



Data Visualization

Course Code: B24DS01SE
BSc Data Science and Analytics
Skill Enhancement Course
Self Learning Material



SREENARAYANAGURU
OPEN UNIVERSITY

SREENARAYANAGURU OPEN UNIVERSITY

The State University for Education, Training and Research in Blended Format, Kerala

SREENARAYANAGURU OPEN UNIVERSITY

Vision

To increase access of potential learners of all categories to higher education, research and training, and ensure equity through delivery of high quality processes and outcomes fostering inclusive educational empowerment for social advancement.

Mission

To be benchmarked as a model for conservation and dissemination of knowledge and skill on blended and virtual mode in education, training and research for normal, continuing, and adult learners.

Pathway

Access and Quality define Equity.

Data Visualization
Course Code: B24DS01SE
Semester- III

Skill Enhancement Course
Bachelor of Science
(Data Science And Analytics)
Self Learning Material
(with Model Question Paper Sets)



SREENARAYANAGURU
OPEN UNIVERSITY

SREENARAYANAGURU OPEN UNIVERSITY

The State University for Education, Training and Research in Blended Format, Kerala



Data Visualization

Course Code: B24BS01SE
Semester- III
Skill Enhancement Course
Bachelor of Science
(Data Science and Analytics)

Academic Committee

Dr. T. K. Manoj
Dr. Smitha Dharan
Dr. Satheesh S.
Dr. Vinod Chandra S.S.
Dr. Hari V. S.
Dr. Sharon Susan Jacob
Dr. Ajith Kumar R.
Dr. Smiju I.S.
Dr. Nimitha Aboobaker

Development of the Content

Deepasree Varma

Review and Edit

Dr. Sabeena K.

Linguistics

Dr. Sabeena K.

Scrutiny

Shamin S. Greeshma P.P.
Sreerekha V.K. Anjitha A.V.
Aswathy V.S. Dr. Kanitha Divakar
Subi Priya Laxmi S.B.N.

Design Control

Azeem Babu T.A.

Cover Design

Jobin J.

Co-ordination

Director, MDDC :

Dr. I.G. Shibi

Asst. Director, MDDC :

Dr. Sajeevkumar G.

Coordinator, Development:

Dr. Anfal M.

Coordinator, Distribution:

Dr. Sanitha K.K.



Scan this QR Code for reading the SLM
on a digital device.

First Edition
December 2025

Copyright
© Sreenarayanaguru Open University

All rights reserved. No part of this work may be reproduced in any form, by mimeograph or any other means, without permission in writing from Sreenarayanaguru Open University. Printed and published on behalf of Sreenarayanaguru Open University by Registrar, SGOU, Kollam.

www.sgou.ac.in



Visit and Subscribe our Social Media Platforms

MESSAGE FROM VICE CHANCELLOR

Dear learner,

I extend my heartfelt greetings and profound enthusiasm as I warmly welcome you to Sreenarayanaguru Open University. Established in September 2020 as a state-led endeavour to promote higher education through open and distance learning modes, our institution was shaped by the guiding principle that access and quality are the cornerstones of equity. We have

firmly resolved to uphold the highest standards of education, setting the benchmark and charting the course.

The courses offered by the Sreenarayanaguru Open University aim to strike a quality balance, ensuring students are equipped for both personal growth and professional excellence. The University embraces the widely acclaimed “blended format,” a practical framework that harmoniously integrates Self-Learning Materials, Classroom Counseling, and Virtual modes, fostering a dynamic and enriching experience for both learners and instructors.

The University is committed to providing an engaging and dynamic educational environment that encourages active learning. The Study and Learning Material (SLM) is specifically designed to offer you a comprehensive and integrated learning experience, fostering a strong interest in exploring advancements in information technology (IT). The curriculum has been carefully structured to ensure a logical progression of topics, allowing you to develop a clear understanding of the evolution of the discipline. It is thoughtfully curated to equip you with the knowledge and skills to navigate current trends in IT, while fostering critical thinking and analytical capabilities. The Self-Learning Material has been meticulously crafted, incorporating relevant examples to facilitate better comprehension.

Rest assured, the university’s student support services will be at your disposal throughout your academic journey, readily available to address any concerns or grievances you may encounter. We encourage you to reach out to us freely regarding any matter about your academic programme. It is our sincere wish that you achieve the utmost success.



Regards,
Dr. Jagathy Raj V. P.

01-12-2025

Contents

BLOCK-1	Introduction to Visualization	1
UNIT 1	Fundamentals of Data Visualization	2
UNIT 2	Visual Perception and Cognitive Principles	11
UNIT 3	Designing for Different Audiences - Understanding Your Audience	25
UNIT 4	Visualization Best Practices and Trends	35
BLOCK-2	Data Visualisation Using Power BI	45
UNIT 1	Data Preparation and Integration in Power BI	46
UNIT 2	Data Modelling and Dax Fundamentals	58
UNIT 3	Creating and Publishing Visual Reports	79
UNIT 4	Power BI Deployment and Security	103
	Model Questionpaper Sets	120



BLOCK-1

Introduction to Visualization

UNIT 1

Fundamentals of Data Visualization

Learning Outcomes

At the end of this unit, the learner will be able to :

- ◆ explain data visualization and importance in communication and analysis
- ◆ trace the historical evolution and identify diverse real-world applications of data visualization
- ◆ apply the core principles of visual design (clarity, precision, and efficiency) to evaluate visualizations
- ◆ explain the role of aesthetics in enhancing the effectiveness of data communication

Prerequisite

We find ourselves in a world generating vast, unmanageable seas of data every second from business transactions and social media interactions to healthcare records and sensor readings. Yet, this raw data, stored in endless tables or spreadsheets, is like a language we cannot understand. It fails to convey meaningful insights to decision-makers.

Imagine a company manager staring at a spreadsheet of monthly sales. They are lost in a dense jungle of numbers, unable to spot the path. The story of their business where they are succeeding, where they are failing is completely hidden.

This is the central problem we face. Now, imagine a well-crafted visualization, such as a bar chart or heat map. This tool acts as a translator, cutting through the noise. It instantly highlights growth areas, performance gaps, or emerging patterns that drive strategic decisions.

This journey from “noise” to “story” is the entire purpose of data visualization. Throughout history, humans have used visuals to simplify complexity, and the

growing importance of this field lies in its power to forge complex data into intuitive, actionable narratives.

To begin this journey, learners must first understand basic data concepts and appreciate the role of clarity and precision. This foundational knowledge prepares us to explore the core principles of visual design, laying the groundwork for mastering this transformative skill.

Keywords

Data Visualization, Visual Design Principles, Information Clarity, Data Communication, Aesthetics in Visualization, Visual Storytelling, Data Representation, Graphical Excellence, Visual Efficiency, Insight Discovery, Data Interpretation, Visual Analytics, Chart Design, Visual Context, Data-Driven Decision Making

Discussion

1.1.1 Fundamentals of Data Visualization

As we've established, the modern digital era presents a paradox: we are data-rich but insight-poor. Raw data, in its tabular or textual form, is a silent monolith. It holds valuable secrets—about market trends, scientific breakthroughs, or public health—but it lacks the ability to reveal them. To make it speak, we must give it a voice.

