Data analysts spend a lot of time organizing data and performing

calculations.

Luckily, there's lots of different tools to help them do just that,

including spreadsheets. In this video we'll take a look at some of the ways data

analysts use spreadsheets to help them with their day to day responsibilities.

Later, you'll get to test out some of these things yourself, but for now,

let's start with a quick look at how data analysts use spreadsheets to do

their jobs.

This will change depending on the work you need to complete.

But here's an overview of a few of the major tasks.

Imagine you work for a construction company.

Your company needs your spreadsheet skills to analyze some data about their expenses,

so you access the appropriate data and add it to your spreadsheet.

We won't cover all the details of this project right now, but

you will get a chance to see lots of spreadsheet features up close and

personal as we move forward.

What do you do with the data now that it's in your spreadsheet? Again,

this will be different for each job, but

you might start by organizing your data with the task you've been given.

For example, you might put your data in a pivot table.

We've talked about pivot tables before in this course. We'll cover

them in more detail later on, but for now, just think of them as

well organized and very useful tables.

Next, you might filter the data in the pivot table.

Sorting and filtering data is a common part of most jobs.

This lets you focus only on the data you'll need for your analysis.

In our example, maybe you only need the expenses for a certain time frame,

like the last three months. After you filtered your data,

you could perform some calculations to learn more about it.

Maybe you need to find out which construction projects ended up costing

the most money.

This is where formulas and functions are really handy.

We'll talk about them in just a bit, but formulas and functions are great for

doing some quick math, especially once you run out of fingers and toes to count on.

Now you've seen some of the ways data analysts are using spreadsheets in their

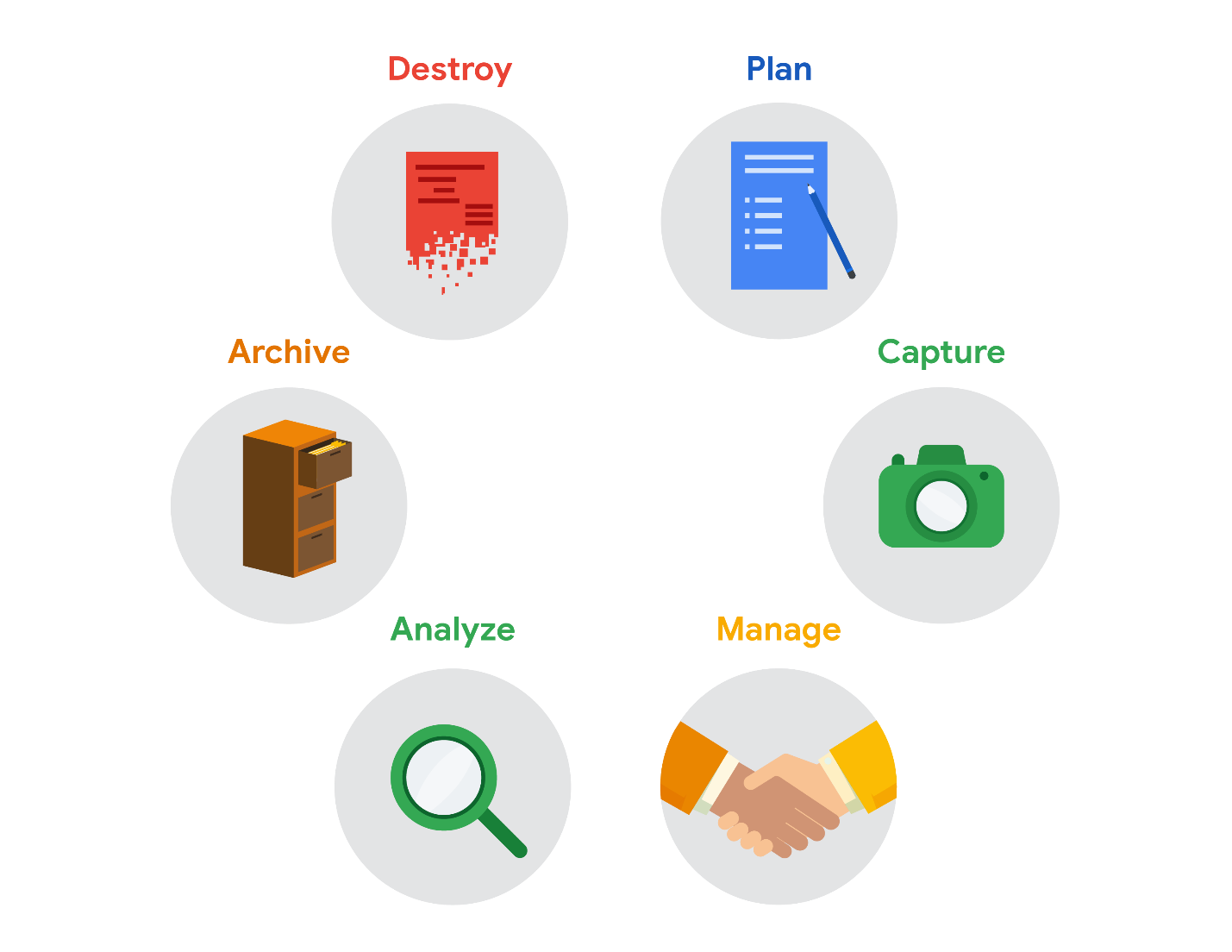
day to day work for a lot of different tasks,

including organizing their data and making calculations. Before you know it we'll

have you working in your own spreadsheets.

Spreadsheets and the data life cycle

To better understand the benefits of using spreadsheets in data analytics, let’s explore how they relate to each phase of the data life cycle: **plan, capture, manage, analyze, archive,** and **destroy**.



* **Plan** for the users who will work within a spreadsheet by developing organizational standards. This can mean formatting your cells, the headings you choose to highlight, the color scheme, and the way you order your data points. When you take the time to set these standards, you will improve communication, ensure consistency, and help people be more efficient with their time.
* **Capture** data by the source by connecting spreadsheets to other data sources, such as an online survey application or a database. This data will automatically be updated in the spreadsheet. That way, the information is always as current and accurate as possible.
* **Manage** different kinds of data with a spreadsheet. This can involve storing, organizing, filtering, and updating information. Spreadsheets also let you decide who can access the data, how the information is shared, and how to keep your data safe and secure.
* **Analyze** data in a spreadsheet to help make better decisions. Some of the most common spreadsheet analysis tools include formulas to aggregate data or create reports, and pivot tables for clear, easy-to-understand visuals.
* **Archive** any spreadsheet that you don’t use often, but might need to reference later with built-in tools. This is especially useful if you want to store historical data before it gets updated.
* **Destroy** your spreadsheet when you are certain that you will never need it again, if you have better backup copies, or for legal or security reasons. Keep in mind, lots of businesses are required to follow certain rules or have measures in place to make sure data is destroyed properly.

