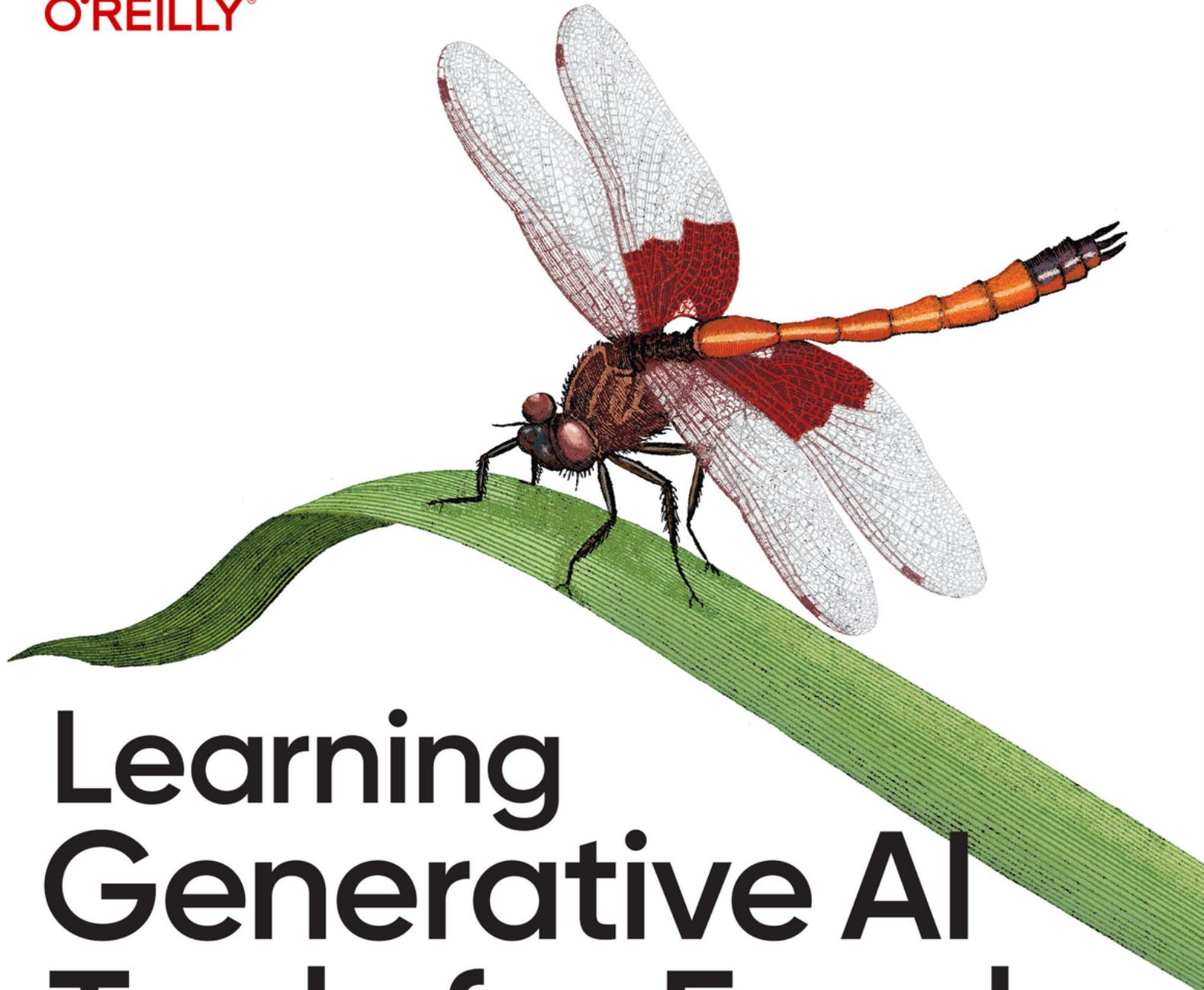


O'REILLY®



Learning Generative AI Tools for Excel

Speed Up Your Everyday Tasks with
Microsoft Excel, Copilot, ChatGPT, and Beyond

Angelica Lo Duca

Learning Generative AI Tools for Excel

Speed Up Your Everyday Tasks with Microsoft Excel, Copilot, ChatGPT, and Beyond

With Early Release ebooks, you get books in their earliest form—the author’s raw and unedited content as they write—so you can take advantage of these technologies long before the official release of these titles.

Angelica Lo Duca



Learning Generative AI Tools for Excel

by Angelica Lo Duca

Copyright © 2025 Angelica Lo Duca. All rights reserved.

Printed in the United States of America.

Published by O'Reilly Media, Inc., 1005 Gravenstein Highway North,
Sebastopol, CA 95472.

O'Reilly books may be purchased for educational, business, or sales promotional use. Online editions are also available for most titles (<http://oreilly.com>). For more information, contact our corporate/institutional sales department: 800-998-9938 or corporate@oreilly.com.

- Acquisitions Editor: Michelle Smith
- Development Editor: Corbin Collins
- Interior Designer: David Futato
- Cover Designer: Karen Montgomery
- Illustrator: Kate Dullea
- May 2026: First Edition

Revision History for the First Edition

- 2025-02-05: First Release
- 2025-03-05: Second Release
- 2025-04-22: Third Release

See <http://oreilly.com/catalog/errata.csp?isbn=9781098190170> for release details.

The O'Reilly logo is a registered trademark of O'Reilly Media, Inc.
Learning Generative AI Tools for Excel, the cover image, and related trade dress are trademarks of O'Reilly Media, Inc.

The views expressed in this work are those of the author and do not represent the publisher's views. While the publisher and the author have used good faith efforts to ensure that the information and instructions contained in this work are accurate, the publisher and the author disclaim all responsibility for errors or omissions, including without limitation responsibility for damages resulting from the use of or reliance on this work. Use of the information and instructions contained in this work is at your own risk. If any code samples or other technology this work contains or describes is subject to open source licenses or the intellectual property rights of others, it is your responsibility to ensure that your use thereof complies with such licenses and/or rights.

978-1-098-19017-0

[FILL IN]

Brief Table of Contents (*Not Yet Final*)

Chapter 1: Breaking the Ice: Introducing AI for Excel (available)

Chapter 2: Getting Started with ChatGPT and Excel (available)

Chapter 3: Speeding Up Common Operations with ChatGPT (available)

Chapter 4: Advanced Operations with ChatGPT (available)

Chapter 5: Getting Started with Copilot (unavailable)

Chapter 6: Data Preparation with Copilot (unavailable)

Chapter 7: Text Transformation with Copilot (unavailable)

Chapter 8: Basic Statistics with Copilot (unavailable)

Chapter 9: Data Analysis with Copilot (unavailable)

Chapter 10: Data Visualization, Reporting, and Storytelling with Copilot
(unavailable)

Chapter 11: A Landscape of the AI Add-Ins in Excel (unavailable)

Chapter 12: Using The OpenAI API in Excel (unavailable)

Appendix A: Solutions to Activities (available)

Appendix B: Exercises for Part 1 (unavailable)

Appendix C: Exercises for Part 2 (unavailable)

Appendix D: Exercises for Part 3 (unavailable)

Chapter 1. Breaking the Ice: Introducing AI for Excel

A NOTE FOR EARLY RELEASE READERS

With Early Release ebooks, you get books in their earliest form—the author’s raw and unedited content as they write—so you can take advantage of these technologies long before the official release of these titles.

This will be the 1st chapter of the final book. Please note that the GitHub repo will be made active later on.

If you have comments about how we might improve the content and/or examples in this book, or if you notice missing material within this chapter, please reach out to the editor at ccollins@oreilly.com.

If you have worked with Microsoft Excel, you’ve probably noticed how tiring and sometimes boring it is to perform repetitive operations such as formatting cells or doing calculations. By searching the web, you have probably found ways to navigate the jungle of data analysis, formulas, and formatting, but you keep encountering new problems that slow down your work. If you’re *not* that familiar with Excel, you may be intimidated by the challenges you’ll undoubtedly face when manipulating data. That has happened to me many times.

When I was 15 I started using Excel to catalog the books on my shelf. At the time, a quarter of a century ago, I was using Microsoft Excel 1995. But later, even after I had been working with Excel for a long time, every time a new problem arose, I had to search the Web or look at books for answers.

Today, everything has changed with the advent of generative AI, which has introduced new ways of working and speeding up tedious tasks. Generative AI is a potent tool that allows you to quickly manipulate and generate data. However, to make the best use of it, you need to know it, learn its limits, and see how far it can help you.

This chapter begins your journey to discover how to integrate generative AI into Excel. This technology is evolving so fast that the AI models coming out today may seem old quickly. However, the *principles* described in this book are always valid, and you can apply them even if the technologies described become obsolete.

So what are we waiting for? Let's go. This chapter covers the following:

- Basic Excel concepts
- Why Excel needs AI
- The benefits of using AI in Excel
- When AI is not helpful in Excel
- Browsing the different AI tools for Excel

Basic Excel Concepts

As you probably know, Microsoft Excel is a very popular tool for processing different types of data. You can use it to display and manipulate tables, build and enrich graphs, extract knowledge, and much more.

Whatever reason you use Excel for, you will surely deal with data. *Data* is the raw element representing an observation or measurement before any interpretation or processing. Data can be numbers or texts. For example, if the number of clouds you see in the sky today is 10, it means that 10 is a data point that represents the number of clouds. Today's date is another data point, and the category *cloud* is another data point. Imagine that every day, you count not only the number of clouds you see in the sky but also the number of birds, which represent a second category, and you organize everything into a table like [Table 1-1](#).

Table 1-1. Sky dataset with the number of clouds and the number of birds organized by day.

Date	Number	Category
2024-09-28	10	Cloud
2024-09-29	12	Cloud
2024-09-30	0	Cloud
2024-09-28	13	Bird
2024-09-29	7	Bird
2024-09-30	9	Bird

What you have there is a dataset. A *dataset* is, very simply, a set of data. In Excel, you can load several datasets, one at a time or all at once.

Before we move forward, let's clarify some Excel definitions you will see throughout the book:

- A *cell* is the basic unit of a table, located at the intersection of a row and a column of the table, where you can enter data or formulas.
- A *worksheet* is a sheet within a work file, organized into rows and columns, where data can be inserted, viewed, and manipulated.
- A *workbook* is a file that contains one or more worksheets. It allows you to organize and manage different data tables within the same document.

- A *range* is a set of selected cells. One *row* is a single horizontal row of cells, while a *column* is a vertical column of cells within a worksheet.
-

NOTE

To illustrate AI in Excel, much of this chapter will have you doing an exercise in Excel. If you know Excel like the back of your hand, this exercise will be child's play.

Consider the dataset on [submarine cables](#) in the United States released by Marine Cadastre and usable under the CC0-1.0 license. The dataset is constantly updated. We will use the September 26, 2024 version, which you can download from the book's [GitHub repository](#) under [01/ch01.csv](#).

This dataset contains the approximate location of commercial and research submarine cables in U.S. waters. Most of these cables are used for communications, while a smaller portion is used for power transmission. In coastal areas, cables are usually buried under the seabed, while in the open sea, they are laid directly on the seabed. Just to get an idea of what we're talking about, [Figure 1-1](#) shows a portion of the data plotted on a US East Coast map. Each line corresponds to a submarine cable.

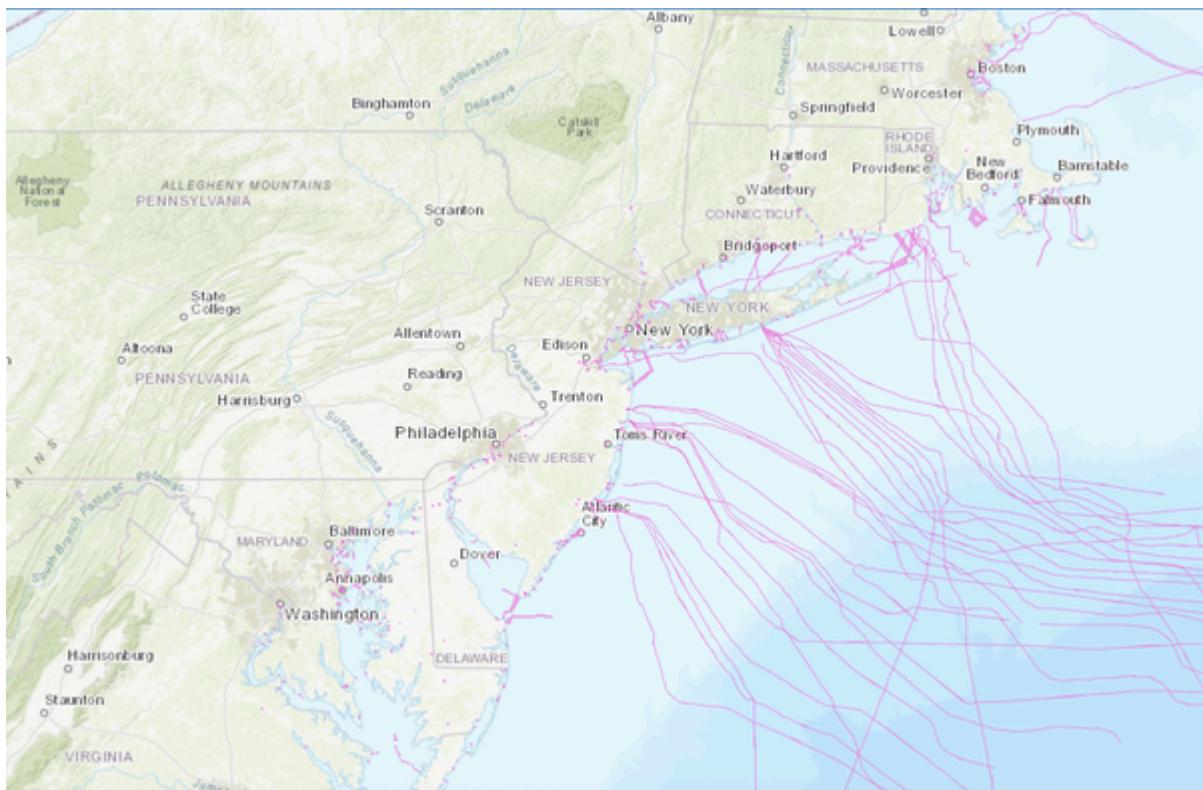


Figure 1-1. A map representation of a portion of the U.S. submarine cable dataset.

All the data shown in [Figure 1-1](#) are organized in a table. [Table 1-2](#) shows a portion of the U.S. submarine cable dataset.

Table 1-2. A portion of the U.S. submarine cable dataset.

owner	cablesystem	region	SHAPE_Length
AT&T Corp	Americas II	East Coast	108697.599067727
GCI Communication Corp	AU-Aleutian Submarine Cable System	Pacific Northwest	1129434.70407587
GU Holdings Inc	Curie	West Coast	1274007.14615703
Eversource Energy	Eversource HEEC Cable	East Coast	10111.1441552573
GU Holdings Inc	Firmina	East Coast	1006081.05355044
GU Holdings Inc	Grace Hopper	East Coast	1493252.13857506

The dataset contains the following fields:

- *Owner*: The entity that owns the cable. It is a company or an organization

- *Cable System*: The full name of the submarine cable system to which the cable belongs
- *Region*: The geographic region where the cable is located
- *SHAPE_Length*: The cable's length
- *Short Name*: The short name of the submarine cable
- *Objectid*: A unique identifier for each object in the dataset
- *SHAPE_Area*: The geographic area occupied by the cable
- *Status*: The current status of the cable.

In the remainder of this section, we'll use this dataset to see how to use Excel for different purposes, including data management, analysis, and visualization. You will see the other cases later in the book.

Excel for Data Management

Data management involves collecting, organizing, storing, and maintaining data while ensuring its accessibility and security. The first step to managing data is to load the dataset as a table. Excel supports various formats, such as CSV, but to fully exploit its potential, here you will save the dataset in XLSX format, which is Excel's proprietary format.

Download the dataset from the GitHub repository of the book under [01/ch01.csv and save it in your local filesystem](#). Then, load it into Excel. If this is your first time using Microsoft Excel, perform the following steps to load the example dataset into Excel:

1. Launch Excel, then select File → Open → Select your file → Open → Choose the dataset from your filesystem → Check Delimited. Click Next → Select Comma as a delimiter, then click Next.
2. For each column, select the format. For now, leave the default values and click Finish.
3. If a warning screen appears about possible information loss, proceed to save the file as an Excel file by clicking Save As. On the screen that opens, select XLSX as the File Format. Then click Save.

Figure 1-2 shows a portion of the dataset loaded into Excel.

	A	B	C	D	E	F
1	owner	cablesystem	region	SHAPE_Length	shortname	objectid
2	AT&T Corp	Americas II	East Coast	108.697.599.067.727	Americas II	1
3	GCI Communication Corp	AU-Aleutian Submarine Cable System	Pacific Northwest	112.943.470.407.587	AU-Aleutian	
4	GCI Communication Corp	AU-Aleutian Submarine Cable System	Pacific Northwest	538.208.941.786.863	AU-Aleutian	
5	GCI Communication Corp	AU-Aleutian Submarine Cable System	Pacific Northwest	907.186.773.386.604	AU-Aleutian	
6	GCI Communication Corp	AU-Aleutian Submarine Cable System	Pacific Northwest	132.612.998.874.575	AU-Aleutian	
7	GCI Communication Corp	AU-Aleutian Submarine Cable System	Pacific Northwest	583.757.167.942.325	AU-Aleutian	
8	GCI Communication Corp	AU-Aleutian Submarine Cable System	Pacific Northwest	914.494.783.021.393	AU-Aleutian	
9	GCI Communication Corp	AU-Aleutian Submarine Cable System	Pacific Northwest	399.667.309.103.789	AU-Aleutian	
10	GCI Communication Corp	AU-Aleutian Submarine Cable System	Pacific Northwest	432.525.541.965.511	AU-Aleutian	
11	GCI Communication Corp	AU-Aleutian Submarine Cable System	Pacific Northwest	331.138.928.353.232	AU-Aleutian	
12	GCI Communication Corp	AU-Aleutian Submarine Cable System	Pacific Northwest	251.466.114.110.333	AU-Aleutian	
13	GU Holdings Inc	Curie	West Coast	127.400.714.615.703	Curie	1.1
14	Eversource Energy	Eversource HEEC Cable	East Coast	101.111.441.552.573	Eversource	
15	GU Holdings Inc	Firmina	East Coast	100.608.105.355.044	Firmina	
16	GU Holdings Inc	Grace Hopper	East Coast	149.325.213.857.506	Grace Hopper	
17	Hawaiian Telcom Inc	Hawaiian Islands Fiber Network	Pacific Islands	97.329.897.991.137	HIFN	
18	NTT America Inc	JUPITER System	West Coast	33.426.018.888.378	JUPITER	
19	GU Holdings Inc	Pacific Light Cable Network	West Coast	16.998.263.576.784	PLCN	1.1.1
20	Orsted	Revolution Wind OSS1	East Coast	166.903.532.402.238	RW-1	

Figure 1-2. A portion of the dataset loaded into Excel.

The dataset is not yet formatted as a table. A *table* in Excel is a tool that helps you organize data more efficiently than a regular range of cells. Using a table has several benefits, such as using column names instead of cell

ranges. As you'll see later in the book, Copilot, the AI tool integrated with Excel, works only with tables.

Select Format as Table from the Ribbon Home tab to turn your raw data into a table. In the next window, select the table layout and leave the columns selected by default. Also, make sure there is a checkmark on the entry "My table has headers." Before we start working, let's give the table a name. Select the Table tab from the Ribbon and change the table's name to MainTable in the Table Name box on the left, as shown in [Figure 1-3](#).

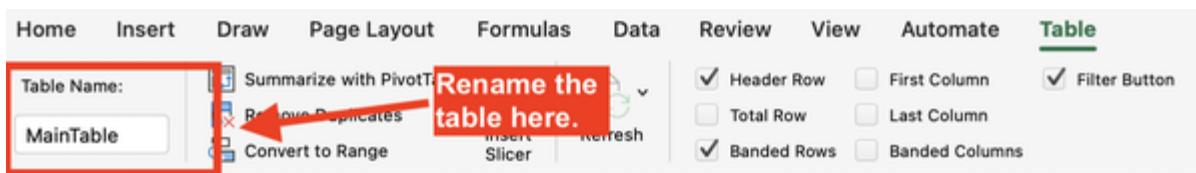


Figure 1-3. How to rename a table in Excel.

At this point, you are ready to work.

NOTE

The Ribbon in Excel is the toolbar at the top of the Excel window. It organizes commands and features into tabs, such as Home, Insert, and Formulas, as shown in [Figure 1-4](#).

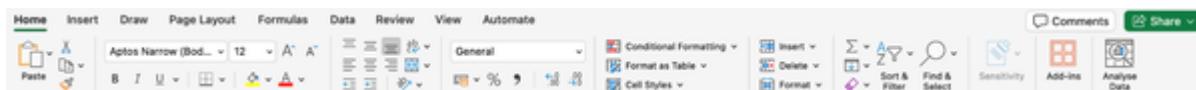


Figure 1-4. The Ribbon in Excel.

Excel for Data Analysis

The second main purpose of using Excel is *data analysis*, which allows you to carry out mathematical and financial calculations. Excel offers various formulas and functions for performing numerical, financial, and statistical calculations, such as sums, averages, and so on.

Let's query the dataset by building a new table with the number of In Service, Planned, and Out of Service cables. We will call this table StatusTable. Many data points are unavailable, so they will be excluded from our calculation. To build the StatusTable:

1. Click in whatever cell out of the main table.
2. Write the data values shown in [Table 1-3](#), one for each cell, as shown in [Figure 1-5](#).

Table 1-3. The data values for the StatusTable.

status	count
In Service	
Planned	
Out of Service	

The screenshot shows an Excel spreadsheet with a main table and a smaller table to its right. The main table has columns for owner, cablesystem, region, SHAPE_Length, shortname, objectid, SHAPE_Area, and status. A red box highlights the status column, and an arrow points from it to a callout box containing the text "Write data values out of the MainTable.". To the right of the main table is a smaller table with two columns: status and count. The status column contains three entries: In Service, Planned, and Out of Service. The count column contains the corresponding counts for each status.

owner	cablesystem	region	SHAPE_Length	shortname	objectid	SHAPE_Area	status
J&T Corp	Americas II	East Coast	#####	Americas II 12	1	#####	In Service
ICL Commun	AU-Aleutian Sub Pacific North	###	#####	AU-Aleutian A	2	#####	In Service
ICL Commun	AU-Aleutian Sub Pacific North	###	#####	AU-Aleutian B	3	#####	In Service
ICL Commun	AU-Aleutian Sub Pacific North	###	#####	AU-Aleutian C	4	#####	In Service
ICL Commun	AU-Aleutian Sub Pacific North	###	#####	AU-Aleutian D	5	#####	In Service
ICL Commun	AU-Aleutian Sub Pacific North	###	#####	AU-Aleutian E	6	#####	In Service
ICL Commun	AU-Aleutian Sub Pacific North	###	#####	AU-Aleutian F	7	#####	In Service
ICL Commun	AU-Aleutian Sub Pacific North	###	#####	AU-Aleutian G	11	#####	In Service
ICL Commun	AU-Aleutian Sub Pacific North	###	#####	AU-Aleutian H	12	#####	In Service
ICL Commun	AU-Aleutian Sub Pacific North	###	#####	AU-Aleutian I	13	#####	In Service
ICL Commun	AU-Aleutian Sub Pacific North	###	#####	AU-Aleutian J	14	#####	In Service
RU Holdings	Curie	West Coast	#####	Curie 1.1	12	#####	In Service
Eversource	Eversource HEE	East Coast	#####	Eversource	13	#####	In Service
RU Holdings	Firmina	East Coast	#####	Firmina	14	#####	In Service
RU Holdings	Grace Hopper	East Coast	#####	Grace Hopper	15	#####	In Service

status	count
In Service	
Planned	
Out of Service	

Figure 1-5. A possible place to insert data values.

- Format as a table by selecting all the cells involved and clicking on Format as Table, as described previously. Then, rename the table to StatusTable ([Figure 1-6](#)).

The screenshot shows a small table titled "StatusTable". It has two columns: "status" and "count". The "status" column contains three entries: "In Service", "Planned", and "Out of Service". The "count" column is currently empty. The table is displayed in a grid format with horizontal and vertical lines.

status	count
In Service	
Planned	
Out of Service	

Figure 1-6. The StatusTable with the possible states.

As a result, you have built a table with the following values:

- The first column contains the status with the three values (In Service, Planned, and Out of Service).

2. The second column contains only the header (named count), which will include the number of items in the MainTable for each status.

To populate the count column values automatically, you must calculate the number of cables in the In Service state using Excel formulas. To access a formula, select it from the Formulas tab of the Ribbon or proceed as follows:

1. Double-click the first cell of the count column and insert the equals (=) symbol. This opens formula mode, where you can enter any formulas you like.
2. You will use a *conditional* count, which counts rows from the status column of the main table only if they satisfy a criterion. Enter the following formula and press Enter:

```
=COUNTIF(MainTable[status]; [@status])
```

The formula means that we use the **COUNTIF** function to count the cells in the status column of the MainTable table only if they meet a particular criterion specified by the function's second parameter. The criterion is this: a cell is counted only if its value matches the current row's value in the status column. In other words, the **[@status]** function refers to the value of the status column in the currently processed row within the table. So, the formula is counting how many times the value of the current row's status column appears in the entire status column of the MainTable table.

To make the formula effective, click the cell where you wrote the formula. A small square appears at the bottom right of the cell. Drag the cell down to the last row of the StatusTable. The fields will be automatically populated with the correct values. [Figure 1-17](#) shows the final result.

status	count
In Service	218
Planned	4
Out of Service	44

Figure 1-7. The StatusTble filled with missing values.

You can achieve the same result as the previous example using pivot tables. A *pivot table* is a table for dynamically summarizing, analyzing, exploring, and presenting data. Follow these steps:

1. Select the Insert tab from the Ribbon and then Pivot Table.
2. In the open window, select the status column (click the letter H to select the entire column) and then click OK.
3. In the box that opens on the right, select the status column and drag it into the Σ Values area. Excel automatically populates the table, adding the states we had not considered in the previous example.

[Figure 1-8](#) shows the newly built pivot table.

Row Labels	Count of status
Abandoned	46
Active	1
In Service	218
Out of Service	44
Planned	4
Under Constructi	1
(blank)	
Grand Total	314

Figure 1-8. The PivotTable for the different statuses.

Excel for Data Visualization

A third use of Excel is *data visualization*, which involves representing data through graphs, such as lines, bar graphs, or pie charts. Follow these steps to draw a bar chart of the pivot table:

1. Click anywhere inside your pivot table.
2. Select the Insert tab on the Excel Ribbon.
3. In the Charts group, click Histograms → 2D.

[Figure 1-9](#) shows the resulting chart.

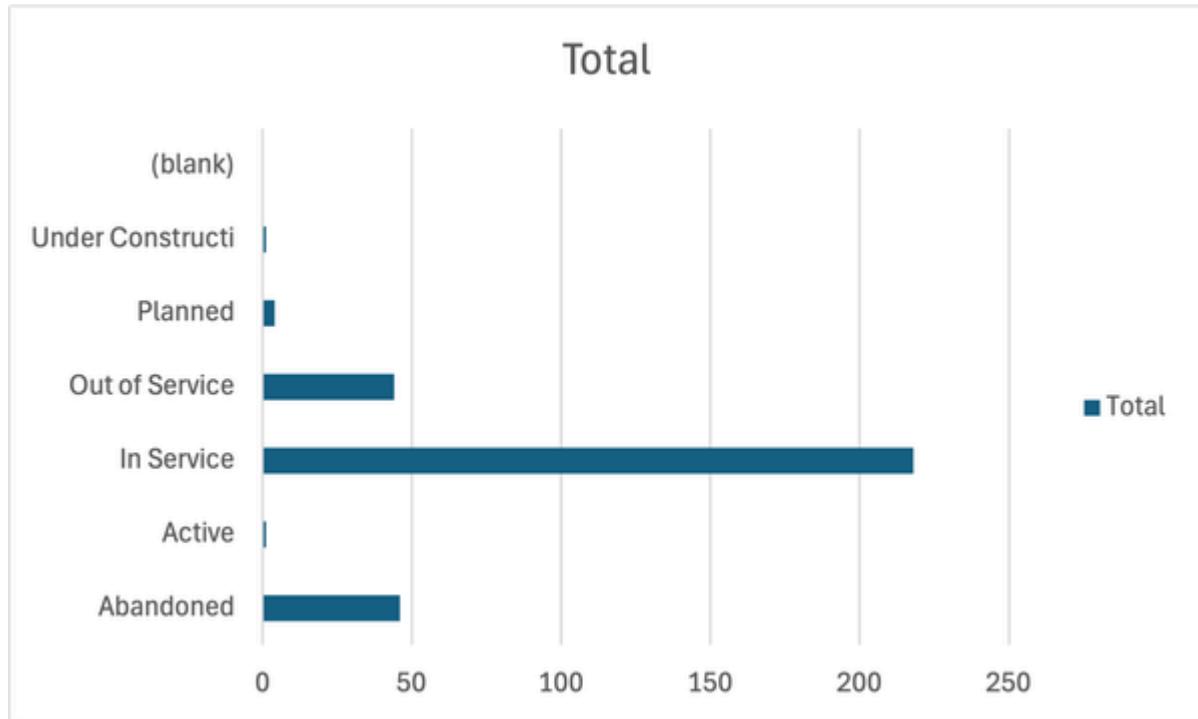


Figure 1-9. The bar chart built from the pivot table.