This transformation, the process of turning raw data into a graphical or pictorial form, is the fundamental goal of data visualization. It is the bridge between a complex dataset and the human mind, enabling quick understanding, effective communication, and better decision-making by illuminating the key trends, correlations, and variations that lie dormant within the data.

A well-designed visualization is a narrative in itself. It guides the viewer's eye to the most important message, presenting information clearly and precisely. This is where the core principles of visual design—clarity, precision, and efficiency—come into play. They are the grammar of this new language, ensuring that the story we tell is not only compelling but also true.

But this “language” wasn't born overnight. Its roots run deep, evolving over centuries from early maps and diagrams to the sophisticated tools we use today. This evolution, driven by innovators in navigation, economics, and public health, is a story of human ingenuity. Today, this honed tool is applied everywhere, from business intelligence dashboards to groundbreaking scientific research, integrating analytical rigor with design aesthetics to deliver impactful insights.



Thus, the fundamentals of data visualization lie in understanding how to become both an analyst and a storyteller. It is a craft that combines analytical thinking with creative design, enabling us to perceive data not as mere numbers, but as compelling visual narratives that drive understanding and inspire action.

1.1.2 Definition and Importance of Data Visualization

Let us, then, formalize our tool. At its heart, **data visualization refers to the process of representing data and information in a graphical or pictorial format**, such as charts, graphs, maps, or dashboards. This is its technical definition.

But this definition only hints at its profound importance. Its *true* value is as a **catalyst for understanding**. Why is this so critical? Because the human brain is wired for visual processing. We can grasp patterns and relationships in a single, well-composed image far more quickly than we can by reading a corresponding table of numbers.

In an environment of massive data-overload, this ability to “grasp essential information at a glance” is a superpower. A well-designed visualization acts as a filter, cutting through the static to highlight key findings. It is the key to unlocking data-driven thinking, where decisions are backed by clear, evidence-based insights rather than assumptions.

Moreover, visualization is the engine of communication. It is indispensable for sharing findings with stakeholders who may not possess technical expertise. A great visualization does more than inform; it *engages*. It transforms a sterile set of numbers into a compelling story that can be shared, remembered, and acted upon, inspiring innovation.

1.1.3 Historical Evolution of Data Visualization

The quest to “see” data is as old as human curiosity itself. The story of data visualization is a fascinating journey, reflecting our continuous effort to make sense of the world. Long before computers, this journey began with the most basic of human needs: to record and interpret information about geography, trade, and astronomy.

The plot thickened during the Renaissance, as science and exploration demanded more systematic visual forms. But the true “heroes” of our story emerge in the late 18th and early 19th centuries. We meet **William Playfair**, often regarded as the father of statistical graphics. Frustrated with dense economic tables, he *invented* the bar chart, line graph, and pie chart, giving economists and policymakers a lens they never had before.

Then, our story takes a dramatic turn with **Florence Nightingale**. She wasn’t just a nurse; she was a passionate statistician and a data-driven reformer. She used innovative “polar area diagrams” to illustrate the causes of mortality during the Crimean War. Her charts were not just academic; they were weapons of persuasion, proving to Parliament that more soldiers were dying from poor sanitation than from battle wounds, leading directly to public health reforms.

At the same time, **Charles Minard** told the tragic, multi-dimensional story of Napoleon’s 1812 Russian campaign in a single, devastating graphic. His famous flow map is a masterpiece, weaving together geography, temperature, and the catastrophic loss of an army into one coherent narrative.

These pioneers walked so that modern technology could run. The 20th century's computing-powered graphics and the 21st's interactive dashboards are all built on the foundations they laid. Today, as we stand on their shoulders, visualization continues to evolve, becoming an even more essential tool for navigating our complex world.

1.1.4 Real-World Applications of Data Visualization

Having traced the origins of our “tool,” we now find it at work, shaping the modern world in countless ways. Its applications are a testament to its power, turning abstract data into a tangible force across every major domain.

In the fast-paced **business and corporate sector**, executives are not analysts; they are captains steering a ship. Interactive dashboards tracking sales, revenue, and customer behavior become the “cockpit” of the company, providing real-time data to navigate market shifts and maintain a competitive edge.

In the critical **healthcare industry**, the stakes are human lives. Epidemiologists become detectives, tracking a disease's footprint on a geographic map. Hospitals use dashboards to monitor patient flow and bed occupancy, making life-saving decisions based on clear, visual data. The COVID-19 pandemic, more than any other event, proved that accessible visuals are vital for guiding public health policy.

Within **education and research**, visualization becomes the tool of enlightenment. Educators use concept maps to simplify complex topics, while scientists employ sophisticated visual models to represent experimental results, revealing the hidden architecture of their findings.

In the **government and public policy** domain, visualization becomes the language of transparency. Open data portals use visual reports on budgets and crime statistics to empower citizens, building trust and enabling policymakers to identify areas needing intervention.

In the **media and communication** industry, it is the engine of the visual narrative. Data journalism organizations use infographics and animated charts to translate complex societal issues into stories that are engaging, understandable, and shareable.

Across the **scientific and technological sectors**, it is the only way to tame the gargantuan datasets generated by complex simulations, artificial intelligence, and environmental monitoring. Finally, in the **social and cultural** domains, it helps us understand the most complex subject of all: ourselves, by visualizing the vast, invisible networks of human behavior.

In every case, the story is the same: data visualization is the essential bridge between unmanageable data and actionable understanding.

1.1.5 Principles of Visual Design: Clarity, Precision and Efficiency

But a powerful tool demands great skill. A visualization *can* lie. It can be confusing. It can fail. To master this craft and wield this tool responsibly, we must adhere to three



core principles that form the foundation of its integrity. They are the “commandments” that separate a true insight from a dangerous misinterpretation.

First, **Clarity**. A visualization must be a beacon, not a fog. Its primary purpose is to convey information in a straightforward and understandable manner. It must answer questions, not create them. Clarity is achieved through a ruthless elimination of “visual clutter”—any ink on the page that doesn’t represent data—and a focus on highlighting the key patterns without overwhelming the audience.

Second, **Precision**. This is the principle of truth. The visualization must *faithfully* and accurately represent the underlying data. A distorted scale, a “cherry-picked” dataset, or an exaggerated representation is a form of deception, whether intended or not. Precision is the ethical backbone of data visualization, ensuring that the story we tell is accurate.

Finally, **Efficiency**. This is the principle of respect for the viewer’s time. An efficient visualization conveys the maximum amount of information with the minimal amount of cognitive effort. It is the “at-a-glance” quality that allows a viewer to extract insights quickly. This is achieved by selecting the right chart for the data and prioritizing the most meaningful information.

