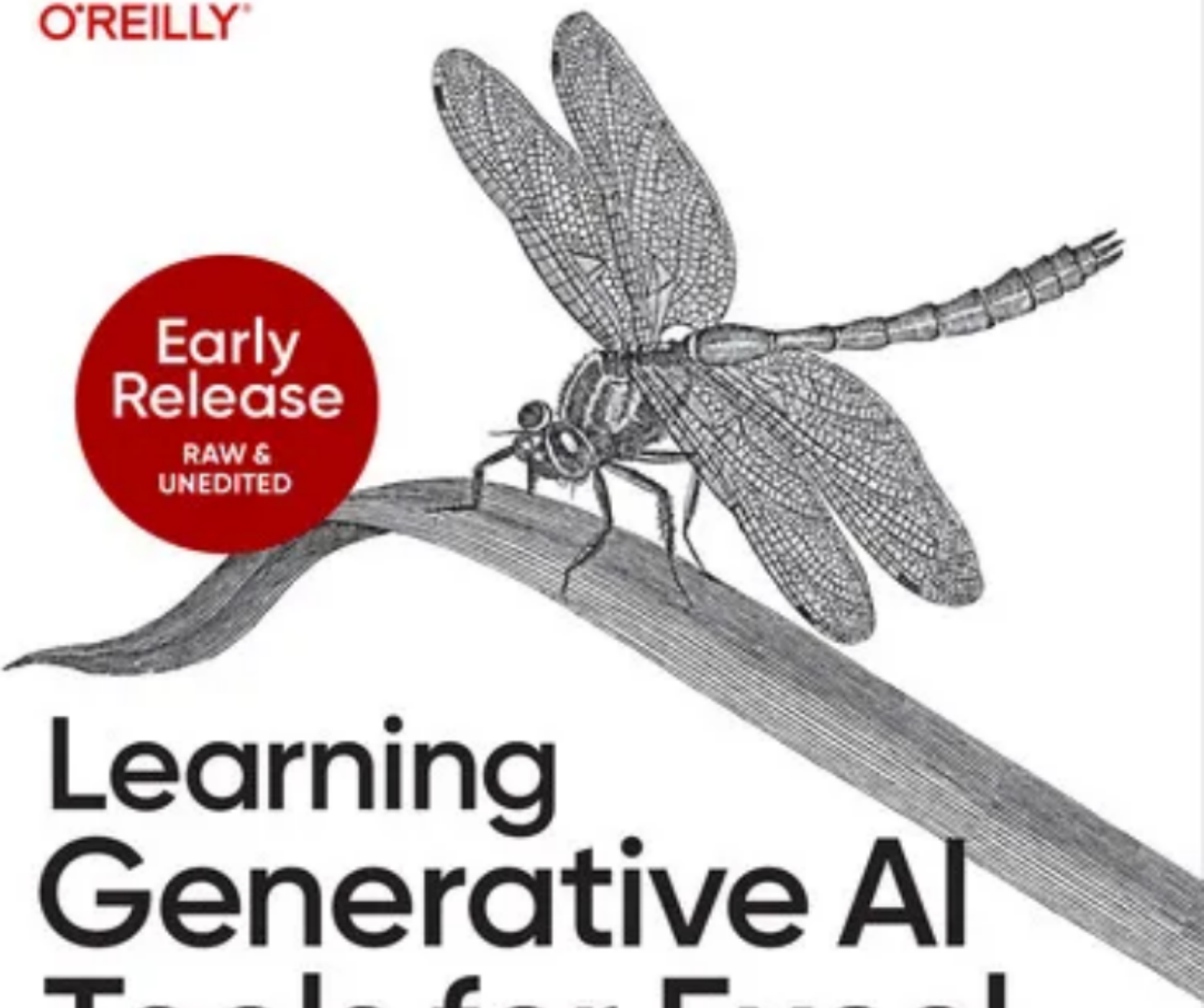
O'REILLY



Early Release

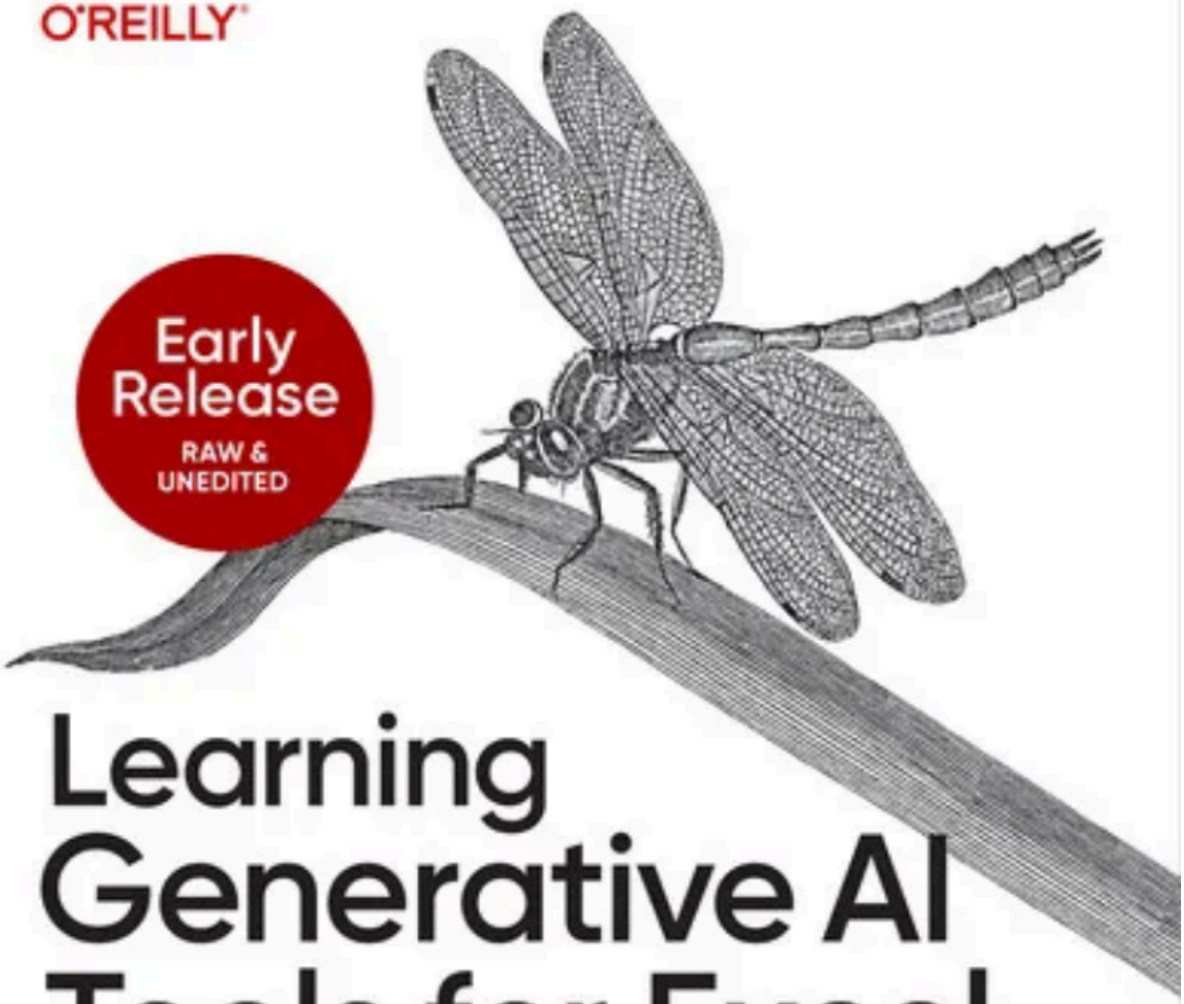
RAW & UNEDITED

Learning Generative AI Tools for Excel

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

Angelica Lo Duca

O'REILLY



Early Release

RAW & UNEDITED

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

O'REILLY®

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

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 (unavailable)

Chapter 4: Advanced Operations with ChatGPT (unavailable)

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)

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 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. [End note]

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.