**Resources**

Spreadsheet shortcuts can help you become more efficient with spreadsheets. If you’d like to learn more, you can explore the collection of [Google Sheets shortcuts](https://support.google.com/docs/answer/181110), or visit the [Microsoft Excel shortcuts](https://support.microsoft.com/en-us/office/keyboard-shortcuts-in-excel-1798d9d5-842a-42b8-9c99-9b7213f0040f) page if you are using Excel. Both of these resources contain a list of spreadsheet shortcuts you can save and reference as you work more with spreadsheets on your own.

We've talked about how spreadsheets are great

for organizing data and performing calculations.

Now, it's time to get our hands

dirty and start building a real spreadsheet.

In this video, I'm going to demonstrate some basic tasks

we know data analysts use spreadsheets

for, including entering and organizing data.

We'll start with a step-by-step process to show

you some tools to organize your data in a spreadsheet.

Consider these steps the basics.

You won't always have to use

them when working with a data set,

but if your data is a bit messy when you get it,

these steps can help you get it ready for analysis.

Let's start by opening a new spreadsheet.

As a data analyst,

you might not start with a blank spreadsheet,

but it's good to know how to do it, just in case.

Start by opening Excel,

Google Sheets or whatever

spreadsheet software you're using,

then select a new blank file.

The first thing you'll want to do when you open

a new spreadsheet is give it a title. Here's a pro tip.

Make your title short, clear,

and have it state exactly what

the data in the spreadsheet is about.

Trust me, it'll make searching for it a lot easier.

Creating a folder on your computer specifically for

spreadsheets and related files

can also make it easier to find them.

For this spreadsheet, it's already saved in our drive.

So we'll open our File menu to click Move.

Play video starting at :1:26 and follow transcript1:26

Then we'll create a new folder,

name it "Population Data,"

Play video starting at :1:36 and follow transcript1:36

and move the spreadsheet there.

Play video starting at :1:41 and follow transcript1:41

Our spreadsheet now has a new home.

This will save you a lot of

unnecessary clicks and headaches

when you look for this file.

There's a few different ways data analysts

get data they work with.

Depending on the job,

you might use data from an open source,

you might be given data to work with

or you might be asked to find your own data.

You'll experience all of these later in the program.

There's a lot of open data sources online,

where data is made available to the public.

For example, we'll use data from worldbank.org,

that's already in the spreadsheet.

The data shows the population of

Latin American and Caribbean countries from 2010-2019.

Let's open this spreadsheet.

Time to get the data ready for analysis.

We'll start by selecting the whole sheet and

making our columns wider by

dragging the boundary of one of the columns.

Play video starting at :2:37 and follow transcript2:37

This will help us see the data clearly,

then we can adjust

any individual columns that need it.

You can make columns wider in other ways as well,

but this will work for now.

The first row of the spreadsheet is for

data attributes or variables.

It's basically labeling the type of data in each column.

Let's make the attributes stand out from the rest of

the rows by selecting it and filling it with color.

We'll also make the labels bold.

Play video starting at :3:10 and follow transcript3:10

If we want to add another data attribute

between two of the other attributes,

we can always add a new column.

Just click on any cell within

a column and use the Insert menu to add a new one.

Play video starting at :3:29 and follow transcript3:29

It will appear next to the column you

originally clicked, pretty simple.

Deleting a column is just as simple.

To delete, right-click in

a cell in the column you want to get rid of.

The steps we're showing may be different

depending on the spreadsheet program you're using,

but should be pretty similar.

Let's add one more thing to our data table: borders.

This can help you see each piece of data more clearly.

To add borders start by clicking the Select

All button at the top left corner of your spreadsheet.

This is like a magic button because you can click it

whenever you need to make

changes to every cell in your spreadsheet.

Then click the Border button in

the menu, and choose the type of borders you want.

To keep our spreadsheets uniform,

we'll choose borders for all cells.

Just like that, we've gone from raw to refined.

Now our spreadsheet is filled with

data and it's nice to look at too.

Using these organization tools before you analyze can

help you focus on the data once you start your analysis.

Now that we've gone over some ways

spreadsheets can be used to organize data,

you're ready to start working on them yourself.

Later you'll learn more about spreadsheets,

including some common errors and how to fix them.

Learn more about spreadsheet basics

Below, you will find a list that covers two types of spreadsheet programs: **Microsoft Excel** and **Google Sheets**. The list includes quick-start guides, tutorials, and more. The examples in this course use Google Sheets, but you can follow along using Excel or any other spreadsheet application. The user interface might be a little different, but it should look and work similarly.



**Microsoft Excel**

* [Office Quick Starts](https://support.microsoft.com/en-us/office/office-quick-starts-25f909da-3e76-443d-94f4-6cdf7dedc51e#ID0EAADAAA=At_work_or_school): Scroll down to the **Downloadable guides** section to download the **Excel Quick Start Guide**: This PDF guide begins with a labeled map of Excel that can guide you through the basic tasks you can accomplish in Excel. For tips on starting and opening Excel, this [Microsoft Support page](https://support.microsoft.com/en-us/office/create-a-new-workbook-ae99f19b-cecb-4aa0-92c8-7126d6212a83?wt.mc_id=otc_excel) will show you how to begin a new workbook.
* [Excel video training](https://support.microsoft.com/en-us/office/excel-for-windows-training-9bc05390-e94c-46af-a5b3-d7c22f6990bb?wt.mc_id=otc_home): This is a collection of step-by-step videos to use all sorts of Excel features, including adding and working within rows, columns, and cells; formatting; using formulas and functions; and adding charts and pivot tables.
* [Sort data in a range or table](https://support.microsoft.com/en-us/office/sort-data-in-a-range-or-table-62d0b95d-2a90-4610-a6ae-2e545c4a4654): This page guides you through all of the steps you will need to sort data by number, text, and color. You’ll also have the option to sort by custom list so that you can customize exactly what you want to sort.
* [Filter data in a range or table](https://support.microsoft.com/en-us/office/filter-data-in-a-range-or-table-01832226-31b5-4568-8806-38c37dcc180e): This article has step-by-step instructions on how to filter an Excel spreadsheet to show only the data you want to see. You can also use built-in comparison operators, such as “greater than” and “top 10” to reveal only the most relevant data.
* [Format a worksheet](https://support.microsoft.com/en-us/office/quick-start-format-a-worksheet-d70f75a2-23e6-4c92-83d6-2f219e4ad42e): The guide will help you select and format your Excel spreadsheet, then change the borders, shading, colors, and text. This can help improve your spreadsheet’s readability.