This simple exercise has demonstrated some of the main features provided by Excel:

- Importing data in CSV format and transforming it into the XLSX format
- Formatting the data as a table
- Naming the table
- Calculating a simple formula
- Using a pivot table
- Creating a chart in Excel

Excel can also be used for other purposes, such as automating processes and creating financial models. Throughout the book, we will explore these aspects and see how to best use AI to manage them.

Exercise

Solve the following exercise using the Submarine Cable dataset:

1. Create a new worksheet by clicking the + symbol at the bottom of your page
2. Rename the worksheet as Pivot Table, using the right-click menu to select Rename.
3. In the new worksheet, create a pivot table showing the number of cables, organized by owner, for each status.

The solution to the exercise is in the GitHub repository of the book, under [01/ch01_solution.xlsx](#).

For a long time I ignored the existence of tables in Excel, working completely “by hand”. “Working by hand” meant exactly this: applying line after line the colors of my template, the formulas and so on. Then, one fine day I discovered tables, and I understood their importance, when I realized that adding or removing rows to the table was very simple: everything updated automatically. If you haven’t tried tables yet, I suggest you do so soon and you’ll see that your Excel life will improve, as mine has.

Why Does Excel Need AI?

So far you have seen Excel can be used for three different purposes: data management, data analysis and data visualization. Indeed, there are other possible uses of Excel as well. [Figure 1-10](#) illustrates some of the many common uses of Excel.

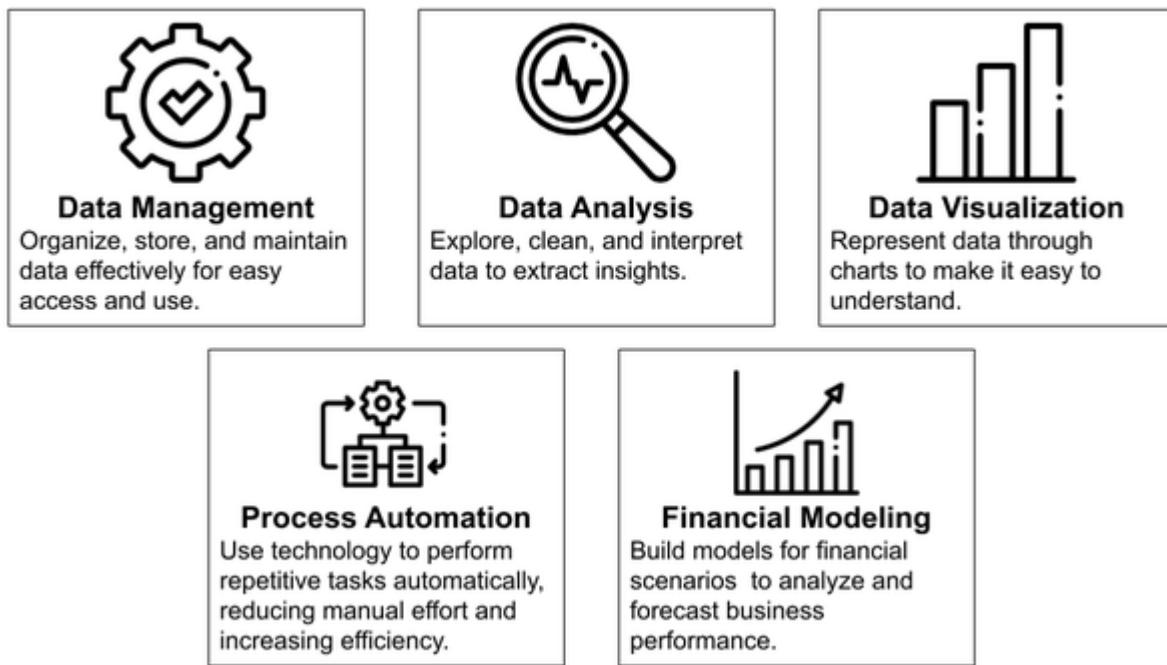


Figure 1-10. Some of the multiple purposes of using Excel.

Here's a bit more about those uses of Excel:

- Data management involves using Excel tables to store and organize data, making it easier to sort and filter. For example, you can create a table with customer information and use filters to organize the data.

- Data analysis involves extracting insights from a dataset, for example, applying and analyzing pivot tables.
- Data visualization involves creating charts, such as trends in sales data over time.
- Process automation automates repetitive tasks, such as using macros to format and update tables.
- Financial modeling creates a financial forecast, for example, to predict future cash flows.

Excel is already a complete data exploration and analysis tool, so extending it with external tools such as AI might seem excessive. However, there are at least two good reasons to integrate AI tools into Excel.

The first reason is the increase in *productivity*, the measure of the efficiency with which you carry out activities. Greater productivity means obtaining more significant results with less resource and time consumption.

Combining Excel and AI helps you do the following:

- Carry out simple and complex tasks or suggest the steps to carry them out if you are new to Excel
- Automate repetitive tasks, such as data cleansing, formatting, or applying formulas, to save time and effort
- Process large quantities of data to extract insights
- Suggest appropriate visualizations for the data

The second reason to integrate AI into Excel concerns improving the *accuracy* of the results—that is, the ability to produce correct and precise results without errors:

- Reduce human errors, which can arise when you enter inputs or calculate values manually
- Recognize anomalies that could go unnoticed by the human eye.

Let's look at an example to see how to combine AI and Excel to improve productivity. Consider the dataset from the previous exercise in which you calculated the number of occurrences for each cable status. Let's do the same exercise again using ChatGPT, one of the most popular AI assistants developed by OpenAI.

Start ChatGPT. If you don't have an account yet, visit <https://chatgpt.com/> and proceed with registration. In this book, I consider the Plus version of ChatGPT, which requires a monthly subscription, but the suggestions given also apply to the free version, even if the results may be slightly worse. Formulating correct and precise instructions is essential to using ChatGPT and Excel correctly. We will see the structure of the various instructions later in the book.

To interact with ChatGPT, you use instructions called prompts. A *prompt* is a text that describes what ChatGPT should do, and *prompt engineering* is

the art and science of defining appropriate and effective prompts. There are several techniques for defining prompts:

Zero-shot prompting

This consists of giving ChatGPT a task without providing examples. Use this technique if you want a quick answer. For example, write:
What is Microsoft Excel?.

One-shot prompting

Provide a specific example to tell ChatGPT how to respond. For example, write: Microsoft Word is a word processing program developed by Microsoft, which allows you to create, edit, and format text documents professionally. Based on the previous definition, give a Microsoft Excel definition.

Few-shot prompting

Give the model multiple examples of how to complete a task. This helps the model better understand the context and format of the desired response. For example, you might provide several examples and then ask for the Microsoft Excel definition.

Instruction-based prompting

Give explicit instructions, often in natural language, to clarify the task. For example, you might say, *Explain what Microsoft Excel is to*

a nine-year-old.

Chain of Thought (CoT) prompting

Guide ChatGPT to think aloud, breaking down the reasoning step by step, before arriving at an answer. This helps ChatGPT solve complex tasks that require logical reasoning. For example, you could prompt: *To explain what Microsoft Excel is, first describe what spreadsheets are, then explain how Excel uses them to organize data, and finally give an example of how it might be used to make a budget.*

Dynamic prompts

Adapt the prompt based on ChatGPT's previous answers. This is an interactive approach in which the prompt evolves as the conversation context changes. For example, if ChatGPT just explained what a spreadsheet is, you might build on that answer and ask, *Based on your explanation, how might a spreadsheet be used to analyze a company's expenses?*

[Figure 1-11](#) summarizes the techniques described.

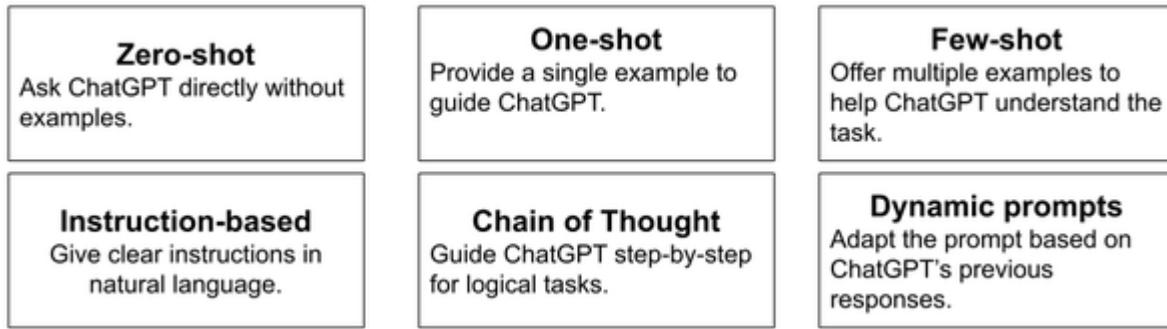


Figure 1-11. The different prompt techniques.

We will see and use these various techniques throughout the book. For now, let's do a test to evaluate ChatGPT's potential.

Excel and AI for Data Management

In the menu at the top left, select the model used by ChatGPT that is best suited to perform the calculations (ChatGPT version 4o, o1, or o1-mini at the time of writing this book). This exercise gives the data management to ChatGPT, so we load the dataset directly into ChatGPT.

Excel and AI for Data Analysis

There are two ways of using Excel and ChatGPT. The first consists of carrying out all the analyses in ChatGPT and exporting the final file. The second is to use ChatGPT as an assistant to suggest formulas or procedures for carrying out tasks. This book uses a mixed approach: in some cases, we will use ChatGPT to conduct the analysis, and in others, we will use it as an assistant. You can choose from time to time, based on the situations that arise, which approach to follow. When we talk about Copilot in the second

part of the book, we will perform all the operations within Excel, given that Copilot is already integrated into Excel.

Let's use the first approach, carrying out all the analysis in ChatGPT and then exporting the final file. Perform the following operations:

1. Load the dataset [ch01.xlsx](#) in ChatGPT, using the paperclip, as shown in [Figure 1-12](#).

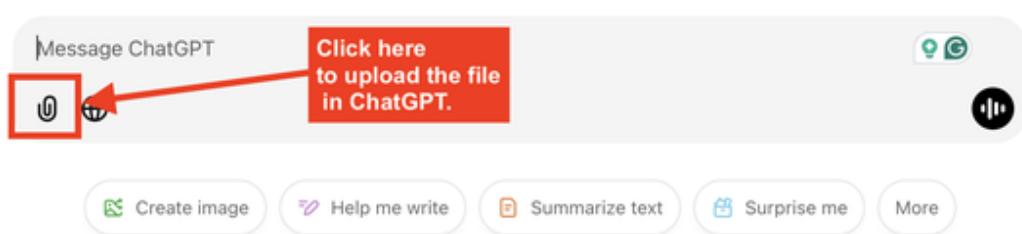


Figure 1-12. How to load a file in ChatGPT.

2. Write the following prompt in the ChatGPT bar (after loading the dataset):

```
Build a pivot table for the column status and c
```

3. Press the Enter key. ChatGPT starts thinking. When it has completed processing, it displays descriptive text and a table similar to the one shown in [Figure 1-13](#).

Status Count Pivot Table		
	status	0
1	Abandoned	46
2	Active	1
3	In Service	218
4	Out of Service	44
5	Planned	4
6	Under Constructi	1

Figure 1-13. The pivot table built by ChatGPT to count different statuses.

The pivot table calculated by ChatGPT is precisely the same as the one built by Excel in [Figure 1-8](#), with a different formatting style.

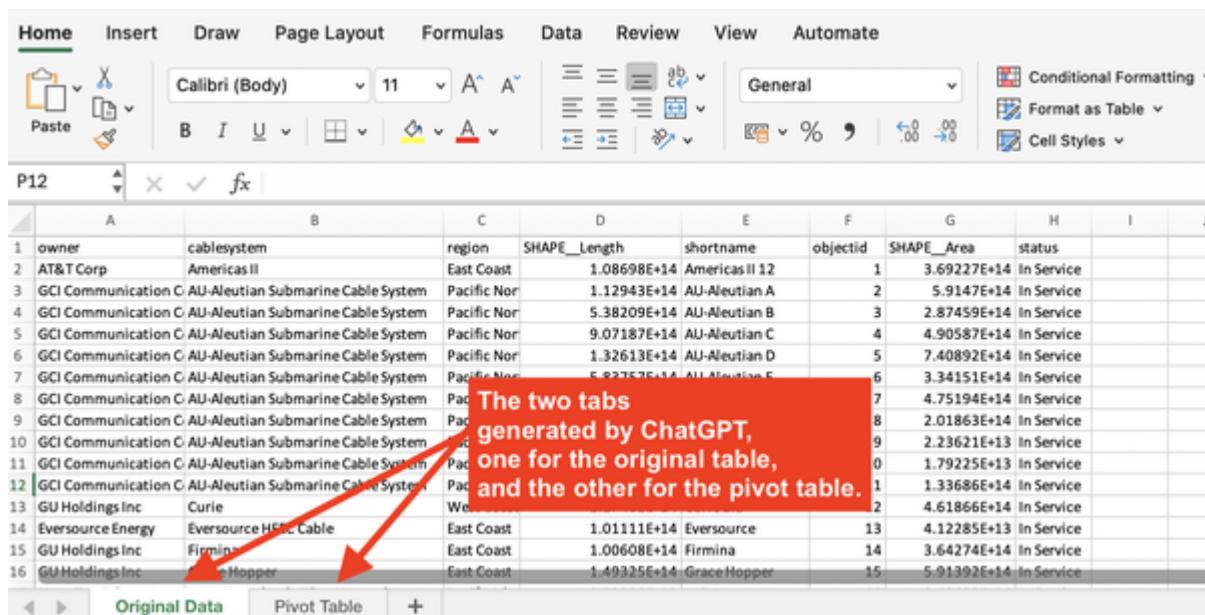
The next step is to export the output as an Excel file. Write the following prompt in ChatGPT, still in the same chat where you wrote the previous prompt:

Format the original table and the pivot table as

With this prompt, we ask ChatGPT to:

- Combine the original table and the generated pivot table in the same Excel file but under two different worksheets.
- Calculate the pivot table without hardwiring the numerical values into the table but applying the formulas to calculate them.
- Use the original table's H (status) column to perform calculations in the pivot table. This statement helps ChatGPT calculate the formula precisely.
- Make the file downloadable in Excel format.

As an output, ChatGPT generates a downloadable file. Download it and open it in Excel ([Figure 1-14](#)).



The two tabs generated by ChatGPT, one for the original table, and the other for the pivot table.

	A	B	C	D	E	F	G	H	I	J
1	owner	cablesystem	region	SHAPE_Length	shortname	objectid	SHAPE_Area	status		
2	AT&T Corp	Americas II	East Coast	1.08698E+14	Americas II 12	1	3.69227E+14	In Service		
3	GCI Communication C	AU-Aleutian Submarine Cable System	Pacific Nor	1.12943E+14	AU-Aleutian A	2	5.9147E+14	In Service		
4	GCI Communication C	AU-Aleutian Submarine Cable System	Pacific Nor	5.38209E+14	AU-Aleutian B	3	2.87459E+14	In Service		
5	GCI Communication C	AU-Aleutian Submarine Cable System	Pacific Nor	9.07187E+14	AU-Aleutian C	4	4.90587E+14	In Service		
6	GCI Communication C	AU-Aleutian Submarine Cable System	Pacific Nor	1.32613E+14	AU-Aleutian D	5	7.40892E+14	In Service		
7	GCI Communication C	AU-Aleutian Submarine Cable System	Pacific Nor	5.92757E+14	AU-Aleutian E	6	3.34151E+14	In Service		
8	GCI Communication C	AU-Aleutian Submarine Cable System	Pac			7	4.75194E+14	In Service		
9	GCI Communication C	AU-Aleutian Submarine Cable System	Pac			8	2.01863E+14	In Service		
10	GCI Communication C	AU-Aleutian Submarine Cable System	Pac			9	2.23621E+13	In Service		
11	GCI Communication C	AU-Aleutian Submarine Cable System	Pac			0	1.79225E+13	In Service		
12	GCI Communication C	AU-Aleutian Submarine Cable System	Pac			1	1.33686E+14	In Service		
13	GU Holdings Inc	Curie	We			2	4.61866E+14	In Service		
14	Eversource Energy	Eversource HFC Cable	East Coast	1.01111E+14	Eversource	13	4.12285E+13	In Service		
15	GU Holdings Inc	Firmina	East Coast	1.00608E+14	Firmina	14	3.64274E+14	In Service		
16	GU Holdings Inc	Grace Hopper	East Coast	1.49325E+14	Grace Hopper	15	5.91392E+14	In Service		

Figure 1-14. The Excel file generated by ChatGPT.

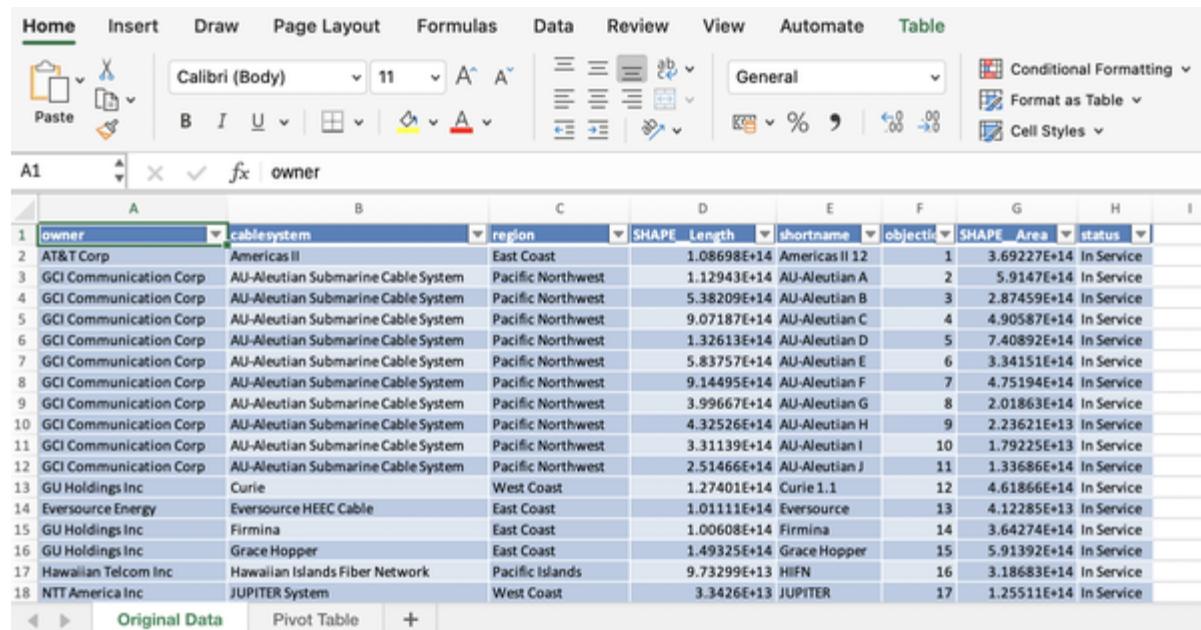
The downloaded file is very similar to the one we produced manually, with some differences:

- The tables are not formatted like Excel tables.
- The pivot table is not a pivot table, as Excel usually calculates it, but it contains formulas to calculate the values correctly.

In any case, it's not a bad starting point. In the same ChatGPT chat, write a new prompt, which is a modified version of the previous one, so that ChatGPT formats the data as Excel tables:

Format the original table and the pivot table as

Download the generated file and open it with Excel. The data is now formatted as tables, as shown in [Figure 1-15](#).



owner	cablesystem	region	shape	length	shortname	objectid	shape	area	status
AT&T Corp	Americas II	East Coast	1.08698E+14	Americas II 12	1	3.69227E+14	In Service		
GCI Communication Corp	AU-Aleutian Submarine Cable System	Pacific Northwest	1.12943E+14	AU-Aleutian A	2	5.9147E+14	In Service		
GCI Communication Corp	AU-Aleutian Submarine Cable System	Pacific Northwest	5.38209E+14	AU-Aleutian B	3	2.87459E+14	In Service		
GCI Communication Corp	AU-Aleutian Submarine Cable System	Pacific Northwest	9.07187E+14	AU-Aleutian C	4	4.90587E+14	In Service		
GCI Communication Corp	AU-Aleutian Submarine Cable System	Pacific Northwest	1.32613E+14	AU-Aleutian D	5	7.40892E+14	In Service		
GCI Communication Corp	AU-Aleutian Submarine Cable System	Pacific Northwest	5.83757E+14	AU-Aleutian E	6	3.34151E+14	In Service		
GCI Communication Corp	AU-Aleutian Submarine Cable System	Pacific Northwest	9.14495E+14	AU-Aleutian F	7	4.75194E+14	In Service		
GCI Communication Corp	AU-Aleutian Submarine Cable System	Pacific Northwest	3.99667E+14	AU-Aleutian G	8	2.01863E+14	In Service		
GCI Communication Corp	AU-Aleutian Submarine Cable System	Pacific Northwest	4.32526E+14	AU-Aleutian H	9	2.23621E+13	In Service		
GCI Communication Corp	AU-Aleutian Submarine Cable System	Pacific Northwest	3.31139E+14	AU-Aleutian I	10	1.79225E+13	In Service		
GCI Communication Corp	AU-Aleutian Submarine Cable System	Pacific Northwest	2.51466E+14	AU-Aleutian J	11	1.33686E+14	In Service		
GU Holdings Inc	Curie	West Coast	1.27401E+14	Curie 1.1	12	4.61866E+14	In Service		
Eversource Energy	Eversource HEEC Cable	East Coast	1.01111E+14	Eversource	13	4.12285E+13	In Service		
GU Holdings Inc	Firmina	East Coast	1.00608E+14	Firmina	14	3.64274E+14	In Service		
GU Holdings Inc	Grace Hopper	East Coast	1.49325E+14	Grace Hopper	15	5.91392E+14	In Service		
Hawaiian Telcom Inc	Hawaiian Islands Fiber Network	Pacific Islands	9.73299E+13	HIFN	16	3.18683E+14	In Service		
NTT America Inc	JUPITER System	West Coast	3.3426E+13	JUPITER	17	1.25511E+14	In Service		

Figure 1-15. The Excel file generated by ChatGPT with formatted tables.

The problem of the pivot table remains. Later in the book, you will see how to deal with a problem like this. For now, I hope I've convinced you that it's at least worth using ChatGPT in Excel.

Excel and AI for Data Visualization

Write the following prompt, always in the same ChatGPT chat:

```
Use Excel tools to build a bar chart for the pivot table.
```

Download the generated file. The results are very similar to those you obtained manually, as shown in [Figure 1-16](#), with the only difference being the pivot table (which is not a classic Excel pivot table).

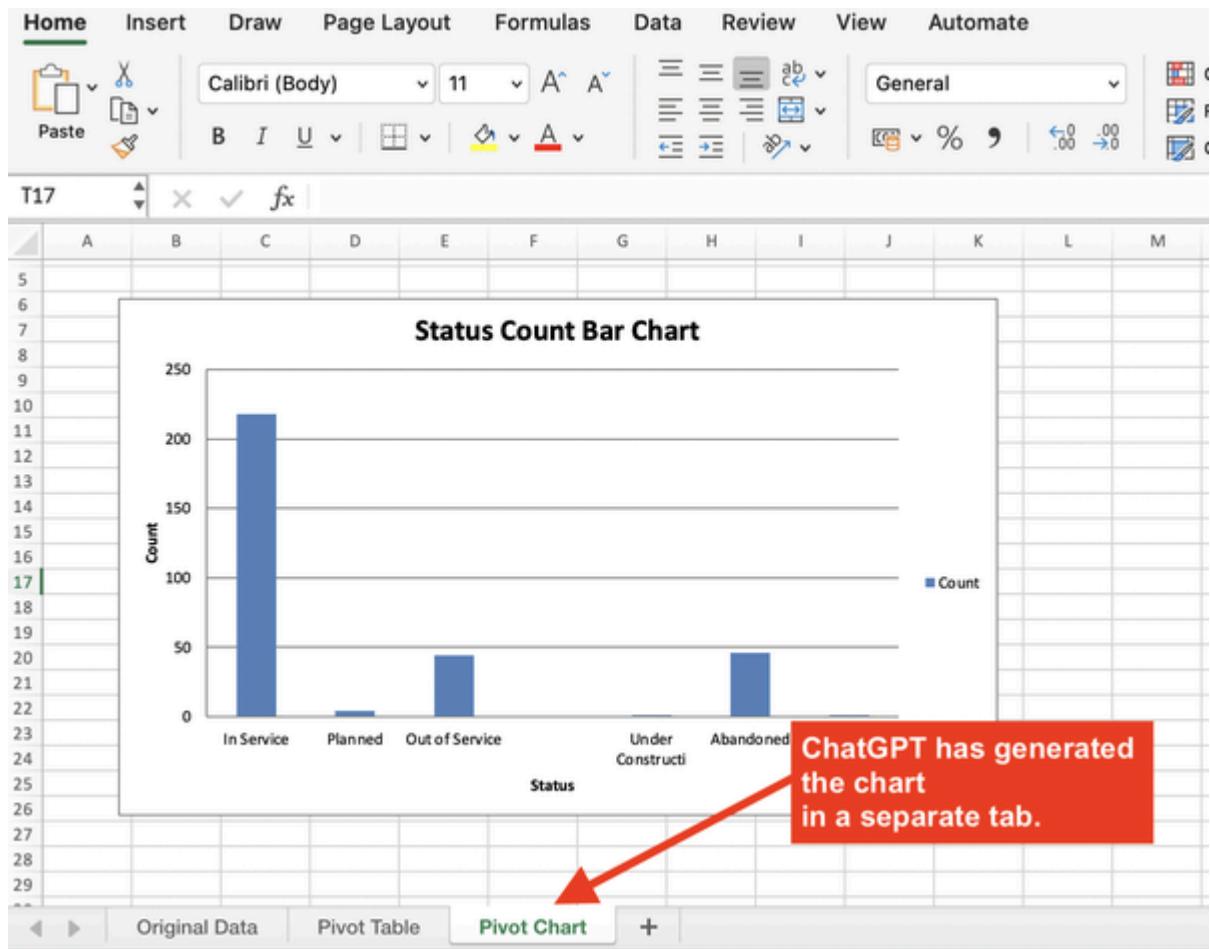


Figure 1-16. The Excel file generated by ChatGPT with the column chart.

With this simple example, you have seen one of ChatGPT's two uses with Excel: automatic data analysis. The other type of use, which we will also see throughout the book, is as an assistant for suggesting formulas or procedures for data analysis.

Having tried this exercise, I hope ChatGPT has not disappointed you, and that I have convinced you to continue reading this book. If you were faster than ChatGPT, it doesn't mean that ChatGPT isn't helpful to you, but rather that you already know Excel well, and the integration with AI will only

further improve your skills, as we'll see later in the book. For now, I guarantee you that using AI in Excel has benefits. Let's look at them.

The Benefits of Using AI in Excel

You may think that you need to use AI as your substitute in the sense that it automatically carries out tasks for you. In reality, in any field in which you use AI, including Excel, AI is always an assistant that can help you carry out your tasks more efficiently and accurately, but it can never wholly replace you. The last word will always be yours, and your task will be to ensure that the AI does not make gross errors, including hallucinations and biases, as we will see later in the book.

NOTE

AI *hallucinations* happen when the AI generates fake information. *Bias* is when the AI produces unfair or one-sided results.

In this book, I will talk about artificial intelligence, but not in the sense of an automatic process that completely replaces humans. We will not be getting into sci-fi robots that make autonomous decisions and ultimately destroy the universe. In this book, you will learn some strategies to use AI to augment *your* intelligence.

AI is nothing other than a significant part of human knowledge contained in a tool. Much of the human effort to understand how the world works is contained in a tool with impressive potential. This is fantastic and very powerful, but also risky. Therefore, the main benefit of using AI in Excel is not so much automating tasks and having the machine carry them out autonomously as it is to tap into a vast available knowledge base, which is mankind's heritage.

Your task is to learn to use this treasure in the best ways possible, and I will try to help you with that throughout the book. It is, therefore, not a question of acquiring concepts to automate more or less complex tasks, but instead a question of developing a mentality that will allow you to think with AI. You will always be the protagonist of the choices and analyses. AI will be your faithful travel companion at your service, helping you speed up tasks and be as precise as possible.

EVALUATING AI RESULTS

To evaluate whether the AI is working well on your data, you can apply supervised learning of the AI without making errors. This involves testing the AI results on a subset of your data. Manually perform your calculations on some of the data using Excel and then repeat the same calculations using AI. Once you have consolidated the prompts for the AI and ensured that they respond correctly, let the AI perform the task on the entire dataset.

When AI Is Not Useful in Excel

AI is a potent tool. However, at least two cases indicate that its use is not yet mature (at least as of the writing of this book):

Large datasets

When working with large datasets, AI can encounter some difficulties regarding performance and accuracy. The analysis of large datasets requires many resources and computational capabilities. An AI model such as ChatGPT could not provide these resources to ordinary users. For this reason, if you still want to use AI with large datasets, split the dataset into small parts that AI can process. For example, you could filter out unnecessary columns or rows from a dataset and provide only the data you want to analyze as input to AI.