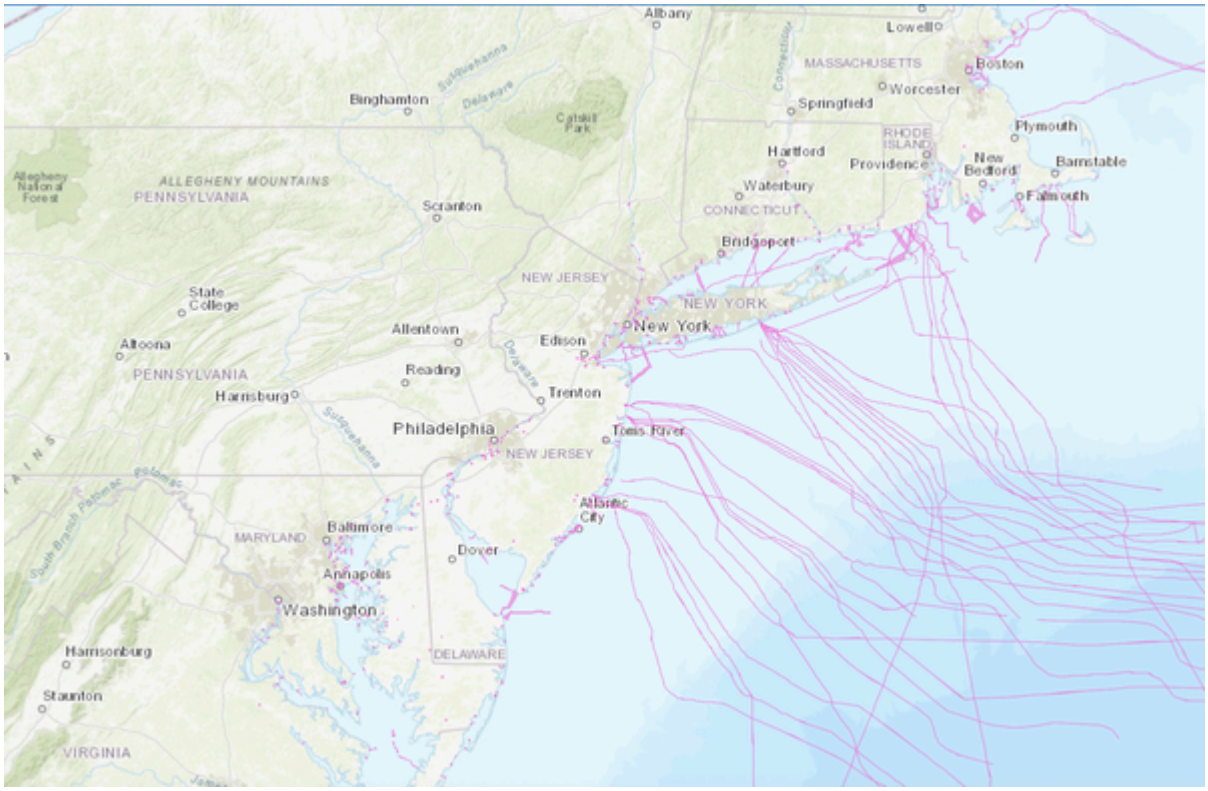


Figure 1-1. A map representation of a portion of the U.S. submarine cable dataset.

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__Lengt
AT&T Corp	Americas II	East Coast	108697.599067
GCI Communication Corp	AU-Aleutian Submarine Cable System	Pacific Northwest	1129434.704075
GU Holdings Inc	Curie	West Coast	1274007.146157
Eversource Energy	Eversource HEEC Cable	East Coast	10111.14415525
GU Holdings Inc	Firmina	East Coast	1006081.053550
GU Holdings Inc	Grace Hopper	East Coast	1493252.138575

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

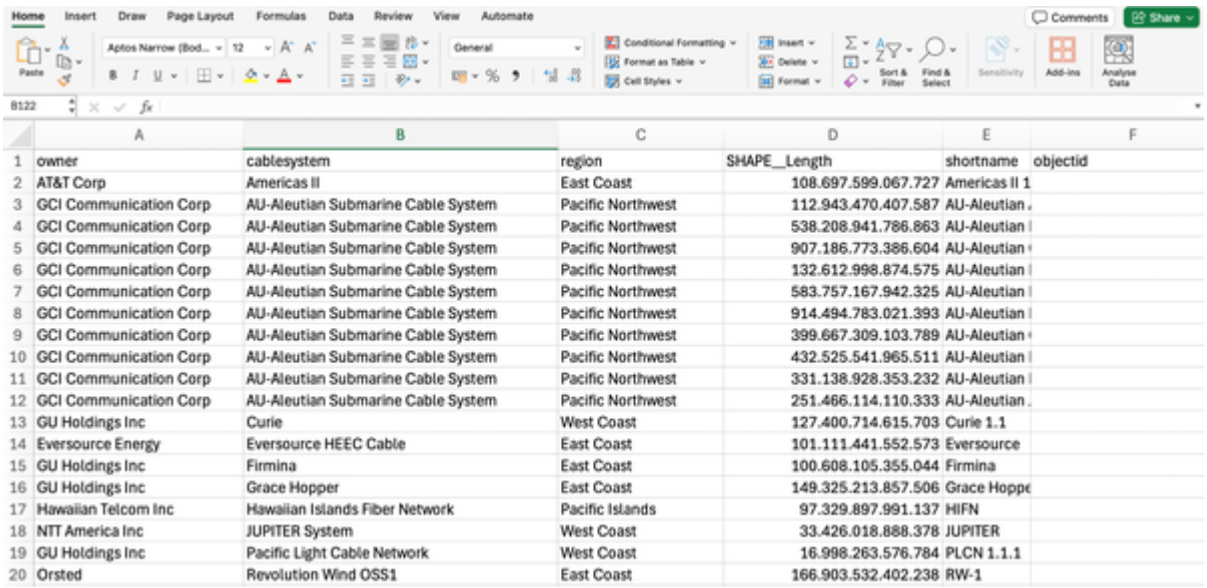
Data management involves collecting, organizing, storing, and maintaining data while ensuring its accessibility and security. The first step to managing data is to load the dataset as a table. Excel supports various formats, such as CSV, but to fully exploit its potential, here you will save the dataset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