**Pro tip**: If you’re searching for information about using customizable options, check out Microsoft’s [Guidelines for organizing and formatting data on a worksheet](https://support.microsoft.com/en-us/office/guidelines-for-organizing-and-formatting-data-on-a-worksheet-90895cad-6c85-4e02-90d3-8798660166e3). This article provides clear methods for creating easy-to-read spreadsheets.

**Google Sheets**

* [Google Sheets cheat sheet](https://support.google.com/a/users/answer/9300022): The cheat sheet puts all the basics of Sheets on a single page for easy reference. Here, you can learn about customizing your spreadsheet and the data inside; working with rows, columns, and cells; sharing your spreadsheet with others; creating different versions and copies of a spreadsheet; and more.
* [Get started with Sheets: Create and import files](https://support.google.com/a/users/answer/9300311?hl=en&ref_topic=9296423): This guide is a step-by-step guide for working with Sheets. You start by learning how to open a spreadsheet, then move on to adding data.
* [Sort and filter your data](https://support.google.com/docs/answer/3540681?co=GENIE.Platform%3DDesktop&hl=en): This resource can help you organize data in Sheets. Use this guide to sort part or all of a spreadsheet. You can sort by text, number, and color. Then, learn how to create filters to show only certain data while hiding the rest. Finally, the article includes information on creating, saving, and removing a filter view.
* [Edit and format a spreadsheet](https://support.google.com/docs/answer/46973?co=GENIE.Platform%3DDesktop&hl=en&oco=0): This will help you make easy-to-read spreadsheets. You will learn how to assign a color, customize borders around cells, and change the appearance of text. If you’d like to give your spreadsheet a theme, you can scroll to the bottom of the page and find how to apply it to parts of your spreadsheet.

**Tip**: Microsoft Excel and Google Sheets are very similar in terms of calculations, formulas, functions, and many other features. But there are some differences, which can make it tricky to switch from one to the other. If you are moving between Excel and Google Sheets, find a quick list of the differences between the two kinds of spreadsheet applications in [Overview: Differences between Sheets and Excel](https://support.google.com/a/users/answer/9331278?hl=en).

## Transcript



Help Us Translate

### **Interactive Transcript - Enable basic transcript mode by pressing the escape key**

You may navigate through the transcript using tab. To save a note for a section of text press CTRL + S. To expand your selection you may use CTRL + arrow key. You may contract your selection using shift + CTRL + arrow key. For screen readers that are incompatible with using arrow keys for shortcuts, you can replace them with the H J K L keys. Some screen readers may require using CTRL in conjunction with the alt key

Play video starting at :: and follow transcript0:00

So far we've covered how to start a new spreadsheet,

enter in data, and make it look

refined and ready for some serious analysis.

Now we'll learn how to perform

calculations in your spreadsheet.

You may need to calculate everything

from sums to averages,

to finding minimum and maximum amounts.

You'll use calculations for

a lot of different kinds of tasks.

In this video, we'll focus on learning the basics

and then do a little math with

some sales data to practice.

Let's talk about formulas first.

You might remember that a formula is a set of

instructions that perform a specific calculation.

Basically, formulas can do the math for you.

Now, they don't only do math,

they can do a lot more.

Soon you'll learn different ways you can use

them throughout the data analysis processes.

Formulas are built on operators which are symbols that

name the type of operation or

calculation to be performed.

For example, a plus sign is a common operator.

The formulas you use as a data analyst will

usually include at least one operator.

Now, let's talk about math expressions or equations.

These can take a lot of different forms,

but you might be familiar with them already.

3 minus 1, 15 plus 8 divided by 2, 846 times 513.

These are all examples of expressions.

Is this bringing back memories of grade school?

Well, back in math class,

you most likely learned to complete an expression by

including an equal sign and the solution.

It's slightly different with spreadsheets.

When you create a formula using

an expression in a spreadsheet,

you start the formula with an equal sign.

For example, if we want to subtract,

we type an equal sign followed by the rest of

the expression without any spaces in the formula.

Now let's try an expression

that's a bit more challenging.

We'll type 31982, then

a hyphen for a minus sign, then 17795.

To calculate, we press "Enter."

You'll most likely use

formulas this way when dealing with

large numbers or expressions with multiple steps.

Here are the operators you will use to complete formulas.

The plus sign for addition,

the minus or hyphen for subtraction,

the asterisk for multiplication,

and the forward slash for division.

The division and multiplication symbols

might be different than what you're used to.

Small changes, but important to keep in mind.

If you already have data in your spreadsheet,

you can use cell references in your formulas instead.

A cell reference is a single cell or

range of cells in

a worksheet that can be used in a formula.

Cell references contain the letter of

the column and the number of the row where the data is.

A range of cells is a collection of two or more cells.

A range can include cells from the same row or column,

or from different columns and rows collected together.

We'll show you an example in an upcoming video.

Now let's apply what we just learned to some sales data.

If we want to add these figures to find

the total sales for the first row of data,

you can click "cell F2".

From there, we'll start with an equal sign and use

the cell references to input values in your expression.

We're starting with cell B2 because the year in

A2 is not a value we want to add to the total.

Play video starting at :3:38 and follow transcript3:38

Then press "Enter."

Just like that,

your total sales has been calculated for you,

but what if you realized one of

the values in your data was wrong?

No problem.

You can change the value in any cell using

the formula and the total will update automatically.

Play video starting at :4:3 and follow transcript4:03

The great thing about using cell references is that they

also automatically update when

a formula is copied to a new cell.

Talk about a time-saver.

Instead of entering the same formula again

for every new set of cell references,

just copy the formula using

the menu or a keyboard shortcut like Control plus C.

Play video starting at :4:26 and follow transcript4:26

Then paste the formula where you want to apply it using

Control plus V. And presto!

The formula updates

all the new cells and values correctly.

Now let's say you also want it to find the average sales.

For this, you create a new formula in a different cell.

Play video starting at :4:57 and follow transcript4:57

To group values in a formula, use parentheses.

This lets your spreadsheet know which values to calculate

together and the order of the operations to be performed.

For example, open parentheses,

then B2 plus C2 plus D2 plus E2,

and close parentheses, then divide

the value of all of this by typing slash four.

You are adding the values in the four cells together

and then using the slash to divide the total by four,

and just like the last one,

we can copy and paste the formula.

Here's another formula you can use if you want to find

the percent change in sales between June and July.

Play video starting at :5:56 and follow transcript5:56

Once a formula calculates the value,

you can then use the percent button

to change the value to a percentage.

When you apply the formula to the other rows,

both the formula and

the percent will automatically update.

That doesn't look like the right answer.

Looks like we've got an error. Don't worry.

Errors can happen at any stage of data analysis,

and that includes when you're using spreadsheets.

A formula has to be air tight.

If there's something wrong with one of

the cell references, it won't work.

So what's our error?

Well, we can see that the value in cell D4 is missing.

