

IN THIS CHAPTER

- » Figuring out the different types of data Power BI can handle
- » Understanding your options for business intelligence tooling
- » Familiarizing yourself with Power BI terminology

Chapter **1**

A Crash Course in Data Analytics Terms: Power BI Style

Data is everywhere — literally. From the moment you awaken until the time you sleep, some system somewhere collects data on your behalf. Even as you sleep, data is being generated that correlates to some aspect of your life. What is done with this data is often the proverbial million-dollar question. Does the data make sense? Does it have any structure? Is the dataset so voluminous that finding what you're looking for is like finding a needle in a haystack? Or is it more like you can't even find what you need unless you have a special tool to help you navigate?

I answer that last question with an emphatic yes, and that's where data analytics and business intelligence join the party. And let's be honest: The party can be overwhelming if data is consistently generating something on your behalf.



Dealing with data isn't always a chore — data can be fun to explore as well. Sometimes it's easy to figure out precisely what is needed to solve a problem, but at other times you need to put on your Sherlock Holmes deerstalker cap. Why? Because the data you're working with may lack structure and meaning. Of course, you're bound to take up tools to help you play the role of detective, evaluator, designer, and curator.

In this chapter, I discuss the different types of data you may encounter along your journey. I review the key terminology that you should become familiar with upfront. Don't worry: It's not like you need to memorize a dictionary. You learn a few key concepts to give you a head start in Power BI and business intelligence. Are you ready to go?

What Is Data, Really?

Ask a hundred people in a room what the definition of data is and you may receive one hundred different answers. Why is that? Because, in the world of business, data means a lot of different things to a lot of different people. So, here's try to get a streamlined response. Data contains facts. Sometimes, the facts make sense; sometimes, they're meaningless unless you add a bit of context.

The facts can sometimes be quantities, characters, symbols, or a combination of sorts that come together when collecting information. The information allows people — and more importantly, businesses — to make sense of the facts that, unless brought together, make absolutely no sense whatsoever.

When you have an information system full of business data, you also must have a set of unique data identifiers you can use so that, when searched, it's easy to make sense of the data in the form of a transaction. Examples of transactions may include the number of jobs completed, inquiries processed, income received, and expenses incurred.

The list can go on and on. To gain insight into business interactions and conduct analyses, your information system must have relevant and timely data that is of the highest quality.



Data isn't the same as information. *Data* is the raw facts. That means you should think of data in terms of the individual fields or columns of data you may find in a relational database or perhaps the loose document (tagged with some descriptors called *metadata*) stored in a document repository. On their own, these items are unlikely to make much sense to you or to a business. And that's perfectly okay — sometimes. *Information* is the collective body of all those data parts, that results in the factoids making logical sense.

Working with structured data

Have you ever opened a database or spreadsheet and noticed that data is bound to specific columns or rows? For example, would you ever find a United States zip code containing letters of the alphabet? Or, perhaps when you think of a first name, middle initial, and last name, you notice that you always find letters in those specific fields. Another example is when you're limited to the number of characters you can input into a field. Think of Y as Yes; N is for No. Anything else is irrelevant.

What I'm describing here is called *structured data*. When you evaluate structured data, you notice that it conforms to a tabular format, meaning that each column and row must maintain an interrelationship. Because each column has a representative name that adheres to a predefined data model, your ability to analyze the data should be straightforward. This embodies your classic relational data system.

If you're using Power BI, you notice that structured data conform to a formal specification of tables with rows and columns, commonly referred to as a *data schema*. In Figure 1-1, you find an example of structured data as it appears in a Microsoft Excel spreadsheet.

FIGURE 1-1:
An example of
structured data.

	A	B	C	D	E	F	G	H
1	Employee ID	First Name	Last Name	Birth Date	Email Address	Mobile Number	Department	Office Location
2	123-45-453	Joe	Smith	1/3/2000	joe.smith@dataco.com	555.421.9051	Data Management	Seattle
3	123-45-459	Bob	Jones	2/14/1974	bob.jones@dataco.com	555.429.9082	Data Management	Seattle
4	123-49-907	Jane	Richards	3/15/1978	jane.richards@dataco.com	555.904.2852	Data Management	Seattle
5	190-90-223	Sally	Frank	2/28/1967	sally.frank@dataco.com	555.229.1804	Accounting	Atlanta
6	229-29-004	Emma	Donaldson	10/21/2002	emma.donaldson@dataco.com	555.867.5309	Marketing	San Francisco



REMEMBER

Whether you're using Power BI for personal analysis, educational purposes, or business support, the most accessible data sources for BI tools are structured. Platforms that offer robust structured data options include Microsoft SQL Server, Microsoft Azure SQL Server, Microsoft Access, Azure Table Storage, Oracle, IBM DB2, MySQL, PostgreSQL, Microsoft Excel, and Google Sheets.