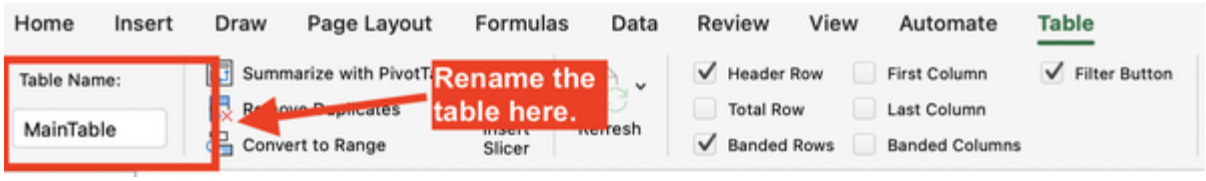


Figure 1-3. Figures 1-3. How to rename a table in Excel.

At this point, you are ready to work.

NOTE

[begin Note icon]

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.

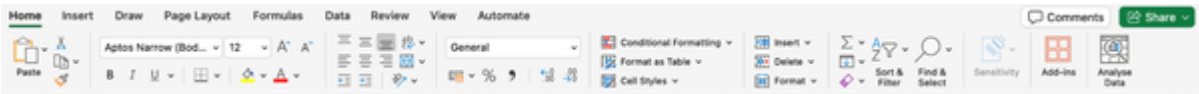


Figure 1-4. The ribbon in Excel.

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	

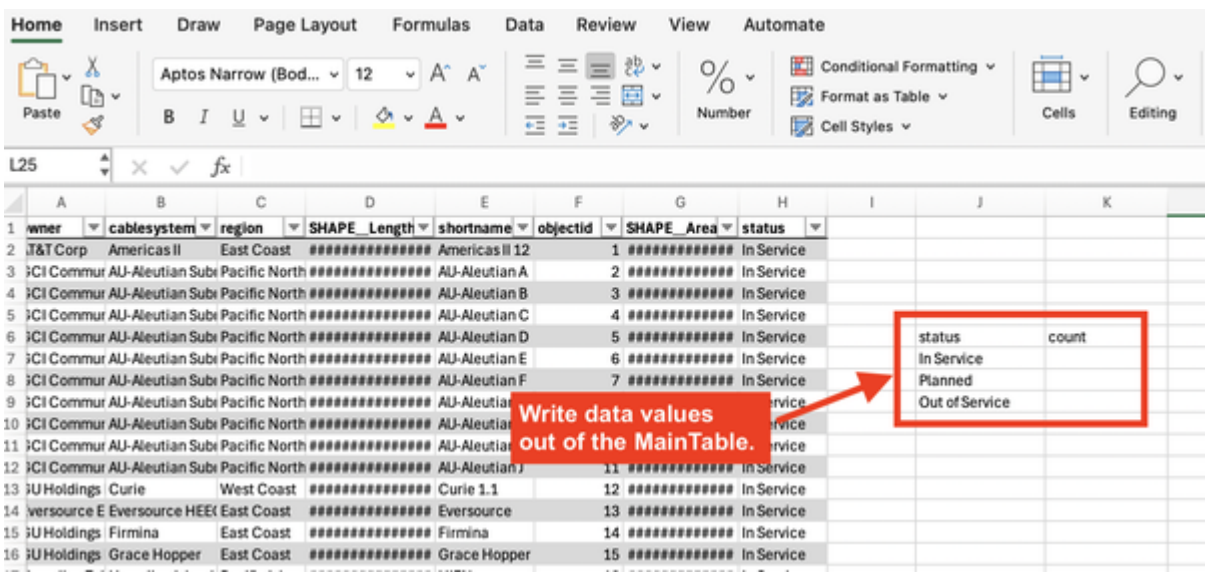


Figure 1-5. A possible place to insert data values.

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).

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:

1. 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 1-7 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 ∑ 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.

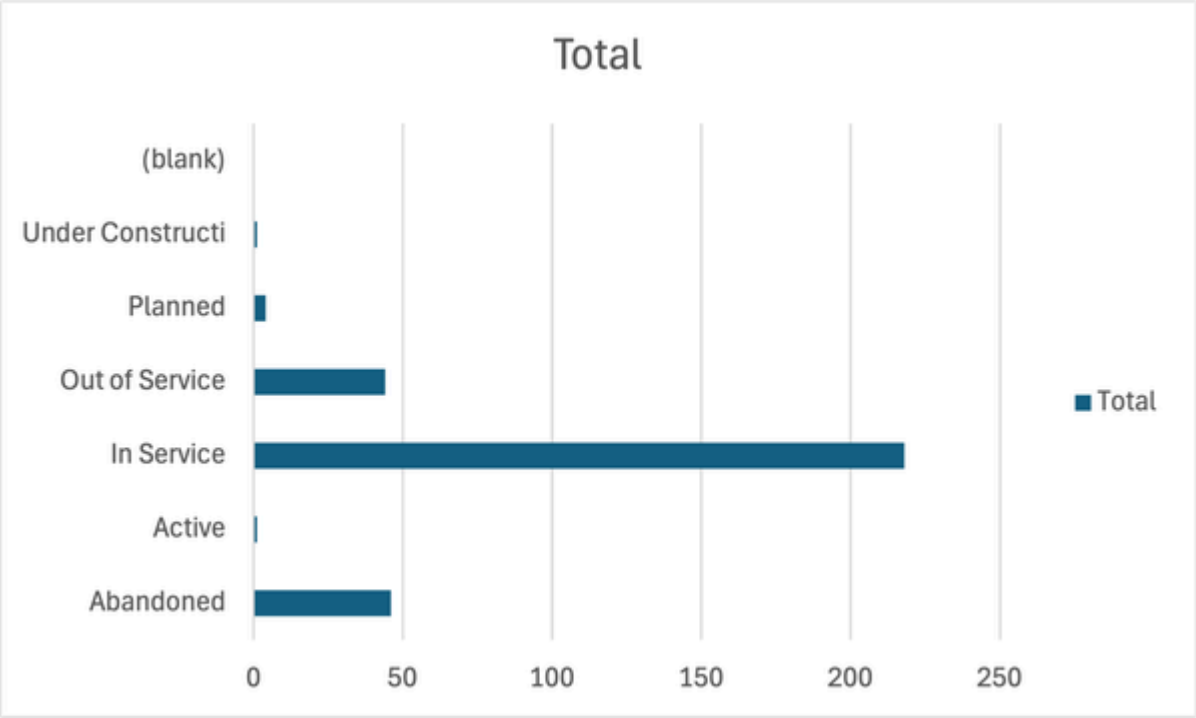


Figure 1-9. The bar chart built from the pivot table.

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 *01/01.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.