It might take some time and research on

your part to find the correct value, but it's worth it.

You want your analysis to be as accurate as possible.

When you do add the value,

the formula takes care of the rest.

Play video starting at :6:49 and follow transcript6:49

That was a lot to take in.

Thanks for staying with me.

You'll be able to apply what you

learned about formulas here and later

in the program to make

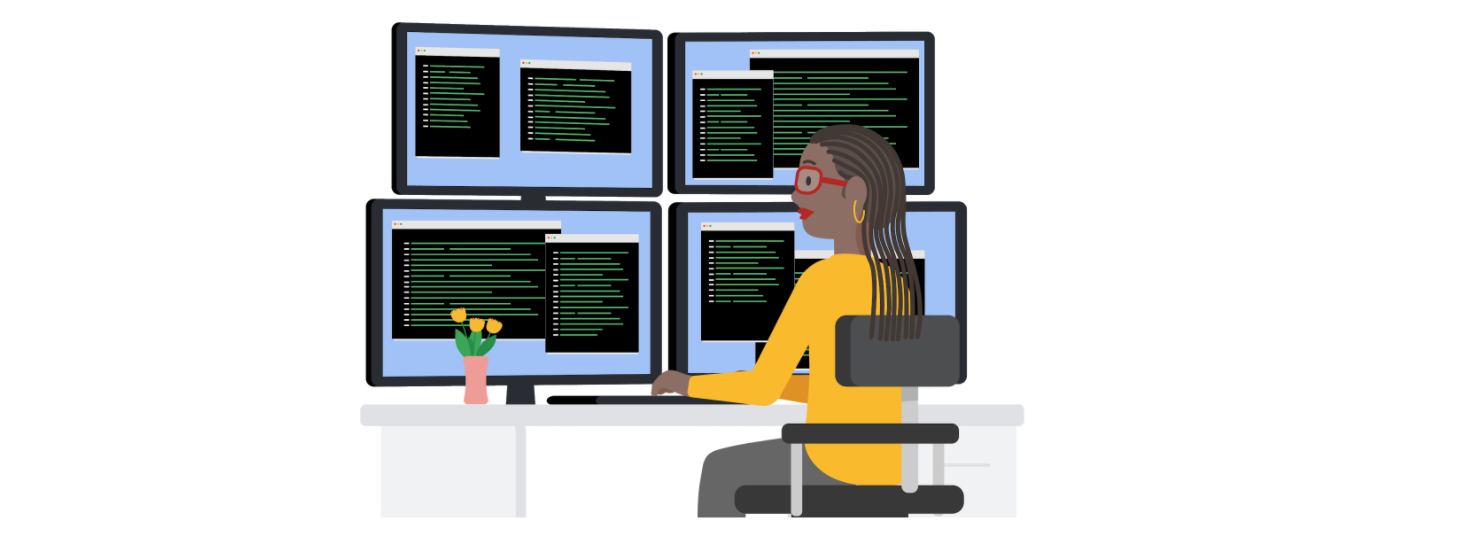
your analysis more efficient and your job,

a little easier, and

soon you'll work in your own spreadsheet.

Happy spreadsheeting.

# Quick reference: Formulas in spreadsheets



You have been learning a lot about spreadsheets and all kinds of time-saving calculations and organizational features they offer. One of the most valuable spreadsheet features is a **formula**. As a quick reminder, a formula is a set of instructions that does a specific calculation using the data in a spreadsheet. Formulas make it easy for data analysts to do powerful calculations automatically, which helps them analyze data more effectively. Below is a  quick-reference guide to help you get the most out of formulas.

## Formulas

### **The basics**

* When you write a formula in math, it generally ends with an equal sign (2 + 3 = ?). But with formulas, they always start with one instead (=A2+A3). The equal sign tells the spreadsheet that what follows is part of a formula, not just a word or number in a cell.
* After you type the equal sign, most spreadsheet applications will display an autocomplete menu that lists valid formulas, names, and text strings. This is a great way to create and edit formulas while avoiding typing and syntax errors.
* A fun way to learn new formulas is just by typing an equal sign and a single letter of the alphabet. Choose one of the options that pops up and you will learn what that formula does.

### **Mathematical operators**

* The mathematical operators used in spreadsheet formulas include:
* Subtraction – minus sign ( - )
* Addition – plus sign ( + )
* Division – forward-slash ( / )
* Multiplication – asterisk ( \* )

### **Auto-filling**

The lower-right corner of each cell has a fill handle. It is a small green square in Microsoft Excel and a small blue square in Google Sheets.

* Click the fill handle for a cell and drag it down a column to auto-fill other cells in the column with the same value or formula in that cell.
* Click the fill handle for a cell and drag it across a row to auto-fill other cells in the row with the same value or formula in that cell.
* If you want to create a numbered sequence in a column or row, do the following: 1) Fill in the first two numbers of the sequence in two adjacent cells, 2) Select to highlight the cells, and 3) Drag the fill handle to the last cell to complete the sequence of numbers. For example, to insert 1 through 100 in each row of column A, enter **1** in cell A1 and **2** in cell A2. Then, select to highlight both cells, click the fill handle in cell A2, and drag it down to cell A100. This auto-fills the numbers sequentially so you don't have to type them in each cell.

### **Absolute referencing**

* Absolute referencing is marked by a dollar sign ($). For example, =$A$10 has absolute referencing for both the column and the row value
* Relative references (which is what you normally do e.g. “=A10”) will change anytime the formula is copied and pasted. They are in relation to where the referenced cell is located. For example if you copied “=A10” to the cell to the right it would become “=B10”. With absolute referencing “=$A$10” copied to the cell to the right would remain “=$A$10”. But if you copied $A10 to the cell below, it would change to $A11 because the row value isn't an absolute reference.
* Absolute references will not change when you copy and paste the formula in a different cell. The cell being referenced is always the same.
* To easily switch between absolute and relative referencing in the formula bar, highlight the reference you want to change and press the F4 key; for example, if you want to change the absolute reference, $A$10, in your formula to a relative reference, A10, highlight $A$10 in the formula bar and then press the F4 key to make the change.

### **Data range**

* When you click into your formula, the colored ranges let you see which cells are being used in your spreadsheet. There are different colors for each unique range in your formula.
* In a lot of spreadsheet applications, you can press the F2 (or Enter) key to highlight the range of data in the spreadsheet that is referenced in a formula. Click the cell with the formula, and then press the F2 (or Enter) key to highlight the data in your spreadsheet.