Together, these three principles form the functional foundation of our craft. When applied thoughtfully, they ensure our visualizations are not just beautiful, but also accessible, accurate, and actionable. But one more element is needed to make them truly powerful.

1.1.6 Role of Aesthetics in Data Visualization

If clarity, precision, and efficiency are the *science* of data visualization, then **Aesthetics** is its *art*. It is the element that elevates a visualization from merely functional to truly **compelling**.

This is not about “decoration” or adding superfluous visual flair. Instead, aesthetics are functional in their own right, influencing both comprehension and engagement. A visually well-designed chart captures attention, facilitates understanding, and helps viewers retain information.

Thoughtful **color** selection is a prime example. It is not an adornment; it is a *tool for comprehension*. Color can be used to highlight critical data points, differentiate categories, and establish a clear visual hierarchy. Poorly chosen colors, on the other hand, can confuse or even mislead.

Similarly, **typography** (fonts, labels) and **layout** (spacing, alignment) are crucial aesthetic elements. Clear, legible text and a balanced, organized layout reduce cognitive load, allowing the viewer to focus on the data’s story, not on deciphering the chart itself.

Aesthetics also build **trust**. A well-designed, professional-looking visual conveys care, attention to detail, and authority. A poorly designed one can, fairly or not, cast doubt on the credibility of the data itself.

Ultimately, aesthetics are what make a visualization **memorable**. They engage the viewer on an intuitive and even emotional level. By carefully integrating color, typography, and layout, we transform data from abstract numbers into a meaningful visual narrative that not only informs but *inspires*.

Recap

- ◆ Data visualization presents information graphically to make complex data easier to understand and interpret quickly.
- ◆ It improves clarity, supports decision-making, and helps identify patterns and trends that are not visible in raw data.
- ◆ Early forms included maps, diagrams, and pictographs used for recording information.
- ◆ Major developments occurred in the 18th–19th centuries through contributors like William Playfair, Florence Nightingale, and Charles Minard.
- ◆ In the modern era, digital dashboards, interactive tools, and big data analytics have transformed visualization.
- ◆ It is widely used in business for intelligence, marketing, and performance monitoring.
- ◆ In healthcare, visualizations help track diseases, analyze patient data, and evaluate treatment outcomes.
- ◆ Education and research use visualization to explain concepts and communicate findings effectively.
- ◆ Government and public policy benefit through transparency, accountability, and public engagement.
- ◆ Media and communication apply visualization for storytelling and data journalism.
- ◆ Science, technology, and social analytics use visualizations to analyze complex and large datasets.
- ◆ Effective visual design requires clarity by focusing on readability and removing unnecessary elements.
- ◆ Precision ensures data is represented accurately without misleading visuals.
- ◆ Efficiency aims to deliver maximum information with minimal visual effort.
- ◆ Aesthetics, including color, layout, typography, and balance, enhance engagement, understanding, and memorability.

Objective Questions

1. What is the best definition of data visualization?
2. What is the primary purpose of data visualization?
3. Who is considered the father of statistical graphics?
4. Florence Nightingale is known for using data visualization to show what?
5. Which principle is NOT included in visual design principles?
6. In visual design, what does clarity ensure?
7. What is the role of aesthetics in data visualization?
8. What is a common application of data visualization in business?
9. Charles Minard is famous for which historical visualization?
10. What does efficiency in visual design refer to?

Answers

1. Graphical representation of data to convey insights
2. To simplify complex data and highlight patterns
3. William Playfair
4. Illustrating causes of mortality during the Crimean War
5. Complexity
6. Removing unnecessary clutter and highlighting key trends
7. Enhancing comprehension, engagement, and memorability
8. Tracking sales and performance metrics using dashboards
9. Napoleon's 1812 Russian campaign
10. Conveying maximum information with minimal visual effort

Assignments

1. Data Visualization Concepts

Scenario:

You are working as a business analyst for a retail company. Your manager provides you with raw sales, customer, and inventory data in multiple spreadsheets. She wants you to present a summary that highlights key trends and actionable insights for the last quarter.

Assignment Task:

- ◆ Explain how you would use data visualization to transform this raw data into meaningful insights.
- ◆ Provide examples of specific visualizations (e.g., bar chart, line chart, heatmap) and describe why each type is suitable for particular kinds of data.
- ◆ Discuss how these visualizations can aid decision-making across different departments such as sales, marketing, and inventory management.

2. Historical Evolution of Data Visualization

Scenario:

Your organization is preparing a training session on the history of data visualization for new analysts. You are asked to create a short report that demonstrates how historical pioneers shaped modern visual practices.

Assignment Task:

- ◆ Describe the contributions of William Playfair, Florence Nightingale, and Charles Minard to data visualization.
- ◆ Analyze how their techniques (e.g., bar charts, polar area diagrams, flow maps) have influenced contemporary visual analytics.
- ◆ Provide examples of how these principles can still be applied in modern dashboards or reports.

3. Principles and Aesthetics in Visual Design

Scenario:

A manager has created a sales performance dashboard that looks visually cluttered and confusing. You are asked to review it and propose improvements.

Assignment Task:

- ◆ Evaluate the dashboard using the principles of clarity, precision, and efficiency. Identify specific areas where these principles are not followed.
- ◆ Explain how aesthetics (such as color choice, spacing, and layout) can be improved to enhance comprehension and engagement.
- ◆ Suggest actionable modifications to the dashboard that would make the information more understandable and visually appealing.

Reference

- ◆ <https://www.dundas.com/resources/blogs/benefits-of-bi/3-examples-of-data-analytics-mking-enterprises-smarter-more-successful>
- ◆ <https://www.justinmind.com/blog/data-visualization-examples-principles/>

Suggested Reading

1. Tufte, E. R. (2001). The Visual Display of Quantitative Information. Cheshire, CT: Graphics Press.
2. Kirk, A. (2016). Data Visualisation: A Handbook for Data Driven Design. London: Sage Publications.
3. Yau, N. (2013). Data Points: Visualization That Means Something. Indianapolis, IN: Wiley.

UNIT 2

Visual Perception and Cognitive Principles

Learning Outcomes

At the end of this unit, the learner will be able to :

- ♦ understand how the human eye and brain perceive visual information and how it affects data interpretation
- ♦ explain Gestalt principles and use them to organize visual elements clearly in charts and graphs
- ♦ design visuals that reduce cognitive load, making them easy to read and understand
- ♦ apply Tufte's rules, such as Data-Ink Ratio, context, and comparison to create clean and effective visualizations
- ♦ identify good practices in data visualization by analyzing examples and using cognitive and perceptual principles to improve clarity and engagement

Prerequisite

Imagine, for a moment, that the human mind is a sophisticated, yet finite, processing unit is a **Grand Central Station** for information. Every single day, from the moment we wake, this station is utterly *bombarded*. We are standing at the nexus of an overwhelming torrent of visual data: the flashing colours of traffic, the endless scroll of digital feeds, the intricate patterns within a complex chart, or the critical numbers on a business dashboard. This flood is constant, immense, and unforgiving.