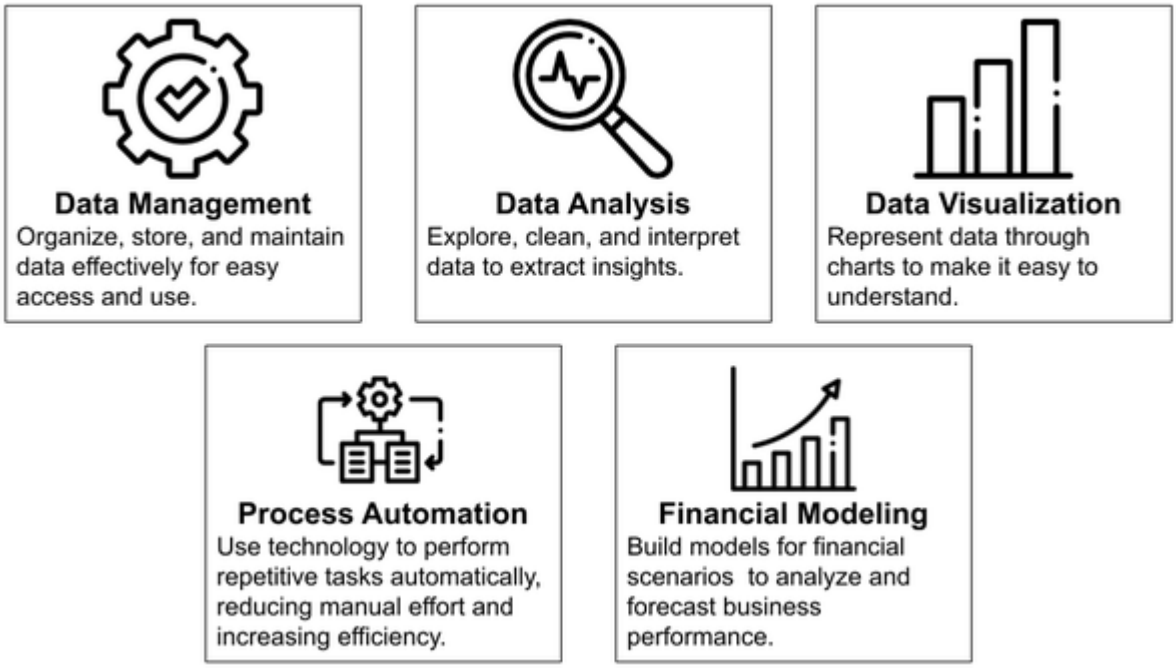


Figure 1-10. The multiple purposes of using Excel.

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.

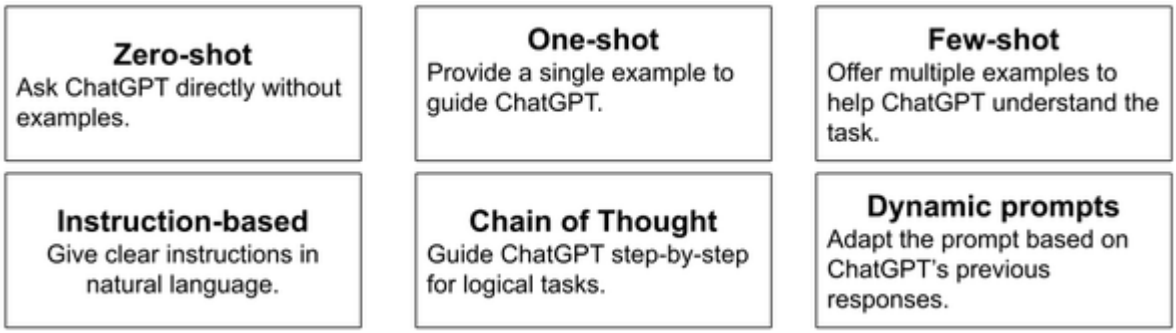


Figure 1-11. The different prompt techniques.

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 version of ChatGPT 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.

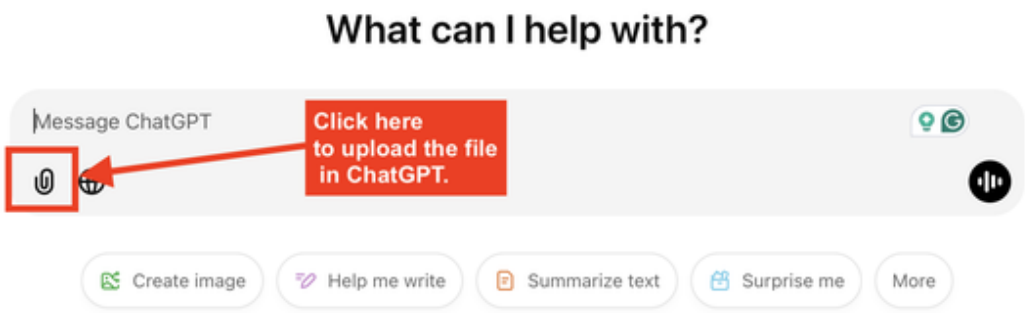


Figure 1-12. Figures 1-12. How to load a file in ChatGPT.

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

```
Build a pivot table for the column status
and count the number of occurrences for each
category.
```

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.

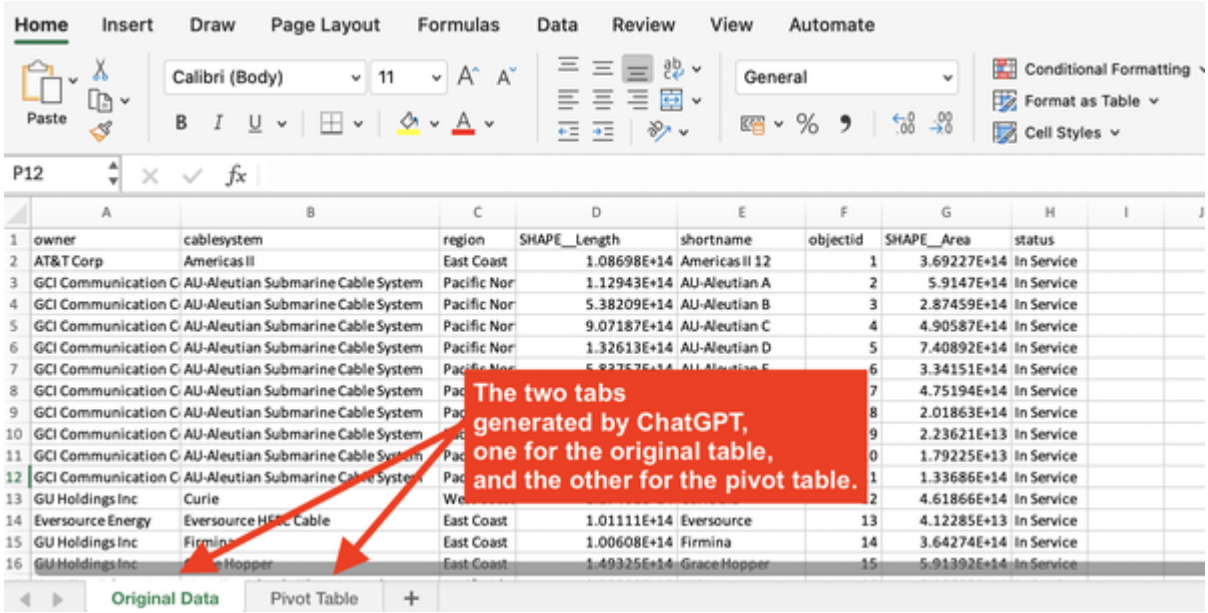
The next step is to export the output as an Excel file. Write the following prompt in ChatGPT, still in the same chat where you wrote the previous prompt:

```
Format the original table and the pivot table
as two worksheets of a single workbook in
Excel. Calculate the pivot table using
formulas instead of raw data. Use column H of
the original table in the pivot table
formulas. Make the output downloadable as a
.xslx file.
```

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).