### **Combining with functions**

* COUNTIF() is a formula and a function. This means the function runs based on criteria set by the formula. In this case, COUNT is the formula; it will be executed IF the conditions you create are true. For example, you could use =COUNTIF(A1:A16, “7”) to count only the cells that contained the number 7. Combining formulas and functions allows you to do more work with a single command.
* Hi and welcome back.
* Recently we've been learning about formulas.
* Sometimes data analysts encounter
* a problem with our formulas and we get an error.
* We've all been there and it can be frustrating.
* But there are solutions,
* that's what we're going to explore in this video.
* One error you may encounter is the DIV error.
* The DIV error happens when a formula is trying to divide
* a value in a cell by zero or by an empty cell.
* In this spreadsheet,
* the percentage Complete values in
* column C are calculated by
* dividing the values in
* the Tasks Completed column by
* the values in the Required Tasks column.
* Notice that column C is
* already formatted as a percentage.
* The DIV error is in cell C4 because we're
* dividing by zero the value in cell A4.
* To avoid this problem,
* we can have this spreadsheet
* automatically enter not applicable
* whenever a cell in column A
* contains a zero that would cause the error.
* To do this, we'll use the IFERROR function.
* If it encounters a DIV error
* caused by a cell that contains the zero,
* the phrase "Not applicable" will be inserted.
* Play video starting at :1:17 and follow transcript1:17
* We can also copy the formula to the rest of the cells in
* column C so it checks for
* any other cells that contain a zero.
* Now let's move on to ERROR.
* In Google Sheets,
* ERROR tells us the formula can't
* be interpreted as it is input.
* This is also known as a parsing error.
* Say we want to tally the number of
* total tasks in column B and C,
* we use the SUM function,
* but the formula equal sum B2 to B6,
* C2 to C6 causes an error.
* Examining it more closely,
* we see that a comma is missing between
* the cell ranges B2 to B6 and C2 to C6.
* We can fix this by inserting a comma between the cell
* ranges to indicate the end of each data item.
* This is called a delimiter,
* which you will learn more about soon.
* Now, the formula can correctly
* calculate the total number of tasks as 25.
* Another type of error is N/A.
* The N/A error tells you that the data
* in your formula can't be found by the spreadsheet.
* Generally, this means the data doesn't exist.
* This error most often occurs
* when using functions such as VLOOKUP,
* which searches for a certain value in a column
* to return a corresponding piece of information.
* Here, we see a master list of nuts and their prices.
* Using VLOOKUP, the spreadsheet finds prices in the list,
* then calculates the prices for
* each store using the assigned markup.
* But we have a N/A error in cells B49 and C49.
* The VLOOKUP formula is correct,
* so what's going on?
* Well, if we look carefully at the name of the nut,
* "almond" has no match in the lookup table,
* the lookup table uses the plural "almonds" instead.
* So we change almond to almonds,
* and with that typo fixed,
* the right prices are filled in.
* Speaking of typos, sometimes
* a typo can cause a NAME error.
* A NAME error can happen when
* a formula's name isn't recognized or understood.
* Suppose we see a NAME error
* in the nut prices spreadsheet.
* If we look carefully,
* the VLOOKUP function in cell B21 is spelled incorrectly,
* it has one extra O;
* this causes a NAME error for
* both the price and
* the resulting markup calculation for the store.
* To fix this error,
* we can delete the extra O in VLOOKUP.
* Play video starting at :4:7 and follow transcript4:07
* Perfect. Sometimes an error
* is caused by inconsistent or wrong data.
* For instance, the NUM error tells us that
* a formula's calculation can't
* be performed as specified by the data.
* The data doesn't make sense for that calculation.
* Here's what I mean.
* Suppose we're working on
* a large construction project using
* a spreadsheet to track
* how many months it takes to reach key milestones.
* We can use the DATEDIF function to
* calculate the number of months
* between start and end dates.
* The function requires the start date
* to be in the first cell
* referenced and the end date
* to be in the second cell referenced.
* In our case, cells B2 and C2 respectively.
* The M represents months,
* as we want this spreadsheet to calculate the number of
* months between our start and end dates.
* But we get a NUM error in cell D6.
* We notice that the end date comes before the start date,
* so the DATEDIF function
* can't calculate the number of months between.
* It's likely the start and end dates
* were interchanged by accident.
* We can request verification of the data to make sure.
* In the meantime, let's reverse the order of
* the cells in the formula to
* temporarily get around the error.
* Now, the result is nine months.
* What if the client's name was accidentally
* inserted into the start date in the spreadsheet?
* You guessed it, we get an error.
* The VALUE error can indicate
* a problem with a formula or referenced cells.
* It's often not clear right away what the problem is,
* so this error might take a little more effort to fix.
* In this case, John Welty was input as the start date,
* making the calculation impossible for
* the DATEDIF function in the cell D6.
* We just replace the text, John Welty,
* with the correct start date of September 1st, 2016.
* Play video starting at :6:18 and follow transcript6:18
* Last is the REF error,
* which often comes up when cells being
* referenced in a formula have been deleted,
* thus making the formula unable
* to perform the calculation.
* Here's a spreadsheet used to calculate
* the number of seats available for a company lunch.
* Let's say the company
* decided not to run the second floor,
* so we delete row 4.
* This results in a REF error when
* calculating the total seats available in cell B5.
* To fix this, we can change the formula to
* add the values in cells B2 and B3.
* Also, in this case,
* we could have prevented
* the REF error by using the SUM function and
* a range of cells instead of adding
* the cell value by direct reference.
* Now, if we delete row 10,
* the SUM function calculates the total seats
* available. There you go.
* We've now fixed some of
* the most common spreadsheet errors.
* When you see them again,
* you'll know what they mean.
* Troubleshooting is a big part of data analysis,
* so being able to find solutions
* is a key skill for data analysts.

When you are new to data analytics—and sometimes even when you aren't—spreadsheet struggles are real. It never feels good when you type in what you are sure is a perfect formula or function, only to get an error message. Understanding errors and how to fix them is a big part of keeping your data clean, so it’s important to know how to deal with issues as they come up, and more importantly, not to get discouraged.

Remember, even the most advanced spreadsheet users come across problems from time to time.



As a follow-up to what you learned in the previous video, here are a few best practices and helpful tips. These strategies will help you avoid spreadsheet errors to begin with, making your life in analytics a whole lot less stressful:

1. Filter data to make your spreadsheet less complex and busy.
2. Use and freeze headers so you know what is in each column, even when scrolling.
3. When multiplying numbers, use an asterisk (\*) not an X.
4. Start every formula and function with an equal sign (=).
5. Whenever you use an open parenthesis, make sure there is a closed parenthesis on the other end to match.
6. Change the font to something easy to read.
7. Set the border colors to white so that you are working in a blank sheet.
8. Create a tab with just the raw data, and a separate tab with just the data you need.