This deluge of visual stimuli is the first great challenge. Before we can master the art of data visualization, we must first truly understand the machinery that receives it. Without a deep comprehension of how the human brain instinctively **perceives, organizes, and ultimately interprets** these visual cues, even the most rigorously accurate and beautifully rendered visualizations become nothing more than high-resolution noise. They will fail to communicate.



The critical lesson lies in recognizing our **cognitive limitations**. The human brain, while powerful, is highly susceptible to *information overload*. This is the moment a complex visual, presented without foresight, transforms from an insightful tool into a confusing, misleading obstacle. It is like shouting critical instructions into a hurricane.

This is where our principles become the **map and compass**. Our goal is not just accuracy, but *intuition*. By systematically applying the timeless principles of visual perception and cognition, by utilizing the innate organizational shortcuts described by the Gestalt school, by purposefully employing techniques that actively reduce the *cognitive load*, and by following the elegant guidelines for clarity set forth by visionaries like Tufte, we can forge a bridge. We can transition from simply displaying data to creating representations that communicate insights seamlessly, guide the observer's attention precisely to what is critical, minimize mental effort, and fundamentally support sound, informed decision-making.

This prerequisite is our essential training: preparing us to move from being passive observers to becoming masterful architects of visual understanding.

Keywords

Visual Perception, Gestalt Principles, Cognitive Load, Data-Ink Ratio, Tufte's Rules, Perceptual Accuracy, Visual Hierarchy

Discussion

1.2.1 Visual Perception and Cognitive Principles Introduction

Our journey begins in the human mind's Visual Control Room. Think of the eye as a sophisticated scanner, capturing light and shape, sending raw electrical signals—a kind of complex digital code to the brain. Visual perception is the incredible interpreter that takes this code and translates it into meaningful reality. It's the brain's genius-level ability to not just see the pixels, but to recognize the underlying patterns, note the subtle differences, and instantly grasp the relationships hidden within the data.

The fundamental truth we must embrace is this: a truly effective visualization is not merely an accurate drawing; it is a successful conversation. It must align perfectly with the brain's innate perceptual mechanisms, acting as a gentle guide rather than a demanding instructor. When done correctly, complex information becomes instantly accessible, and the cognitive strain the mental fatigue of figuring out what you are looking at is minimized.

Decoding the Mind's Grammar: The Gestalt Principles : To become this successful conversationalist, we must first learn the brain's native language. This language is captured by the Gestalt Principles. Imagine these principles as the default “grammar rules” the brain uses to impose order on chaos. They describe the universal, unconscious shortcuts humans use to group and interpret visual elements.

When you see a cluster of data points positioned closer together (Proximity), your brain instantly labels them as a related group.

When a set of lines flows naturally in the same direction (Continuity), the brain anticipates them belonging to a single trend.

These principles—Proximity, Similarity, Continuity, Closure, and Figure-Ground—are the foundational building blocks. By designing charts where proximity clearly signals relationships (like clustering bars in a grouped chart), we are simply whispering the data's story in a language the viewer already understands. To truly grasp this, we will dive into a diagram that vividly illustrates how these powerful effects work, turning abstract rules into tangible visual experiences.

The Scarcity of Attention: Managing Cognitive Load: Next, we confront the brain's greatest limitation: Cognitive Load. This is the measurable mental effort required to process the information presented. Imagine a narrow intellectual pipeline; if you try to force too much data through at once—a cluttered chart with too many colours, labels, and unnecessary flourishes—the pipeline clogs. The viewer is overwhelmed, leading to confusion, misinterpretation, or, worst of all, the complete overlooking of the crucial insight.

Our mission as visualization designers is to be efficient curators. We reduce this load by ruthlessly simplifying the visual environment, by establishing consistent colour schemes that act as reliable navigational markers, and by strategically highlighting only the most important data points. We will examine a comparative table, starkly contrasting high vs. low cognitive load visualizations, to feel the difference between a cluttered mental space and one of immediate clarity.

The Classics of Clarity: Tufte's Design Philosophy: Finally, we turn to the masters. Edward Tufte's principles offer the ultimate guidance for achieving clarity and elegance. His core philosophy centers on ensuring that every element on the page serves the data. His famous Data-Ink Ratio challenges us to become minimalist sculptors: remove every drop of non-essential ink, every needless border, shadow, or decoration, until only the data itself remains in sharp focus.

But clarity is not enough; the data must be meaningful. We must provide Context, ensuring the viewer grasps the story behind the numbers, and enable Comparison, allowing for effective analysis of differences and trends. To appreciate the power of these rules, we will study classic, iconic visualizations like Minard's tragic chart of Napoleon's campaign or Florence Nightingale's revolutionary coxcomb diagrams to see these rules applied in practice, transforming historical data into enduring masterpieces of communication.

This unit, therefore, is our training ground. By mastering how the human brain perceives patterns, skillfully reduces information overload, and instinctively interprets visual cues, we move beyond mere chart-making. We become designers capable of creating visualizations that are not just accurate, but deeply intuitive, highly engaging, and ultimately, profoundly effective.

1.2.2 Visual Perception Basics

We have established that the eye is a scanner and the brain is a decoder. Now, let us solidify the fundamental rules of this decoding process—the essential operating manual for creating clarity from data. These are the built-in, non-negotiable processing mechanisms that dictate how every viewer will engage with your work.

1.2.2.1 The Instinct for Difference

The human visual system is fundamentally a difference-detector. It is not designed to admire sameness; it is designed to pinpoint variation. This instinct is the engine of insight.

- ◆ When you vary the Size or Orientation of an element, you are instantaneously signaling a difference in value or a shift in trend. A sudden spike in a line graph, for instance, leverages our sensitivity to position and alignment to scream “Pay attention! An anomaly is here!”
- ◆ When you use Color, you are exploiting our sensitivity to hue to build instantaneous categories. A data point in a contrasting colour is instantly recognized as an outlier or a member of a separate cluster.

1.2.2.2 The Four Pillars of Visual Coding

Every successful chart you design relies on one or more of these four core perceptual pillars:

1. **Colour Perception :** The Signal Flag Colour is the ultimate organizational tool, used to categorize or emphasize. It is the signal flag you raise to guide the viewer’s attention. But be warned: irresponsible use of colour (too many hues, non-sequential palettes) is the fastest way to confusion. It can accidentally create false categories or, in the case of vision deficiencies, create an invisible wall around your data.
2. **Shape and Size :** The Quantitative Scale This pillar dictates that magnitude is best communicated through size. When we see a larger circle in a bubble chart, our interpretation is immediate and accurate: more value. When we vary the shape, we are offering distinct visual identities for different variables.
3. **Position and Alignment :** The Comparison King The most accurate and least mentally taxing way to compare data is by placing it along a shared axis. Bar charts and scatter plots are perceptual champions precisely because they rely on the brain’s flawless ability to compare relative position and alignment. This makes the act of comparison subconscious and instant.