	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				6	3.34151E+14	In Service
8	GCI Communication C	AU-Aleutian Submarine Cable System				7	4.75194E+14	In Service
9	GCI Communication C	AU-Aleutian Submarine Cable System				8	2.01863E+14	In Service
10	GCI Communication C	AU-Aleutian Submarine Cable System				9	2.23621E+13	In Service
11	GCI Communication C	AU-Aleutian Submarine Cable System				0	1.79225E+13	In Service
12	GCI Communication C	AU-Aleutian Submarine Cable System				1	1.33686E+14	In Service
13	GU Holdings Inc	Curie				2	4.61866E+14	In Service
14	Eversource Energy	Eversource HEX 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.

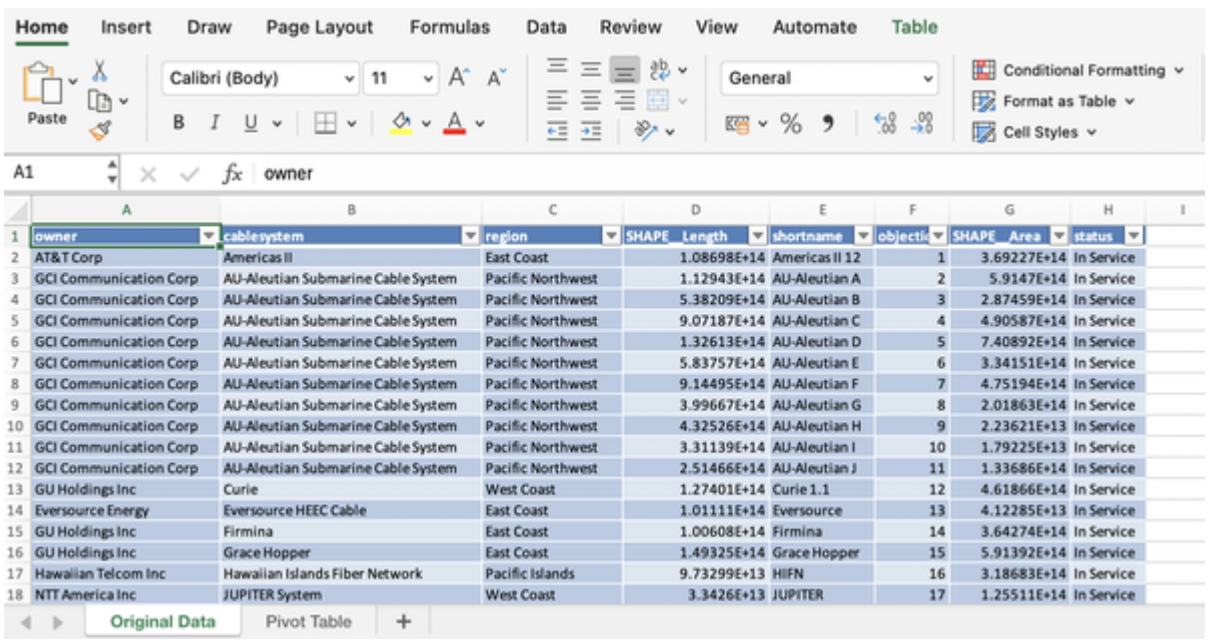
The downloaded file is very similar to the one we produced manually, with some differences:

- 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 two Excel tables into two worksheets of a
single workbook in Excel. Calculate the pivot
table using formulas instead of raw data. Use
column H of the original table in the pivot
table formulas. Make the output downloadable
as a .xslx file.
```

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. Add the chart as an additional
worksheet of the previous workbook. Make the
output downloadable as a .xslx file.
```