Now that you have learned some basic ways to avoid errors, you can focus on what to do when that dreaded pop-up does appear. The following table is a reference you can use to look up common spreadsheet errors and examples of each. Knowing what the errors mean takes some of the fear out of getting them.

| **Error** | **Description** | **Example** |
| --- | --- | --- |
| **#DIV/0!** | A formula is trying to divide a value in a cell by 0 (or an empty cell with no value) | =B2/B3, when the cell B3 contains the value 0 |
| **#ERROR!** | (Google Sheets only) Something can’t be interpreted as it has been input. This is also known as a parsing error. | =COUNT(B1:D1 C1:C10) is invalid because the cell ranges aren't separated by a comma |
| **#N/A** | A formula can't find the data | The cell being referenced can't be found |
| **#NAME?** | The name of a formula or function used isn't recognized | The name of a function is misspelled |
| **#NUM!** | The spreadsheet can't perform a formula calculation because a cell has an invalid numeric value | =DATEDIF(A4, B4, "M") is unable to calculate the number of months between two dates because the date in cell A4 falls after the date in cell B4 |
| **#REF!** | A formula is referencing a cell that isn't valid | A cell used in a formula was in a column that was deleted |
| **#VALUE!** | A general error indicating a problem with a formula or with referenced cells | There could be problems with spaces or text, or with referenced cells in a formula; you may have additional work to find the source of the problem. |

If you are working with Microsoft Excel, an interactive page, [How to correct a #VALUE! error](https://support.microsoft.com/en-us/office/how-to-correct-a-value-error-15e1b616-fbf2-4147-9c0b-0a11a20e409e), can help you narrow down the cause of this error. You can select a specific function from a drop-down list to display a link to tips to fix the error when using that function.

## Pro tip: Spotting errors in spreadsheets with conditional formatting

Conditional formatting can be used to highlight cells a different color based on their contents. This feature can be extremely helpful when you want to locate all errors in a large spreadsheet. For example, using conditional formatting, you can highlight in yellow all cells that contain an error, and then work to fix them.

### **Conditional formatting in Microsoft Excel**

To set up conditional formatting in Microsoft Excel to highlight all cells in a spreadsheet that contain errors, do the following:

1. Click the gray triangle above row number 1 and to the left of Column A to select all cells in the spreadsheet.
2. From the main menu, click **Home**, and then click **Conditional Formatting** to select **Highlight Cell Rules > More Rules**.
3. For Select a Rule Type, choose **Use a formula to determine which cells to format**.
4. For Format values where this formula is true, enter **=ISERROR(A1)**.
5. Click the **Format** button, select the Fill tab, select yellow (or any other color), and then click **OK**.
6. Click **OK** to close the format rule window.

To remove conditional formatting, click Home and select Conditional Formatting, and then click Manage Rules. Locate the format rule in the list, click Delete Rule, and then click OK.

### **Conditional formatting in Google Sheets**

To set up conditional formatting in Google Sheets to highlight all cells in a spreadsheet that contain errors, do the following:

1. Click the empty rectangle above row number 1 and to the left of Column A to select all cells in the spreadsheet. In the [Step-by-step in spreadsheets](https://www.coursera.org/learn/ask-questions-make-decisions/lecture/lpuHf/step-by-step-in-spreadsheets) video, this was called the Select All button.
2. From the main menu, click **Format** and select **Conditional Formatting** to open the Conditional format rules pane on the right.
3. While in the Single Color tab, under Format rules, use the drop-down to select **Custom formula is,** enter **=ISERROR(A1)**, select yellow (or any other color) for the formatting style, and then click **Done**.

To remove conditional formatting, click Format and select Conditional Formatting, and then click the Trash icon for the format rule.

## Spreadsheet error resources

To learn more and read about additional examples of errors and solutions, explore these resources:

* [**Microsoft Formulas and Functions**](https://support.microsoft.com/en-us/office/formulas-and-functions-294d9486-b332-48ed-b489-abe7d0f9eda9?ui=en-US&rs=en-US&ad=US#id0eaabaaa=errors)**:** This resource describes how to avoid broken formulas and how to correct errors in Microsoft Excel. This is a useful reference to have saved in case you run into a specific error and need to find solutions quickly while working in Excel.
* [**When Your Formula Doesn’t Work: Formula Parse Errors in Google Sheets**](https://www.benlcollins.com/spreadsheets/formula-parse-error/): This resource is a guide to finding and fixing some common errors in Google Sheets. If you are working with Google Sheets, you can use this as a quick reference for solving problems you might encounter working on your own.

With some practice and investigative determination, you will become much more comfortable handling errors in spreadsheets. Each error you catch and fix will make your data clearer, cleaner, and more useful.

Formulas are a great way to become

more efficient when using spreadsheets,

especially when you add shortcuts

like copying and pasting, into the mix.

As you progress as a data analyst,

you'll most likely learn

more shortcuts to help your process.

But now it's time to move on to functions.

While they're closely related to formulas,

they're not exactly the same.

By the end of this video,

you'll understand the difference and

know when to use them both.

In the world of spreadsheets

a function is a preset command that

automatically performs a specific process

or task using the data.

You might remember some of the shortcuts

we learned that can be used with formulas.

Think of functions as the most useful of the shortcuts.

The good news is a lot of

spreadsheet functions have names

that tell you what they do.

There are tons of functions out there.

As you continue to work with spreadsheets,

you'll find that you use certain ones a lot,

and others, rarely or not at all.

For now, let's take a look

at some of the functions that we can

apply to our sales data from the previous video.

We'll start with total sales.

Let's use the SUM function for this in cell F2.

The first steps are pretty

similar to what we did in the last video.

First, we'll select the cell

where we want the calculation to appear.

Type equals, then add the word SUM as our function.

One of the great things about functions

is they don't always need operators,

like a plus sign for addition.

In this case, after the open parentheses,

you can go ahead and select

the range of cells you're adding.

A colon between the cell references

shows that you're using a range.

In this case, the range includes cells from the same row.

After the closed parentheses, we press Enter.

Just like that, our total sales number appears.

Just like the formula we used before,

functions can be copied and pasted

into other cells in the same column.

Play video starting at :2:7 and follow transcript2:07

But let's undo that step so that you can

see another way to copy a function or formula.

Spreadsheets have something called a fill handle.

It's a little box that appears in

the lower right-hand corner when you click on a cell.

If you rest your cursor on the box,

you can then drag the fill handle to

the other boxes in the same row or column.

Any formula or function in that cell will

automatically be added to the cells you fill plus,

the fill handle will update the formula so

the cell references match

the row of the columns of the cells you fill.

Play video starting at :2:52 and follow transcript2:52

This means the formula is calculated based

on the data in each separate row or column.

Filling won't work for every situation,

but it's still a pretty great trick.

Now let's find the average sale for

each month using the AVERAGE function.

Play video starting at :3:20 and follow transcript3:20

Different functions perform different calculations,

but they work in the same way.

Keep in mind, not every calculation

you'll come across has its own function to help you.