4. **Gestalt Cues** : The Brain's Search for Wholeness This final pillar reminds us that the brain is an active constructor of meaning. It doesn't just receive scattered elements; it actively seeks to group related items (Proximity, Similarity) and complete incomplete stories (Closure). Your job as a designer is to respect and leverage these cues, ensuring the viewer's natural instinct for order leads them directly to the intended conclusion.

When your visualization aligns with these inherent perceptual tendencies, the data narrative becomes effortless to follow. But when you create a cluttered battlefield of overlapping elements or chaotic colours, you force the brain to abandon its natural process. You increase the cognitive burden and guarantee that the crucial patterns—the very reason for creating the visualization—will remain tragically overlooked. For a clear demonstration of this difference, observe Fig 1.2.1 Cluttered vs. Clear Bar Chart: Visualizing the Impact of Effective Design, which illustrates how visual noise obscures insight.

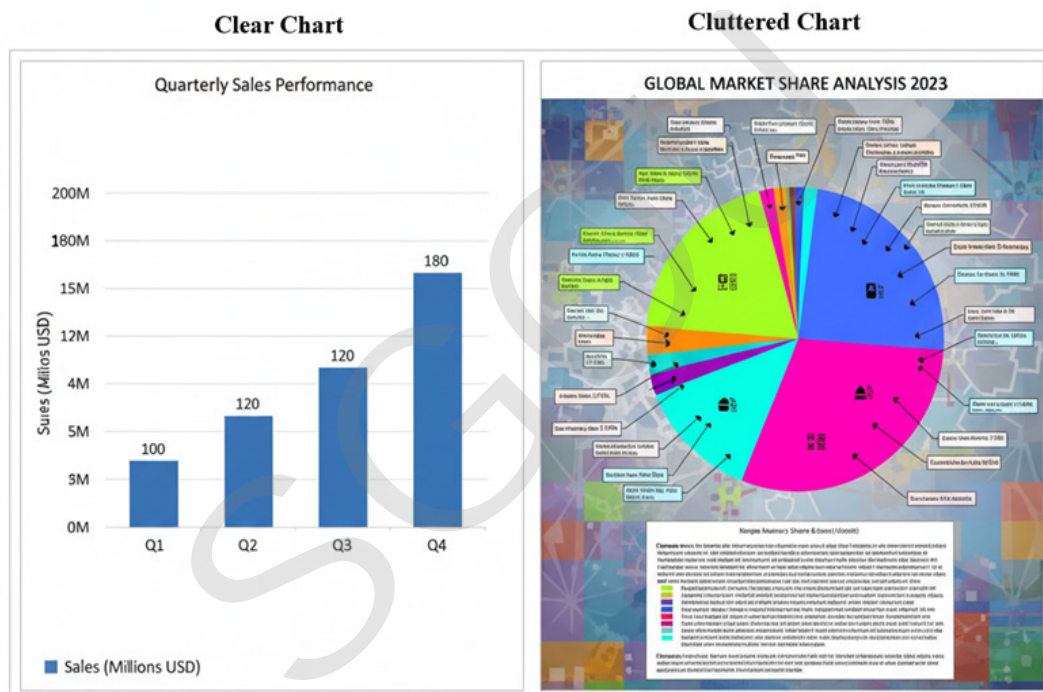


Fig. 1.2.1 Cluttered vs. Clear Bar Chart: Visualizing the Impact of Effective Design

1.2.3 Gestalt Principles

If the eye delivers raw data, and perception is the translator, then the Gestalt Principles are the brain's innate architect's blueprint—the hardwired set of instructions it uses to bring order to visual chaos. Originating from the profound insights of Gestalt psychology, these principles reveal a fundamental truth: the whole is perceived as greater, and often different, than the mere sum of its parts. They are the secret keys to ensuring your audience can intuitively group, distinguish, and interpret complex data without ever feeling the unnecessary friction of cognitive effort.

Let's unpack these powerful organizational instincts:



1.2.3.1 Proximity: The Rule of Nearness

Imagine data points as individuals in a crowded room. The principle of Proximity states that those standing closest together are immediately perceived as being related—they belong to the same conversation or group. In visualization, this is the simple yet powerful tool that defines clustered bar charts or logical groupings in a network diagram. By physically bringing related data elements near each other, we instantly, and correctly, signal their relationship to the viewer.

1.2.3.2 Similarity: The Common Uniform

When elements share a similar characteristic—be it the same colour, size, or shape—the brain automatically assigns them to the same category. They are wearing the same uniform. This principle of Similarity is critical for navigating complex datasets, allowing the viewer to effortlessly track a specific segment or category across an entire chart, immediately revealing trends and patterns that might otherwise be hidden in the data's complexity.

1.2.3.3 Continuity: The Unbroken Path

The brain prefers a smooth narrative. Continuity is our tendency to follow lines, curves, and sequences as uninterrupted paths, even if they pass through other elements. This is why line graphs are so effective for showing trends over time. The viewer instinctively perceives the connected line as a continuous journey, making the relationship and movement between time points simple to grasp. We dislike abrupt breaks; we prefer the smooth, logical flow.

1.2.3.4 Closure: The Mind's Completion

The brain is a tireless pattern-completer. The principle of Closure allows us to recognize a familiar shape or pattern even when the information is incomplete or fragmented. It's the visual instinct to 'fill in the gaps' to perceive a complete whole. This can be subtly leveraged in visualization to simplify designs, allowing the viewer's mind to do the work of completion, which often makes the design feel cleaner and more elegant.

1.2.3.5 Figure-Ground: The Spotlight Effect

Figure-Ground is the battle for attention—the fundamental ability to distinguish the object of focus (the figure) from the rest of the visual scene (the ground). Effective use of contrast, bold color, and strategic whitespace ensures that the critical data points and focal elements leap forward, standing out clearly against the background noise. If this principle is ignored, the data points blend into the axes and gridlines, and the message is lost in a visual mush.

1.2.3.6 Common Fate: Synchronized Movement

This principle is the heartbeat of dynamic data. Common Fate dictates that elements that move or change together are perceived as belonging to a single group or representing a singular, coordinated trend. In interactive or animated charts, this is a powerful

device: watching a cluster of elements rise together signals a shared destiny, making the communication of evolving trends and patterns immediate and impactful (For a summary of these key organizational instincts, see Fig 1.2.2 Gestalt Principles of Visual Organization.).

By consciously embedding these Gestalt principles into our visualization design, we stop fighting the brain and start collaborating with it. We drastically reduce the cognitive tax on the viewer, dramatically enhance comprehension, and transform the communication of data into an intuitive, efficient, and ultimately, far more impactful experience.