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

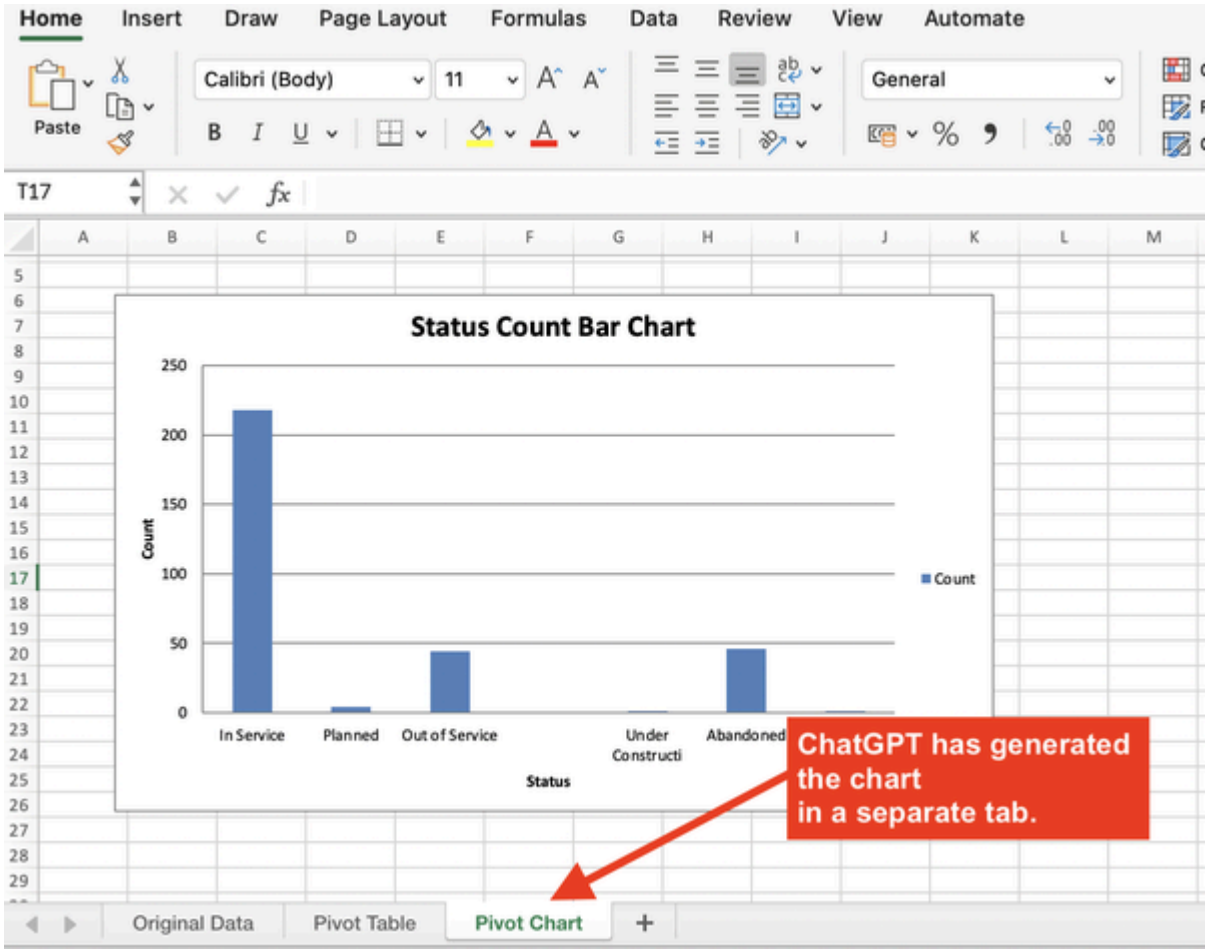


Figure 1-16. The Excel file generated by ChatGPT with the bar chart.

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 is 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 moreetc. It allows you to automate repetitive tasks such as data cleansing, applying complex formulas, and generateing 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 facts 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 saw 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.

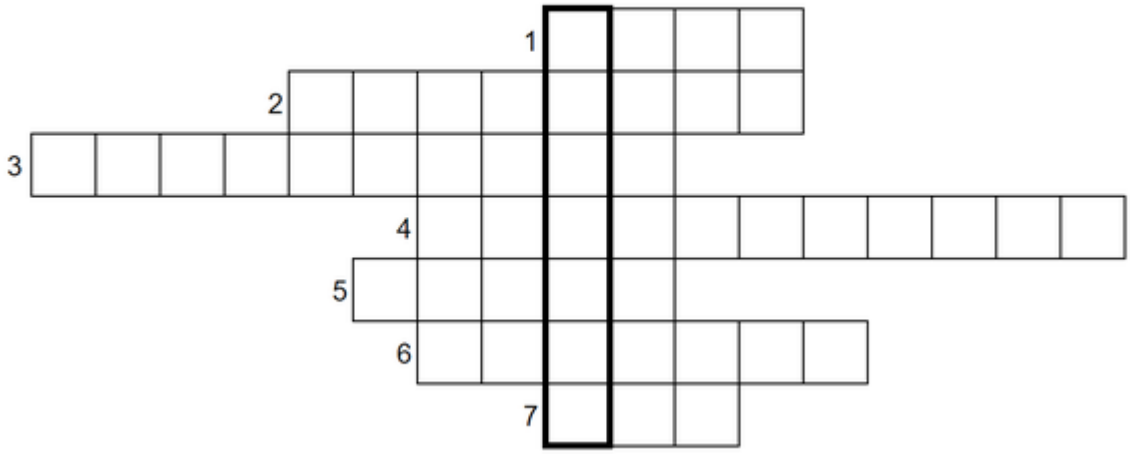


Figure 1-17. Crossword puzzle to reveal the tool you will use next

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. A set of selected cells.
5. The built-in AI tool provided by Excel. You'll use it in the second part of this book.
6. 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*.

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 2-1.

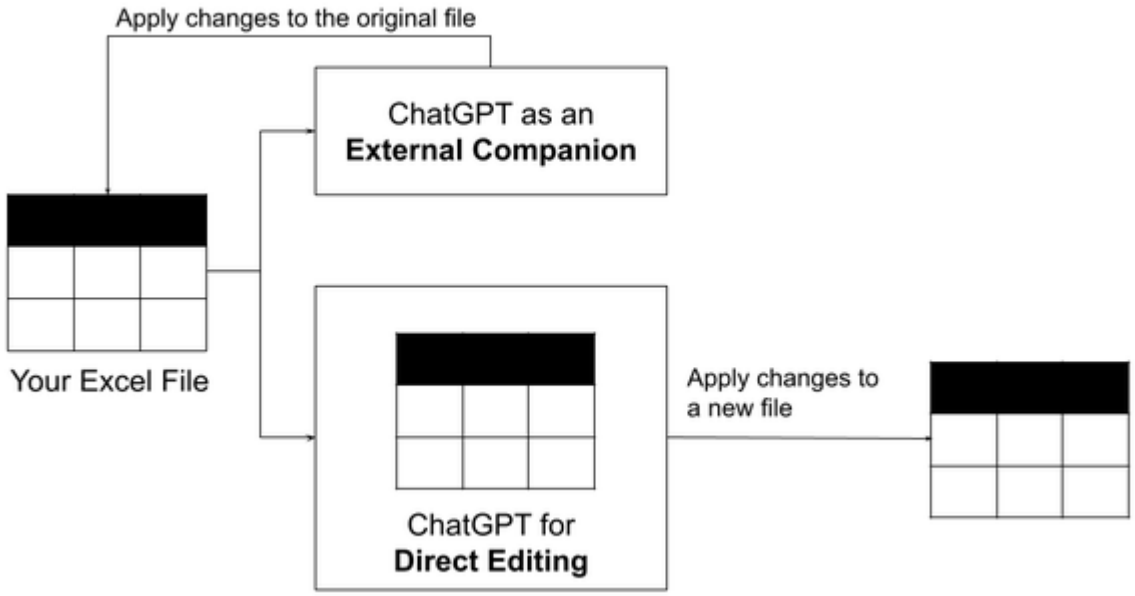


Figure 2-1. The two possible uses of ChatGPT in Excel.

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.

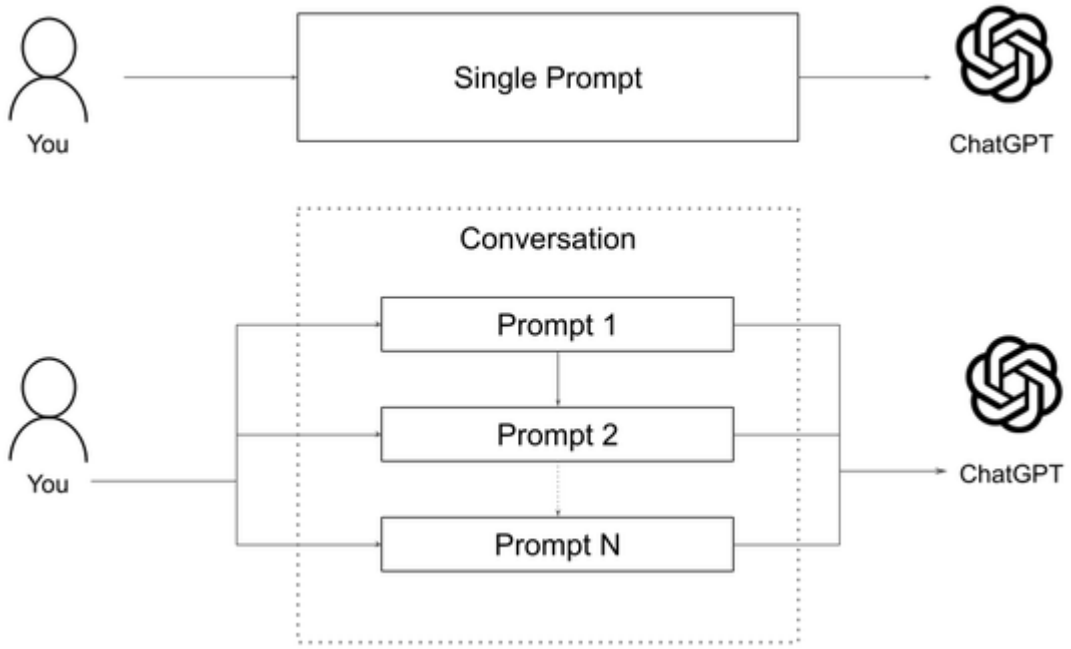


Figure 2-2. The two ways of communicating with ChatGPT.