For example, to find

the percent change in sales between June and July,

you'd use the same formula you used in an earlier video.

Play video starting at :3:47 and follow transcript3:47

Let's say you're asked to find

the lowest monthly sales in this data set.

There's a function for that.

It's called the MIN function,

which stands for minimum.

Here's how it works.

Say you need to find the lowest monthly sales

for the whole set.

Play video starting at :4:8 and follow transcript4:08

All you have to do is set up the function.

Then after the open parenthesis,

select the values from all three rows.

Play video starting at :4:21 and follow transcript4:21

This might be important information

for your stake holders.

Let's add color to the cell with that value,

in your data set to make it stand out.

In this case, click on cell D2 and then fill color icon,

which looks like a paint can,

then choose a color.

I'll use yellow here.

You can follow the same steps for

the highest sales by using the,

wait for it, MAX function.

Play video starting at :5:1 and follow transcript5:01

Looks like we have an error message.

What could be wrong?

We forgot to include

an open parentheses after the function.

No worries, it's a quick fix.

Play video starting at :5:19 and follow transcript5:19

But this is a good reminder to continually check

the format of your functions and

formulas as you use them.

We'll learn more about Error messages

and how to work with them later.

That's better. Now we'll add

color to the cell with the highest sales too.

Play video starting at :5:37 and follow transcript5:37

This is just one way to highlight key data.

You'll find out about some others later.

You've now had a peek at some ways you

can add and organize data in a spreadsheet.

You've also seen how powerful formulas and

functions can be when applied to real world data.

As a data analyst,

this is just the beginning of

your experience with spreadsheets.

You'll soon find out

how much more spreadsheets have to offer.

In the meantime, you're free to

practice some of these formulas,

functions, and other processes on your own.

It can be fun to experiment,

and see all that spreadsheets can do.

Soon, you will switch from

spreadsheets to structured thinking.

The data analytics pieces are starting to fit together.

Exciting stuff is coming right up. So stick around.

# Quick reference: Functions in spreadsheets

As a quick refresher, a function is a preset command that automatically performs a specific process or task using the data in a spreadsheet. Functions give data analysts the ability to do calculations, which can be anything from simple arithmetic to complex equations. Use this reading to help you keep track of some of the most useful options.

## Functions

### **The basics**

* Just like formulas, start all of your functions with an equal sign; for example =SUM. The equal sign tells the spreadsheet that what follows is part of a function, not just a word or number in a cell.
* After you type the equal sign, most spreadsheet applications will display an autocomplete menu that lists valid functions, names, and text strings. This is a great way to create and edit functions while avoiding typing and syntax errors.
* A fun way to learn new functions is by simply typing an equal sign and a single letter of the alphabet. Choose one of the options that pops up and learn what that function does.

### **Difference between formulas and functions**

* A formula is a set of instructions used to perform a calculation using the data in a spreadsheet.
* A function is a preset command that automatically performs a specific process or task using the data in a spreadsheet.

### **Popular functions**

A lot of people don’t realize that keyboard shortcuts like cut, save, and find are actually functions. These functions are built into an application and are amazing time-savers. Using shortcuts lets you do more with less effort. They can make you more efficient and productive because you are not constantly reaching for the mouse and navigating menus. Use these links to discover the most popular shortcuts, for [Chromebook](https://support.google.com/chromebook/answer/183101?hl=en), [PC](https://support.microsoft.com/en-us/windows/keyboard-shortcuts-in-windows-dcc61a57-8ff0-cffe-9796-cb9706c75eec), and [Mac](https://support.apple.com/en-us/HT201236).

### **Auto-filling**

The lower-right corner of each cell has a fill handle. It is a small green square in Microsoft Excel and a small blue square in Google Sheets.

* Click the fill handle for a cell and drag it down a column to auto-fill other cells in the column with the same formula or function used in that cell.
* Click the fill handle for a cell and drag it across a row to auto-fill other cells in the row with the same formula or function used in that cell.

### **Relative, absolute, and mixed references**

* Relative references (cells referenced without a dollar sign, like A2) will change when you copy and paste the function into a different cell. With relative references, the location of the cell that contains the function determines the cells used by the function.
* Absolute references (cells fully referenced with a dollar sign, like $A$2) will not change when you copy and paste the function into a different cell. With absolute references, the cells referenced always remain the same.
* Mixed references (cells partially referenced with a dollar sign, like $A2 or A$2) will change when you copy and paste the function into a different cell. With mixed references, the location of the cell that contains the function determines the cells used by the function, but only the row or column is relative (not both).
* In spreadsheets, you can press the F4 key to toggle between relative, absolute, and mixed references in a function. Click the cell containing the function, highlight the referenced cells in the formula bar, and then press F4 to toggle between and select relative, absolute, or mixed referencing.

### **Data ranges**

* When you click a cell that contains a function, colored data ranges in the formula bar indicate which cells are being used in the spreadsheet. There are different colors for each unique range in a function.
* Colored data ranges help prevent you from getting lost in complex functions.
* In spreadsheets, you can press the F2 key to highlight the range of data used by a function. Click the cell containing the function, highlight the range of data used by the function in the formula bar, and then press F2. The spreadsheet will go to and highlight the cells specified by the range.

### **Data ranges evaluated for a condition**

**COUNTIF** is an example of a function that returns a value based on a condition that the data range is evaluated for. The function counts the number of cells that meet the criteria. For example, in an expense spreadsheet, use COUNTIF to count the number of cells that contain a reimbursement for "airfare."

For more information, refer to:

* [Microsoft Support's page for COUNTIF](https://support.microsoft.com/en-us/office/countif-function-e0de10c6-f885-4e71-abb4-1f464816df34)
* [Google Help Center's documentation for COUNTIF](https://support.google.com/docs/answer/3093480?hl=en) where you can copy a sheet with [COUNTIF examples](https://docs.google.com/spreadsheets/d/1PYoKCYZAkWSaMBsiTyvxZzCCt2WQ-QKOC763RWHMB7c/template/preview) (click "Use Template" if you click the COUNTIF link provided on this page)

## Conclusion

There are a lot more functions that can help you make the most of your data. This is just the start. You can keep learning how to use functions to help you solve complex problems efficiently and accurately throughout your entire career.

## Keyboard shortcuts

You can save these functions for future reference. Feel free to download a PDF version of the functions below:

Earlier I told you that carefully

defining a business problem can

ultimately save time, money, and resources.

All of this is achieved through structured thinking.

Structured thinking is the process of

recognizing the current problem or situation,

organizing available information,

revealing gaps and opportunities,

and identifying the options.

In other words, it's a way of being super prepared.

It's having a clear list of what

you are expected to deliver,

a timeline for major tasks and activities,

and checkpoints so the team knows you're making progress.