Low-quality data

If you provide incomplete, noisy, or poorly structured data as input to AI, the quality of results can negatively impact model performance, leading to incorrect predictions or distorted interpretations. AI will have difficulty providing valuable and accurate outputs if data is missing, duplicated, or contains systematic errors. This is the case to repeat the classic maxim *garbage in, garbage out*, which means you can't extract a significant output from a bad input.

Despite these limitations, AI's potential continues to grow, and as technologies and infrastructure advance, improvements in these areas are expected.

Browsing Among the Different AI Tools for Excel

The introduction of AI into Excel has opened up new opportunities to simplify working with data, improve efficiency, and reduce human errors. In recent years, numerous add-ins have been developed that integrate AI directly into Excel, each with different characteristics to meet specific needs. Many of them may require a subscription. This section looks at some of the significant AI tools for Excel:

- *Numerous.ai*: [Numerous.ai](#) is an add-in that uses AI to make predictions, clean data, and generate new insights. It can summarize, categorize, classify, and cleanse data and texts. Numerous.ai defines a specific function called `AI()`, which you can use to prompt ChatGPT directly in your worksheet.
- *GPT for Excel*: [GPT for Excel](#) is an add-in that uses AI for data cleaning and classification, text generation and summarization, and translations. It defines two main functions:
 - `GPTINTERACT()` takes as input instructions and the data to process.

- `GPTPREDICT()` takes some input example data, outputs example data, and data to be processed.
- *AutoPilot*: [AutoPilot](#) is an add-in that integrates different AI models in Excel, such as ChatGPT, Google Gemini, Anthropic Claude, and more. It allows you to automate repetitive tasks such as data cleansing, applying complex formulas, and generating graphs.
- *PowerGPT*: [PowerGPT](#) is an add-in that uses AI to answer questions from your data, decipher any formula with explanations in plain English, generate formulas, create SQL queries from simple descriptions, generate regular expressions, and so on.
- *AI-aided Formula Editor*: The [AI-aided Formula Editor](#) supports creating and optimizing complex formulas in Excel. It can also explain and format formulas. The command `write to cell` enables you to copy the generated formula into a cell quickly.
- *Sally Suite*: [Sally Suite](#) is an add-in that uses AI to automate repetitive processes, create sheets, write functions, filter data, create or recommend charts, translate sheets, perform data analysis, and even help generate Python code.
- *Ghost Writer*: [Ghost Writer](#) is an add-in for generating sample data, translating sheets, writing prompts, generating formulas, and automating documentation. It is also available for Microsoft Word and PowerPoint, which can generate reports, write documents, create inline comments on your content, and more.

The various add-ins are constantly updated, so new features that are not described in this book may become available. I always suggest you check each tool's official page before installing it.

Which tool(s) should you choose? It depends on your needs. This book does not use any specific add-ins (unless explicitly stated), but we will rely on ChatGPT in the first part and Copilot in the second part. I haven't talked much about Copilot so far because It is not an add-in but a tool integrated with Excel. We will see this later in the book.

NOTE

I would like to end this chapter with a small side note. When I was little, like many children, I had an imaginary friend to whom I told all my fears and into whom I projected all my childhood fears, doubts, and uncertainties. Unfortunately, since many years have passed since my childhood, I no longer remember the name of this imaginary friend. Today, however, I can say that I have a new friend, not exactly imaginary, but almost: ChatGPT (or AI in general), with whom I can interact in various ways. In this book, we'll see how to combine AI and Excel, but I've even read posts on social media from people who use AI as a therapist. Well, in this book we will not use ChatGPT and AI as a therapist, but you can think of AI as your new imaginary friend, which unlike the one from your childhood, responds to your requests and is of great help to you. All you need to do is discover the key to open the door to this new imaginary friend.

Summary

Our journey together to discover the use of AI in Excel has just begun. In this chapter, you've seen why it makes sense to use AI in Excel and went through a brief example demonstrating it. In the next chapter, you will learn how to use ChatGPT within your Excel sheets best.

- Integrating AI into Excel simplifies data management, analysis, visualization, and more tasks, improving overall efficiency.
- AI in Excel improves productivity and accuracy. AI allows you to perform tasks faster and more precisely. It also reduces human errors, ensuring more accurate and reliable results.
- Despite its many benefits, AI is not yet fully mature when processing massive datasets without advanced computational resources.
- Even the most powerful AI algorithm only works with quality data. Incomplete or noisy data can lead to misleading predictions or biased analyses.
- AI in Excel should be seen as an assistant to improve performance, not a complete replacement for human work. You should supervise the results generated by the AI to avoid errors.

Activity

Solve the crossword puzzle in Figure 1-17. After completing it, you'll read the name of the AI tool you'll use in the next chapter. The solution is in [Appendix A](#).

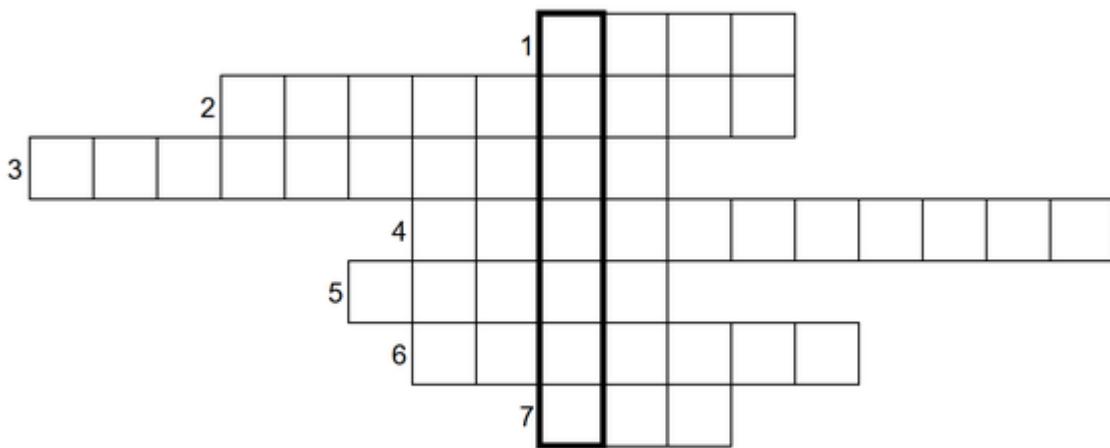


Figure 1-17. Crossword puzzle to reveal the tool you will use next

Definitions:

1. The basic unit of a table, located at the intersection of a row and a column of the table, where you can enter data or formulas.
2. A sheet within a work file, organized into rows and columns, where data can be inserted, viewed, and manipulated.
3. An add-in for Excel providing the `AI()` function to prompt ChatGPT directly.

4. An add-in for Excel providing the `GPTINTERACT()` function, which takes as input instructions and the data to process.
5. A set of selected cells.
6. The built-in AI tool provided by Excel. You'll use it in the second part of this book.
7. A section in the Ribbon that groups related commands in Excel.

Chapter 2. Getting Started with ChatGPT and Excel

A NOTE FOR EARLY RELEASE READERS

With Early Release ebooks, you get books in their earliest form—the author’s raw and unedited content as they write—so you can take advantage of these technologies long before the official release of these titles.

This will be the 2nd chapter of the final book. Please note that the GitHub repo will be made active later on.

If you have comments about how we might improve the content and/or examples in this book, or if you notice missing material within this chapter, please reach out to the editor at ccollins@oreilly.com.

In high school, I had to memorize trigonometry formulas by heart to solve math exercises. These were also quite complex formulas that contained sines and cosines, additions and subtractions, and so on. After a lot of work, I managed to learn them perfectly and was able to solve the trigonometry exercises perfectly. Periodically, however, I had to review the formulas aloud—otherwise, I would forget them. I was 17 at the time, and my brain was very trained. More than twenty years have passed since then, and I

remember very little of those trigonometry formulas. Our brains tend to forget what we don't use, especially if it involves mnemonic formulas.

The same thing happens with all available Excel functions. It is impossible to remember all the functions that Excel makes available to carry out specific calculations. You may have a more or less general idea about the name of the function to use, but you often don't remember the syntax or the specific name. ChatGPT can help you with this, as it can tell you not only which function to apply each time but also how to apply it.

In this chapter, you'll see how to use ChatGPT to generate the formula or function you need. You will see a strategy for communicating with ChatGPT to make it generate precisely the output you want: not something general but specific and valid for your dataset. When I talk about *communicating*, I mean the literal meaning of transmitting a message to someone. The goal is to ensure that ChatGPT understands what you are talking about, generating output as close as possible to what you want it to develop. So, it's more like you want to communicate with ChatGPT. This communication strategy with ChatGPT involves organizing the conversation as a story, with its typical parts: the context, the main point, and the end. We will test this strategy for calculating essential Excel functions, such as calculation, logic, and text functions.

In this chapter, you will learn more about:

- Introducing ChatGPT for Excel
- The structure of a conversation
- Basic functions

Let's start with the first aspect: ChatGPT for Excel.

Introducing ChatGPT for Excel

There are two ways to combine ChatGPT and Excel, as shown in [Figure 2-2](#).

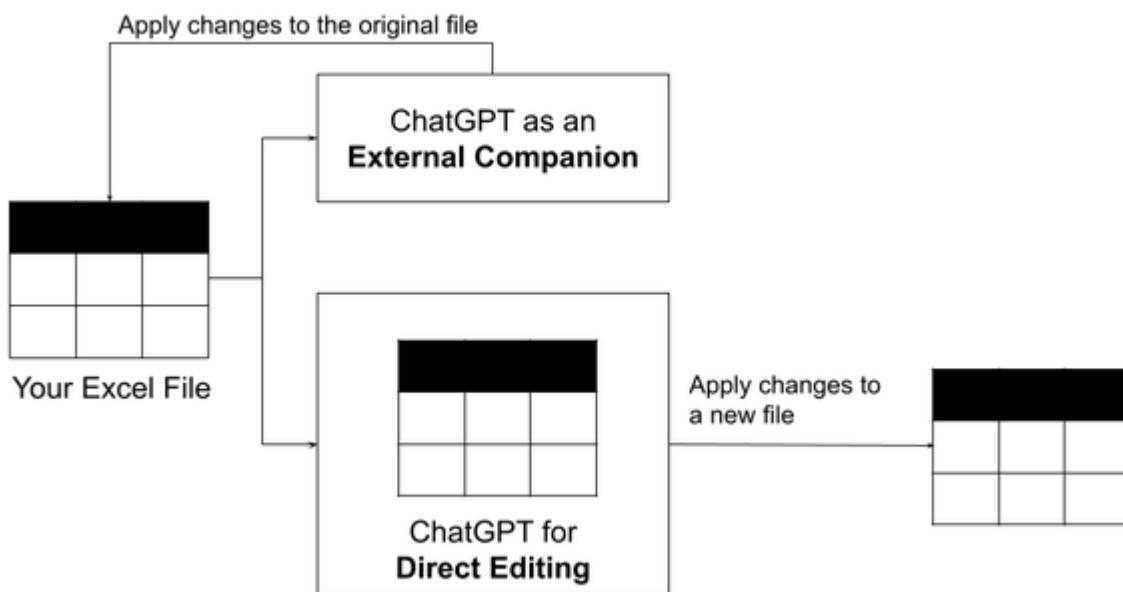


Figure 2-1. The two possible uses of ChatGPT in Excel.

In the first case, you can use ChatGPT as a *companion* or external tutor: ask specific questions about your dataset, and ChatGPT will help you answer

them. Once you have the answers, apply them to your Excel file. This working method is helpful if you need clarification about how to write formulas or which functions to use, or if you need help answering the questions your data contains.

The second way to use ChatGPT is as a *direct editing tool*. In this case, you load the file directly into ChatGPT and work directly inside ChatGPT. In the end, you download the file with the changes already made by ChatGPT.

WARNING

Remember that the suggestions provided by ChatGPT may not always be correct, so you should always check the suggestions or proceed iteratively to gradually achieve the desired result.

Whether you use ChatGPT as an assistant or as a direct editor, your job is to communicate with ChatGPT so that it understands what you want to talk about and helps you respond best. You can communicate with ChatGPT in two ways, as shown in Figure 2-2.

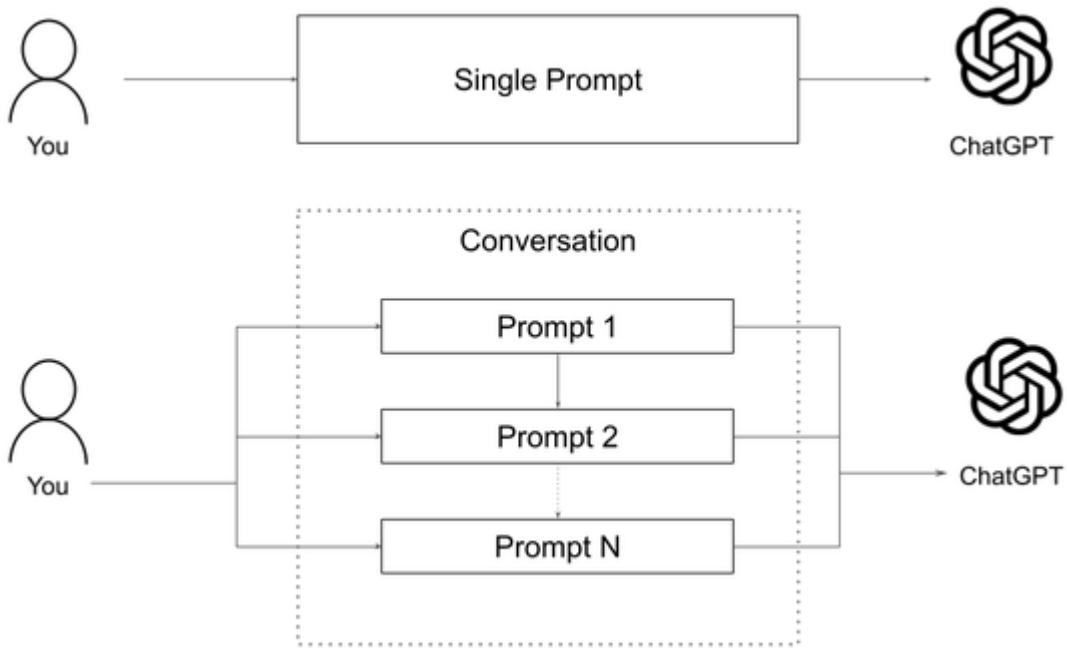


Figure 2-2. The two ways of communicating with ChatGPT.

A *single prompt* is a single input used to directly specify a request or question. It does not consider any previous context unless that context is made explicit in the prompt itself. For example: *Explain the formula for calculating the mean in Excel.*

A *conversation* is a set of multiple prompts connected to each other. It is an interaction between the user and ChatGPT that considers the session's previous context, allowing for a continuous and coherent dialogue. Furthermore, it enables you to refine or expand a previously provided answer or ask subsequent questions about the ongoing conversation. For example, you might start with *Write the formula to calculate the average in Excel* and then ask, *Explain the formula.*

The rest of this chapter focuses on using ChatGPT as an external assistant and working in conversation mode. However, everything I say about conversations also applies to prompts. At this point, we are ready to see the structure of a conversation (or a single prompt).

The Structure of a Conversation

Imagine that ChatGPT is a real interlocutor to whom you must say something. One important way humanity has communicated since ancient times is through stories. A *story* is defined primarily by a plot. You can organize your ChatGPT conversation as a simple story, where you are the narrator and ChatGPT is your audience, as shown in [Figure 2-3](#).

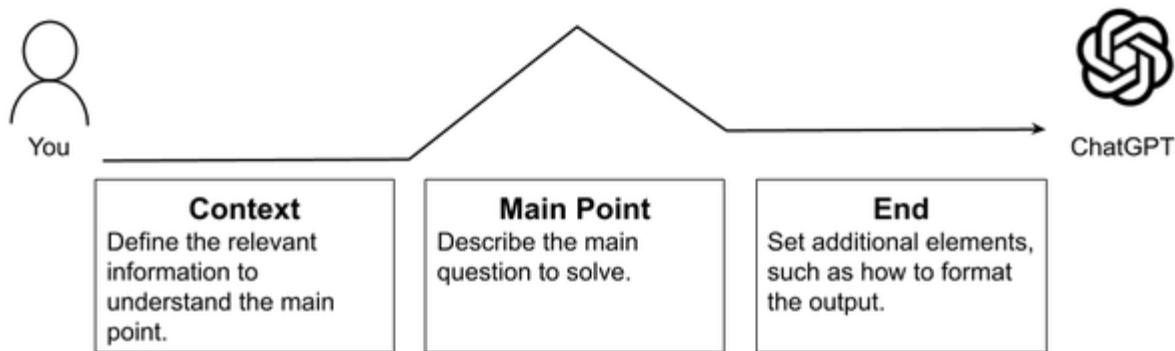


Figure 2-3. How to organize a conversation with ChatGPT in the form of a story.

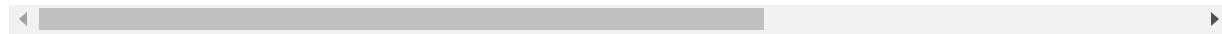
Each story begins by describing the *context*, the set of all the information necessary to ensure that the story's audience (in your case, ChatGPT) understands what you are talking about. Then you present your *main point*

to ChatGPT, such as calculating a formula or a specific function. Finally, you can ask ChatGPT to format the result in a certain way or to tell you why it used a specific formula.

Let's take a practical example to understand how to organize a conversation like a story. Suppose you have the dataset of [Amazon Top 50 Bestselling Books 2009 - 2019](#) available on Kaggle under the CC0: Public Domain license. The dataset contains 550 books categorized into fiction and nonfiction. [Table 2-1](#) shows an excerpt of the dataset.

Table 2-1. An extract from the dataset: the Amazon Top 50 Bestselling Books 2009-2019 dataset.

Name	Author	User Rating	Reviews	Price
10-Day Green Smoothie Cleanse	JJ Smith	4.7	17350	8
11/22/63: The Novel	Stephen King	4.6	2052	22
12 Rules for Life: An Antidote to Chaos	Jordan B. Peterson	4.7	18979	15
1984 (Signet Classics)	George Orwell	4.7	21424	6
5,000 Awesome Facts (About Everything!) (National Geographic Kids)	National Geographic Kids	4.8	7665	12



The dataset contains the following columns:

- *Name*: Name of the book
- *Author*: The author of the book
- *User Rating*: Amazon user rating. A number between 1 and 5
- *Reviews*: Number of written reviews on Amazon
- *Price*: The price of the book (as of 13/10/2020)
- *Year*: The Year(s) it ranked on the bestsellers
- *Genre*: Whether fiction or nonfiction.

You can download the dataset directly from Kaggle in CSV format or from the book's GitHub repository in XLSX format under [02/books.xlsx](#). In the latter case, the dataset is already formatted as an Excel table and named **books**.

Let's imagine we want to ask ChatGPT how to calculate the average price across all books. We structure the prompt like a story: context, main point, and end. Let's start with the context.

Setting the Context

The *context* corresponds to a text that contains all the information ChatGPT needs to frame your dataset. The context defines the following:

1. The topic you are discussing
2. The structure of the dataset with column names

3. Guidelines on how to work in Excel, such as specifying the table name

As an alternative to points 2 and 3, you can directly upload the Excel file to ChatGPT. However, I suggest passing the table structure and name if your file contains sensitive or large data.

TIP

ChatGPT, by default, uses your input and conversations to train and improve the AI models on which ChatGPT is based. You can disable data sharing and make your data private by following these steps:

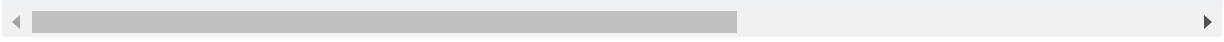
1. On the ChatGPT web page, click on your profile icon at the top right and then select Settings from the menu.
2. Inside the settings, look for the section Privacy and Data Settings and deactivate the option Data Sharing.
3. Be sure to save any changes if required.

Turning this option off will prevent ChatGPT from using your data used to improve models. This way, you can upload your datasets to ChatGPT without the risk of sharing sensitive content inappropriately.

In the case of the **books** dataset, organize the context into three parts.

First, define the topic you are dealing with:

Consider a dataset relating to the Amazon Top 50



Then specify the dataset structure with column names:

The dataset has the following columns:
Name - Name of the book
Author - The author of the book
User Rating - Amazon user rating. A number between 1 and 5.
Reviews - Number of written reviews on Amazon
Price - The price of the book (as of 2020/10/13)
Year - The Year(s) it ranked on the bestsellers
Genre - Whether fiction or non-fiction.

Finally, provide some guidelines on how to work in Excel with your dataset:

Suppose you import this dataset into Excel in a tabular format.

In total, the context prompt part should look like this:

Consider a dataset relating to the Amazon Top 50 books.
Name - Name of the book
Author - The author of the book
User Rating - Amazon user rating. A number between 1 and 5.
Reviews - Number of written reviews on Amazon
Price - The price of the book (as of 2020/10/13)
Year - The Year(s) it ranked on the bestsellers
Genre - Whether fiction or non-fiction.
Suppose you import this dataset into Excel in a tabular format.

This prompt is the first message of a more extended conversation, or you can aggregate it with the rest of the story, which we'll see shortly, and build a single prompt. Point to the ChatGPT website and give it the previous context prompt as input. To open the ChatGPT page follow the steps described below:

1. Connect on <https://chatgpt.com/>, login, and start a new conversation.
2. Copy the context prompt defined previously and check the result.

ChatGPT will respond with suggestions on what operations or types of analysis you can perform using this dataset. Out of pure curiosity, I fed this prompt to ChatGPT five times, starting a new chat each time, and I noticed that ChatGPT more or less suggests always doing the same operations, even if applied to different dataset columns. Operations include calculation of statistics, filtering and sorting operations, analysis, and visualizations of specific data.

Once you've defined the context, you can specify the main point of the prompt or conversation. We'll explore this further in the next section.

Defining the Main Point

The story's *main point* is the problem that ChatGPT needs to solve, or more generally, the question you want to ask ChatGPT. The structure of the main prompt comprises four main elements:

The action

This indicates the operation you want ChatGPT to perform. It's formulated clearly and directly. It includes verbs that describe what you want to achieve (e.g., *Write*, *Explain*, *Analyze*).

Objective

This describes your request's final result or purpose, that is, what you want to achieve through the action. It must be specific and linked to the problem to be solved or the task to be completed.

Additional details

These define additional information that has not been described in the context and should help ChatGPT to understand the request, such as available data, relevant parameters, or restrictions. This helps personalize the response by making it more appropriate.

The level of complexity

This specifies the necessary level of detail in the response, adapting it to the user experience. For example, it can be basic (for beginners) or advanced (for advanced users).

In some cases, as we will see shortly, you can omit some elements. For example, after defining the context as explained above, use this main point

prompt to write the formula in Excel to calculate the average price of books by genre:

- Action: Write
- Objective: the formula in Excel to calculate the average price of books.
- Additional details: -
- Level of complexity: -

Here is a possible output produced by ChatGPT (plus some explanatory text):

```
=AVERAGE(books[Price])
```

Explanation:

* books[Price]: Refers to the Price column in the books table.
* AVERAGE: Calculates the mean value of the numbers in the Price column.

As you can see, ChatGPT has generated a formula adapted to your dataset because it contains exactly the names of the books table and the columns (Genre and Price) that you previously set in the context.

NOTE

ChatGPT returns the formula in English by default. However, if Excel, in your case, is set in another language, you should also indicate the language in which Excel is used for ChatGPT. For example, if you use the Italian version of Excel, you can write the following prompt:

```
Write the formula for the Italian version of Excel to calculate the average price of books.
```

In this case, the generated formula is: `=MEDIA(books[Price])`.

You can adapt the prompt structure based on your use of ChatGPT.

Generally, you can structure a main point prompt differently based on the different uses of ChatGPT, as shown in [Figure 2-4](#).

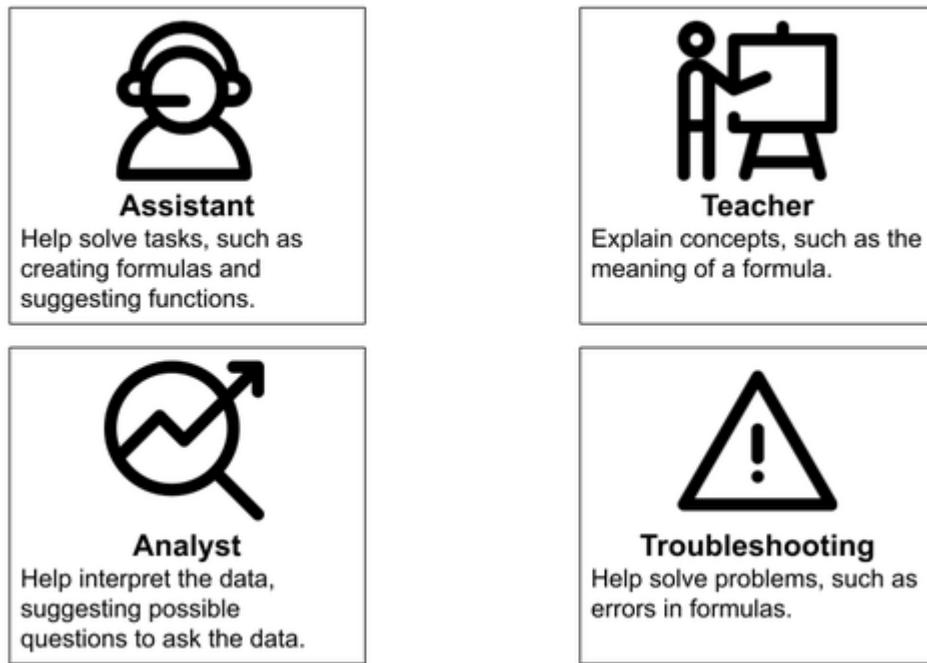


Figure 2-4. The roles ChatGPT can play in helping you answer your questions.

- As an *assistant*, ChatGPT can guide you through using Excel, helping you understand how to create formulas, use functions, or customize spreadsheets to fit your needs. For example, you could ask ChatGPT: *Write the formula in Excel to calculate the average price of books by genre.*
- As a *teacher*, ChatGPT can clarify complex Excel concepts, such as macros, pivot tables, and data analysis, making them more understandable. For example, you could ask ChatGPT *how pivot tables work to analyze the total number of reviews for each year in the dataset.*
- As an *analyst*, ChatGPT supports you in interpreting data, suggesting ways to visualize information through graphs, or identifying trends and patterns. For example, you might ask: *What are the five books with the highest average rating, considering only fiction books?*

- For *troubleshooting*, ChatGPT can help you fix common problems in Excel, such as formula errors, formatting problems, or difficulties importing data. For example, you could ask: *What is the error in this [formula]?*

Throughout the book, we will see how to define the various questions based on the task you must carry out in Excel. For now, it's enough to know that there are four types of main points.

The main prompt's contents change depending on the ChatGPT role, as described in [Table 2-2](#).

Table 2-2. The contents of the main prompt based on the ChatGPT role.

Role	Action	Objective	Additional Details	Level
				Context
Assistant	Clear and direct request. Verbs: Write, Create, Show, Calculate, Define	Get a formula or function to complete a specific task.	Specific context, such as columns or functions.	Not
Teacher	Request a detailed explanation or step-by-step guide. Verbs: Explain, Show, Describe, Guide, Illustrate	Understanding a complex concept or learning to use a function.	Add a practical example or data relevant to the concept explained	Specific level, beginner to advanced
Analyst	Request analysis or	Extract meaningful	Specific context, such as	Not but

Role	Action	Objective	Additional Details	Level
	visualization suggestions. Verbs: Analyze, Suggest, Identify, Visualize	trends, patterns, or insights from your data.	as columns.	help spec char
Troubleshooting	Request for assistance to resolve a specific problem. Verbs: Solve, Correct, Identify, Help, Explain	Request for assistance to resolve a specific problem.	Provide details such as the formula or context of the problem.	Not

Role	Action	Objective	Additional Details	Level
Assistant	Ask for a step-by-step guide.	Get a detailed response.	None	Intermediate
Teacher	Ask for a formula to calculate a specific value.	Get a formula and its explanation.	None	Intermediate
Analyst	Ask for a report on market trends.	Get a comprehensive market analysis report.	None	Advanced
Troubleshooting	Ask for troubleshooting steps for a software error.	Get a list of troubleshooting steps and their outcomes.	None	Intermediate

The table shows a structured guide for defining prompts for use with ChatGPT based on four main roles: Assistant, Teacher, Analyst, and Troubleshooting. For each role type, the table shows the structure of the main point prompt and an example of its use.

Once you have defined the prompt's main point, you can proceed to its final part, which corresponds to the end of the story.

Setting the End

The end of the story tells ChatGPT what to do after performing the main task. You have three main options. You can ask ChatGPT the following:

- What steps are required to generate the response in Excel? For example, if we ask ChatGPT to create a formula, we can now ask what the steps are to insert it into Excel.
- List alternative solutions.

- Calculate the answer to the question directly in an Excel file to export as output. You can only use this solution if you initially loaded the Excel file containing the entire dataset or a portion of it.