Looking at unstructured data

Unstructured data is ambiguous, having no rhyme, reason, or consistency whatsoever. Pretend that you're looking at a batch of photos or videos. Are there explicit data points that one can associate with a video or photo? Perhaps, because the file itself may consist of a structure and be made of some metadata. However, the byproduct itself — the represented depiction — is unique. The data isn't replicable; therefore, it's unstructured. That's why video, audio, photos, text files, and social media posts are considered unstructured data.

Adding semi-structured data to the mix

Semi-structured data does have some formality, but it isn't stored in a relational system, and it has no set format. Fields containing the data are by no means neatly organized into strategically placed tables, rows, or columns. Instead, semi-structured data contains tags that make the data easier to organize and manage in some form of hierarchy. Nonrelational data systems or NoSQL databases are best associated with semi-structured data, in which the programmatic code, often serialized, is driven by the technical requirements. There is no hard-and-fast coding practice.

Common examples of semi-structured data include JSON files from web APIs, XML documents, and data from cloud applications. Power BI can work with these formats, making it possible to analyze data from modern web services and applications alongside your traditional databases and spreadsheets. This flexibility makes Power BI particularly valuable in today's business environment, where data comes from many different sources and formats.

Looking Under the Power BI Hood

The Power BI platform combines cloud-based apps and services to help you organize, collect, manage, and analyze your data. Whether you're working with a few hundred rows or millions of records, Power BI can handle datasets of various sizes and reveal patterns and trends that may not be obvious from the raw data. Unlike a tool such as Microsoft Excel, Power BI can simultaneously connect to and evaluate multiple data sources. The sources don't need to be structured like a spreadsheet, either. They can include unstructured and semi-structured data. After connecting to these data sources and processing them, Power BI helps you create visually compelling outputs such as charts, graphics, reports, dashboards, and KPIs (key performance indicators).

As you see throughout this book, Power BI isn't just a single application. It has desktop, online, and mobile components.



REMEMBER

Across the Power BI platform, you are certain at some point to encounter one (or more) of the following products:

- » **Power Query:** A data connection tool you can use to transform, combine, and enhance data across several data sources
- » **Power Pivot:** A data modeling tool
- » **Power BI Visualizations:** Built-in tools for creating interactive charts, graphs, maps, and other visuals

- » **Power Q&A:** An AI-powered engine that allows you to ask questions about your data and receive responses using plain language.
- » **Copilot in Power BI:** An AI assistant that helps create reports, write DAX formulas, and answer questions about your data
- » **Power BI Desktop:** A free application that brings together data connection, modeling, and visualization capabilities in a single interface
- » **Power BI Service:** A cloud-based platform for collaboration, sharing, and distributing reports and dashboards with others

In the following sections, I explore each product's core functionality.

Posing questions with Power Query

Power Query is the capability that allows Power BI to create data connections and transform data among many different data sources, from databases like SQL Server and Oracle to cloud services, spreadsheets, and web APIs. Power Query also serves as your trusty helper to clean and shape (or *transform*) your raw data before analysis, handling tasks like removing empty rows, changing data types, and combining information from multiple sources. The transformed data can then be used in your Power BI reports and dashboards, enabling you to work more easily with information that may otherwise be messy or difficult to analyze. For the data analyst and engineer, these tools are invaluable.

Power Pivot is the Power BI data modeling tool that helps you create relationships between data sources, craft models of varying complexities, and build calculations for your analysis. With Power Pivot, you can connect tables from various sources and define how they relate, which enables you to analyze data across multiple datasets. For example, a sales manager can link customer records from a CRM system with revenue data from an ERP system to quickly see which customer segments are driving the most profit. Power Pivot uses a formula language called DAX (Data Analysis Expressions) to create custom calculations, but don't worry — Power BI includes many built-in calculations, so you can get started without writing formulas. I cover DAX in Chapters 15 and 16.