In this video, we'll look at

how structured thinking helps us save time and effort,

but also makes our job as data analysts easier

because it allows us to better

understand the work we are doing.

In the business world,

it's common for teams to spend hours of

valuable time trying to solve an important problem,

only to end up back where they started.

Not only is the initial problem not resolved,

but they've spent hours not resolving it.

This outcome negatively affects you,

your team, and the organization as a whole.

But it can usually be prevented.

Many times the situation is

a result of not fully understanding the issue.

Structured thinking will help you

understand problems at a high level so

that you can identify areas that need

deeper investigation and understanding.

The starting place for

structured thinking is the problem domain,

which you might have remembered from earlier.

Once you know the specific area of analysis,

you can set your base and lay out

all your requirements and

hypotheses before you start investigating.

With a solid base in place,

you'll be ready to deal with any obstacles that come up.

What kind of obstacles?

Well, let's say you're

asked to predict the future value of

an apartment building based on a given dataset.

You have hundreds of variables and

every one is crucial to your analysis.

But what if one variable accidentally gets left out,

like square footage, for example?

You'd have to go back and redo all your hard work.

That's because missing variables can

lead to inaccurate conclusions.

Another way that you can practice structured thinking and

avoid mistakes is by using a scope of work.

A scope of work or SOW is an agreed-

upon outline of the work

you're going to perform on a project.

For many businesses, this

includes things like work details,

schedules, and reports that the client can expect.

Now, as a data analyst,

your scope of work will be a bit more

technical and include

those basic items we just mentioned,

but you'll also focus on things

like data preparation, validation,

analysis of quantitative and qualitative datasets,

initial results, and maybe

even some visuals to really get the point across.

Let's bring a scope of work to

life with a simple example.

Say a couple has hired a wedding planner.

We'll focus on just one task, the wedding invitations.

Here's what might be in scope of work: deliverables,

timeline, milestones, and reports.

Let's break down just one of these, deliverables.

The wedding planner and

couple will need to decide on the invitation,

make a list of people to invite, collect their addresses,

print the invitations, address the envelopes,

stamp them, and mail them out.

Now let's check out the timelines.

You'll notice the dates and

the milestones which keep us on track.

Finally, we have the reports,

which give our couple some peace of mind by

telling them when each step is complete.

A scope of work can be a simple but powerful tool.

With a solid scope of work,

you'll be able to address any confusion,

contradictions, or questions about the data up-

front and make sure

these sneaky setbacks don't stand in your way.

This is a simple example

of what a scope of work might look like.

But later, you'll be able to practice building your own.

Next up in our scope,

we'll check out setbacks from

a different angle by learning

the importance of contextualizing data and avoiding bias.

Looking forward to sharing some cool insights with you.

The importance of context

**Context** is the condition in which something exists or happens. Context is important in data analytics because it helps you sift through huge amounts of disorganized data and turn it into something meaningful. The fact is, data has little value if it is not paired with context.



Understanding the context behind the data can help us make it more meaningful at every stage of the data analysis process. For example, you might be able to make a few guesses about what you're looking at in the following table, but you couldn't be certain without more context.

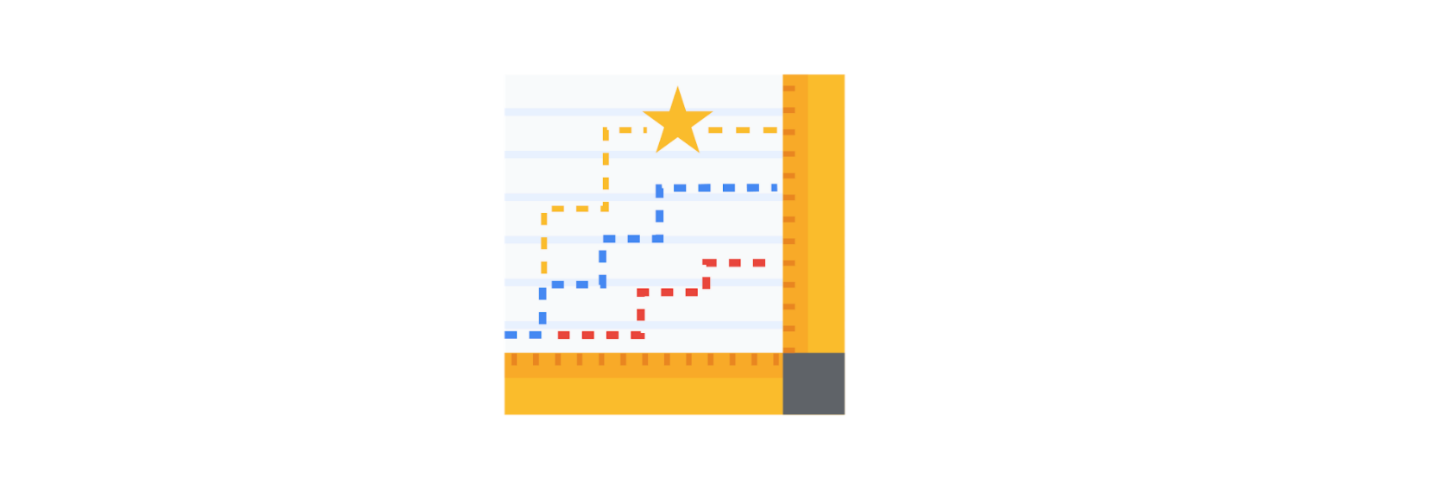
|  |  |
| --- | --- |
| 2010 | 28000 |
| 2005 | 18000 |
| 2000 | 23000 |
| 1995 | 10000 |

On the other hand, if the first column was labeled to represent the years when a survey was conducted, and the second column showed the number of people who responded to that survey, then the table would start to make a lot more sense. Take this a step further, and you might notice that the survey is conducted every 5 years. This added context helps you understand why there are five-year gaps in the table.

| **Years (Collected every 5 years)** | **Respondents** |
| --- | --- |
| 2010 | 28000 |
| 2005 | 18000 |
| 2000 | 23000 |
| 1995 | 10000 |

**Context can turn raw data into meaningful information.** It is very important for data analysts to contextualize their data. This means giving the data perspective by defining it. To do this, you need to identify:

* Who: The person or organization that created, collected, and/or funded the data collection
* What: The things in the world that data could have an impact on
* Where: The origin of the data
* When: The time when the data was created or collected
* Why: The motivation behind the creation or collection
* How: The method used to create or collect it



Understanding and including the context is important during each step of your analysis process, so it is a good idea to get comfortable with it early in your career. For example, when you collect data, you’ll also want to ask questions about the context to make sure that you understand the business and business process. During organization, the context is important for your naming conventions, how you choose to show relationships between variables, and what you choose to keep or leave out. And finally, when you present, it is important to include contextual information so that your stakeholders understand your analysis.