Returning to our example prompt, we can ask ChatGPT to list the steps for entering the formula into Excel using the following prompt (always within the same conversation that contains the context and main point):

```
List the steps to include the proposed formula in
```

ChatGPT generates a detailed list of steps. Here, I report only the top-level steps:

```
Step 1: Open the Dataset
Step 2: Convert Data to a Table (Optional)
Step 3: Select a Cell for the Formula
Step 4: Enter the Formula
Step 5: Verify the Result
Optional: Format the Result
Step 6: Save the File
```

At this point, you have learned how to communicate with ChatGPT. Organize your conversation like a story: context, main point, and end. You can structure your chat as a single prompt to feed to ChatGPT or as three separate prompts within the same conversation: one for the context, one for the main point, and one for the end.

TIP

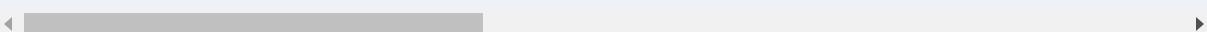
I suggest using three separate prompts in the same conversation so you can reuse the same context for multiple questions or main points.

So far, we have seen the generic structure of a prompt in ChatGPT. The more knowledge you have about the problem you want to solve, the better the results that will be produced. The next exercise looks at a practical example to see how best to structure prompts.

Exercise: Structuring Prompts

Consider again the dataset on Amazon Top 50 Bestselling Books 2009 - 2019. Start a new conversation with ChatGPT and write the following context prompt (defined above):

Consider a dataset relating to the Amazon Top 50
Name - Name of the book
Author - The author of the book
User Rating - Amazon user rating. A number between 1 and 5.
Reviews - Number of written reviews on Amazon
Price - The price of the book (as of 2020/10/13)
Year - The Year(s) it ranked on the bestsellers
Genre - Whether fiction or non-fiction.
Suppose you import this dataset into Excel in a 1000x1000 grid.



Then write the main point prompt for each of these tasks:

1. *Task 1:* Calculate the highest price
2. *Task 2:* Explain how to create a pivot table to analyze the total sum of reviews for each year and sort the results in ascending or descending order.
3. *Task 3:* Find the error in this formula: =AVERAGEIF(books[User Rating]>4.5, books[Price]) .

Let's proceed with the resolution of the exercise, starting with the first task.

Task 1

Use ChatGPT as an assistant and write the following prompt:

Write the formula in Excel to calculate the highest price of books.

ChatGPT should respond with the following output:

To calculate the highest price of books in the Price column, you can use the MAX function:
=MAX(books[Price])
Explanation:
books is the name of the table.
[Price] refers to the column named Price within the table.
MAX is the function that returns the highest value in a range.
Place this formula in any cell, and it will display the highest price of the books.

By pasting the formula suggested by ChatGPT into the Excel sheet, we get the result shown in Figure 2-5.

Name	Author	User Rating	Reviews	Price	Year	Genre
10-Day Green Smoothie Cleanse	JJ Smith	4.7	17350	\$8	2016	Non Fiction
11/22/63: A Novel	Stephen King	4.6	2052	\$22	2011	Fiction
12 Rules for Life: An Antidote to Chaos	Jordan B. Peterson	4.7	18979	\$19	2018	Non Fiction
1984 (Signet Classics)	George Orwell	4.7	21424	\$6	2017	Fiction
5,000 Awesome Facts (About Everything!) (National Geographic Kids)	National Geographic Kids	4.8	7665	\$12	2018	Non Fiction
A Dance with Dragons (A Song of Ice and Fire)	George R. R. Martin	4.4	12643	\$11	2011	Fiction
A Game of Thrones / A Clash of Kings / A Storm of Swords / A Feast of Crows / A Dance with Dragons	George R. R. Martin	4.7	19735	\$30	2014	Fiction
A Gentleman in Moscow: A Novel	Anja Towles	4.7	19699	\$15	2017	Fiction
A Higher Loyalty: Truth, Lies, and Leadership	James Comey	4.7	5983	\$3	2018	Non Fiction
A Man Called Ove: A Novel	Fredrik Backman	4.6	23848	\$8	2016	Fiction
A Man Called Ove: A Novel	Fredrik Backman	4.6	23848	\$8	2017	Fiction
A Patriot's History of the United States: From Columbus's Great Discovery to the War on Terror	Larry Schweikart	4.6	460	\$2	2010	
A Stolen Life: A Memoir	Jaycee Dugard	4.6	4149	\$22	2011	
A Wrinkle in Time (Time Quintet)	Madeleine L'Engle	4.5	5153	\$5	2018	
Act Like a Lady, Think Like a Man: What Men Really Think About Love, Relationships, Intimacy, and Commitment	Steve Harvey	4.6	5013	\$17	2009	

Figure 2-5. Figures 2-5. The result of Task 1.

Task 2

Use ChatGPT as a teacher and write the following prompt:

Explain how to create a Pivot Table in Excel to analyze the book data.

In my case, ChatGPT produced the following output, with a simple explanation of the steps to follow to create the required Pivot Table:

Here is a step-by-step guide to creating a Pivot Table:
Step 1: Select Your Data
Click anywhere inside the books table to ensure a single data source.
Go to the Insert tab on the Ribbon and click PivotTable.
Step 2: Set Up the Pivot Table
In the Create PivotTable dialog:
Ensure the correct table or range (books) is selected.
Choose where to place the Pivot Table:

Select New Worksheet (recommended) or Existing Worksheet. Click OK.

Step 3: Add Fields to the Pivot Table

In the PivotTable Fields there are:

Drag the Year field to the Rows area. This will create a new row for each year.

Drag the Reviews field to the Values area. By default, it will sum the reviews for each year.

Step 4: Sort the Results

Click the dropdown arrow in the Row Labels section.

Choose Sort by... and select either Smallest to Largest or Largest to Smallest.

Step 5: Optional - Format and Analyze

If desired, format the numbers in the Pivot Table.

Right-click any number in the Sum of Reviews column.

Select Number Format and choose a format like Number.

Review the table to see the total reviews for each year.

If you follow the instructions provided by ChatGPT step by step, you should get the table shown in Figure 2-6.

Row Labels	Sum of Reviews
2009	235,506.00
2010	273,981.00
2011	405,041.00
2012	654,546.00
2013	654,907.00
2014	792,997.00
2015	711,669.00
2016	709,800.00
2017	644,420.00
2018	696,521.00
2019	794,917.00
Grand Total	6,574,305.00

Figure 2-6. Figures 2-6. The Pivot Table produced as a result of Task 2.

Task 3

Use ChatGPT to troubleshoot and write the following prompt:

Find and explain the error in the provided formula.

ChatGPT responds by explaining that the formula contains several errors.

After explaining the correct syntax of the AVERAGEIF function, it provides the correct formula:

```
=AVERAGEIF(books[User Rating], ">4.5", books[Price])
```

Paste the formula into an Excel cell to obtain the result shown in Figure 2-7.

The screenshot shows an Excel spreadsheet with a table of book data. The table has columns for Name, Author, User Rating, Reviews, Price, Year, and Genre. The formula bar at the top shows the formula =AVERAGEIF(books[User Rating], ">4.5", books[Price]). A red box highlights the formula in the formula bar, and another red box highlights the range "books[User Rating], ">4.5" in the formula. A red arrow points from the text "While you are inserting the formula, the used columns are highlighted in the table. Note the different colors." to the highlighted range in the formula bar. Another red arrow points from the text "Paste your formula into a random cell." to the formula bar.

Name	Author	User Rating	Reviews	Price	Year	Genre
1. 50-Day Green Smoothie Cleanse	JJ Smith	4.7	8	2016	Non Fiction	
2. 11/22/63: A Novel	Stephen King	4.8	2052	22	2011	Fiction
3. 52 Rules for Life: An Antidote to Chaos	Jordan B. Peterson	4.7	38079	15	2018	Non Fiction
4. 1984 (Signet Classics)	George Orwell	4.7	25434	6	2017	Fiction
5. 5,000 Awesome Facts (About Everything!) (National Geographic Kids)	James Kormier	4.6	7665	12	2019	Non Fiction
6. A Dance with Dragons (A Song of Ice and Fire)	George R.R. Martin	4.6	12643	11	2011	Fiction
7. A Game of Thrones / A Clash of Kings / A Storm of Swords (Song of Ice and Fire)	George R.R. Martin	4.7	39738	30	2014	Fiction
8. A Gentleman in Moscow: A Novel	Amor Towles	4.7	18699	15	2017	Fiction
9. Higher Loyalty: Truth, Lies, and Leadership	James Comey	4.7	5983	3	2018	Non Fiction
10. A Man Called Ove: A Novel	Fredrik Backman	4.6	23848	8	2016	Fiction
11. A Man Called Ove: A Novel	Fredrik Backman	4.6	23848	8	2016	Fiction
12. A Patriot's History of the United States: From Columbus's Great Discovery to the War on Terror	Larry Schweikart	4.6	400			
13. A Stolen Life: A Memoir	Joyce Maynard	4.6	4149			

Figure 2-7. Figures 2-7. The result of Task 3.

WARNING

If your computer uses a comma (,) as a decimal separator instead of a period (.), the provided formula may return an error or unexpected results. This is because Excel interprets numeric values differently depending on your system locale.

To solve the problem, you can change the general settings of your computer, or you can adapt the formula to the local decimal point. For example, if you use a comma as a decimal separator, write the formula as follows:

```
=AVERAGEIF(books[User Rating]; ">4,5"; books[Price])
```

Note that the period in the value 4.5 is replaced with the comma 4,5, and the argument separator changes from “,” to “;”.

Challenge: What information must you add to the prompt context so that ChatGPT automatically generates the formula using your local decimal point?

Now that we have consolidated the conversation structure with ChatGPT, let's examine the concept of formulas and functions in Excel and see how we can use ChatGPT to speed up their creation.

Basic Functions and Formulas

I have always been passionate about mathematics, particularly the study of functions and formulas. I remember with pleasure that when I was in high

school, my mathematics teacher explicitly prevented me from answering the questions she asked the class about which function to use in a certain situation because she knew I already knew the answer. My passion for functions and formulas has spanned the years, and I have carried it to this day.

There is a fundamental difference between the concept of function and formula:

- *Functions* are predefined operations that simplify complex or repetitive calculations, such as adding numbers, calculating averages, or searching for specific values (for example, `SUM`, `AVERAGE`).
- *Formulas* are user-defined expressions that perform calculations or manipulations on data. They always start with an equal sign (`=`) and can combine operators, constants, and references (such as `=A1+A2`).

This section discusses the main types of functions with practical examples using the `books` dataset. You will also see how ChatGPT can help you use these functions to build basic formulas.

You can group the functions provided by Excel into the following six macro-categories:

- Calculation: Perform mathematical operations like average, sum, maximum, and minimum
- Counting: Count cells, values, or specific conditions

- Logical: Evaluate conditions or perform actions based on conditions
- Search: Find specific values or positions within data
- Text Manipulation: Modify or analyze text strings
- Date and Time: Handle operations involving dates and times.

Let's examine each group of functions individually and consider how ChatGPT can help you define them faster based on your data.

Calculation

Use the *calculation functions* to perform basic mathematical and statistical operations, such as adding, averaging, rounding, or finding the maximum and minimum value in a set of data. Examples:

- `=SUM(A1:A10)` : Calculates the sum of the values in cells A1 to A10.
- `=AVERAGE(A1:A10)` : Calculates the average values in a range.
- `=MAX(A1:A10)` : Returns the maximum value in a range.
- `=ROUND(A1, 2)` : Rounds the value in A1 to two decimal places.

You can use ChatGPT as an assistant to generate the right formula for your data. Start the conversation with ChatGPT by defining the context as described above and then define the main point as follows:

- Action: Write
- Objective: the formula in Excel to calculate [calculation function] [column].

- Additional Details: Use the table [table name].
- Level of complexity: -

For example, if you want to calculate the total sum of all reviews in the books table, use the following prompt:

```
Write the formula in Excel to calculate the sum of reviews in the books table.
```

ChatGPT responds by indicating the following formula:

```
=SUM(books[Reviews])
```

Exercise

After defining the context for the books table, write the prompt for the following tasks:

1. Calculate the average of the reviews in the books table.
2. Calculate the sum of the prices of all the books in the books table.
3. Calculate the average price of books in the "Fiction" category.
4. Identify the highest price among books in the "Non-fiction" category.
5. Calculate the minimum price of books in the "Mystery" category.

Test every output produced by ChatGPT in Excel.

Counting

Use the *counting functions* to count cells based on certain criteria, such as counting non-blank cells, counting cells that meet a specific condition, or multiple criteria at once. Examples:

- `=COUNTA(A1:A10)` : Count all non-blank cells.
- `=COUNTIF(A1:A10, ">5")` : Count cells with values greater than 5.
- `=COUNTIFS(A1:A10, ">5", B1:B10, "Fiction")` : Count cells that match multiple criteria.

You can use ChatGPT as an assistant to generate the right formula for your data. Start the conversation with ChatGPT by defining the context as described above and then define the main point as follows:

- Action: Write
- Objective: the formula in Excel to count cells [criterium].
- Additional details: Use the table [table name].
- Level of complexity: -

For example, if you want to count all books with reviews greater than 4, write the following prompt:

Write the formula in Excel to count cells with reviews greater than 4.

ChatGPT generates the following output:

```
=COUNTIF(books[Reviews], ">4")
```

Exercise

After defining the context for the books table, write the prompt for the following tasks:

1. Determine how many books belong to the “Fiction” genre
2. Count the number of books that have a price of 0
3. Check how many books have received more than 1000 reviews
4. Count how many times the author “J.K. Rowling” appears in the Author column
5. Counts the total number of books in the table, regardless of specific values.

Test every output produced by ChatGPT in Excel.

Logical

Use the *logical functions* to perform conditional operations returning different results based on true or false conditions. Examples:

- `=IF(A1>10, "High", "Low")` : Returns “High” if A1 is greater than 10, otherwise “Low”.
- `=AND(A1>10, B1<5)` : Returns TRUE if both conditions are satisfied.
- `=OR(A1>10, B1<5)` : Returns TRUE if at least one of the conditions is satisfied.

Search

Use the *search functions* to find specific values within a range of data.

Examples:

- `=VLOOKUP(1001, A1:D10, 3, FALSE)` : Looks for the value “1001” in the first column and returns the value of the third column.
- `=INDEX(A1:A10, 5)` : Returns the value in the fifth row of the range.
- `=MATCH(50, A1:A10, 0)` : Find the position of the value “50” in the range.

You can use ChatGPT as an assistant to generate the right formula for your data. Start the conversation with ChatGPT by defining the context as described above and then define the main point as follows:

- Action: `Write`

- Objective: the formula in Excel to search [criterium].
- Additional details: Use the table [table name].
- Level of complexity: -

For example, if you want to find the genre of the book *10-Day Green Smoothie Cleanse*, write the following prompt:

Write the formula in Excel to search the Genre for "10-Day Green Smoothie Cleanse".

ChatGPT generates the following output:

```
=VLOOKUP("10-Day Green Smoothie Cleanse", books, 7, FALSE)  
Explanation:  
"10-Day Green Smoothie Cleanse": The value you are searching for.  
books: The name of the table.  
7: The column index for Genre in the books table  
FALSE: Ensures an exact match.
```

Exercise

After defining the context for the books table, write the prompt for the following tasks:

1. Find the author of the book *The Hunger Games*
2. Find the genre of the book with the maximum price

3. Find the price of the book *The Great Gatsby*
4. Find the year of the book with the maximum number of reviews
5. Find the name of the book at the lowest price.

Test every output produced by ChatGPT in Excel.

Text Manipulation

Use *text formulas* to manipulate and manage text strings. They allow you to merge text, extract specific characters, and determine the length of a string, improving the processing of textual data. Examples:

- `=CONCAT(A1, B1)` : Combine the text in cells A1 and B1.
- `=LEFT(A1, 5)` : Extracts the first 5 characters of the text in A1.
- `=ONLY(A1)` : Returns the length of the text in A1.

You can use ChatGPT as an assistant to generate the right formula for your data. Start the conversation with ChatGPT by defining the context as described above and then define the main point as follows:

- Action: `Write`
- Objective: `the formula in Excel to [text manipulation] [involved columns].`
- Additional details: `Use the table [table name].`
- Level of complexity: -

For example, if you want to extract the first name of each author, write the following prompt:

Write the formula in Excel to extract each author's first name.

ChatGPT generates the following output:

```
=LEFT(books[@Author], FIND(" ", books[@Author]))
```

Explanation:
books[@Author]: Refers to the current row's value
FIND(" ", books[@Author]): Finds the position of the first space character
LEFT(books[@Author], FIND(" ", books[@Author])) - Extracts the text from the start of the cell up to the first space character.

Exercise

After defining the context for the `books` table, write the prompt for the following tasks:

- Remove extra spaces from a title
- Extract the last name of the author
- Transform the title to lowercase
- Combine the book title and author
- Find the length of a book title.

Test every output produced by ChatGPT in Excel.

Date and Time

Use *date* and *time* formulas to work with dates and times, allowing you to get the current date and time or extract specific parts of a date. Examples:

- `=TODAY()` : Returns the current date.
- `=YEAR(A1)` : Extracts the year from a date.
- `=DATEDIF(A1, B1, "d")` : Calculate the difference between two dates in days.

You can use ChatGPT as an assistant to generate the right formula for your data. Start the conversation with ChatGPT by defining the context as described above and then define the main point as follows:

- Action: `Write`
- Objective: `the formula in Excel to calculate [date operation]`.
- Additional Details: `Use the table [table name]`.
- Level of complexity: `-`

For example, if you want to extract the year from today's date, write the following prompt:

```
Write the formula in Excel to calculate the year
```

ChatGPT generates the following output:

```
=YEAR(TODAY())
```

Exercise

After defining the context for the books table, write the prompt for the following tasks:

1. Calculate the number of years that have passed since a book was published
2. Find books published in the last 5 years
3. Calculate the difference in days between today and a book year
4. Calculate the age of a book in months
5. Check if a book was published in a leap year

Test every output produced by ChatGPT in Excel.

THE FOREST FOR THE TREES

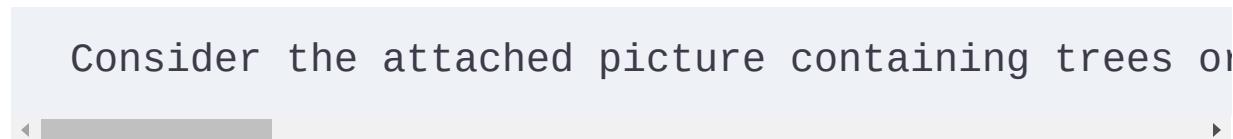
At the beginning of the chapter I told you about my passion for mathematics since I was a child. Over the years, to this passion I have added a passion for data and everything it represents. When you start loving data, you'll see it everywhere. Just a few days ago, during a car trip, I was sitting in the passenger seat, and my husband was driving. As the car rolled along the highway, I admired the passing landscape. At a certain point, something strange appeared: a hill with trees lined up and arranged in order. And I noticed how similar those trees were to the rows and columns of a table.

To immortalize the moment, I took the photo shown in Figure 2-8.



Figure 2-8. The photo I took during my trip.

Out of curiosity, after arriving at my destination, I gave Figure 2-8 as input to ChatGPT and asked it to extract the number of rows and columns of trees, organize them in the form of a table, and format them as an Excel file. I used this prompt:



The result was amazing. ChatGPT counted 20 tree columns and 10 tree rows and organized them into an Excel file, as shown in Figure 2-9.

	I	J	K	L	M	N	O	P	Q	R	S	T
1	Col_9	Col_10	Col_11	Col_12	Col_13	Col_14	Col_15	Col_16	Col_17	Col_18	Col_19	Col_20
2	Tree	Tree	Tree	Tree	Tree	Tree	Tree	Tree	Tree	Tree	Tree	Tree
3	Tree	Tree	Tree	Tree	Tree	Tree	Tree	Tree	Tree	Tree	Tree	Tree
4	Tree	Tree	Tree	Tree	Tree	Tree	Tree	Tree	Tree	Tree	Tree	Tree
5	Tree	Tree	Tree	Tree	Tree	Tree	Tree	Tree	Tree	Tree	Tree	Tree
6	Tree	Tree	Tree	Tree	Tree	Tree	Tree	Tree	Tree	Tree	Tree	Tree
7	Tree	Tree	Tree	Tree	Tree	Tree	Tree	Tree	Tree	Tree	Tree	Tree
8	Tree	Tree	Tree	Tree	Tree	Tree	Tree	Tree	Tree	Tree	Tree	Tree
9	Tree	Tree	Tree	Tree	Tree	Tree	Tree	Tree	Tree	Tree	Tree	Tree
10	Tree	Tree	Tree	Tree	Tree	Tree	Tree	Tree	Tree	Tree	Tree	Tree
11	Tree	Tree	Tree	Tree	Tree	Tree	Tree	Tree	Tree	Tree	Tree	Tree
12												

Figure 2-9. The table built by ChatGPT based on the picture shown in Figure 2-8.

Counting the rows and columns by hand, I deduced that there are exactly 20 columns and about 10 rows. I say *about* because it's difficult to extract the exact number of rows. However, the result provided by ChatGPT was impressive.

From this example you can see the potential of ChatGPT, which can extract tables even from images. And it is precisely this, as well as other concepts, that we will deal with in the next chapter.

Summary

We have reached the end of this chapter, and you have now seen some basic strategies for using ChatGPT to generate appropriate prompts for your dataset. In the next chapter, you will learn how to speed up common operations in ChatGPT.

- ChatGPT can simplify complex tasks like generating functions, guiding data analysis, and resolving errors efficiently.
- Prompts are organized as narratives with three key parts:
 - *Context*: Define the dataset and requirements.
 - *Main Point*: Specify the task or question.
 - *Conclusion*: Outline the desired outcome or next steps.
- ChatGPT can play four different roles while assisting you:
 - *Assistant*: Generate formulas and solve specific tasks.
 - *Teacher*: Explain complex concepts and guide step-by-step processes.
 - *Analyst*: Identify patterns, trends, and insights in data.
 - *Troubleshooter*: Diagnose and correct formula errors or functionality issues.

- Excel essential functions can be grouped into six categories: calculation, counting, logical, search, text manipulation, and date and time.
- ChatGPT can help you to build customized formulas for your specific dataset.

Activity

Complete the following text using the keywords indicated below. The solution is in [Appendix A](#).



In this chapter, you learned how to structure an effective _____ to best use _____ in Excel, organizing it into three fundamental parts: context, main point, and end. The _____ defines the data and information necessary to frame the problem, the_____ describes the action required, while the _____ indicates how to use the result.

ChatGPT can be used as an_____ to perform various tasks, such as generating formulas for _____ (for example, summing or averaging) or performing search operations with functions like _____. This combination of strategies helps you improve productivity, automate complex calculations, and quickly resolve problems in your datasets.

Chapter 3. Speeding Up Common Operations with ChatGPT

A NOTE FOR EARLY RELEASE READERS

With Early Release ebooks, you get books in their earliest form—the author’s raw and unedited content as they write—so you can take advantage of these technologies long before the official release of these titles.

This will be the 3rd chapter of the final book. Please note that the GitHub repo will be made active later on.

If you have comments about how we might improve the content and/or examples in this book, or if you notice missing material within this chapter, please reach out to the editor at ccollins@oreilly.com.

Everyone has to deal with some tedious, time-consuming, and repetitive tasks. One of these for me is making the bed every morning, not only mine (and my husband’s) but also my children’s. You must do it for the decorum of the house, but it’s very dull, takes time, and is repetitive. To solve the problem, I could hire a housekeeper and ask them to arrange the beds for me. This is a bit like what ChatGPT does in the digital world. You can request it to become your housekeeper, and to carry out tedious and

repetitive tasks for you. If there were a ChatGPT that could make beds, I would definitely buy it (of course, only at a reasonable cost).

Similarly, when you use Excel, you often have to deal with tedious, time-consuming, and repetitive tasks. You want to speed them up as much as possible. These include, for example, the creation of templates, filling tables with patterns defined as dates or progressive numbers, and even enriching a dataset with values external to the dataset. Well, ChatGPT can be your housekeeper.

In Chapter 2, you saw that ChatGPT can be a companion in your activities, an assistant, or a secretary—call it what you want. The concept is that ChatGPT assists you while carrying out your tasks. In this chapter, you will use ChatGPT in a new way, as a housekeeper, to carry out tedious, repetitive, and time-consuming tasks for you.

This chapter covers the following topics using ChatGPT:

- Creating templates
- Autofill
- Data enrichment

Let's start with creating templates.

Creating Templates

A *template* in Excel is a predefined file with a ready-made structure and format for specific tasks, such as budget management or inventory. Usually, a template contains pre-configured tables, formulas, and graphs. When I think of a template in Excel, I refer to the templates I see when creating a new workbook, such as budget planners and homework logs. You can make your personal templates by creating an Excel workbook with the layout, formatting, and formulas you want, saving it as an Excel Template, and putting it in the appropriate folder. Once you've created the template, you can use it to create new workbooks based on it, so you don't have to reinvent the wheel every time.

Using a template offers several advantages, such as saving time, achieving consistency in format and style between different documents, and reducing errors. In essence, a template can speed up repetitive tasks such as creating reports, invoices, or calendars by including predefined layouts, formatting, and formulas.

To create your template in ChatGPT, write a prompt with the story structure you learned about in Chapter 2: context, main point, and end. There are two application cases:

- Create the template from scratch, providing ChatGPT with all the details

- Create the template from a (handwritten) sketch, which you provide to ChatGPT as input.

Let's analyze both cases, starting with creating the template from scratch.

Creating Templates from Scratch

When you create a template from scratch, you must provide ChatGPT with all the necessary information. Start a conversation with ChatGPT by following the flow described in [Figure 3-1](#).

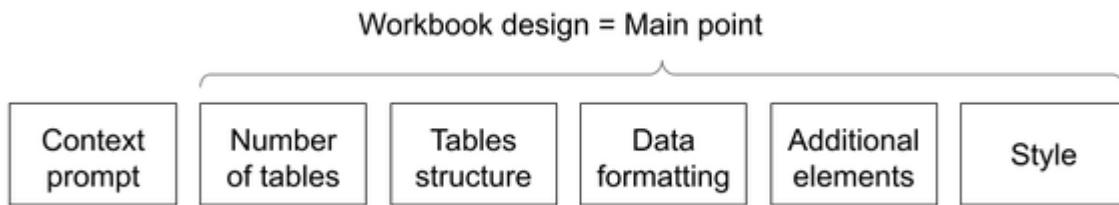


Figure 3-1. The flow to follow to generate an Excel template using ChatGPT.

The prompt follows the story structure, where the final stage of the story is the creation of the Excel file, already included in the context prompt:

1. After logging in to <https://chatgpt.com>, start a new conversation, defining the context, which contains the fact that you want to create a template:

I want to create an Excel template. I'll provide

2. Specify the structure of the workbook, such as the number of worksheets, the number of tables, and the structure of the tables with columns. For example, you can create a single worksheet containing a single table using the following prompt:

Create a single worksheet containing a table with:

- Date
- Category
- Description
- Amount
- Payment Method

Format the table as an Excel table and name it

Add 10 empty rows to the table.

At this point, ChatGPT has generated a file in .xlsx format, which you can download. Next, open the file. [Figure 3-2](#) shows the result.

A	B	C	D	E	F
1	Date	Category	Description	Amount	Payment Method
2					
3					
4					
5					
6					
7					
8					
9					
10					
11					
12					
13					

Figure 3-2. The Excel file generated by ChatGPT after the second step.

3. Define how to format the data, as shown in the following prompt:

Format cells Number as follows:
 - Column A as a date (YYYY/MM/DD)
 - Column B as text
 - Column C as text
 - Column D as currency with symbol \$
 - Column E as text

Download the file generated by ChatGPT. To test the formatting correctness, enter the following information in the first row of the table:

2025-01-08, MyCategory, MyDescription, 3, Cash

The date refers to 2025, 8th January. Excel automatically formats the date and amount, as [Figure 3-3](#) shows. Once you have tested the formatting correctness, you can delete the content of the first row.

	A	B	C	D	E	F
1	Date	Category	Description	Amount	Payment Method	
2	2025/01/08	MyCategory	MyDescription	\$3.00	Cash	
3						
4						
5						
6						
7						
8						
9						
10						
11						
12						

Figure 3-3. The Excel file generated by ChatGPT after data validation.