Mastering Power BI Visualizations

Power BI visualizations help you see your data in a way that makes sense. For example, instead of scanning a long list of numbers, a bar chart can quickly show which department spends the most on travel. A map can display where most service requests come from so that you can spot regional patterns right away.

The key is that these visuals are interactive. You can click, filter, and drill into the details to answer specific questions. Later in this book, I walk you through the complete set of chart types available in Power BI and talk about when each one works best.

Interpreting data with Power Q&A

One of the biggest challenges for many users is data interpretation. Say, for example, that you've built this incredible data model using Power Pivot. Your data sample is significant in size, meaning that you need some way to make sense of all the data you've deployed in the model. But is the model useful? Possibly not, unless you have a way to interpret the data. That's why Microsoft created Power Q&A, an AI-powered feature that allows you to ask questions about your data using plain English and get instant visual answers.

Power Q&A comes with AI assistance to help you ask better questions by suggesting queries based on your data. For example, you may type, "How many customers purchased red shirts in our Chicago store last month?" From there, Power Q&A can suggest follow-up questions such as "Which day of the week had the most sales?" or "How does this compare to last year?" Of course, this only works if your dataset includes those details. Power Q&A can't generate insights from data you don't have. When your data model is well-designed, these AI-driven prompts can guide you toward insights you may not think to ask on your own.

Guiding the way with Copilot

At first glance, Q&A and Copilot may seem to be the same thing, seeing as though both are AI assistants. Here's how they differ: Copilot in Power BI is an AI assistant that helps you learn and use Power BI itself, whereas Power Q&A allows you to explore and analyze your data. Although Power Q&A helps you discover what's in your data by asking questions like "How many red shirts sold last month," Copilot enables you to figure out how to work with Power BI as a tool. For example, if you need to create a calculation, Copilot can write DAX formulas when you describe what you want to calculate. It can also suggest the best visualization types for your data and even help create entire reports based on your requirements. Whether you're wondering "How do I calculate year-over-year growth" or "What's the best chart to show regional sales trends," Copilot provides step-by-step guidance to help you accomplish your goals. This makes it particularly valuable for beginners who know what they want to achieve but aren't sure how to use Power BI's features to get there.

Power BI Desktop

Power BI Desktop is a free application that combines Power Query, Power Pivot, Power BI Visualizations, and Power Q&A. Using Power BI Desktop, you can

complete all your business intelligence activities under one umbrella — from connecting to data sources and building models to creating reports and dashboards. Power BI Desktop is where you do all your development work on your local computer. Microsoft updates Power BI Desktop monthly with new features and improvements, so you can always access the latest capabilities.

Power BI Service

Accessible at <https://app.powerbi.com> from any device with an Internet connection, Power BI Service allows you to publish the work you create in Power BI Desktop and share it with colleagues across your company.

Power BI Service is the cloud platform where your reports and dashboards come to life for your organization. You can access it at <https://app.powerbi.com> from any device with an Internet connection, but you'll also need an organizational or educational email account to sign in. This means personal email addresses (like Gmail, Yahoo, or Outlook.com) won't work for publishing and sharing reports in Power BI Service.

Once you're set up, you can publish the reports you've built in Power BI Desktop, share them securely with colleagues, and even schedule automatic refreshes so your data stays up to date. Administrators can manage permissions to ensure that the right people see the right information, keeping your data protected.



REMEMBER

Your Power BI license determines which sharing, refresh, and collaboration features you can access.

Knowing Your Power BI Terminology

Whether Microsoft or another vendor creates it, every product you come across has its own terminology. It may seem like a foreign language, but if you visit a vendor's website and do a simple search, you're sure to find a glossary that spells out what all their mysterious terms mean.

Microsoft, unsurprisingly, has its own glossary for Power BI as well. In Microsoft Power BI-speak, some concepts resonate across vendors no matter who you are. For example, all vendors have reports and dashboards as critical concepts. Of course, not all vendors adopt Microsoft's practice and call dataflows a type of workflow. But even though they have their own names for these features, the features generally work the same way.