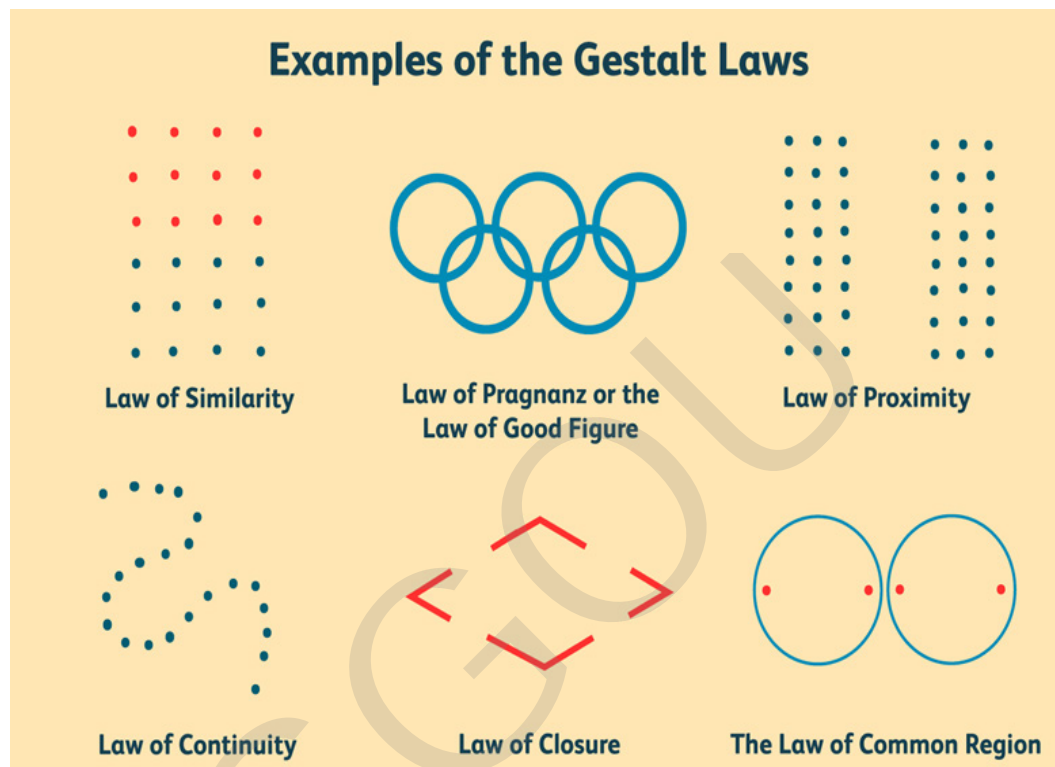


Fig 1.2.2 Gestalt Principles of Visual Organization

1.2.4 Cognitive Load and Reduction Techniques

Imagine the human mind is not just a processor, but a pipeline with a strictly limited bandwidth. Cognitive load refers to the mental effort required to force information through that pipeline. Every unnecessary element, every cluttered label, and every confusing colour choice acts as a piece of grit, slowing the flow and increasing the processing effort. Our core objective here is to prevent the cognitive pipeline from clogging.

Visualizations that demand excessive mental gymnastics—those that are overly complex or densely cluttered—do not communicate; they overwhelm. This immediately leads the viewer to one of two detrimental outcomes: either they entirely misinterpret the data, or they suffer from insight fatigue, causing them to completely overlook the crucial patterns you intended to show.

1.2.4.1 The Designer's Duty: Curating for Clarity

Our duty as visualization designers is to be meticulous curators of attention and active reducers of cognitive load. We achieve this through strategic, mindful simplification, ensuring that the brain's resources are dedicated only to understanding the data, not deciphering the design.

Key strategies for effective load reduction include:

- ◆ **Simplification and Minimalism:** Ruthlessly remove any visual element that does not directly carry data or provide essential context. Unnecessary borders, excessive gridlines, or decorative 3D effects are simply taxing the viewer for no gain.
- ◆ **Consistency and Predictability:** Use a consistent colour scheme and logical layout across all related visuals. When the viewer knows that a certain colour always represents "Sales," their brain doesn't have to waste effort re-categorizing it on every new chart.
- ◆ **Strategic Highlighting:** Use visual intensity (e.g., brighter colours, bolding, contrast) to purposefully highlight the most important data points. This acts as an immediate visual signpost, directing the viewer's limited attention precisely to the crucial insight and bypassing the need to scan every data point equally.
- ◆ **Effective Grouping:** Leverage the Gestalt principles (especially Proximity and Similarity) to ensure data elements that belong together are seen together. This organizational clarity reduces the effort needed to mentally link related pieces of information.

To truly appreciate this difference, consider the stark contrast between a dashboard choked with fifty different colours and labels versus one that uses a refined palette to emphasize three key performance indicators. This side-by-side comparison is often the most powerful lesson in design.

By actively managing cognitive load, we transform our visualizations from challenging puzzles into intuitive narratives, ensuring the audience's mental energy is spent analyzing the data's meaning, not struggling with the chart's mechanics.

1.2.5 Tufte's Rules: Data-Ink Ratio, Context, and Comparison

Having mastered the brain's perception (Gestalt) and managing its limitations (Cognitive Load), we arrive at the principles of Edward Tufte, the visionary who taught us that the clarity and intellectual integrity of a visualization are its ultimate aesthetic. Tufte's work is not about decoration; it is about establishing a rigorous standard for maximum data density and minimal distraction. These guidelines are our final polish, ensuring our visualizations are not just accurate, but honest, clear, and profoundly efficient.

1.2.5.1 The Data-Ink Ratio: Eliminating the Noise

Tufte's most famous command is the Data-Ink Ratio. Imagine you have a fixed amount of 'ink' to spend on your design. This principle dictates that every single drop of ink—every line, border, background, and tick mark—must serve the data itself.

$$\text{Data – Ink Ratio} = \frac{\text{Data – Ink}}{\text{Total Ink Used}}$$

This forces us to become visual minimalists: we must ruthlessly remove all non-essential ink (chart junk) that fails to communicate a piece of data. Removing shadows, redundant labels, or decorative frames ensures that the viewer's eye is focused exclusively on the signal, not the noise. The goal is to maximize the data's visibility and impact.

1.2.5.2 Providing Context: The Meaning Behind the Number

A number, isolated, is merely a fact; with context, it becomes an insight. Tufte stresses that effective visualizations must always provide the intellectual setting required for viewers to understand the meaning behind the numbers and trends. This includes providing clear labels, necessary scales, units of measure, and explanatory annotations. We are not just showing data; we are enabling a narrative, ensuring the viewer can grasp the significance of a trend relative to history, goals, or external events.

1.2.5.3 Comparison: The Engine of Analysis

Analysis is fundamentally about evaluating differences. The principle of Comparison emphasizes that a visualization must be designed to facilitate the effective analysis of differences between data points. Whether this is achieved through placing related graphs adjacent to one another (small multiples) or by clearly aligning elements on a single baseline, the goal is to make the act of noticing a difference, a change, or a correlation effortless and immediate.