4. Add more details to the worksheet. For example, you can add a formula that automatically totals the Amount field, as shown in the following prompt:

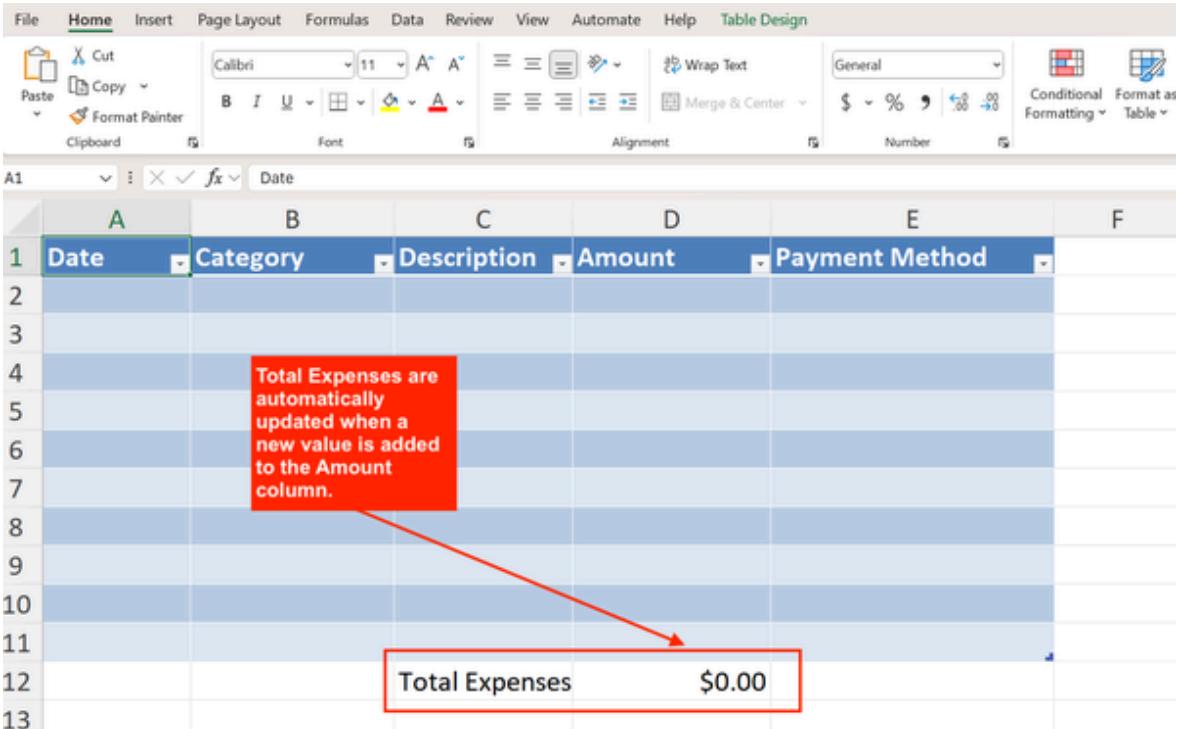
Add a row below Amount and out of the table for

[Figure 3-4](#) shows the Excel file generated by ChatGPT.

NOTE

This method calculates a total value using a cell formula. The goal of this example was to show that you can use ChatGPT to insert formulas outside of tables and enhance your worksheet with various formulas. Alternatively, you can use the Total Row feature, which is available for Excel tables.

Go to Table Design and check the box for Total Row. Excel will add a row at the bottom of the table with built-in functions. You can click on the cells in the total row to choose other functions from a drop-down menu, such as Sum, Average, Count, and more.



The screenshot shows an Excel spreadsheet with a table of expenses. The table has columns for Date, Category, Description, Amount, and Payment Method. A red callout box points to the Amount column, stating: "Total Expenses are automatically updated when a new value is added to the Amount column." A red arrow points from this callout to the Total Expenses cell in the bottom row. The bottom row is highlighted with a red border and contains the text "Total Expenses" and "\$0.00". The Excel ribbon is visible at the top, showing the Home tab is selected. The table design ribbon tab is also visible.

Date	Category	Description	Amount	Payment Method
1				
2				
3				
4		Total Expenses are automatically updated when a new value is added to the Amount column.		
5				
6				
7				
8				
9				
10				
11		Total Expenses	\$0.00	
12				
13				

Figure 3-4. The Excel file generated by ChatGPT after adding a formula to the worksheet.

5. At this point, you can apply a style, defining colors, borders, and so on.

For example, you can use the following prompt:

Apply the following style to the table MainTab

First row:

- Set the text to bold
- Set the text color to black
- Set the background color to light grey
- Align text to the center

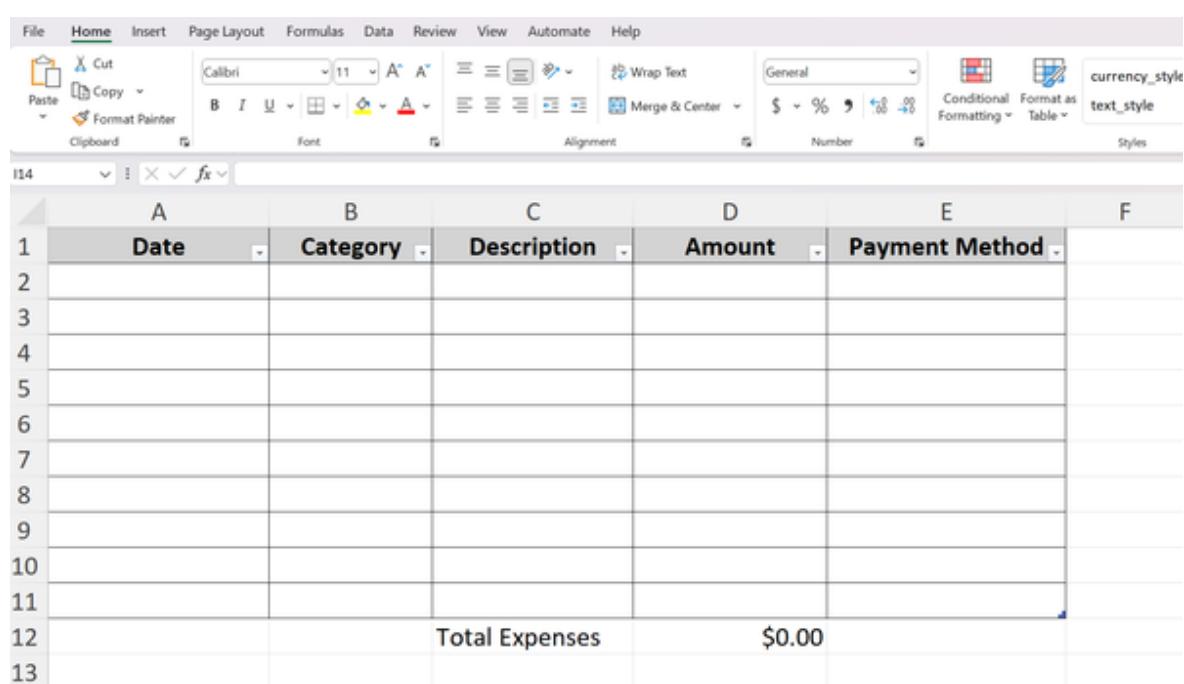
Other rows:

- Align text to the left for text columns and right for the date column
- Set the background color to white for all the rows

Borders:

- Apply thin black borders

Figure 3-5 shows the file generated by ChatGPT after applying the style.



The screenshot shows a Microsoft Excel spreadsheet with the ribbon menu at the top. The table 'MainTab' is visible, containing columns for Date, Category, Description, Amount, and Payment Method. The first row has bold black text and a light grey background. The other rows have standard black text and white backgrounds. All cells have thin black borders. The formula bar at the bottom shows the formula `=Total Expenses`.

	Date	Category	Description	Amount	Payment Method
1					
2					
3					
4					
5					
6					
7					
8					
9					
10					
11					
12			Total Expenses	\$0.00	
13					

Figure 3-5. The Excel file generated by ChatGPT after applying your style.

WARNING

In some cases, ChatGPT may not generate the correct template. This is because the models behind ChatGPT are not perfect. To resolve this issue, apply one of the following suggestions:

- Rerun the exact same prompt several times until you get the desired result. For example, you can use the prompt: “Please try again”.
 - Change the prompt slightly
 - If you lose the cell’s formatting, rerun the prompt, asking ChatGPT to apply specific formatting.
-

Exercise 1

Write prompts for ChatGPT to generate the table in [Figure 3-6](#). Format the ISBN field as XXX-X-XXXXX-XXX-X, and the calculation of Total Books as the count of all the books in the table.

The screenshot shows a Microsoft Excel spreadsheet with a table. The table has columns labeled A through F. Columns A, B, C, D, and E have dropdown arrows indicating they are filterable. Row 1 contains the column headers: Title, Author, Publisher, Year, and ISBN. Row 12 contains the formula =COUNT(A2:A11) in cell A12, which evaluates to 0. The cell containing the formula is also bolded.

	A	B	C	D	E	F
1	Title	Author	Publisher	Year	ISBN	
2						
3						
4						
5						
6						
7						
8						
9						
10						
11						
12	Total Books				0	
13						

Figure 3-6. The table to generate using ChatGPT.

You can find the solution containing the prompt in the book's GitHub repository as [03/exercise1_solution.txt](#).

So far, you have seen how to organize a conversation with ChatGPT step by step so that it generates the desired output. You've probably noticed how tiring it is to follow the conversation with ChatGPT. It would be great if you could give all the conversation prompts in one go and then let ChatGPT digest them gradually. Well, I have some good news for you. This is possible, which is exactly what we will see in the next section.

Automatically Generating a Template from Scratch

To automatically generate a template from scratch, write all the prompts in a single file, and then feed the file to ChatGPT. The only care you need to take is to define an opening prompt accompanying the file loaded into ChatGPT. The opening prompt specifies exactly how ChatGPT should behave when executing the instructions contained in the file.

For example, you can define the following opening prompt:

Open the attached file and identify the first instruction.

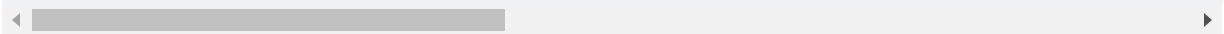
Carefully read the first instruction and follow its steps.
Complete the first instruction fully before moving on.
After completing the first instruction, move on to the next one.

Ensure that each instruction is executed separately.
If you encounter an unclear instruction or one that does not make sense,
Provide a detailed explanation of the issue.

Once all instructions are completed, ensure the following:
- Every request has been implemented as specified.
- All required formatting and styles are properly applied.

After completing all steps, confirm that the process is complete.
Share the generated file for review.

Execute each instruction as a separate unit, without dependencies.
Follow a clear and logical sequence, ensuring each step builds upon the previous one.



Attach to the opening prompt the *template_instructions.txt* file formatted as follows:

[STEP] Prompt 1

[STEP] Prompt 2

[STEP] Prompt N

You can test this technique directly using the file found in the book's GitHub repository, as [03/template_instructions.txt](#). This file contains the instructions described in the "Creating Templates from Scratch" section, preceded by the keyword [STEP], as shown in the following prompt:

[STEP] I want to create an Excel template. I'll p

[STEP] Create a single worksheet containing a tab

- Date
- Category
- Description
- Amount
- Payment Method

Format the columns as an Excel table and assign t

[STEP] ...

Using this technique can be time-consuming for ChatGPT, so don't worry if ChatGPT doesn't provide the answer immediately.

TIP

Applying this technique, I've noticed that in most cases, ChatGPT fails to correctly format the table as an Excel table (despite it being explicitly asked in the instructions), so if this happens, manually pass it the following prompt after it has executed all the prompts contained in the template:

Format the generated table as an Excel Table.

After this statement, ChatGPT formats the table correctly.

CHATGPT AS A FRIEND

I have tested this automatic technique many times with different templates and almost always obtained the same result: the tables were not formatted as Excel tables. Only after one or two trials did ChatGPT format the file correctly. I racked my brain, trying to understand why this happened. I changed the order of the prompts and the prompt text, but the result was still the same. Until I had a flash of genius, or as Gary Klein calls it in one of his books, *creative desperation*: I asked ChatGPT why it didn't consider the instruction I had passed on to it. The response was shocking, beyond all expectations.

I gave it this prompt as input:

```
Explain why you didn't apply the format the table
```

ChatGPT gave me the following response:

```
I did not apply the instruction to format the ta
```

```
Difference between cell ranges and Excel tables:
```

```
Library used (openpyxl): In Python, the openpyxl
```

```
If you want, I can immediately correct this and u
```

Aside from the first answer, which might seem bizarre, the second answer explains the technique ChatGPT uses to generate Excel files. In practice, ChatGPT transforms the prompts into Python code (Python is a programming language) and executes the code using the openpyxl library, which is precisely the Python library for generating Excel files. Anyway, the performance and the quality of results depend on the chosen model. The Python library used for generating Excel files may depend on the model. Some models use the XlsxWriter Python library, and others use the openpyxl library.

Creating Templates from Sketches

At the end of chapter 2, I told you about the photo with the trees and the use of ChatGPT to extract a table from that photo. I also promised you that in the next chapter (this one), you would see how to create Excel tables directly from photos. Well, now it's time to talk about it.

Imagine that you have made notes on a piece of paper relating to the expenses you have incurred recently, as shown in [Figure 3-7](#) (The figure is deliberately faded and not optimal for reasons I will explain shortly).

Date	Category	Description	Amount	Payment Method
2024/12/12	Grocery	Weekly Grocery	\$ 43.00	Cash
2024/12/12	Presents	Present for Giulia	\$ 12.00	Credit Card
Total Amount				\$ 55.00

Figure 3-7. A photo of a handwritten table

Now, imagine that you want to transcribe your paper table (and the total amount row) into an Excel sheet. The traditional approach involves manual transcription from paper to computer. But by using ChatGPT, you can speed up the operation.

You can proceed in two ways. In the first mode, proceed using the ChatGPT web interface and follow these steps:

1. Before connecting to the ChatGPT web interface, take a piece of paper and write a table by hand. For example, you can copy the table in

[Figure 3-7](#) (typed out in [Table 3-1](#)). Also, add the Total Amount row, as shown at the bottom of [Figure 3-7](#).

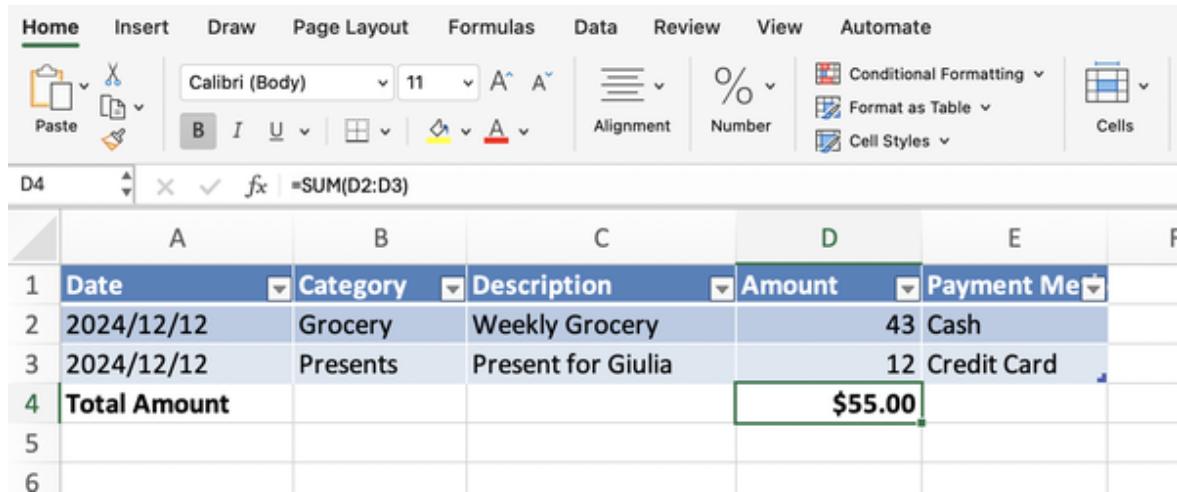
Table 3-1. An example of a table

Date	Category	Description	Amount	Per M
2024/12/12	Grocery	Weekly	\$43.00	Ca
		Grocery		
2024/12/12	Presents	Present for Giulia	\$12.00	Ci

2. Take a photo of your table and the total amount, as shown in the original photo of [Figure 3-7](#). It doesn't matter whether the photo is excellent as long as it is legible, like the one in [Figure 3-7](#). Don't worry about lights, shadows, or taking perfect photos.
3. Upload the photo to ChatGPT via the upload button and use the following prompt:

Format the attached table as an Excel file and Follow this workflow to build the Excel file:
Populate columns and rows
Format as an Excel Table except for the row Total
Calculate the Total Amount using a formula (Sum)

4. Download the Excel file produced by ChatGPT. You should get something very similar to [Figure 3-8](#).



The screenshot shows a Microsoft Excel spreadsheet with a table of transactions. The table has columns for Date, Category, Description, Amount, and Payment Method. Row 4 is highlighted in green and contains the formula '=SUM(D2:D3)' in the formula bar, indicating a calculated total.

	A	B	C	D	E	F
1	Date	Category	Description	Amount	Payment Me	
2	2024/12/12	Grocery	Weekly Grocery	43	Cash	
3	2024/12/12	Presents	Present for Giulia	12	Credit Card	
4	Total Amount			\$55.00		
5						
6						

Figure 3-8. The Excel file produced by ChatGPT

5. As a bonus, you can also ask ChatGPT to add more consistent lines. This is particularly useful if you want to generate example files for other cases. Use this prompt, continuing the conversation with ChatGPT:

Populate this table with 15 consistent rows.

[Figure 3-9](#) shows a possible result generated by ChatGPT.

	A	B	C	D	E
1	Date	Category	Description	Amount	Payment Method
2	2024/12/13	Entertainment	Streaming Subscription	41.38	Cash
3	2024/12/06	Presents	Anniversary Gift	94.45	Cash
4	2024/12/03	Presents	Anniversary Gift	78.08	Cash
5	2024/12/10	Grocery	Vegetables	63.8	Credit Card
6	2024/12/07	Entertainment	Movie Ticket	63.54	Credit Card
7	2024/12/13	Bills	Internet Bill	71.97	Debit Card
8	2024/12/01	Bills	Internet Bill	58.3	Cash
9	2024/12/11	Bills	Internet Bill	70.63	Debit Card
10	2024/12/09	Entertainment	Movie Ticket	94.22	Cash
11	2024/12/12	Dining	Fast Food	41.93	Debit Card
12	2024/12/12	Dining	Restaurant Dinner	20.76	Cash
13	2024/12/07	Dining	Fast Food	25.56	Debit Card
14	2024/12/16	Bills	Electricity Bill	13.8	Cash
15	2024/12/12	Grocery	Vegetables	78.4	Debit Card
16	2024/12/14	Dining	Fast Food	38.06	Credit Card
17	Total Amount			\$854.88	
18					

Figure 3-9. The result of this prompt: Populate this table with 15 consistent rows.

If you have followed all the steps described above, you have noticed a minor complication. You probably took the photo with your smartphone, uploaded it to the cloud, downloaded it from your computer, and finally uploaded it to ChatGPT. This operation is cumbersome and slow.

There is an alternative way to proceed: the ChatGPT app for smartphones! You can download the official ChatGPT app from your favorite app store, log in to your account, and have ChatGPT within reach of your smartphone. Everything I have said so far about the web interface applies to the smartphone app. So you can decide from time to time whether to use the web interface or the app.

In this specific case, I still suggest you use the app when creating the template from a photo. After logging into the app, you have an interface available where you can directly take the picture, as shown in [Figure 3-10](#).

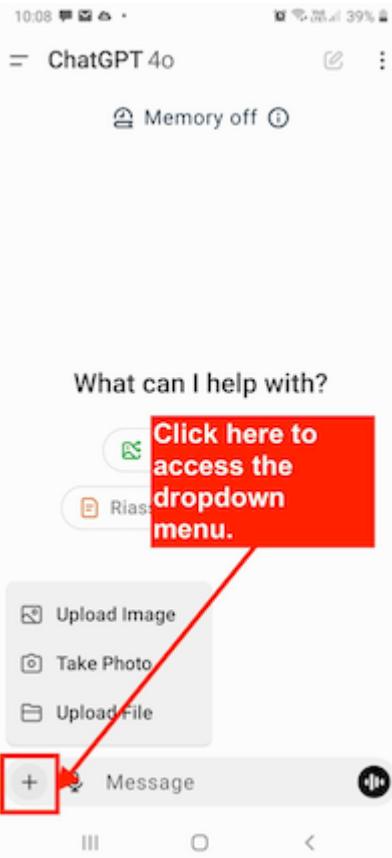


Figure 3-10. The ChatGPT smartphone app interface with the ability to upload photos.

After taking the photo directly in ChatGPT, perform steps 3-4 described previously. If you don't want to download the Excel file to your smartphone, you can continue the conversation via the web interface.

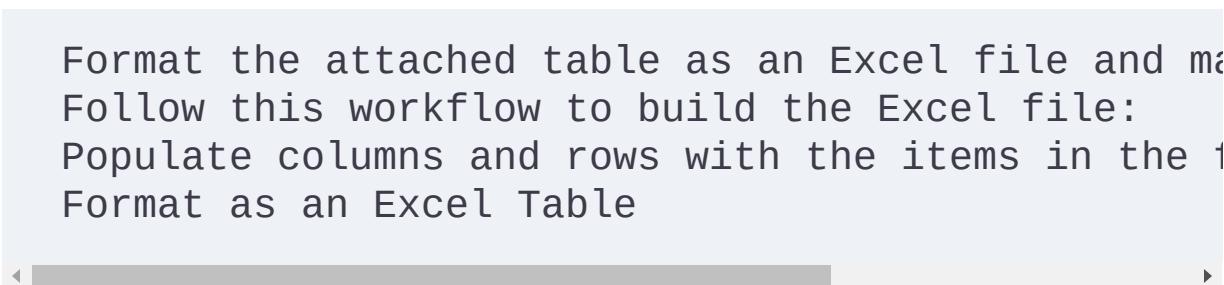
TIP

An interesting aspect of ChatGPT is that it may store conversations, which you can access via a bar on the left. These conversations are available in both the app and the web version, so you can start a conversation in the app and continue it in the web interface and vice versa. This option is not available if you start a temporary chat.

Exercise 2

If you haven't recycled it, find the latest supermarket flyer you received in the mail and select a page of your choice. Take a photo and input it to ChatGPT using the following prompt:

Format the attached table as an Excel file and make it available for download.
Follow this workflow to build the Excel file:
1. Scan the page for items and their details.
2. Create a new Excel sheet.
3. Populate columns and rows with the items in the following structure:
4. Format as an Excel Table



I tried this exercise, and the result was quite good. ChatGPT organized the items into a table with the structure shown in Table 3-2 and identified almost all items on the page.

Table 3-2. The table structure generated by ChatGPT from the flyer page.

Category	Product Name	Size/Quantity	Price (€)

From the experience of this exercise, I always advise you to check the correctness of the results produced by ChatGPT.

At this point, you should have a pretty clear idea of how to use ChatGPT to generate templates, either from scratch or from a photo or image. In the next section, you will see the concept of Autofill and how to use ChatGPT to speed up this operation.

Autofill

Autofill in Excel is a function that allows you to automatically fill a series of cells with values, formulas, or patterns based on a predefined template. You can use Autofill to complete number sequences, days of the week, months, formulas, and more.

Suppose you have the table shown in Table 3-3 ([available](#) in the book's GitHub repository as [*03/autofill.xlsx*](#).

Table 3-3. An example table

#	Date	Day of Week
1	2024-01-01	
2	2024-01-02	
3	2024-01-03	

The first column is a progressive number, the second is a progressive date in the format YYYY-MM-DD, and the third is the week's day. It is currently empty, but you would like it filled with the day of the week corresponding to the date. In the example, the first date should correspond to Monday, the second to Tuesday, and so on, given that January 1, 2024, was a Monday.

Suppose you want to extend the table using Autofill to have dates up to January 31, 2024. Let's see how to perform this operation first manually using only Excel and then with ChatGPT.

To proceed manually, do the following:

1. Select the first three dates and move the cursor to the bottom right, where there is a small square called Fill Handle, as shown in [Figure 3-11](#).

	A	B	C	D
1	#	Date	Day of Week	
2	1	2024-01-01		
3	2	2024-01-02		
4	3	2024-01-03		
5				
6				

Figure 3-11. How to activate Autofill in Excel (step 1)

2. Move the mouse cursor until it becomes a black cross, then click and drag down a few rows until the date 2024-01-08 appears, as shown in [Figure 3-12](#).

The screenshot shows the Microsoft Excel interface with the Home tab selected in the ribbon. The font is set to Calibri (Body) at size 11. The formula bar displays the date 01/01/2024. A table is visible on the worksheet, with columns labeled A, B, C, D, and E. Rows 1 through 4 are populated with data: Row 1 has columns #, Date, and Day of Week; Row 2 has value 1 and date 2024-01-01; Row 3 has value 2 and date 2024-01-02; Row 4 has value 3 and date 2024-01-03. The cell B10 contains the date 2024-01-08. A red callout box with a white border and red text is positioned over the range B2:B4, containing the instruction "Drag until the label is 2024-01-31." A red arrow points from the text in the callout box to the bottom-right corner of the selected range.

	A	B	C	D	E
1	#	Date	Day of Week		
2	1	2024-01-01			
3	2	2024-01-02			
4	3	2024-01-03			
5					
6					
7					
8					
9					
10		2024-01-08			
11					

Figure 3-12. How to activate Autofill in Excel (step 2)

Release the cursor. All dates are populated automatically through 2024-01-08, as shown in [Table 3-4](#).

Table 3-4. The result of the Autofill process

#	Date	Day of Week
1	2024-01-01	
2	2024-01-02	
3	2024-01-03	
	2024-01-04	
	2024-01-05	
	2024-01-06	
	2024-01-07	
	2024-01-08	

Exercise 3

Repeat the same operations just described for the # column containing the progressive number and the Day of Week column. To use the Autofill function in the latter case, remember to write the values for at least the first two lines.

At this point, you are ready to repeat the same operations using ChatGPT. Remember that each prompt must be structured into three parts: context,

main point, and end. In our example, we directly load the *autofill.xlsx* dataset into ChatGPT, and this forms our context. Then we ask ChatGPT to fill in the missing values (main points). Basically we use ChatGPT to perform the autofill for us automatically. Finally, we tell ChatGPT to format the output as an Excel table, maintaining the style of the original file.

TIP

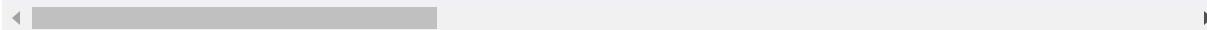
Using ChatGPT to perform Autofill is particularly useful for populating a dataset with many values. If you want to fill a few values, you can still use ChatGPT, but the advantage over the manual solution is minimal.

Use the following prompt:

Consider the table attached.

Using Autofill, populate all the table columns w:

Format the resulting table as an Excel table and



ChatGPT performs the operations and then generates the file in Excel format to download. [Figure 3-13](#) shows a portion of the file generated by ChatGPT.

The screenshot shows a Microsoft Excel spreadsheet titled "Sheet1". The table has three columns: "#", "Date", and "Day of Week". The data starts from row 1 and continues to row 12. The "Day of Week" column uses the standard short-form abbreviations for days of the week. The entire table is selected, indicated by a green border around the cells.

	A	B	C	D
1	#	Date	Day of Week	
2	1	2024-01-01	Monday	
3	2	2024-01-02	Tuesday	
4	3	2024-01-03	Wednesday	
5	4	2024-01-04	Thursday	
6	5	2024-01-05	Friday	
7	6	2024-01-06	Saturday	
8	7	2024-01-07	Sunday	
9	8	2024-01-08	Monday	
10	9	2024-01-09	Tuesday	
11	10	2024-01-10	Wednesday	
12	11	2024-01-11	Thursday	

Figure 3-13. The file generated by ChatGPT after running the Autofill

In some cases, the file generated by ChatGPT may not contain data formatted as a table and with the same layout as the original file. In this case, continue the conversation with ChatGPT, repeating the last part of the prompt:

Format the resulting table as an Excel table and



At most, after a couple of tries, ChatGPT should generate the correctly formatted file.

As you've seen, you can use Autofill to automatically fill the rows of a table with predefined values or sequences. If you have a lot of values to fill out, use ChatGPT to speed up this operation. But what if you want to extend your data with data external to your source dataset? You can still use ChatGPT. Let's see how in the next section.

Data Enrichment

Data enrichment consists of extending an original dataset with data from other sources. You can extend a dataset with geographic data, descriptions, or other information from external sources, such as the web, or internal sources, such as other datasets from the same source. To extend a dataset with other data, the dataset must not contain duplicates and have at least one column to be used as a unique identifier.

Let's take an example. Suppose you have a workbook with two tables organized into two separate worksheets. You can download the workbook from the book's GitHub repository as [03/data_enrichment.xlsx](#). The first

table, called Books (available in the Books tab), contains book titles and status (whether it has been read, unread, or reading), as shown in Table 3-5.

Table 3-5. The table contained in the Books tab in the data_enrichment.xlsx example file

Title	Status
Effective Data Storytelling	Read
Become a Great Data Storyteller	To read
Storytelling with Data	To read
The Data Storyteller's Handbook	Reading

The second table, called Info (available in the Info tab), contains book titles, authors, and years of publication, as shown in Table 3-6.

Table 3-6. The table contained in the Info tab in the data_enrichment.xlsx example file

Title	Author	Year
The Data Storyteller's Handbook	Kat Greenbrook	2023
Storytelling with Data	Cole Nussbaumer Knaflic	2015
Effective Data Storytelling	Brent Dykes	2019
Become a Great Data Storyteller	Angelica Lo Duca	2025

Suppose you want to extend the first table with the information contained in the second. The two tables share a common column, Title, which will act as a unique identifier for matching between the two tables.

Use the function `VLOOKUP()` to extend the first table with the data from the second. The function `VLOOKUP()` has the following syntax:

```
=VLOOKUP(lookup_value, table_array, col_index_nur
```

lookup_value

The value you want to look for in the first column of the `table_array` table.

`table_array`

The table or range of cells to search for the value. The first column of this range is where `VLOOKUP()` searches for the `lookup_value`.

`col_index_num`

The number of the column (in the table range) from which to take the value to be returned. The first column always has the number 1.