TIP

Microsoft has done a pretty good job of trying to stick with mainstream names for critical concepts. Nevertheless, some of the more advanced product features specific to AI or machine learning and security adopt the specialized terminology of Microsoft products such as Microsoft Entra ID or Azure Machine Learning.

Capacities

The first concept you must be familiar with is *capacities*, which are central to Power BI. Capacities are the sum total of system resources needed for you to complete a project you create in Power BI. Resources include the storage, processor, and memory required to host and deliver Power BI projects. Both the type and quantity of your data can affect the capacity required to deliver your projects.

When you use Power BI, you must consider two types of capacity: shared and dedicated. A *shared* capacity allows you to share resources with other Microsoft end users. *Dedicated* capacities fully commit resources to you alone. Shared capacity is available for both free and paying Power BI users, and dedicated capacity requires a Power BI Premium subscription.

Workspaces

Workspaces are a means of collaborating and sharing content with colleagues. Whether it's personal or intended for collaboration, any workspace you create is created on capacities. Think of a workspace as a container that allows you to manage the entire lifecycle of dashboards, reports, workbooks, datasets, and dataflows in the Power BI Service environment. (Figure 1-2 shows the My Workspace, which is a personal Power BI workspace.)



REMEMBER

My Workspace isn't the only type of workspace available. You also have the option to collaborate. If you want to collaborate, you have no choice but to upgrade to a Power BI Pro, Premium Per User, or Premium Per Capacity plan. Features that come with collaboration include the ability to create and publish Power BI-based dashboards, reports, workbooks, datasets, and apps with a team.

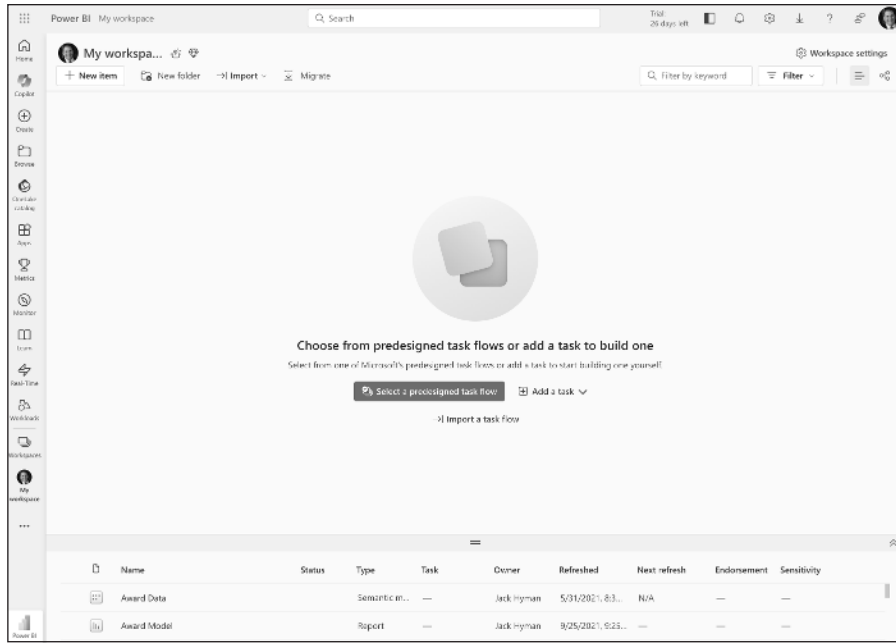


REMEMBER

To upload the work you've created using Power BI Desktop, or to manipulate the work online without collaborating with anyone, you can use My Workspace — this is all that is necessary. You only *require* the use of Power BI Free. As soon as you want to collaborate with others, you need to upgrade to a Power BI Pro, Premium Per User, or Premium Per Capacity subscription.

So now you know that your work is stored in a workspace. Next question: What happens with the data in that workspace? When you work with Power BI data, you primarily deal with two key concepts: datasets and reports. A dataset contains your data and any relationships or calculations you've defined. Think of it as the

FIGURE 1-2:
My Workspace in
Power BI Service.



foundation that holds all your information. Reports are then built on top of data-sets to create the visualizations you see. The Power BI Service Workspace is where both assets are managed behind the scenes, so you can focus on analyzing and presenting your information.

