair algorithmic bias.
* There are gaps in many big data business solutions.

Now for the good news! Here are some **benefits** that come with big data:

* When large amounts of data can be stored and analyzed, it can help companies identify more efficient ways of doing business and save a lot of time and money.
* Big data helps organizations spot the trends of customer buying patterns and satisfaction levels, which can help them create new products and solutions that will make customers happy.
* By analyzing big data, businesses get a much better understanding of current market conditions, which can help them stay ahead of the competition.
* As in our earlier social media example, big data helps companies keep track of their online presence—especially feedback, both good and bad, from customers. This gives them the information they need to improve and protect their brand.

**The three (or four) V words for big data**

When thinking about the benefits and challenges of big data, it helps to think about the three Vs: **volume, variety,** and **velocity.** Volume describes the amount of data. Variety describes the different kinds of data. Velocity describes how fast the data can be processed. Some data analysts also consider a fourth V: **veracity.** Veracity refers to the quality and reliability of the data. These are all important considerations related to processing huge, complex datasets.

| **Volume** | **Variety** | **Velocity** | **Veracity** |
| --- | --- | --- | --- |
| The amount of data | The different kinds of data | How fast the data can be processed | The quality and reliability of the data |

**---------------------------------------------------------------------------------------------**

## 

## **Terms and definitions for Course 2, Module 2**

**Algorithm:** A process or set of rules followed for a specific task

**Big data:** Large, complex datasets typically involving long periods of time, which enable data analysts to address far-reaching business problems

**Dashboard:** A tool that monitors live, incoming data

**Data-inspired decision-making:** The process of exploring different data sources to find out what they have in common

**Metric:** A single, quantifiable type of data that is used for measurement

**Metric goal:** A measurable goal set by a company and evaluated using metrics

**Pivot chart:** A chart created from the fields in a pivot table

**Pivot table:** A data summarization tool used to sort, reorganize, group, count, total, or average data

**Problem types:** The various problems that data analysts encounter, including categorizing things, discovering connections, finding patterns, identifying themes, making predictions, and spotting something unusual

**Qualitative data:** A subjective and explanatory measure of a quality or characteristic

**Quantitative data:** A specific and objective measure, such as a number, quantity, or range

**Report:** A static collection of data periodically given to stakeholders

**Return on investment (ROI):** A formula that uses the metrics of investment and profit to evaluate the success of an investment

**Revenue:** The total amount of income generated by the sale of goods or services

**Small data:** Small, specific data points typically involving a short period of time, which are useful for making day-to-day decisions

**---------------------------------------------------------------------------------------------**