`range_lookup (optional)`

Specify the match type. If set to `TRUE` (or omitted), it looks for an approximate match (values should be sorted in ascending order). If set to `FALSE`, it searches for an exact match.

The function returns the value in the column specified by `col_index_num`. Apply the function `VLOOKUP()` to the Books table as follows:

1. Select cell C2 in the Books worksheet, as in [Figure 3-14](#).

Select cell C2.

	A	B	C	D
1	Title	Status		
2	Effective Data Storytelling	Read		
3	Become a Great Data Storyteller	To read		
4	Storytelling with Data	To read		
5	The Data Storyteller's Handbook	Reading		
6				
7				
8				

Figure 3-14. How to enrich data using VLOOKUP() (step 1)

2. Write the following formula to select the author:

```
=VLOOKUP( [@Title], Info[#All], 2, FALSE )
```

Then press the Enter key. [Figure 3-15](#) shows the result.

	A	B	C
1	Title	Status	Column1
2	Effective Data Storytelling	Read	Brent Dykes
3	Become a Great Data Storyteller	To read	Angelica Lo Duca
4	Storytelling with Data	To read	Cole Nussbaumer Knaflic
5	The Data Storyteller's Handbook	Reading	Kat Greenbrook
6			
7			
8			

Figure 3-15. How to enrich data using VLOOKUP() (step 2)

3. Rename the column to Author.

4. Repeat the operation for the Year column (using a formula). [Figure 3-16](#) shows the final result. You can find the solution in the book's GitHub repository as [*03/data_enrichment_manual.xlsx*](#).

	A	B	C	D	E
1	Title	Status	Author	Year	
2	Effective Data Storytelling	Read	Brent Dykes	2019	
3	Become a Great Data Storyteller	To read	Angelica Lo Duca	2025	
4	Storytelling with Data	To read	Cole Nussbaumer Knaflic	2015	
5	The Data Storyteller's Handbook	Reading	Kat Greenbrook	2023	
6					

Figure 3-16. The final result after data enrichment

So far, you have seen how to extend the dataset manually. This example required that I already have the enriched data in a separate table. However, you may not have the data immediately, and want to extract it from an external source, such as the web. In this case, ChatGPT can help you.

To demonstrate how to extract data from the web, we'll use a simplified version of the dataset, which contains only the Books table, and we will ask ChatGPT to fill it using the data it knows. You can find this simplified file version in the book's GitHub repository as

[*03/data_enrichment_simplified.xlsx*](#).

First, let's define a prompt that consists of three parts: context, main point, and end. As a context, we directly use the *data_enrichment.xlsx* dataset,

which we load into ChatGPT using the appropriate button to load data.

Let's start the conversation with ChatGPT by writing the following prompt:

Consider the table Books in the first worksheet of the attached Excel file.

ChatGPT generates the result shown in [Figure 3-17](#), but it is not formatted as an Excel file. Furthermore, the table does not contain some data because it is not available in the GPT-4o model used by ChatGPT in this experiment.

Title	Author	Year of Publication	Publisher
Effective Data Storytelling	Brent Dykes	2019	Wiley
Become a Great Data Storyteller	Angelica Lo Duca	2025	Wiley

Figure 3-17. The table enriched by ChatGPT does not contain all the required information

WARNING

In some cases, ChatGPT can *hallucinate*, meaning it generates incorrect results.

Always remember to check the results produced by ChatGPT before using them!

NOTE

In some cases, ChatGPT may not allow you to download the generated file because it remains temporarily stored in memory. If you attempt to download the file after a long inactivity, it may have already been deleted and is no longer available.

In other cases, ChatGPT might encounter issues during the file generation or download process, such as a timeout, a server error, or a browser-related restriction. To avoid losing access to the file, download it as soon as it's created.

To solve the problem, ask ChatGPT to repeat the operation by searching for information on the Web. Select the “Search the web” button in the ChatGPT interface, as shown in [Figure 3-18](#).



Figure 3-18. The Search button in the ChatGPT web interface

Once the button is enabled, write this prompt:

Repeat the operation.

ChatGPT now generates a new table, as shown in [Figure 3-19](#).

Title	Author	Year of Publication	Publisher
Effective Data Storytelling	Brent Dykes	2019	Wiley
Become a Great Data Storyteller	Angelica Lo Duca	2025	Wiley
Storytelling with Data	Cole Nussbaumer Knaflic	2015	Wiley
The Data Storyteller's Handbook	Kat Greenbrook	2023	Rogue Kororā

Figure 3-19. Using the Search function, the enriched table contains all the required information

The table is not yet downloadable as an Excel file. So, ask ChatGPT to format the table as an Excel file, keeping the original formatting and data present in the source file:

Format the resulting table as an Excel Table. Keep the original styling.

At this point, the file is ready, as shown in [Figure 3-20](#).

Title	Status	Author	Year of Pub	Publisher
Effective Data Storytelling	Read	Brent Dykes	2019	Wiley
Become a Great Data Storyteller	To read	Angelica Lo Duca	2025	Wiley
Storytelling with Data	To read	Cole Nussbaumer Knaflic	2015	Wiley
The Data Storytelling's Handbook	Reading	Kat Greenbrook	2023	Rogue Kororā

Figure 3-20. The final file generated by ChatGPT

Compared to the original file, ChatGPT did not maintain the original theme of the table but still enriched the data correctly.

From this example, we learned the following about ChatGPT:

- It automates most operations, such as calculations and data enrichment.
 - Sometimes you need to ask it several times to format the table as a table and keep the original style.
 - It doesn't always maintain the original table's theme and style.
-

TIP

Sometimes, ChatGPT fails to generate the file to download and proposes instructions to create the file by yourself. In this case, you can ask ChatGPT to retry, for example, using the following prompt: Please try again. Alternatively, you can follow the instructions provided by ChatGPT and apply them by hand to your original Excel file.

Exercise 4

Write instructions for ChatGPT to populate [Table 3-7](#) with the country name and geographic coordinates.

Table 3-7. The table to enrich.

City	Country	Latitude	Longitude
Pisa			
New York			
Sydney			
Paris			

The Excel file containing the table in the book's GitHub repository is available as [03/exercise4.xlsx](#), and the solution is here:

[03/exercise4_solution.txt](#).

NOTE

I conclude this chapter with a personal reflection. Last night, I had a moment of pure contemplation: I stopped to observe my nine-year-old daughter Giulia while she played the piano. She was enchantingly graceful. The melody that filled the room was so sweet and enveloping that it touched the deepest strings of the soul. At that moment, a thought took shape in my mind. ChatGPT, upon closer inspection, is like a piano: it has infinite keys and can create extraordinary melodies. However, just as a piano remains silent and inert without someone who can play it, ChatGPT remains unused if you don't approach it with the proper prompts. It is human interaction, awareness, and creativity that give it voice and make it extraordinary.

Summary

In this chapter, you learned how to speed up some common operations using ChatGPT. In the next chapter, you will learn how to use ChatGPT to perform advanced operations.

- A template is a predefined file with a ready-made structure and format for specific tasks. You can use ChatGPT to speed up template creation. Structure the prompt following a precise order: first, define the context and then the design of the workbook, with the number of tables and their structure, data validation, additional elements and style.
- With ChatGPT, you can create a template from scratch or use a photo of a hand-made sketch. You can even ask ChatGPT to extract Excel tables from flyers or other images.
- Autofill in Excel is a useful function for automatically filling a series of cells with values, formulas, or patterns based on a predefined template. ChatGPT can speed up the auto-filling process if the number of data to insert is large.
- Enriching a dataset involves extending an original dataset with data from other sources. Use ChatGPT's Search function to enrich a dataset with information extracted from the web.

Activity

Solve the word puzzle in [Figure 3-21](#) by circling all the words you see. In the end, you will discover something about Autocomplete by reading the letters that are not circled, which will now form words. You can find the solution in [Appendix A](#).



Figure 3-21. The crossword puzzle to solve

Words to search for:

AUTOFILL

CHATGPT

CONTEXT

END

EXCEL

MAINPOINT

PROMPT

ROW

TABLE

TEMPLATE

WEB SEARCH

Autofill

Chapter 4. Advanced Operations with ChatGPT

A NOTE FOR EARLY RELEASE READERS

With Early Release ebooks, you get books in their earliest form—the author’s raw and unedited content as they write—so you can take advantage of these technologies long before the official release of these titles.

This will be the 4th chapter of the final book. Please note that the GitHub repo will be made active later on.

If you have comments about how we might improve the content and/or examples in this book, or if you notice missing material within this chapter, please reach out to the editor at ccollins@oreilly.com.

When I started programming in the early 2000s, fixing a bug in the code was a real challenge. If something wasn’t working, you didn’t even know where to start. It could take hours, sometimes days, of trying to check line by line in the hope of finding that bug in the code. Then, suddenly, a colleague or friend would arrive and, with a quick glance, identify the problem and point it out to you as if it were the most obvious thing in the world. And you, between disbelief and gratitude, thanked them heartily.

Then came the Internet, Web 2.0, and programming forums. If a code generated problems, all you had to do was search on a search engine, and

sooner or later, you would find someone who had had the same problem and had already solved it. Debugging times have shortened dramatically.

Today, with ChatGPT, everything is even faster. Just paste the code and ask where the error is. You'll have the solution in a few minutes, sometimes in a few seconds. When I think about twenty years ago, when I could have spent entire days on a single bug, the progress is impressive. But it doesn't end there. With generative AI, you no longer have to write code; ChatGPT does it directly. Just describe the problem, and it returns the code ready to use, both for programming in languages such as Python, Java, or C# and for creating advanced and efficient formulas in Excel. And that's exactly what you'll see in this chapter. More precisely, this chapter covers the following topics:

- Using ChatGPT to analyze a dataset
- Using ChatGPT for PivotTables
- Combining ChatGPT and macros
- Using ChatGPT to write VBA code

Let's start with the first one, using ChatGPT to analyze a dataset.

Using ChatGPT to Analyze a Dataset

The previous chapter showed how to use AI to help you write or understand formulas. However, having formulas is not enough, and knowing how to

use it is not enough. You also need to use them to perform the proper data analysis. For example, you might have a beautiful dataset and know all the Excel formulas, but you have no idea what to ask of that dataset and what formula to use to extract something important from your data. Well, AI and, in particular, ChatGPT can help you with this. In this section, you'll see how AI can analyze the dataset and generate advanced formulas. To analyze the dataset, follow the flow described in [Figure 4-1](#).

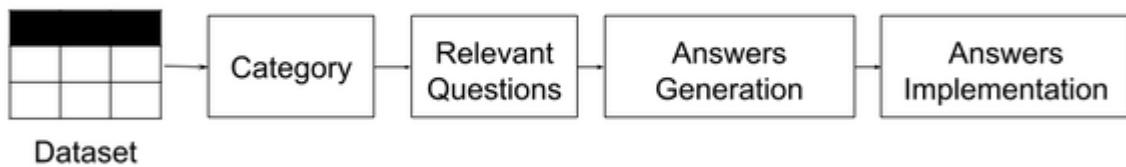


Figure 4-1. The workflow for analyzing a dataset.

A typical workflow in more detail might go like this:

1. Start with your dataset in Excel. Extract a part containing the column names and possibly some example lines (or use the entire dataset if it is not too large).
2. Upload it to ChatGPT. (Alternatively, as seen in previous chapters, you could provide ChatGPT context about the dataset.)
3. Ask ChatGPT to extract the category from your dataset. Examples of categories include finance, environment, health, society, and so on. This step is essential to help ChatGPT take the next step.

4. Ask ChatGPT to define relevant questions for your dataset. For example, Table 4-1 shows some possible questions to ask about the data based on the categories to which it belongs.

Table 4-1. A list of possible questions based on the data category.

Category	Requests
Health	<p>What are the main risk factors for a particular disease?</p> <p>How do diagnoses vary over time and space?</p> <p>How effective is a specific treatment?</p>
Finance	<p>What are the trends in the financial markets?</p> <p>What factors influence stock prices?</p> <p>What is the average return on an investment portfolio?</p>
Environment	<p>What are the primary sources of pollution in a region?</p> <p>How does air quality change over time?</p> <p>What measures can reduce the environmental impact?</p>
Social media	<p>What are the trends in online conversations?</p> <p>How does user engagement vary across different platforms?</p> <p>What is the general sentiment on a particular topic?</p>
Education	<p>What are the factors that influence student achievement?</p>

Category	Requests
	<p>How does the school dropout rate vary?</p> <p>What teaching strategies improve learning?</p>
E-commerce	<p>Which products are best selling in a given period?</p> <p>What factors influence purchasing behavior?</p> <p>Which marketing strategies are most effective?</p>
Transport	<p>What are the most common traffic patterns in a city?</p> <p>What are the main factors causing transport delays?</p> <p>How do we optimize routes to reduce fuel consumption?</p>
Energy	<p>What are the most efficient energy sources?</p> <p>How does energy consumption vary over time?</p> <p>What strategies can reduce energy waste?</p>
Manufacturing	<p>What are the most common causes of failures in production lines?</p> <p>How to improve assembly line efficiency?</p> <p>Which materials offer the best value for money?</p>
Sport	<p>What factors influence athlete performance?</p> <p>Which training strategies are most effective?</p> <p>How to predict the risk of accidents?</p>

5. Once you've identified the questions, ask ChatGPT to answer them using Excel formulas.
6. Tell ChatGPT to implement the responses into your original dataset and make it downloadable.

Let's apply this workflow to a full example.

Dataset

Consider the synthetic dataset available in the book's GitHub repository at [04/sales_dataset.xlsx](#) and shown in [Figure 4-2](#).

1	Transaction ID	Date of Sale	Product ID	Product Name	Quantity Sold	Unit Price	Total Sale Amount	Payment Method	Store Location	Customer Type
2	T1000	2024-02-01	P203	Laptop	3	1200	3600	Cash	Houston	Returning
3	T1001	2024-02-02	P200	Tablet	4	300	1200	PayPal	New York	Returning
4	T1002	2024-02-03	P200	Tablet	1	1200	1200	PayPal	Miami	New
5	T1003	2024-02-04	P202	Headphones	5	1200	6000	Cash	Miami	VIP
6	T1004	2024-02-05	P201	Headphones	5	300	1500	Credit Card	Houston	New
7	T1005	2024-02-06	P205	Headphones	2	1200	2400	Cash	New York	New
8	T1006	2024-02-07	P202	Laptop	2	500	1000	PayPal	Miami	New
9	T1007	2024-02-08	P200	Smartphone	3	300	900	Debit Card	New York	Returning
10	T1008	2024-02-09	P203	Laptop	1	1200	1200	Debit Card	Houston	New
11	T1009	2024-02-10	P200	Headphones	3	800	2400	Debit Card	Chicago	New
12	T1010	2024-02-11	P200	Laptop	4	800	3200	Credit Card	Chicago	New
13	T1011	2024-02-12	P203	Smartwatch	1	300	300	Credit Card	Chicago	Returning
14	T1012	2024-02-13	P200	Tablet	1	800	800	Credit Card	Los Angeles	VIP
15	T1013	2024-02-14	P202	Laptop	5	300	1500	Debit Card	Chicago	New
16	T1014	2024-02-15	P204	Headphones	4	1200	4800	Credit Card	Miami	New
17	T1015	2024-02-16	P205	Headphones	3	500	1500	Cash	Houston	Returning
18	T1016	2024-02-17	P204	Laptop	5	500	2500	Credit Card	Miami	VIP
19	T1017	2024-02-18	P202	Smartphone	3	150	450	PayPal	Houston	New

Figure 4-2. An extract from the sales dataset.

The dataset shows product sales and contains the following fields:

Transaction ID

Unique identifier for each sale

Date of Sale

Date of the transaction (YYYY-MM-DD)

Product ID

Unique identifier for the product sold

Product Name

Name or description of the product

Quantity Sold

Number of units sold

Unit Price

Selling price per unit

Total Sale Amount

Total revenue from the sale (Quantity × Unit Price - Discounts)

Payment Method

Cash, Credit Card, PayPal, etc.

Store Location

City, region, or country where the sale occurred

Customer Type

New, Returning, VIP, etc.

Extract the Category and Relevant Questions

Start a new conversation with ChatGPT, load the dataset, and execute the following prompt:

Consider the attached dataset describing an Excel spreadsheet containing sales data for a retail store. The dataset includes columns for Product ID, Product Name, Category, Store Location, Sales Date, Payment Method, and Customer Type (New, Returning, VIP). Generate a series of questions that can be used to analyze this data.

In our present case, ChatGPT associated the dataset with the Sales/Retail/Business category and generated the following questions:

1. Which product generates the highest revenue?
2. What are the trends in sales over time?
3. Which store location has the highest total sales?
4. What is the distribution of payment methods used by customers?
5. How does customer type (New, Returning, VIP) impact sales and revenue?

WARNING

By repeating the operation several times and starting a new conversation each time, you'll notice that the questions change slightly, but overall, they are always the same.

TIP

You can ask ChatGPT to generate five more questions by writing the following prompt:

List 5 more questions.
Repeat to get five more questions each time.

Answers Generation

Now, ask ChatGPT to generate the answers to the questions using Excel.

Use the following prompt:

Write the Excel formula or steps to answer each question.

You specifically ask ChatGPT not to use PivotTables because, as we will see later in this chapter, in its current version (as of this writing in early 2025), ChatGPT can still not apply PivotTables to an Excel file independently. However, if at the time you are reading this, you use a version of ChatGPT that also manages PivotTables, feel free to delete the text “Don’t use PivotTables” from the prompt. For each answer, ChatGPT lists the steps to implement it in Excel. For example, in this case, to implement the answer to question 1, *Which product generates the highest revenue?* ChatGPT generated the following response:

To calculate total revenue for each product, use

=SUMIFS(MainTable[Total Sale Amount], MainTable[Product Name], A2)

- Replace A2 with the product name cell reference.
- Copy the formula for all unique product names.

Steps:

- List unique product names in a new column.
- Use the SUMIFS formula to sum Total Sale Amount for each product.
- Identify the highest value using MAX or by sorting the results.

First, ChatGPT proposes the formula to calculate the answer and then provides the steps to implement it. In the example, ChatGPT proposes the use of the function **SUMIFS**, which sums the values in a range that satisfy multiple criteria. The syntax is as follows:

SUMIFS(sum_range, criteria_range1, criteria1, [criteria_range2, criteria2], ...)

Let's unpack that:

sum_range

The range of values to sum.

criteria_range1

The first range to apply a condition to.

criteria1

The condition to be met in criteria_range1.

criteria_range2, criteria2 (optional)

Additional conditions to filter the sum.

In the example generated by ChatGPT, the function **SUMIFS** sums the values of the Total Sale Amount column of the MainTable table, but only for the rows in which the value of the Product Name column corresponds to the value present in cell A2. ChatGPT, therefore, does not directly answer the question *Which product generates the highest revenue?*, but selects the relevant data (and processes it) to help you answer the question more easily.

Answers Implementation

At this point, ask ChatGPT to implement the answers it gave in the original file. Use the following prompt:

Add a sheet to the original file for Question 1 r

In my case, ChatGPT generated an Excel file, which you can download from the book's GitHub repository at [04/sales_dataset_q1.xlsx](#). [Figure 4-3](#) shows the portion of the generated file related to Question 1. Interestingly, the values in the Total Revenue column are calculated as formulas.

The screenshot shows a Microsoft Excel interface. The ribbon at the top has tabs for File, Home, Insert, Page Layout, Formulas, Data, Review, View, Automate, Help, and Acrobat. The Home tab is selected. Below the ribbon is a toolbar with icons for Cut, Copy, Paste, Format Painter, and Clipboard. The font section shows 'Calibri 11' and alignment options. The number section shows currency (\$), percentage (%), and other symbols. The active cell is P13. The worksheet contains the following data:

	A	B	C	D	E
1	Product Name	Total Revenue			
2	Laptop	13000			
3	Tablet	9100			
4	Headphones	28850			
5	Smartphone	1350			
6	Smartwatch	1100			
7					
8	ChatGPT has generated a new worksheet to answer Question 1.				
9					
10					
11					
12					

The tab bar at the bottom shows 'Sheet1' and 'Top Revenue Product', with a red arrow pointing to the 'Top Revenue Product' tab. The status bar at the bottom right shows 'Page: 1 / 1'.

Figure 4-3. The worksheet generated by ChatGPT in response to the question of which product generates the highest revenue?

To answer the question *Which product generates the highest revenue?* You need to do a little extra work. Proceed as follows:

1. With the mouse, select all the non-empty cells of the Top Revenue Product worksheet, and from the Home tab of the Ribbon, select Format as Table, choose a format, and confirm the range and headings in a dialog box ([Figure 4-4](#)).

A screenshot of the Microsoft Excel interface. The ribbon at the top shows the 'Home' tab is selected. In the 'Styles' group, the 'Format as Table' button is highlighted with a red box and a red arrow pointing to it from the left. Below the ribbon, a table is displayed with columns labeled 'Product Name' and 'Total Revenue'. The rows contain data for a Laptop, Tablet, Headphones, Smartphone, and Smartwatch. The entire table is selected, indicated by a green border.

Product Name	Total Revenue
Laptop	13000
Tablet	9100
Headphones	28850
Smartphone	1350
Smartwatch	1100

Figure 4-4. Select the non-blank cells of the worksheet that answer Question 1 and format them as a table.

2. Select the arrow at the right of the Total Revenue column and click Sort Largest to Smallest ([Figure 4-5](#)). Then click OK, if needed.

	A	B	C
1	Product Name	Total Revenue	
2	Laptop		
3	Tablet		
4	Headphones		
5	Smartphone		
6	Smartwatch		
7			

Figure 4-5. Select the arrow to open the drop-down menu. Select Sort Largest to Smallest.

3. The first product in the table answers the question *Which product generates the highest revenue?* In our case, the product that produces the most revenue is Headphones, as shown in [Figure 4-6](#).

	A	B	C
1	Product Name	Total Revenue	
2	Headphones	28850	
3	Laptop	13000	
4	Tablet	9100	
5	Smartphone	1350	
6	Smartwatch	1100	
7			

Figure 4-6. The first row of the table answers the question which product generates the highest revenue?

From the procedure shown to answer the first question, we can make some reflections:

1. Use ChatGPT to build the framework for answering your data questions.
2. Use your knowledge and experience to answer the questions based on what ChatGPT generated.

ChatGPT, therefore, automates basic and repetitive operations, but the final answer is always up to you.

After you answer the first question, ask ChatGPT to answer the remaining questions, one at a time. Use the following prompt to answer the second question:

Repeat for Question 2. Use the Excel file you have

Exercise 1

Check if the answer ChatGPT provided to question 2, *What are the trends in sales over time?*, is correct. If necessary, refine your answer. In my case, you can find the answer to question 2 in the book's GitHub repository at [04/sales_dataset_q2.xlsx](#). As in the answer to question 1, in this case ChatGPT generated the framework to answer the question but has not responded definitively. For example, you can complete the answer by inserting a graph showing the sales trend.

Follow these steps:

1. Turn non-blank cells into a table (as described in the section “Answers Implementation”).
2. Insert a graph visually showing the trend: select the Insert tab from the Ribbon → Insert Line or Area Chart → Line [Figure 4-7](#).

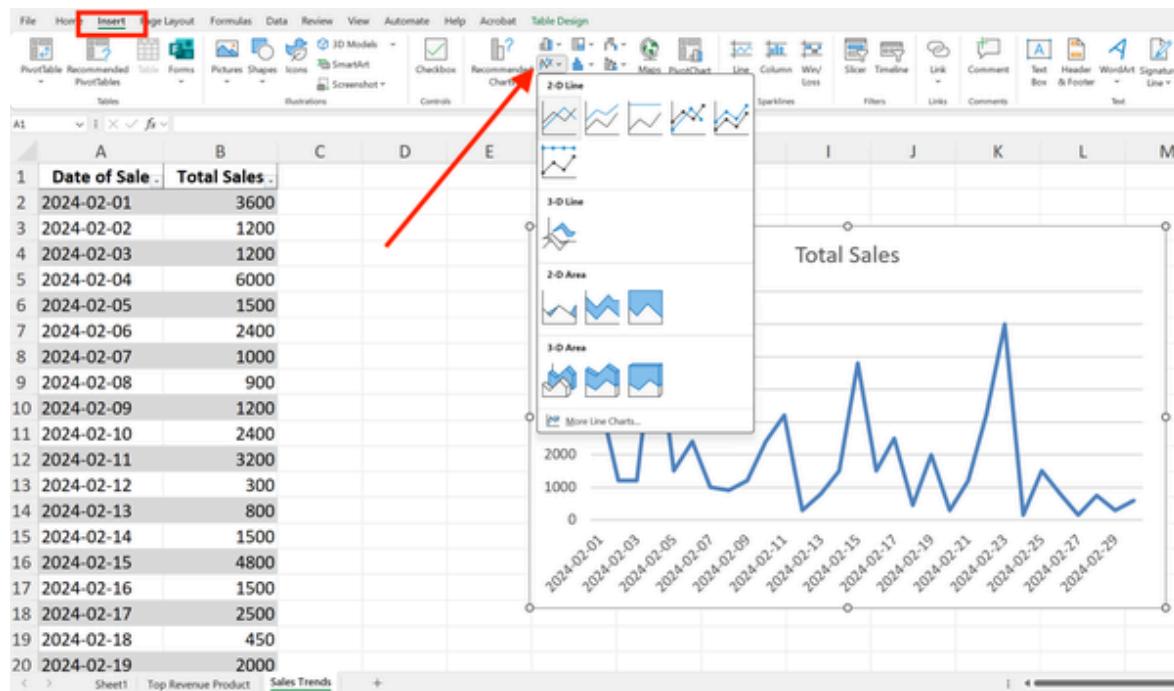


Figure 4-7. Insert a graph to answer the question What are the trends in sales over time?

TIP

Challenge: Ask ChatGPT to generate the graph and insert it into the output file.

Extend to All Questions

After answering Question 2, repeat the process for Questions 3, 4, and 5.

For example, to address Question 3, enter the prompt:

Repeat for Question 3. Use the Excel file you have created.

ChatGPT adds a worksheet to the file whenever it answers a question. You can download the files generated in my case from the book's GitHub.

repository at [04/sales dataset q2.xlsx](#), [04/sales dataset q3.xlsx](#), [04/sales dataset q4.xlsx](#), and [04/sales dataset q5.xlsx](#). Qx indicates the question number. The file relating to question Qx contains all answers from 1 to x. [Figure 4-8](#) shows the structure of the final file ([sales dataset q5.xlsx](#)), which answers all five questions.

The screenshot shows a Microsoft Excel spreadsheet titled 'Customer Analysis'. The table has two columns: 'Customer Type' and 'Total Revenue'. The data is as follows:

	Customer Type	Total Revenue
1	Returning	20950
2	New	22050
3	VIP	10400
4		
5		
6		
7		
8		
9		

A red callout box with the text 'Each question has a separate worksheet.' points to the tabs at the bottom of the Excel window: 'Sheet1', 'Top Revenue Product', 'Sales Trends', 'Store Sales', 'Payment Methods', and 'Customer Analysis'. The 'Customer Analysis' tab is currently selected.

Figure 4-8. ChatGPT answered each question using a separate worksheet.

[Table 4-2](#) shows the questions and the process ChatGPT followed to answer them in my case.

Table 4-2. Questions and answers relating to the sales dataset.

Question	Answer	Used Formula
Which product generates the highest revenue	List unique product names and use <code>SUMIFS</code> to calculate the total revenue for each.	<code>=SUMIFS(MainTable[Total Sale Amount], MainTable[Product Name], A2)</code>
What are the trends in sales over time?	List unique dates and use <code>SUMIFS</code> to calculate total sales per date.	<code>=SUMIFS(MainTable[Total Sale Amount], MainTable[Date of Sale], A2)</code>
Which store location has the highest total sales?	List unique store locations and use <code>SUMIFS</code> to calculate total sales per store.	<code>=SUMIFS(MainTable[Total Sale Amount], MainTable[Store Location], A2)</code>
What is the distribution of payment methods used by customers?	List unique payment methods and use <code>COUNTIF</code> to count occurrences.	<code>=COUNTIF(MainTable[Payment Method], A2)</code>

Question	Answer	Used Formula
How does customer type impact sales and revenue?	List unique customer types and use SUMIFS() to calculate total revenue per type.	=SUMIFS(MainTable[Total Sale Amount], MainTable[Customer Type], A2)

In most cases, ChatGPT proposed using the **SUMIFS()** function. You could have achieved a similar result using PivotTables. However, there are differences between **SUMIFS()** and PivotTables. [Table 4-3](#) shows some pros and cons of **SUMIFS()** vs. a PivotTable.

Table 4-3. A quick comparison between SUMIFS() and a PivotTable.

	PROS	CONS
SUMIFS()	Automatic update without having to update the table manually	The formula SUMIFS() can become complex and difficult to manage in sheets with many conditions.
PivotTable	<p>It allows you to analyze and summarize large amounts of data quickly.</p> <p>It can calculate sums, averages, counts, and other operations without writing complex formulas.</p>	It is not dynamic in real time: if the data changes, it must be updated manually.

Both are useful and allow you to achieve similar results. It is up to you, based on the pros and cons, to evaluate which one to use.

At the time of writing this chapter, ChatGPT can't add PivotTables automatically to an Excel file. However, you can still use ChatGPT as a teacher to learn how to build a specific PivotTable. Let's see how in the next section.