Dataflows and datasets

Power BI works with two main types of data containers: datasets and dataflows. A *dataset* is like a digital filing cabinet that holds your data, relationships, and calculations — it's what you'll work with most often when creating reports. A *dataflow* is a more advanced feature that helps prepare and clean large amounts of data before it becomes a dataset. Think of dataflows as a preprocessing step for complex data scenarios.



TIP

Most users starting with Power BI primarily work with datasets. Dataflows are helpful when dealing with large, complex data sources that need significant cleaning and transformation before analysis, as you see when you dig into Part 3 of this book.



REMEMBER

After you manipulate the data on your own, you must publish the data you created in Power BI to Power Bi Service. Otherwise, it remains on your desktop and only you can access the file. Microsoft assumes that you intend to share the data among users. If you want to share a dataset, you must have a Power BI Pro, Premium Per User, or Premium Per Capacity license.

Reports

Data can be stored in a system indefinitely and remain idle, but what good is it if no one ever looks at it to understand what it means? That’s where Power BI reports come in. A report transforms your raw data into visual stories that make sense at a glance.

For example, imagine you’re a sales manager who needs to understand your team’s performance. You can create a Power BI report that shows monthly sales trends in a line chart, top-performing products in a bar chart, and regional performance on a map — all on different pages of the same report. Whether you’re working with hundreds of sales records or thousands, Power BI helps you see patterns and insights that would be impossible to spot in spreadsheet rows.

Power BI reports translate your data into one or more pages of visualizations — line charts, bar charts, pie charts, treemaps, and dozens of other options. You can view your data at a high level to spot big trends, or drill down into specific details when something catches your attention. Creating a report is straightforward: connect to your data source (like an Excel file, database, or cloud service), choose the information you want to display, and select the best visualization to tell that part of your data’s story. The result is a report (see Figure 1-3) that transforms rows and columns of data into meaningful insights you can use.

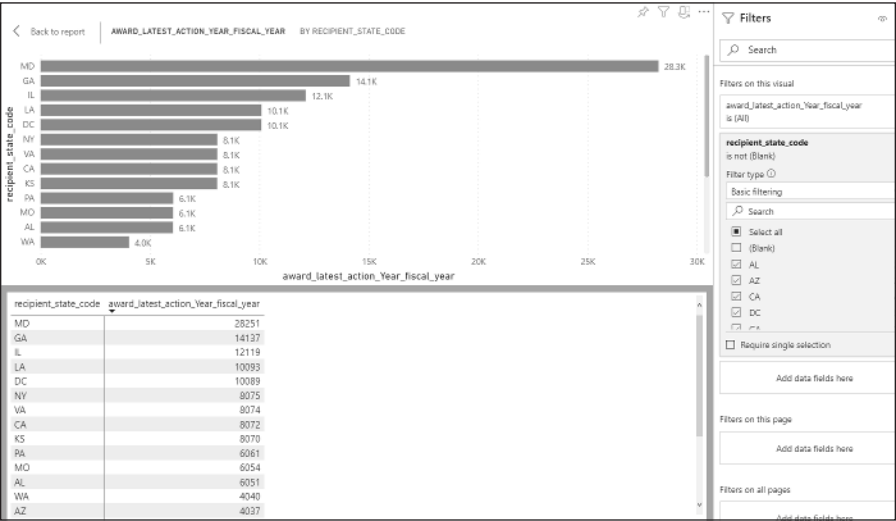


FIGURE 1-3:
A sample
Power BI report.



TIP

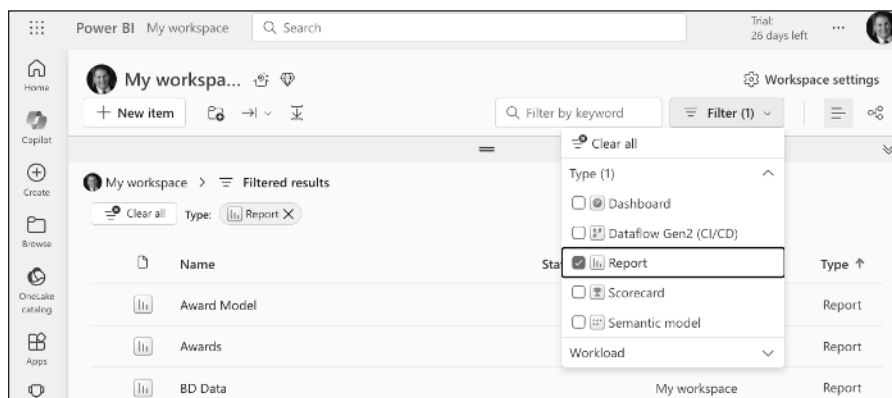
Power BI offers two Report view modes: Reading view and Editing view. When you open a report, it opens in Reading view. You can switch to Editing view to modify a report if granted Edit permissions. Your ability to edit depends on your role in the workspace and your Power BI license.