1.2.5.4 Enduring Examples of Excellence

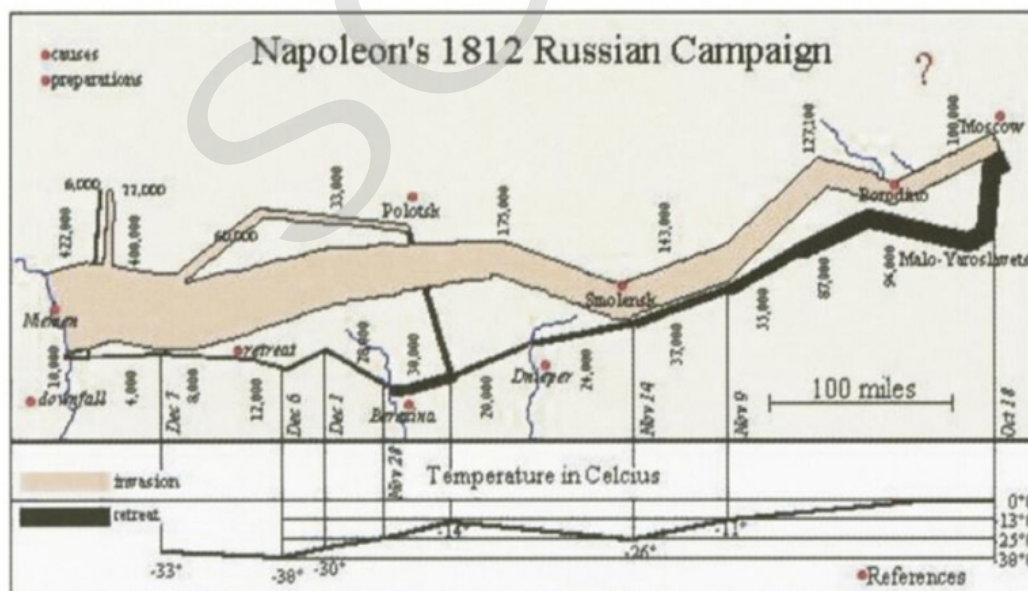


Fig 1.2.3 Minard's 1812 Campaign: A Tuftean Masterpiece of Data Visualization

To truly grasp the power of Tufte's ideals, we look to history's masterpieces. Charles Minard's map of Napoleon's Russian Campaign, for example, elegantly visualizes five variables—the army's size, its location, temperature, date, and direction—in a single, complex, yet instantly understandable graphic (Fig 1.2.3 Minard's 1812 Campaign: A Tuftean Masterpiece of Data Visualization). Likewise, Florence Nightingale's coxcomb diagrams transformed statistics on mortality into a compelling, visual argument for sanitation reform (Fig 1.2.4 Florence Nightingale's Coxcomb Diagram: The Causes of Mortality in the Crimean War). These are not merely charts; they are demonstrations of how high data-density, combined with clarity and purpose, can yield profound, enduring insights.

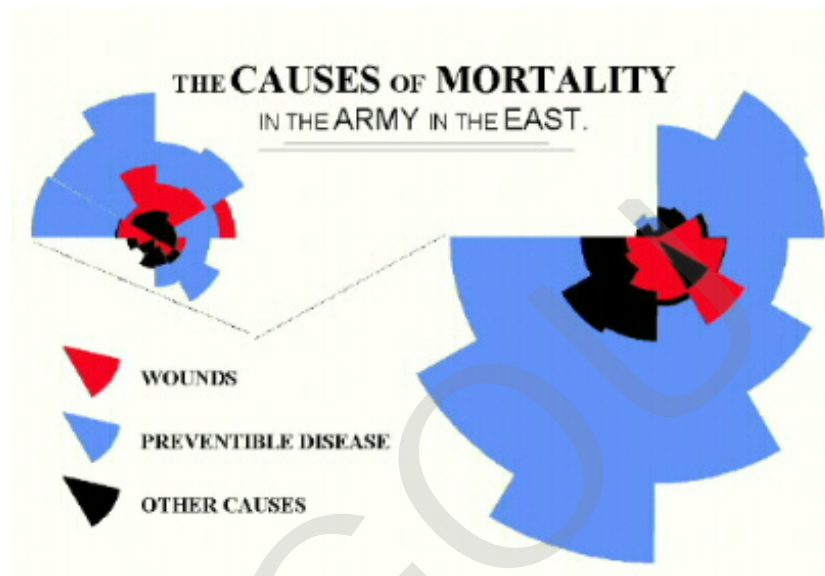


Fig 1.2.4 Florence Nightingale's Coxcomb Diagram:
The Causes of Mortality in the Crimean War

Recap

- ◆ Visual perception explains how the brain interprets visual features such as color, shape, size, position, and spatial relationships.
- ◆ Gestalt principles like proximity, similarity, continuity, closure, figure-ground, and common fate help organize and understand visual elements.
- ◆ Cognitive load refers to the mental effort needed to process information, and reducing it improves understanding.
- ◆ Cognitive load can be reduced by simplifying visuals, using consistent design, grouping related data, and highlighting key information.
- ◆ Tufte's rules emphasize minimizing unnecessary elements (Data-Ink Ratio), providing context, and enabling clear comparison in visualizations.

- ◆ Famous examples such as Minard's chart of Napoleon's campaign and Nightingale's coxcomb diagrams show effective, clear, and meaningful data visualization.

Objective Questions

1. What is the primary definition of Cognitive Load in data visualization design?
2. Which Gestalt principle causes a viewer to perceive elements that are close together as related?
3. What is the main objective of maximizing Tufte's Data-Ink Ratio in a visualization?
4. How does high Cognitive Load interfere with the viewer's ability to extract insights from a complex visualization?
5. Using varying circle sizes in a bubble chart to represent market value leverages which aspect of visual perception?
6. A line graph lacks labels for the Y-axis units. Which of Tufte's principles is violated?
7. Fading effect that highlights only the top 10% of data points represents which Cognitive Load reduction technique?
8. Which Gestalt principle separates the main data (figure) from the background (ground)?
9. The combined use of simplification, consistency, and highlighting aims to achieve what goal?
10. A bar chart with 3D effects, heavy gridlines, and shadows violates which core principles?

Answers

1. The mental effort required by the viewer to process and interpret the information presented
2. Proximity
3. To focus the viewer's attention on the data by removing non-essential elements
4. It overwhelms the viewer, causing misinterpretation or overlooking important patterns
5. Shape and size
6. The principle of Context
7. Strategic highlighting
8. Figure-Ground
9. Reducing Cognitive Load to enhance comprehension and efficiency
10. Data-Ink Ratio and Cognitive Load Reduction

Assignments

1. The Design Challenge (Gestalt and Cognitive Load)

Scenario:

You are tasked with designing a dashboard to show weekly sales performance across four different regional teams (North, South, East, West). The current data is presented in a single, large table containing 50 rows and 15 columns.

