

Data cleaning is a must

Begin cleaning data

Cleaning data in spreadsheets

- Video:** Data-cleaning features in spreadsheets

8 min
- Video:** Optimize the data-cleaning process

14 min
- Reading:** Workflow automation

20 min
- Video:** Different data perspectives

10 min
- Video:** Even more data-cleaning techniques

6 min
- Practice Quiz:** Hands-On Activity: Clean data with spreadsheet functions

2 questions
- Reading:** Learning Log: Develop your approach to cleaning data

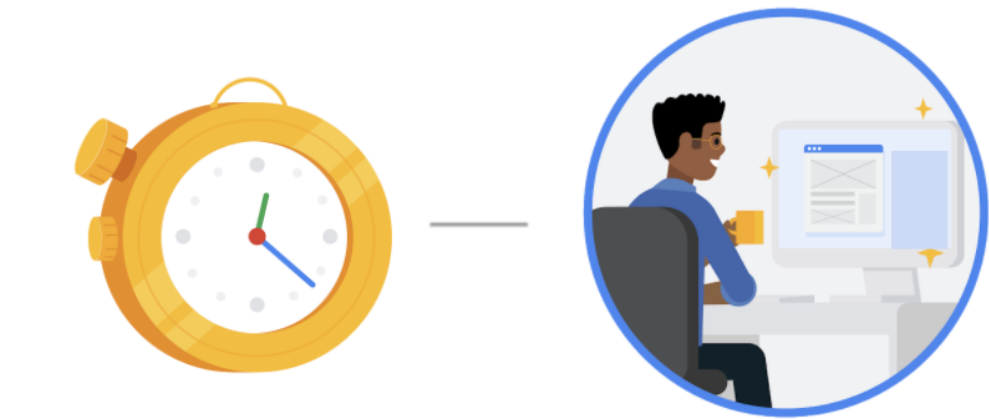
20 min
- Practice Quiz:** Test your knowledge on cleaning data in spreadsheets

3 questions

Weekly challenge 2

Workflow automation

In this reading, you will learn about workflow automation and how it can help you work faster and more efficiently. Basically, workflow automation is the process of automating parts of your work. That could mean creating an event trigger that sends a notification when a system is updated. Or it could mean automating parts of the data cleaning process. As you can probably imagine, automating different parts of your work can save you tons of time, increase productivity, and give you more bandwidth to focus on other important aspects of the job.



What can be automated?

Automation sounds amazing, doesn't it? But as convenient as it is, there are still some parts of the job that can't be automated. Let's take a look at some things we can automate and some things that we can't.

Task	Can it be automated?	Why?
Communicating with your team and stakeholders	No	Communication is key to understanding the needs of your team and stakeholders as you complete the tasks you are working on. There is no replacement for person-to-person communications.
Presenting your findings	No	Presenting your data is a big part of your job as a data analyst. Making data accessible and understandable to stakeholders and creating data visualizations can't be automated for the same reasons that communications can't be automated.
Preparing and cleaning data	Partially	Some tasks in data preparation and cleaning can be automated by setting up specific processes, like using a programming script to automatically detect missing values.
Data exploration	Partially	Sometimes the best way to understand data is to see it. Luckily, there are plenty of tools available that can help automate the process of visualizing data. These tools can speed up the process of visualizing and understanding the data, but the exploration itself still needs to be done by a data analyst.
Modeling the data	Yes	Data modeling is a difficult process that involves lots of different factors; luckily there are tools that can completely automate the different stages.

More about automating data cleaning

One of the most important ways you can streamline your data cleaning is to clean data where it lives. This will benefit your whole team, and it also means you don't have to repeat the process over and over. For example, you could create a programming script that counted the number of words in each spreadsheet file stored in a specific folder. Using tools that can be used where your data is stored means that you don't have to repeat your cleaning steps, saving you and your team time and energy.

More resources

There are a lot of tools out there that can help automate your processes, and those tools are improving all the time. Here are a few articles or blogs you can check out if you want to learn more about workflow automation and the different tools out there for you to use:

- Towards Data Science's [Automating Scientific Data Analysis](#)
- MIT News' [Automating Big-Data Analysis](#)
- TechnologyAdvice's [10 of the Best Options for Workflow Automation Software](#)

As a data analyst, automation can save you a lot of time and energy, and free you up to focus more on other parts of your project. The more analysis you do, the more ways you will find to make your processes simpler and more streamlined.

Mark as completed