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.

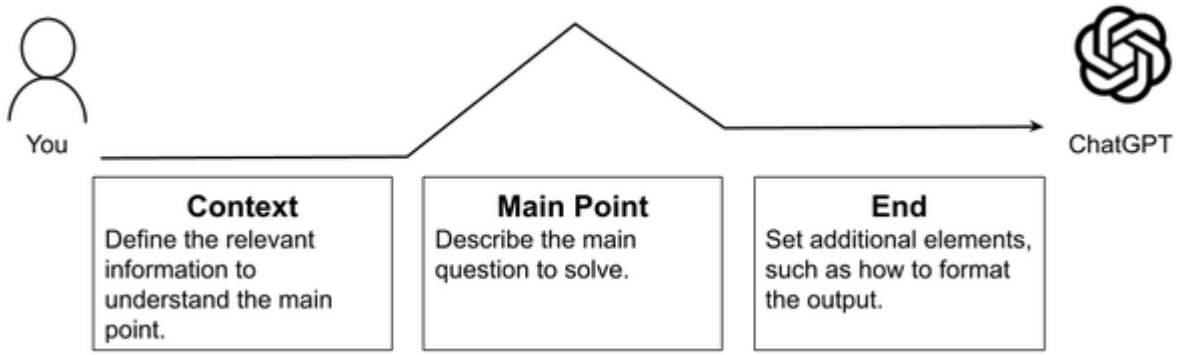


Figure 2-3. How to organize a conversation with ChatGPT in the form of a story.

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

Name	Author	User Rating	Reviews	Pric
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

Name	Author	User Rating	Reviews	Pric
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 Bestselling Books 2009 - 2019.`

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 table named books and the column names corresponding to the dataset columns.

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

> *Consider a dataset relating to the Amazon Top 50 Bestselling Books 2009 - 2019. 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.*
>
> *Suppose you import this dataset into Excel in a table named books and the column names corresponding to the dataset columns.*

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 table named books.`
- `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.

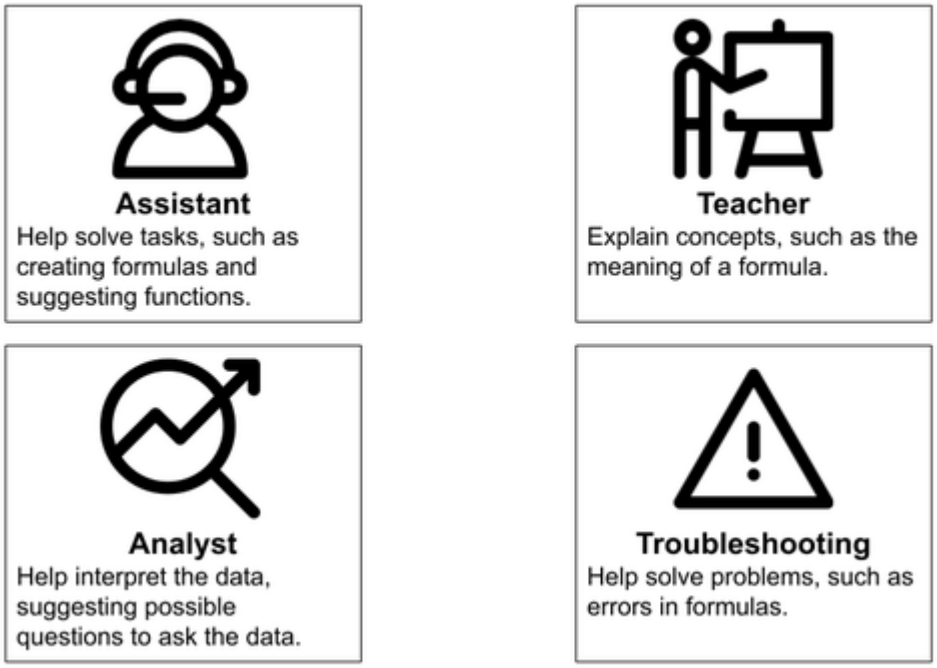


Figure 2-4. The roles ChatGPT can play in helping you answer your questions.

- 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
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.
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

Role	Action	Objective	Additional Details
Analyst	Request analysis or visualization suggestions. Verbs: Analyze, Suggest, Identify, Visualize	Extract meaningful trends, patterns, or insights from your data.	Specific context, such as columns.
Troubleshooting	Request for assistance to resolve a specific problem. Verbs: Solve, Correct, Identify,	Request for assistance to resolve a specific problem.	Provide details such as the formula or context of the problem.

Role	Action	Objective	Additional Details
	Help, Explain		

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 Excel.
```

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 Bestselling Books 2009 - 2019. 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.`*
>
> *Suppose you import this dataset into Excel in a table named books and the column names corresponding to the dataset columns.*

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. Use the Price column from the books table.
```

ChatGPT should respond with the following output:

To calculate the highest price of books in the Price column from the books table, use the following formula in Excel:

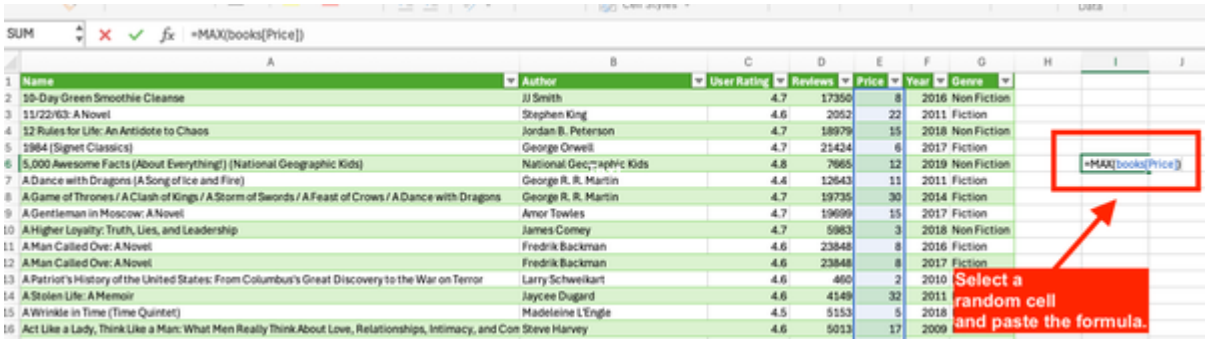
=MAX(books[Price])

Explanation:

- *books is the name of the table.*
- *[Price] refers to the column named* **Price** *within the books table.*
- *MAX is the function that returns the highest value in the specified range.*

Place this formula in any cell, and it will display the maximum price of the books in the table.