Using ChatGPT for PivotTables

A PivotTable in Excel is a tool that helps you quickly summarize and analyze data. You can use PivotTables to easily organize, group, and calculate data (like sums or averages) by dragging and dropping fields, making it easier to find patterns and insights. Use ChatGPT as a teacher to describe the steps to implement a PivotTable.

Consider the following example to understand how to proceed. Implement the answer to question 1, *Which product generates the highest revenue?*, using PivotTables. Continuing the conversation in the same chat from which you downloaded the various Excel files with the answers, ask ChatGPT for the steps to implement Question 1 using PivotTables. Use the following prompt:

Describe the steps to implement question 1 using

In my case, ChatGPT provided the following response:

Step 1: Select Your Data

1. Open your Excel file containing the dataset
2. Click anywhere inside the dataset (MainTable)

Step 2: Insert a PivotTable

1. Go to the Insert tab on the Excel ribbon.
2. Click PivotTable.
3. In the *Create PivotTable* window:
 - Ensure the Table/Range is set to MainTable
 - Choose New Worksheet for the PivotTable location
 - Click OK.

Step 3: Configure the PivotTable Fields

1. In the PivotTable Fields panel, drag and drop:
 - "Product Name" into the Rows area.
 - "Total Sale Amount" into the Values area.
2. Ensure that the "Total Sale Amount" field is selected.

Step 4: Sort to Identify the Top Product

1. Click any value in the Total Sale Amount column.
2. Go to the Data tab and click Sort.
3. Choose to Sort by Total Sale Amount in Descending order.
4. The top product will now be the first entry.

Step 5: Apply Formatting (Optional)

1. Click on the Total Sale Amount column.
2. Go to the Home tab → Click Number Format → \$.
3. Adjust column widths for better readability.

Step 6: Identify the Product with the Highest Revenue

- The product with the highest total revenue will be at the top of the list.
- Optionally, use Conditional Formatting (Home tab) to highlight the top product.

By following the steps described by ChatGPT, you get the PivotTable shown in [Figure 4-9](#), which is very similar to the table in Figure 4-6, although the formatting is much better (thanks to steps 5 and 6).

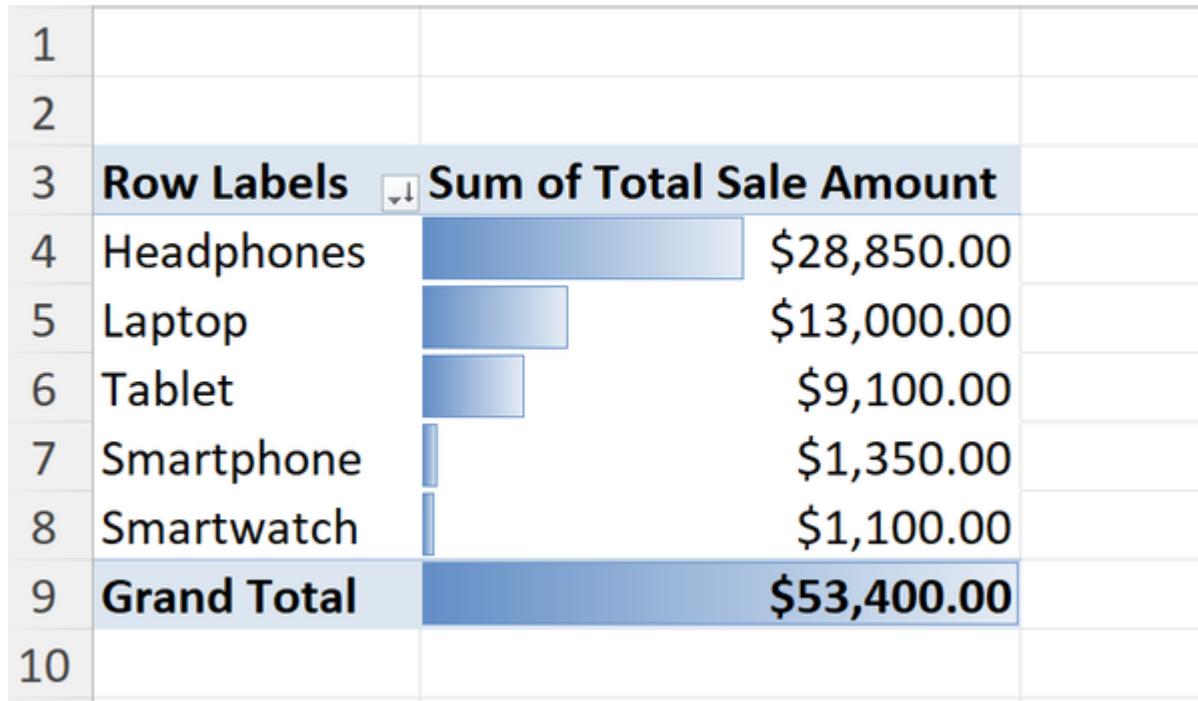


Figure 4-9. The PivotTable implemented using the steps suggested by ChatGPT.

In providing the answer, ChatGPT also added the following bonus:

Bonus: Create a Chart

1. Click anywhere inside the PivotTable.
2. Go to the **Insert** tab → Select **Column Chart**.
3. The chart will visually display the revenue

With this bonus, ChatGPT explains how to add a chart. You obtain the graph shown in [Figure 4-10](#) by following the steps described.

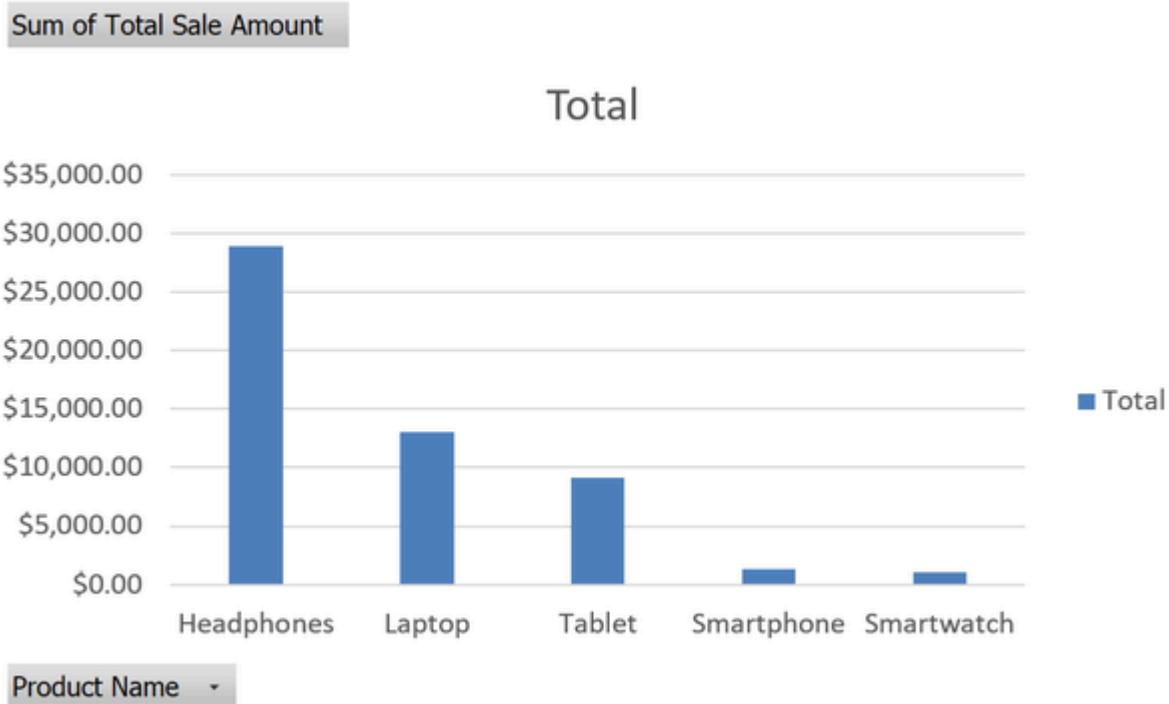


Figure 4-10. The column chart relating to the PivotTable of Question 1.

So far, you have seen how to use ChatGPT to analyze a dataset and, as a teacher, help build PivotTables. However, ChatGPT can go further by generating macros and VBA code. Let's see how in the next section.

Macros and VBA

A *macro* in Excel is a set of automated commands that allow you to quickly perform repetitive tasks, such as formatting data, creating reports, or performing complex calculations. You can write macros using the VBA (Visual Basic for Applications) programming language or record macros without writing code.

Before using macros, you must activate them following these steps (in Windows), if not already activated:

1. Open Excel and go to File → Options.
2. Select Trust Center in the left menu.
3. Click Trust Center Settings.
4. In the window that opens, select Macro Settings.
5. Choose the Disable VBA macros with notification option (recommended).
6. Click OK to confirm your changes.

If you use Mac OS, follow these steps: Open Excel and go to Excel → Preferences → Security → Disable all macros with notification.

At this point, customize the Ribbon to display the Macros commands.

Follow these steps to enable the Developer tab, which also contains macros:

1. Choose File → Options (Windows) or Excel → Preferences (Mac OS).
2. Select Customize Ribbon (Windows) or Ribbon & Toolbar (Mac OS)
3. Under Customize the Ribbon, check the Developer option and click OK
(Figure 4-11)

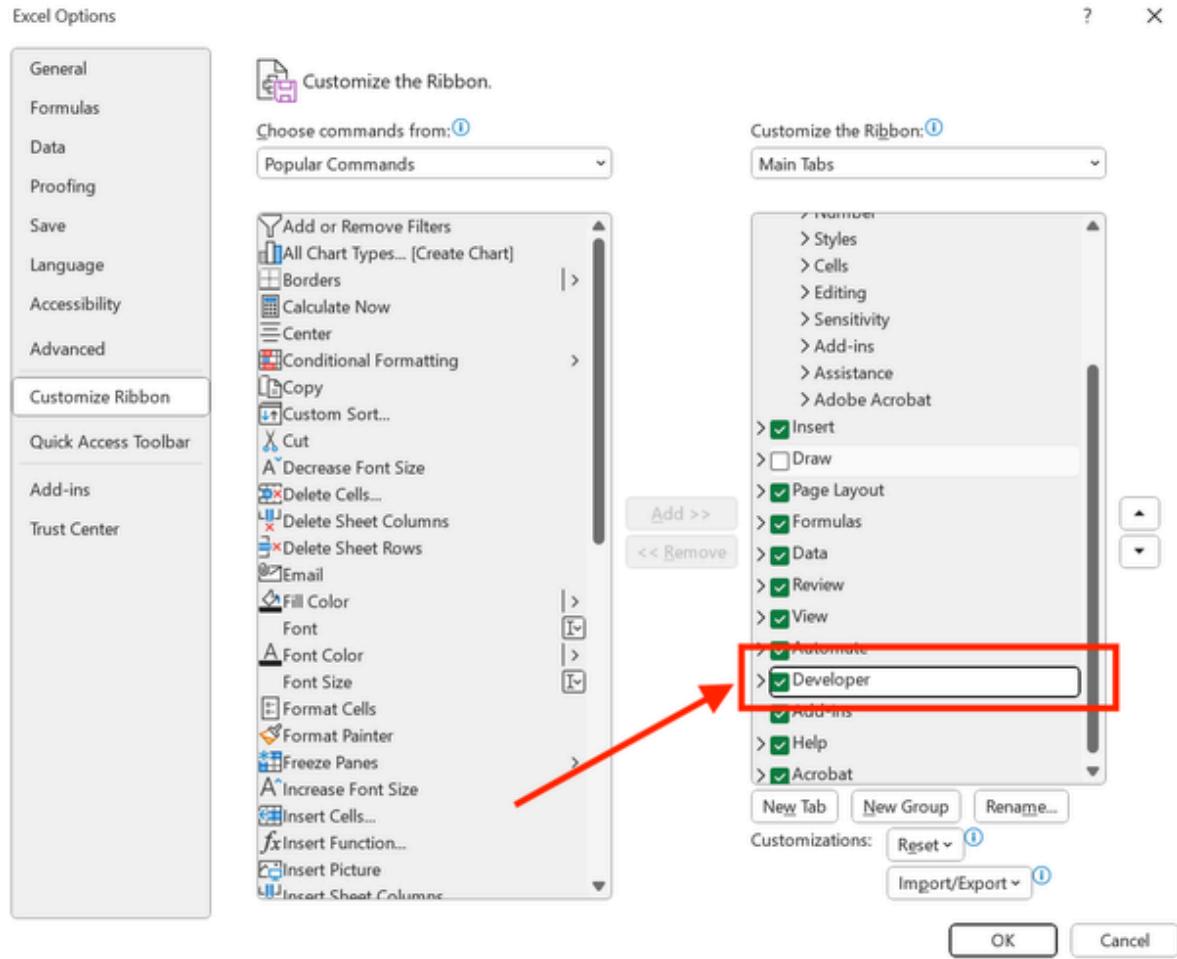


Figure 4-11. Select Customize Ribbon → Main Tabs → Developer to add the Developer tab to the Ribbon.

Once you have performed these operations, you will find the Developer tab in the Ribbon, which should appear as shown in [Figure 4-12](#) (in Windows) or [Figure 4-13](#) (in Mac OS).

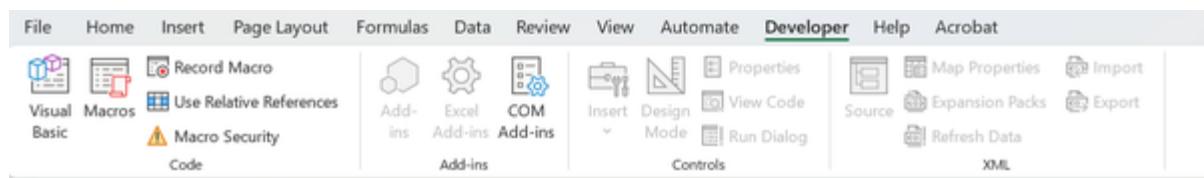


Figure 4-12. The Developer tab of the Ribbon (Windows)

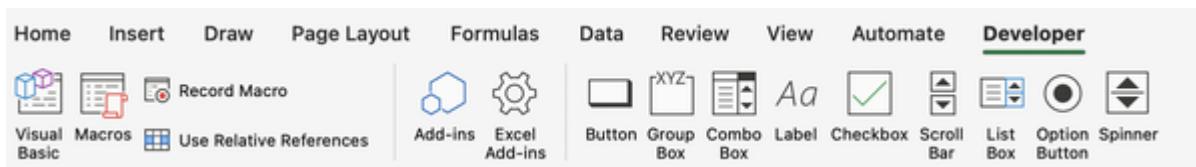


Figure 4-13. The Developer tab of the Ribbon (Mac OS)

At this point, you are ready to use macros. We will look at two strategies. The first is to record a macro, and the second is to write the VBA code to generate the macro. Let's start with the first step: recording a macro.

Recording a Macro

Let's use an example to understand how macros work. Open the file [04/sales_dataset_q5.xlsx](#), which results from the previous exercise (Figure 4-8). Inside the workbook, in addition to the main worksheet containing the main dataset, you will find five worksheets, one for each question. Each worksheet includes ranges of values not formatted as Excel tables (Figure 4-6). Let's now write a macro to transform the data contained in each worksheet into a table. We record the macro without writing code in VBA for a single worksheet and then apply it to all other worksheets.

To record a macro that converts the second worksheet into a table and formats the revenue column as currency in Excel, follow these steps:

1. Navigate to the second worksheet, Top Revenue Product.
2. Go to the Developer tab and click Record Macro. In the Record Macro dialog box:

- a. Enter a Macro Name (e.g., ConvertToTable).
 - b. Choose where to store the macro (This Workbook).
 - c. Optionally, assign a shortcut key (e.g., Ctrl+Shift+T). Using the Shift command prevents you from overriding the Ctrl+t shortcut that converts the selected cells into a table.
 - d. Add a description (e.g., “Converts a cell range into a table and formats revenue as currency”).
 - e. Click OK to start recording.
3. Click cell A2 on the second worksheet. Press Ctrl + T to open the Create Table dialog.
- a. Ensure the “My table has headers” checkbox is checked.
 - b. Click OK to create the table.
4. Identify the column containing revenue (e.g., Total Revenue, column B) and manually select all the rows for that column in the table.
5. Go to the Home tab, and in the Number group, click the dropdown and select Currency. Adjust the currency symbol and decimal places using the Format Cells option if needed.
6. Go back to the Developer tab and click Stop Recording.
7. Save the workbook as a Macro-Enabled Workbook (.xlsm format) to retain the macro. You can find the file in the GitHub repository of the book under [04/sales_dataset_q5.xlsm](#).

[Figure 4-14](#) shows the result after recording the macro: the data is formatted as an Excel table, and the Total Revenue column is formatted as currency.

The screenshot shows a Microsoft Excel interface. At the top, there's a ribbon toolbar with various icons for clipboard operations like Paste, Copy, and Paste Special, along with font and style options for Calibri, size 11, bold, italic, underline, and alignment. Below the ribbon is a formula bar with the cell reference F17. Underneath is a table with six rows of data. The first row contains the headers 'Product Name' and 'Total Revenue'. The data rows are: 'Laptop' (\$13,000.00), 'Tablet' (\$9,100.00), 'Headphones' (\$28,850.00), 'Smartphone' (\$1,350.00), 'Smartwatch' (\$1,100.00), and an empty row 7.

	A	B	C	D
1	Product Name	Total Revenue		
2	Laptop	\$13,000.00		
3	Tablet	\$9,100.00		
4	Headphones	\$28,850.00		
5	Smartphone	\$1,350.00		
6	Smartwatch	\$1,100.00		
7				

Figure 4-14. The data is formatted as an Excel table after recording the macro.

If you were to apply the macro as it is to the other worksheets, though, you would get the wrong result, as shown in [Figure 4-15](#). As you can see, only the first six rows have been converted into an Excel table, while the others have remained unchanged. This is because the macro was recorded considering a 6-row table. You don't need to try this out yourself. It's just to show you what would happen.

The screenshot shows a Microsoft Excel interface with the following details:

- Ribbon:** File, Home, Insert, Page Layout, Formulas, Data, Review, View.
- Table Contextual Tab:** Table Name: Table2_1, Summarize with PivotTable, Remove Duplicates, Convert to Range, Insert Slicer, Export, Refresh, External Ta.
- Cell Selection:** Cell B1 is selected, labeled "Total Sales".
- Table Structure:**

	A	B	C
1	Date of Sale	Total Sales	
2	2024-02-01	\$3,600.00	
3	2024-02-02	\$1,200.00	
4	2024-02-03	\$1,200.00	
5	2024-02-04	\$6,000.00	
6	2024-02-05	\$1,500.00	
7	2024-02-06	\$2,400.00	
8	2024-02-07	\$1,000.00	
9	2024-02-08	\$900.00	
10	2024-02-09	\$1,200.00	
11	2024-02-10	\$2,400.00	

Figure 4-15. If you apply the macro recorded for worksheet 2 to the other worksheets, you get an incorrect result because the macro records the specific values of the worksheet in question.

To make the macro work correctly for other worksheets, you must generalize it, removing specific references to the data in worksheet 2. To do this, manually modify the macro code. ChatGPT helps us with this. You'll use this procedure: open the VBA editor to access the macro, copy the macro code, paste it into ChatGPT, and instruct it to generalize. Follow these steps:

1. Go to the Developer tab, click Macros, select ConvertToTable, and click Edit to open the VBA Editor.
2. Once the Editor VBA is opened, you will see the macro code in the module window ([Figure 4-16](#)).

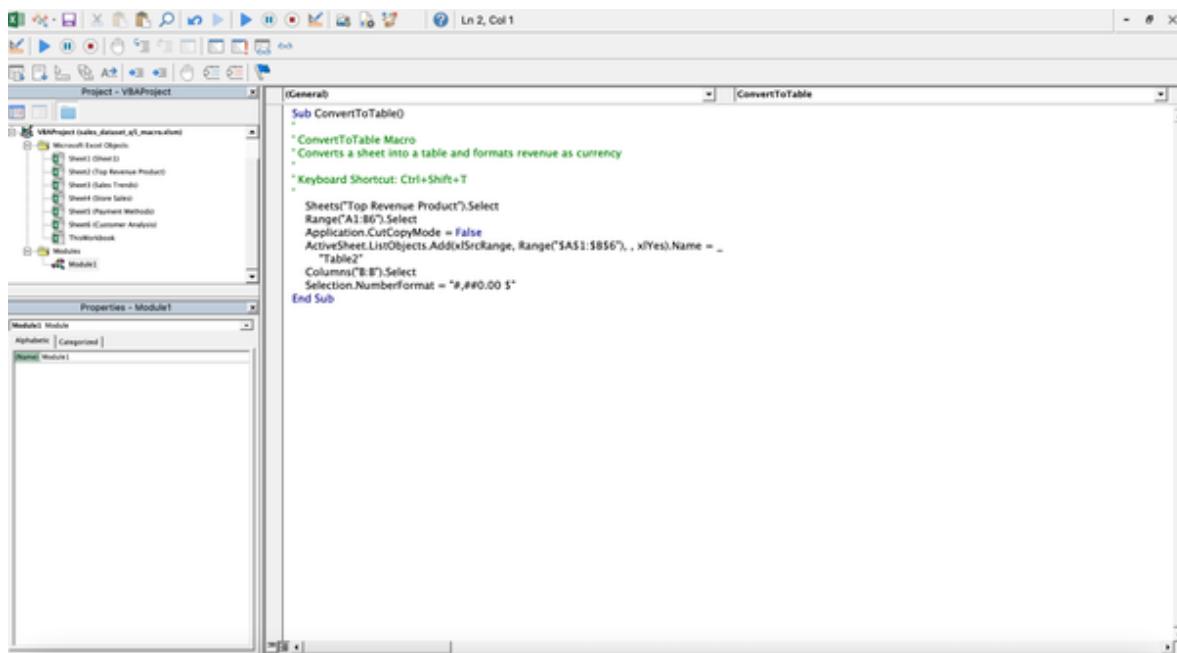


Figure 4-16. The Editor VBA in Excel.

Going into detail, recording the macro generated the following VBA code similar to this:

```
Sub ConvertToTable()
'
' ConvertToTable Macro
' Converts a sheet into a table and formats rever
'
' Keyboard Shortcut: Ctrl+Shift+T
'

    Range("A2").Select
    Application.CutCopyMode = False
    ActiveSheet.ListObjects.Add(xlSrcRange, Range(
        "Table2")
    Range("Table2[Total Revenue]").Select
    Selection.NumberFormat = "General"
    Selection.NumberFormat = "[ $$-en-US]#,##0.00"
End Sub
```

The macro defines the subroutine `ConvertToTable` using the keyword `Sub`. In VBA, subroutines are delimited by keywords `Sub` (to get started) and `End Sub` (to finish). In VBA, comments are identified by the apostrophe (''). Any text that follows an apostrophe in the same line is ignored by the compiler and is useful for documenting the code and improving its readability.

After the comments header, the cell A2 is selected with this instruction:

```
Range("A2").Select
```

This selection selects the data to use. The next instruction is this:

```
Application.CutCopyMode = False
```

It turns off the mode cut/copy, preventing any copy/paste operations from interfering with the macro.

At this point, the macro creates a table (ListObject) using this command:

```
ActiveSheet.ListObjects.Add(xlSrcRange, Range("$A$1:$B$6"), xlYes)
```

This line converts the range A1:B6 in a table, assigning it the name Table2. The parameter xlYes indicates that the first row of the range contains column headings.

To format the data correctly, the macro selects the entire column Total Revenue and formats it as currency:

```
Range("Table2[Total Revenue]").Select  
Selection.NumberFormat = "[[$$-en-US]#,##0.00"
```

Finally, the macro ends with the statement End Sub.

The macro as defined contains specific references to worksheet 2 (such as the range of cells in the table), so when it is applied to another sheet it does not produce the desired results. To fix this, you must generalize the macro. You can do this manually, or you can use ChatGPT. For best results, use the

ChatGPT model optimized for reasoning (at the time of writing this chapter, this model is GPT-o1). To select a specific model in ChatGPT, log in to the web interface and click the name of the current template at the top left. A drop-down menu opens with all available model ([Figure 4-17](#)). Select the model for reasoning (GPT-o1 in my case).

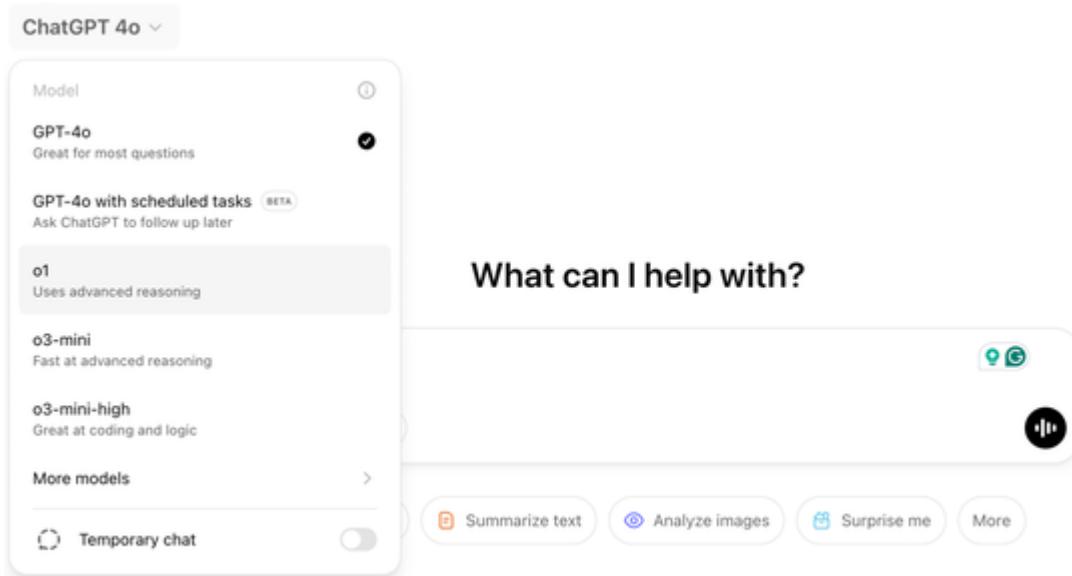


Figure 4-17. The drop-down menu in the ChatGPT web interface allows you to choose the model to use.

Now, ask ChatGPT to modify the VBA code by removing all specific references to the worksheet. Use the following prompt:

```
Modify the following macro as follows: replace the
Sub ConvertToTable()
'
' ConvertToTable Macro
' Converts a sheet into a table and formats rever
'
```

```

' Keyboard Shortcut: Ctrl+Shift+T
'
    Range("A2").Select
    Application.CutCopyMode = False
    ActiveSheet.ListObjects.Add(xlSrcRange, Range("Table2"))
    "Table2"
    Range("Table2[Total Revenue]").Select
    Selection.NumberFormat = "General"
    Selection.NumberFormat = "[ $$-en-US]#,##0.00"
End Sub

```

ChatGPT returns a new macro code, also commenting on the points it changed. In my case, it produced the following code:

```

Sub ConvertToTable()
    '
    ' ConvertToTable Macro
    ' Converts a sheet into a table and formats it
    '
    ' Keyboard Shortcut: Ctrl+Shift+T
    '

    Dim ws As Worksheet
    Dim rng As Range

    ' Work with the active sheet
    Set ws = ActiveSheet

    ' Identify the range of all nonempty cells
    Set rng = ws.UsedRange

    ' Clear any copy/paste mode
    Application.CutCopyMode = False

```

```
' Create a table from the UsedRange. The tab  
ws.ListObjects.Add(SourceType:=xlSrcRange, So  
  
' Format column B (2nd column) as currency  
ws.Columns("B:B").NumberFormat = "#,##0.00 $"  
  
End Sub
```

Replace the original code of the ConvertToTable macro inside the VBA Editor with the code generated by ChatGPT. If the VBA editor does not highlight any lines in red it means that the code does not contain errors. Otherwise, the code generated by ChatGPT is incorrect. In my case, ChatGPT generated a syntax error code, as shown in [Figure 4-18](#).

WARNING

If you replace the entire code, including the Sub and End Sub lines, you'll need to manually put the keyboard shortcut back in. You'll need to do this every time you replace the code. To restore the shortcut key, select Macros from the Developer tab. Then select the ConvertToTable macro and click Options. A new dialog box opens. Re-enter the shortcut key Ctrl+Shift+T and click OK. In the original dialog box, click Run if you want to run the macro or Cancel if you don't want to run it.



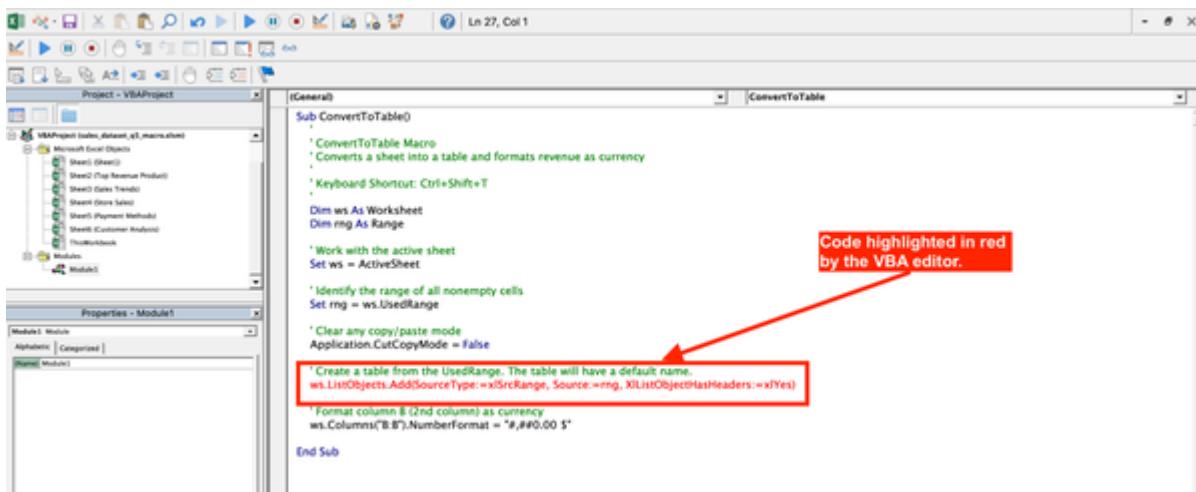


Figure 4-18. The error highlighted by the VBA Editor.