TECHNICAL
STUFF

Users with administrative, member, or contributor roles in a workspace can create and edit reports. Other users can view and interact with these reports but cannot make changes to them. Reports in a workspace can be found under the Reports tab, as shown in Figure 1-4. Reports can contain multiple pages and are built using data from one or more datasets.

FIGURE 1-4:
The Reports
tab in Power
BI Desktop.



Dashboards

If you've had any experience with Power BI, you already know that it's a highly visual tool. In line with its visual nature, the Power BI dashboard brings your data story to life. If you want to take all the pieces of your data puzzle and capture a moment, you use the dashboard. Think of it as a blank canvas. As you build your reports, widgets, tiles, and KPIs over time, you pin the ones you like to the dashboard to create a single visualization. The dashboard represents the large dataset that you feel covers your topic. As such, it can help you make decisions, support you in monitoring data, or allow you to drill down in your dataset by applying different visualization options.

To access a particular dashboard, you must first open a workspace. Then all you need to do is click the Dashboards tab. Remember that every dashboard represents a customized view of an underlying dataset. To locate your dashboards, go to your My Workspace (see Figure 1-5) and then choose Dashboards to see what's available.

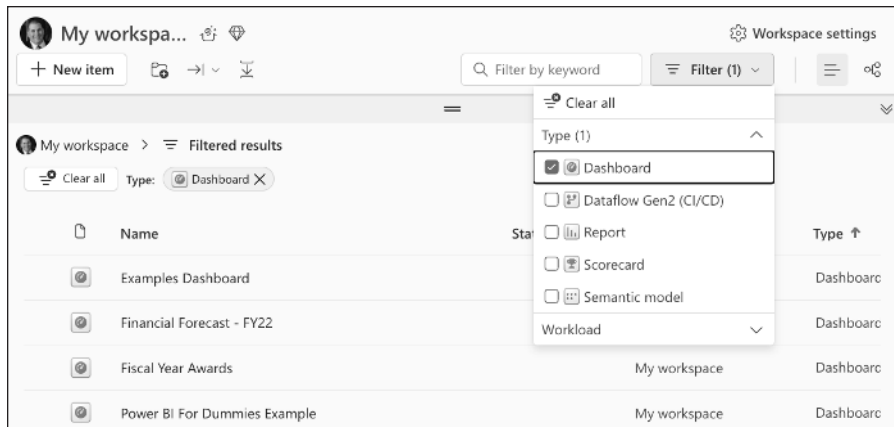


FIGURE 1-5:
Locating your
dashboards.



REMEMBER

If you own a dashboard, you have permission to edit it. Otherwise, you only have read-only access. You can share a dashboard with others, but they may not be able to save any changes. Keep in mind, however, that if you want to share a dashboard with a colleague, you need, at minimum, a Power BI Pro license. (For more on the ins and outs of licensing, see Chapter 3.)

Navigation pane

In this chapter, I talk about many of the must-know concepts in Power BI, but I want to cover one more important element — the Navigation pane. The Navigation pane is your road map to everything in Power BI Service. All the capabilities I discuss in this chapter — workspaces, dashboards, reports, datasets — are accessed through the Navigation pane (see Figure 1-6). You'll use the Navigation pane to move between workspaces and find the Power BI content you want to work with.

Your Navigation pane options are endless. For example, a user such as yourself can

- » Expand and collapse the Navigation pane
- » Open and manage your favorite content with the help of the Favorites option
- » View and open the most recently visited section of content



FIGURE 1-6:
The Navigation
pane.

Although Power BI Desktop has its own interface for building reports locally, the Navigation pane explicitly refers to the online Power BI Service experience, where you'll collaborate, share, and manage your published content.