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	15	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	2019	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	Amor 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	32	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 Con	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 total sum of reviews for each
year and sort the results in ascending or
descending order. Use the Reviews column as
the values and the Year column as rows. Use a
base level of complexity.
```

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 in Excel to analyze the total sum of reviews for each year, using the Reviews column as values and the Year column as rows:

Step 1: Select Your Data

1. *Click anywhere inside the books table to ensure all columns (Year, Reviews, etc.) are recognized.*
2. *Go to the* **Insert** *tab on the Ribbon and click* **PivotTable**.

Step 2: Set Up the Pivot Table

1. *In the Create PivotTable dialog:*
   a. *Ensure the correct table or range (books) is selected.*
   b. *Choose where to place the Pivot Table:*
   c. *Select New Worksheet (recommended) or Existing Worksheet.*
   d. *Click* **OK**.

Step 3: Add Fields to the Pivot Table

1. *In the* **PivotTable Fields** *there are:*

a. *Drag the Year field to the* **Rows** *area. This will list all the years in the rows of the Pivot Table.*
b. *Drag the Reviews field to the* **Values** *area. By default, Excel will sum the Reviews.*

Step 4: Sort the Results

1. *Click the dropdown arrow in the* **Row Labels** *section of the Pivot Table.*
2. *Choose* **Sort by...** *and select either* **Smallest to Largest** *or* **Largest to Smallest** *to sort the total reviews by year in ascending or descending order.*

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** *(with thousands separators) for clarity.*

- `Review the table to see the total reviews for each year sorted as specified.`

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 and suggest the correct version to calculate the average price of books with a user rating higher than 4.5:
=AVERAGEIF(books[User Rating]>4.5, books[Price]))
```

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.

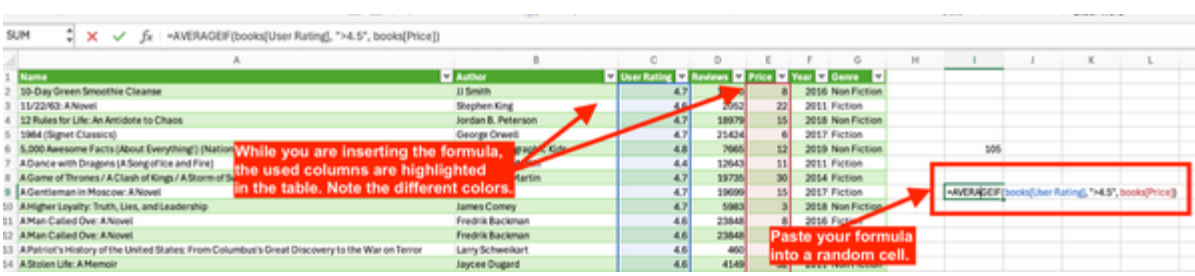


Figure 2-7. The result of Task 3.

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. Use the table books.
```

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`

3. `Calculate the average price of books in the "Fiction" genre.`
4. `Identify the highest price among books in the "Non-Fiction" genre.`
5. `Calculate the minimum price of books in the "Non-Fiction" genre.`

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 review > 4. Use the table books.
```

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 the book "10-Day Green Smoothie Cleanse". Use the table books.
```

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 in the first column of the books table.`
- `books: The name of the table.`
- `7: The column index for Genre in the books table (assuming Genre is the 7th column).`
- `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 in the Author column. Use the table books.
```

ChatGPT generates the following output:

```
=LEFT(books[@Author], FIND(" ",
books[@Author]) - 1)
```

```
Explanation:
```

- `books[@Author]: Refers to the current row's value in the Author column of the books table.`
- `FIND(" ", books[@Author]): Finds the position of the first space in the author's name.`
- `LEFT(books[@Author], FIND(" ", books[@Author]) - 1): Extracts all characters from the start of the name up to (but not including) the first space, giving the first name.`

Exercise

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

1. Remove extra spaces from a title
2. Extract the last name of the author
3. Transform the title to lowercase

4. Combine the book title and author
5. 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 from today's date. Use the table books.
```

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.

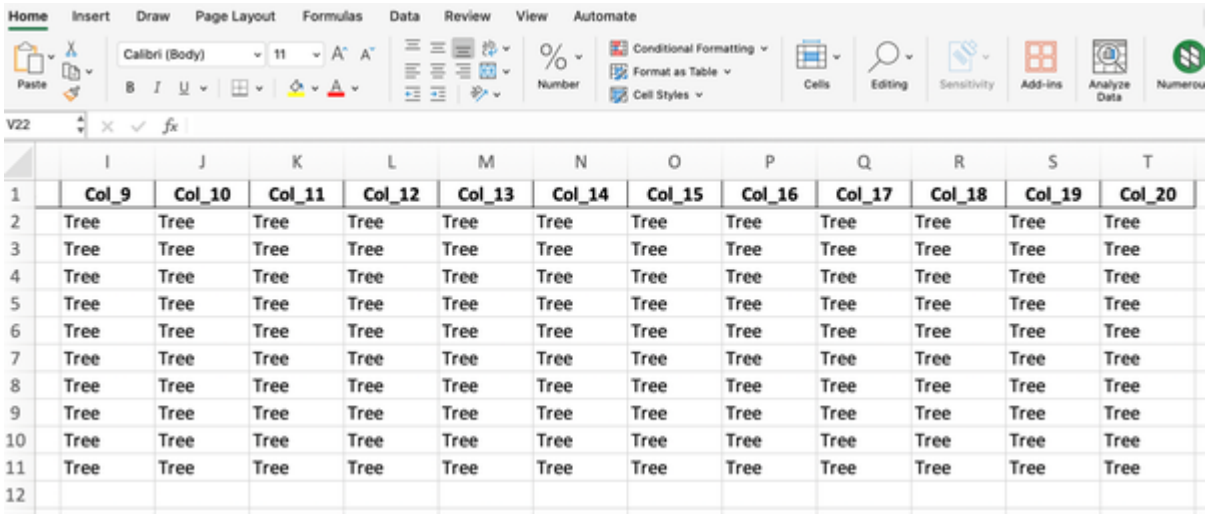
To immortalize the moment, I took the photo shown in Figure 2-8.

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:

```
Consider the attached picture containing trees
organized in rows and columns. Build a table
with as many rows and columns as the number of
tree rows and columns in the picture. Format
the resulting table as an Excel Table
downloadable as an XLSX file.
```

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.



	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.

ASSISTANT - CALCULATION - CHATGPT - CONTEXT - END - MAIN POINT - PROMPT - VLOOKUP

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.