To resolve the error, continue the conversation with ChatGPT by typing the following prompt:

Rewrite and optimize the code by fixing the following error:

Replace the error with the specific error you got. ChatGPT generates a new version of the code. In my case, it generated the following code:

```
Sub ConvertToTable()
    '
    ' ConvertToTable Macro
    ' Converts the active sheet's used range into a table
    ' and formats revenue as currency in Column E
    ' Keyboard Shortcut: Ctrl+Shift+T
    '
    Dim ws As Worksheet
    Dim rng As Range
```

```
' Work with the active sheet
Set ws = ActiveSheet

' Identify the range of all nonempty cells
Set rng = ws.UsedRange

' Clear any copy/paste mode
Application.CutCopyMode = False

' Create a table from the UsedRange.
' By using HasHeaders:=xlYes, Excel will try
' The table will have a default name (e.g.,
ws.ListObjects.Add SourceType:=xlSrcRange, _
                    Source:=rng, _
                    HasHeaders:=xlYes

' Format column B as currency
ws.Columns("B:B").NumberFormat = "#,##0.00 $"
End Sub
```

Paste the code into the VBA Editor. In this case, it contains no syntax errors, so we can analyze it to see if it contains no conceptual errors.

ChatGPT eliminated the reference to specific cells, selecting all non-empty cells:

```
' Identify the range of all nonempty cells
Set rng = ws.UsedRange
```

Next, it removed the reference to the specific table name. Conceptually, the macro seems correct, so we can save it and apply it to the Sales Trends worksheet to see if it works.

1. Click the Sales Trends worksheet.
2. Select the Developer tab in the Ribbon and then Macros.
3. Select the ConvertToTable macro and click Run.

Running the macro, in my case, the VBA compiler gives the following error:

```
Compile error: Named argument not found
```

The error is related to the line of code:

```
ws.ListObjects.Add SourceType:=xlSrcRange, _  
    Source:=rng, _  
    HasHeaders:=xlYes
```

Use ChatGPT again to fix the problem. Continue the conversation by writing the following prompt containing the error generated by the compiler:

```
I got this error: Compile error: Named argument r
```

ChatGPT tries to fix the problem by generating a new code version. In my case, it generated the following code:

```
Sub ConvertToTable()
    '
    ' ConvertToTable Macro
    ' Converts the active sheet's used range into
    ' a table and formats revenue as currency in Column B
    ' Keyboard Shortcut: Ctrl+Shift+T
    '

    Dim ws As Worksheet
    Dim rng As Range

    ' Work with the active sheet
    Set ws = ActiveSheet

    ' Identify the range of all nonempty cells
    Set rng = ws.UsedRange

    ' Clear any copy/paste mode
    Application.CutCopyMode = False

    ' Create a table from the UsedRange using pos
    ' 1. SourceType (xlSrcRange)
    ' 2. Source (rng)
    ' 3. LinkSource (omit or blank)
    ' 4. XlListObjectHasHeaders (xlYes or xlNo)
    ws.ListObjects.Add xlSrcRange, rng, , xlYes

    ' Format column B as currency
    ws.Columns("B:B").NumberFormat = "#,##0.00 $"
End Sub
```

Quickly check the generated code. Since everything looks correct, paste it into the VBA Editor and save it. Now, you are ready to apply the macro to a new worksheet. Follow the steps described below to apply the macro to the Sales Trends worksheet:

1. Select the Sales Trends worksheet.
2. Select Macros from the Developer tab.
3. Select the macro ConvertToTables in the dialog box.
4. Click Run.

At this point, the data contained in the Sales Trends worksheet is correctly converted into a table and formatted as required. Reapply the macro to the other worksheets (except the Payment Methods worksheet, which does not contain monetary data).

TIP

Use ChatGPT not only to modify the code of an existing macro but also to generate and debug the code.)

From the example described we can extract a general workflow ([Figure 4-19](#)) that you can always apply to generalize the macros you record.

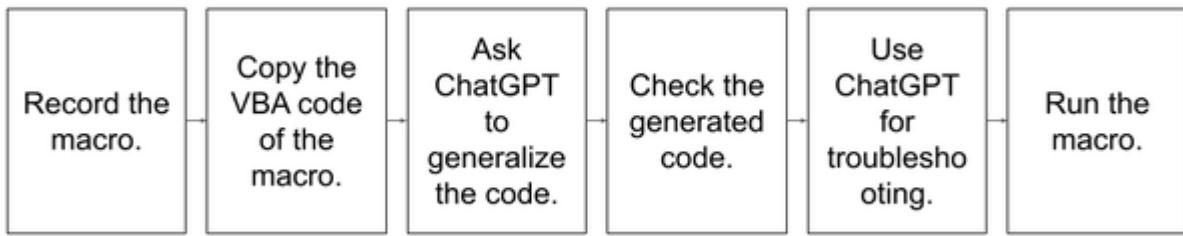


Figure 4-19. The workflow to generalize the recorded macros.

First record the macro, then copy the VBA code of the macro and ask ChatGPT to generalize the code, eliminating all parts that specifically refer to a particular worksheet. Quickly check if the code generated by ChatGPT is correct and test it in Excel. If there are problems, use ChatGPT to troubleshoot. Finally, run the macro.

You can further generalize the workflow by directly asking ChatGPT to generate the macro in VBA, without recording it manually. Let's see how to proceed in the next section.

Using ChatGPT for Macro Generation

In previous chapters you have seen the importance of defining the context to ensure that ChatGPT generates the correct result. You also saw that you can define the context in two ways:

1. Describe the structure of your dataset in words, specifying the name of the columns and possibly also the name of the Excel table, if the dataset is formatted as an Excel table.

2. Provide ChatGPT with the dataset (or a portion thereof) you want to operate on as input directly. This second option is faster but is less suitable if you are dealing with sensitive data.

Even if you want to have ChatGTP generate macro code directly, you need to specify the context. As in the other cases, organize your conversation with ChatGPT in the form of a story: context, main part, and conclusions. The context contains the structure of your table, the main part the instructions for the macro, and the conclusions where you want to save the results of your macro.

Let's take an example. Consider the dataset again [04/sales_dataset_q5.xlsx](#). Suppose you want to add a new worksheet to your workbook containing a weekly sales report for each product. Use the following prompt with context, main body, and conclusions to ask ChatGPT to generate a macro that creates the report:

Consider an Excel table named MainTable representing sales data:
- Date of Sale,
- Product ID,
- Product Name,
- Quantity Sold,
- Unit Price,
- Total Sale Amount,
- Payment Method,
- Store Location,
- Customer Type.

Write a VBA macro to analyze the dataset and generate a report.

1. Group by Week and Product: Aggregate the data.
2. Create a New Report Sheet: Save the results.
3. Format the Report: Convert to a Table.

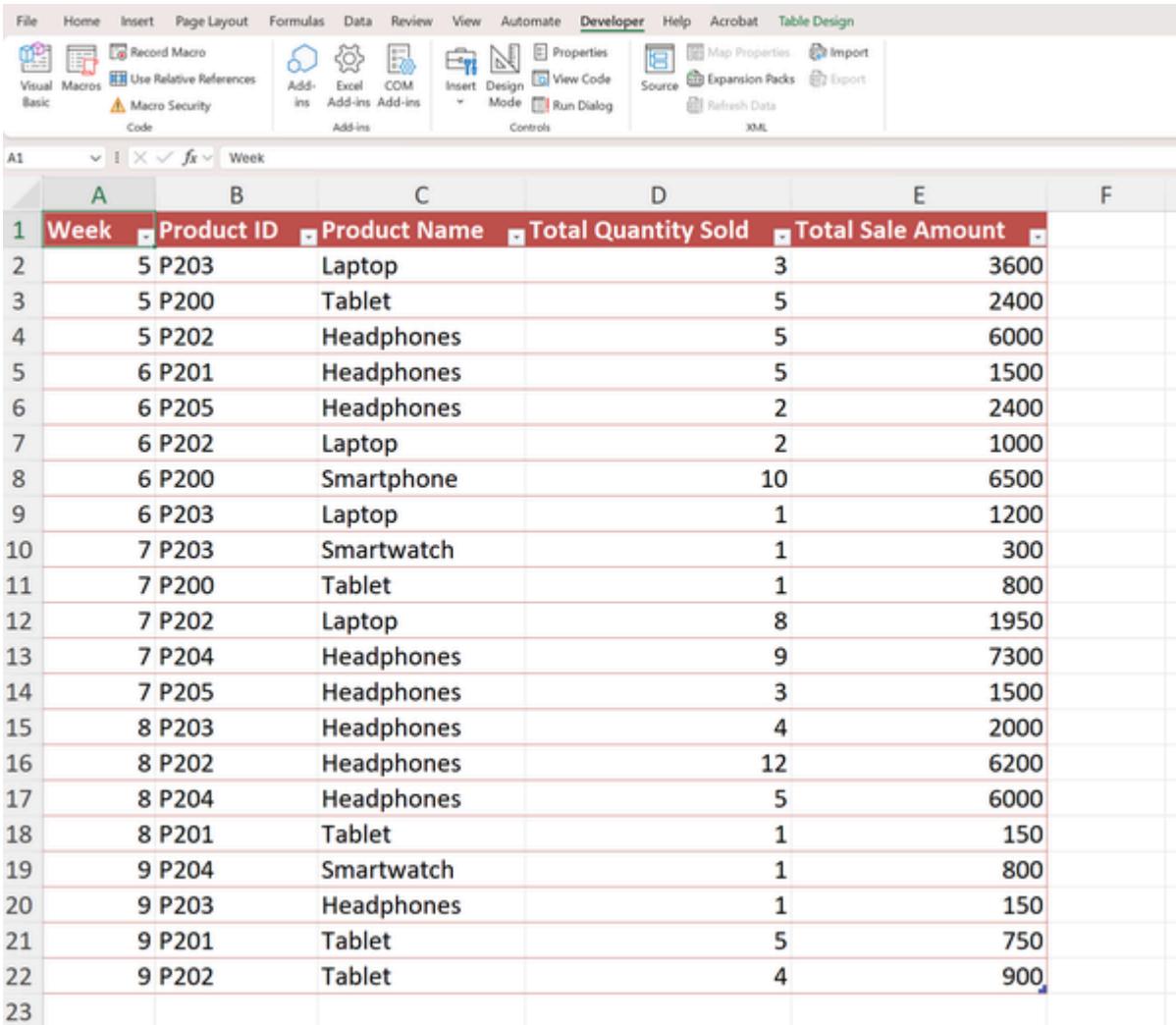
Use ChatGPT's advanced reasoning model (in my case, o1). ChatGPT generates the macro code and a short comment. In my case, ChatGPT generated quite an extended code, which you can find in the book's GitHub repository at [04/create weekly sales report macro.txt](#). The VBA code analyzes the sales data in a list (`ListObject`) called MainTable in Excel, grouping them by week and by product. First, create or overwrite a new sheet called Sales Report. Then, use a dictionary (`Scripting.Dictionary`) to add the quantity sold and the total amount of each product, divided by week (calculated with the function `WeekNum`). Finally, write the results to the Sales Report sheet and convert them into a formatted table.

Add the macro to the Excel file as follows:

1. Click the Developer tab on the Ribbon, and in the Code group click Visual Basic.
2. In the Visual Basic Editor, look for the Project pane on the left (it lists all open workbooks and their objects). Right-click VBAProject (`sales_dataset_q5.xlsx`).
3. Select Insert → Module. A new blank module will appear in the code window (usually titled Module1 by default).

4. Paste the code generated by ChatGPT. Click Save if the file is already in the xlsm format. Otherwise, click Save As, give a name to your file, and choose Excel Macro Enabled Workbook as File Format. Also, close the VBA window.

Run the macro by selecting Macros in the Developer tab and then followed by Run. A new worksheet named Sales Report is added to the workbook, as shown in [Figure 4-20](#).



The screenshot shows the Microsoft Excel ribbon with the 'Developer' tab selected. Below the ribbon, there's a toolbar with icons for Visual Basic, Macros, Record Macro, Use Relative References, Add-ins, Excel Add-ins, COM Add-ins, Insert, Design Mode, Properties, View Code, Run Dialog, Controls, Source, Refresh Data, Map Properties, Expansion Packs, Import, Export, and XML. The main area of the screen displays a worksheet titled 'Sales Report'. The table has columns labeled 'Week', 'Product ID', 'Product Name', 'Total Quantity Sold', and 'Total Sale Amount'. The data is as follows:

Week	Product ID	Product Name	Total Quantity Sold	Total Sale Amount
2	5 P203	Laptop	3	3600
3	5 P200	Tablet	5	2400
4	5 P202	Headphones	5	6000
5	6 P201	Headphones	5	1500
6	6 P205	Headphones	2	2400
7	6 P202	Laptop	2	1000
8	6 P200	Smartphone	10	6500
9	6 P203	Laptop	1	1200
10	7 P203	Smartwatch	1	300
11	7 P200	Tablet	1	800
12	7 P202	Laptop	8	1950
13	7 P204	Headphones	9	7300
14	7 P205	Headphones	3	1500
15	8 P203	Headphones	4	2000
16	8 P202	Headphones	12	6200
17	8 P204	Headphones	5	6000
18	8 P201	Tablet	1	150
19	9 P204	Smartwatch	1	800
20	9 P203	Headphones	1	150
21	9 P201	Tablet	5	750
22	9 P202	Tablet	4	900

Figure 4-20. The result obtained by applying the macro.

Exercise 2

Consider again the dataset [*04/sales dataset q5.xlsx*](#), write a prompt that solves the following problem: write a VBA macro that analyzes the dataset and automatically highlights high and low sales based on a user-defined threshold.

You can find the solution in the book's GitHub repository at
[*04/exercise3_prompt.txt*](#).

Conclusions on Part 1

We have reached the end of this book's first part, where you saw how to combine ChatGPT and Excel to improve efficiency in data analysis, process automation, and information management. We have explored different aspects of the integration between ChatGPT and Excel, each with its specifics and advantages.

We began our journey with an overview of Excel's role in data management, analysis, and visualization, highlighting how AI can enhance these capabilities. ChatGPT helps you reduce errors, speed up repetitive tasks, and improve productivity, especially for those working with large amounts of data. Subsequently, you learned how to interact with ChatGPT to obtain effective responses using prompt engineering techniques. You've seen how to structure a conversation with ChatGPT to get formulas, functions, and practical tips to improve your daily Excel work. Then you learned how ChatGPT can simplify and automate recurring tasks in Excel, such as entering and formatting data, calculating key metrics, and creating custom reports. Finally, you explored advanced features, such as using ChatGPT to generate complex formulas, analyze datasets, create graphs, and even write VBA macros. You have seen how ChatGPT can support you in debugging and optimizing code, improving the quality of analyses and decision-making processes.

Part 1 of this book provided a solid foundation on leveraging ChatGPT to enhance the use of Excel, automating repetitive tasks and making analyses faster and more intuitive. Part 2 focuses on Copilot, another powerful AI integration with Excel, which offers even more advanced tools for analyzing and managing data.

I want to conclude my thoughts about Part 1 with a personal reflection. ChatGPT is a powerful tool capable of speeding up work and simplifying many tasks. But there is a real risk that it will make us unaccustomed to reasoning, learning, and questioning information. How many times do we take what it suggests to us as good without stopping to verify it?

That already happened to me with some data journalism students who had entrusted their project to a developer, paying him handsomely. The result? Impeccable work, but they hadn't learned anything. They couldn't explain how it worked; in the end, their project wasn't theirs. In addition, they couldn't then apply what they would have learned elsewhere, so the problem goes beyond the actual project.

The same happens with ChatGPT. We can use it to generate code, write texts, or solve problems, but if we don't take the time to verify, understand, and reason, we risk becoming just passive executors, incapable of truly growing.

Let's not let convenience turn into laziness, or let laziness translate into a loss of quality. ChatGPT should be an ally, not a substitute for critical thinking. For this reason, my advice is always the same: check, experiment, question. Only in this way can we make the best use of technology without stopping learning.

Summary

In this chapter, you learned how to use ChatGPT to analyze a dataset, define macros, and write VBA code.

- ChatGPT helps you speed up data analysis in Excel by automatically identifying dataset categories, suggesting relevant questions, and generating advanced formulas to extract insights more effectively.
- Following a structured workflow, you can load a dataset, obtain key questions, generate answers via Excel formulas, and implement them directly in the file, automating complex operations without advanced programming knowledge.
- Using ChatGPT for VBA macros allows you to create, generalize, and fix code efficiently, reducing errors and making repetitive tasks faster and more customizable to specific spreadsheet needs.
- Through automation with macros and VBA, you can generate weekly reports, format data dynamically, and highlight key information,

improving data organization and analysis with a more structured and automated approach.

Activity

Solve the crossword puzzle shown in [Figure 4-21](#). The answer is given in [Appendix A](#).

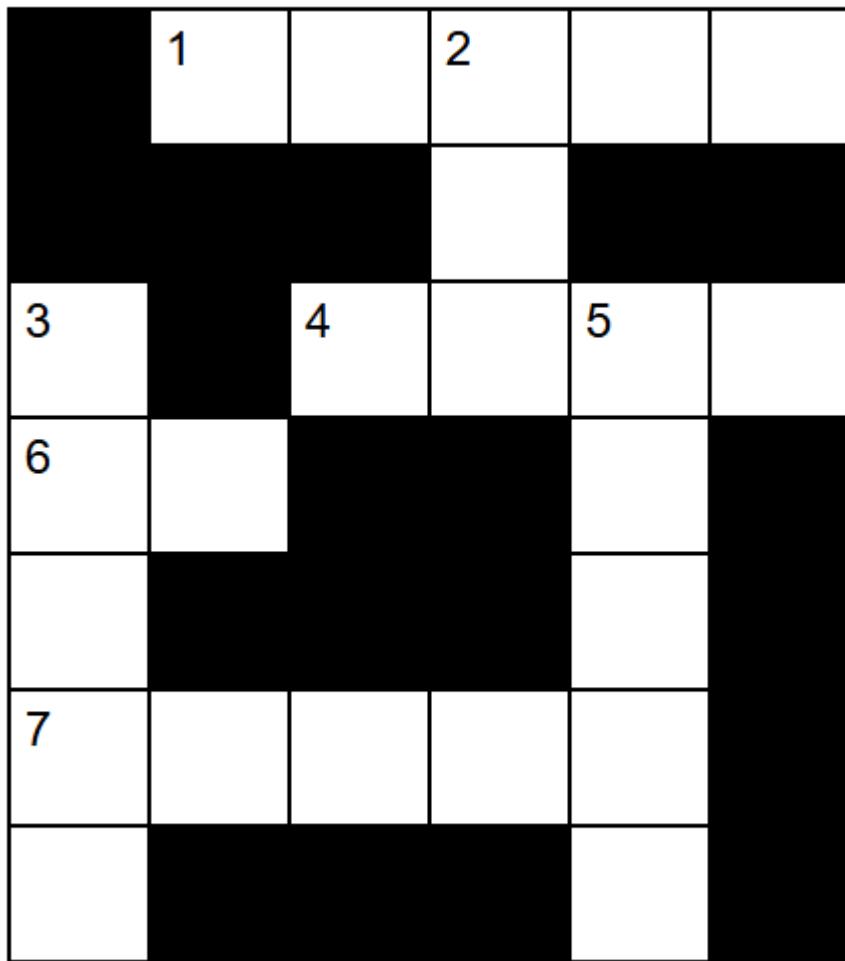


Figure 4-21. Crossword to solve

Across

1 - A type of table in Excel used to summarize and analyze data dynamically.

4 - Raw facts and numbers that can be analyzed or processed in a spreadsheet.

6 - Technology that enables machines to simulate human intelligence and decision-making.

7 - Widely used spreadsheet software developed by Microsoft.

Down

2 - A programming language used in Excel to automate tasks and create macros.

3 - Transactions where goods or services are exchanged for money.

5 - A structured arrangement of data in rows and columns within a spreadsheet.

Appendix A. Solutions to Fun Activities

A NOTE FOR EARLY RELEASE READERS

With Early Release ebooks, you get books in their earliest form—the author’s raw and unedited content as they write—so you can take advantage of these technologies long before the official release of these titles.

This will be Appendix A of the final book. Please note that the GitHub repo will be made active later on.

If you have comments about how we might improve the content and/or examples in this book, or if you notice missing material within this chapter, please reach out to the editor at ccollins@oreilly.com.

Chapter 1

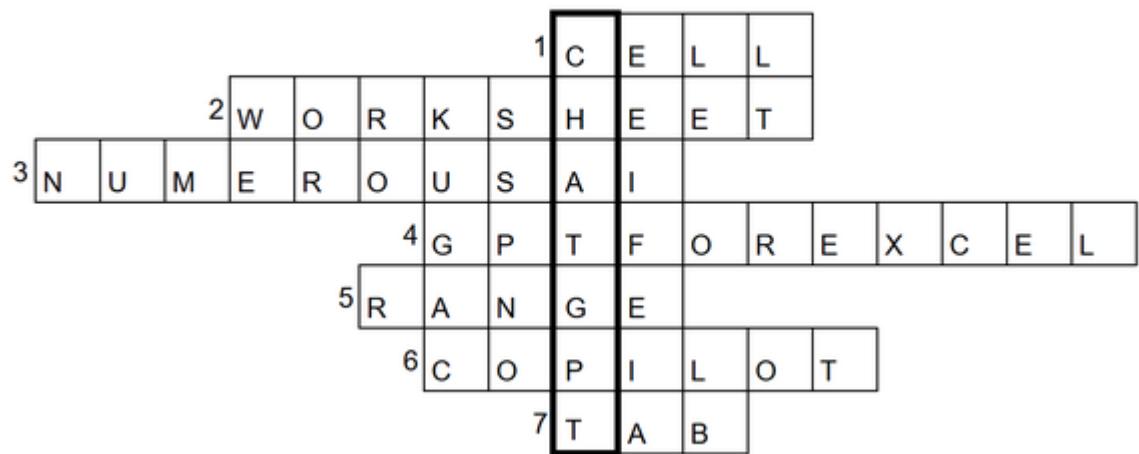


Figure A-1. Placeholder

Chapter 2

In this chapter, you learned how to structure an effective PROMPT to best use CHATGPT in Excel, organizing it into three fundamental parts: context, main point, and end. The CONTEXT defines the data and information necessary to frame the problem, the MAIN POINT describes the action required, while the END indicates how to use the result.

ChatGPT can be used as an ASSISTANT to perform various tasks, such as generating formulas for CALCULATION (for example, summing or averaging) or performing search operations with functions like VLOOKUP. This combination of strategies helps you improve productivity, automate complex calculations, and quickly resolve problems in your datasets.

Chapter 3

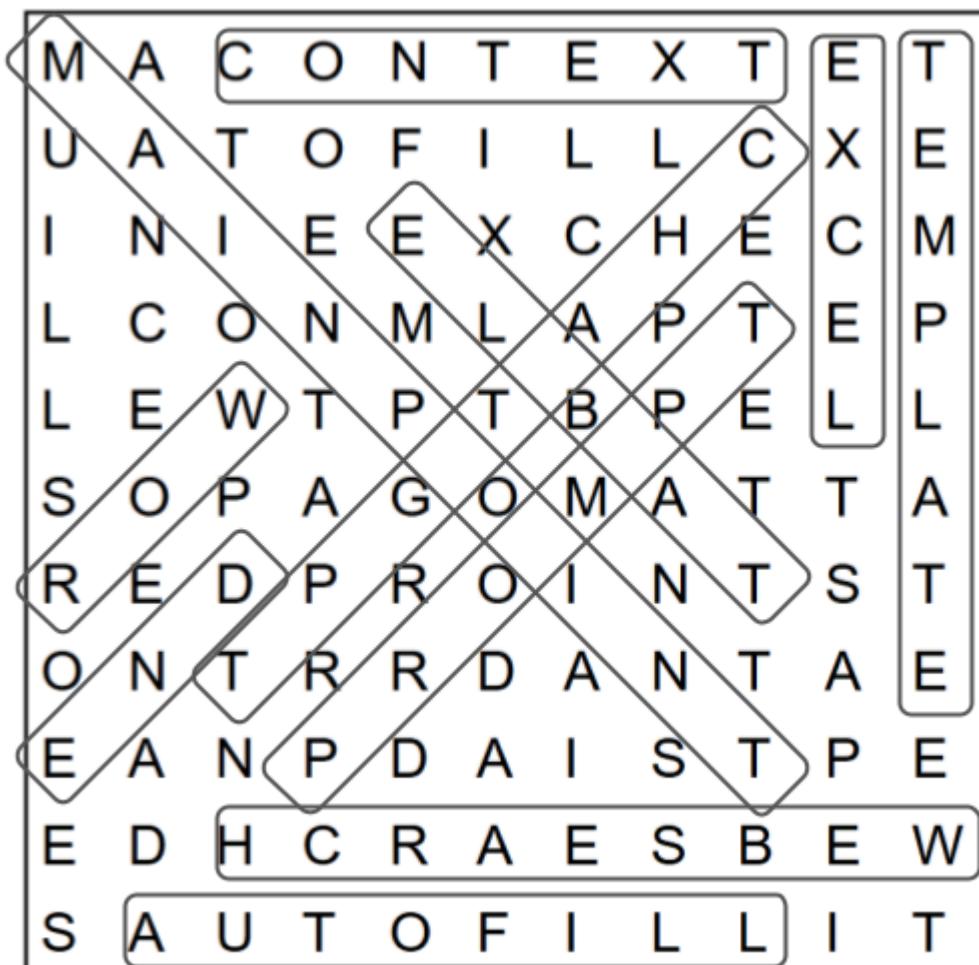


Figure A-2. Placeholder

Autofill

AUTOFILL IN EXCEL COMPLETES PATTERNS OR DATA AND AI SPEEDS IT.

Chapter 4

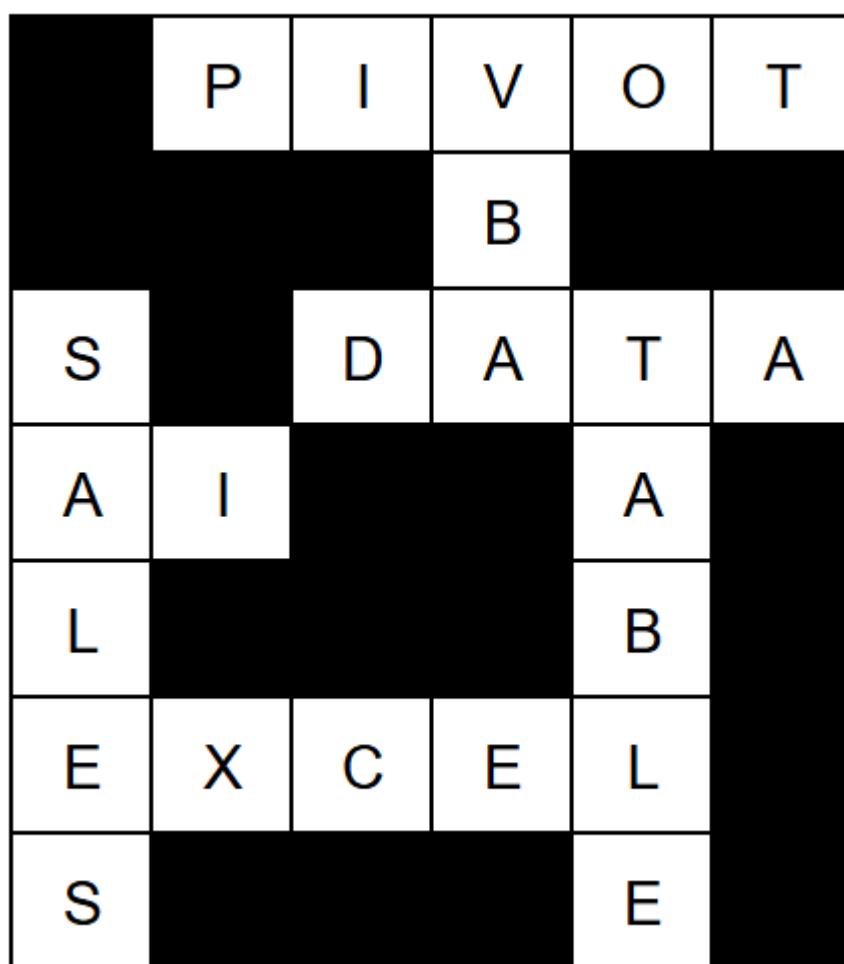


Figure A-3. Placeholder