Assignment Task:

Describe a new visual layout for this dashboard that utilizes the Gestalt principles of Proximity and Similarity to effectively structure the data and minimize the overall Cognitive Load.

- ◆ Specify how you would use Proximity to organize the information.
- ◆ Specify how you would use Similarity (e.g., color or shape) to highlight team performance.

- ◆ Explain how these two design choices successfully reduce the mental effort required for the manager to quickly identify the best and worst performing teams.

2. The Tufte vs. Clutter Breakdown

Scenario:

Examine a visualization that has been deliberately created with high “chart junk”. This visualization features a heavy 3D effect on its bars, a detailed background pattern, bright contrasting colours for non-data elements (axes, gridlines), and unnecessary borders.

Assignment Task:

Conduct a brief analysis (approximately 250 words) on how this visualization fails based on both Edward Tufte’s Data-Ink Ratio and the goal of Cognitive Load Reduction.

3. The Principle Trade-Off

Scenario:

You are designing a series of 12 small line graphs (known as “small multiples”) to compare monthly sales trends for 12 different products. The design team is debating between two options:

- ◆ Option A: Use a unique, bright color for the line of *each* of the 12 products across all 12 graphs.
- ◆ Option B: Use a single, muted gray color for all 12 lines, but highlight only the trend of the *focus product* on each specific graph using a bright, consistent color (e.g., red).

Assignment Task:

Evaluate which option provides the better balance for data visualization effectiveness.

Justify your choice by evaluating the trade-offs between Option A and Option B in terms of:

1. Gestalt Principle of Similarity: Which option leverages it better, and why?
2. Cognitive Load: Which option poses a higher risk of overwhelming the viewer, and why?
3. Strategic Highlighting: Which option is superior for focusing attention and delivering quick insight?

Reference

- ◆ <https://flowingdata.com/category/tutorials/>
- ◆ <https://www.geeksforgeeks.org/data-visualization/mastering-tufte-data-visualization-principles/>

Suggested Reading

1. Tufte, E. R. (2001). The Visual Display of Quantitative Information. Cheshire, CT: Graphics Press.
2. Ware, C. (2013). Information Visualization: Perception for Design. Burlington, MA: Morgan Kaufmann.
3. Few, S. (2012). Show Me the Numbers: Designing Tables and Graphs to Enlighten. Burlingame, CA: Analytics Press.

UNIT 3

Designing for Different Audiences - Understanding Your Audience

Learning Outcomes

At the end of this unit, the learner will be able to :

- ◆ identify the needs and expectations of different stakeholders for data visualization
- ◆ tailor visualizations to suit the audience's knowledge level and decision-making needs
- ◆ apply storytelling principles to convert data into clear, engaging narratives
- ◆ select and design appropriate visualization types to communicate insights effectively
- ◆ evaluate audience-focused visualizations for clarity, effectiveness, and impact

Prerequisite

We have built our beautiful, technically sound visualization, a masterpiece of data, Gestalt, and low cognitive load. But imagine standing on a stage, ready to deliver a profound, perfectly crafted speech... to an empty room. This is the danger of ignoring the Audience Protocol.

It is an enduring truth in this field: Data alone does not tell the whole story; the audience completes it.

Before a single pixel is placed, we must recognize that every person interprets and values information differently, depending entirely on their role, expertise, and core objectives. For example, the Business Executive is the busy pilot needing a quick glance at the altimeter and fuel gauge a high-level summary of Key Performance Metrics (KPIs). Conversely, the Data Analyst is the flight engineer, requiring detailed charts, granular insights, and the ability to drill down to investigate every subtle trend and anomaly.



Without this deep consideration, even the most accurate and elegant visualization is at risk of failure. It may simply speak the wrong language or deliver the wrong message, like handing a highly technical wiring diagram to someone who only needs the summary of the day's profits.

From Data to Narrative: The Power of Story

Understanding your audience means recognizing their knowledge level (how much detail they can handle), their priorities (what decision they need to make), and the context in which they will use the data (a quick mobile glance versus a deep boardroom review). Tailoring our visualizations to meet these specific needs is the guarantee of clarity, relevance, and engagement.

Crucially, this is where the *art* of storytelling transforms the *science* of visualization. By intentionally integrating narrative principles, we stop creating static data pictures and start guiding the viewer through an insightful journey. The visualization becomes a compelling story that highlights patterns, emphasizes key insights, and inexorably supports informed decision-making.

Our final goal is to establish a powerful, unbreakable connection between the three pillars: the Data, the Audience, and the Story. When these three align, the designer creates not just an informative visualization, but a compelling, memorable, and ultimately, actionable tool.

Keywords

Audience Analysis, Stakeholder Needs, Data Storytelling, Narrative Visualization, Visualization, Tailoring, Insight Communication, Decision Support, Contextual Relevance, Engagement Strategies.

Discussion

1.3.1 Designing for Different Audiences – Understanding Your Audience

When we begin exploring the world of data visualization, we quickly recognize an essential truth: data by itself is never the whole story. The true challenge is not simply analyzing or plotting data correctly, but communicating it effectively. Creating a meaningful visualization requires us to understand that we are not designing for ourselves; we are designing for someone else a learner, a decision-maker, or a stakeholder each with unique expectations, prior knowledge, and decision-making responsibilities. If we attempt to use one universal visualization for everyone, we risk losing the core message entirely.

However, when we tailor visualizations carefully to suit a particular audience, we ensure that the message becomes clear, relevant, and actionable.

Designing a visualization is much like teaching a class. If I were to explain a complex concept to different groups a primary school student, a college student, and a domain expert each explanation would look completely different, though the underlying idea remains the same. In the same way, data visualizations must respect the needs of the audience. A visualization that works for a technical analyst will not work for a senior executive, just as a highly detailed schematic is unhelpful in a five-minute strategic meeting.

1.3.1.1 Identifying the Stakeholder's Needs

Understanding the audience begins with seeing from their point of view. We must identify who will be consuming the visualization, determine their specific objectives, and understand how they typically interpret information. This step ensures comprehension and usability.

Executives typically focus only on the destination and the speed at which the organization is moving. They prefer a high-level overview rather than detailed exploration. They look for the broad picture that helps them evaluate performance through summary indicators, trend lines, and quick signals of opportunity or risk. Executives do not concern themselves with granular data or the mechanics behind the results; their priority is clear, immediate insight.

Analysts, however, require charts that allow exploration. They work deeply with data, test hypotheses, inspect outliers, and validate assumptions. They need detailed visualizations such as scatter plots, distribution charts, diagnostic diagrams, and interactive dashboards that let them drill into lower levels of detail. While an executive wants a map showing only the destination, the analyst wants every road, turn, hill, and valley on the way.

Recognizing the difference between these needs is crucial. A visualization that is too detailed for an executive or too shallow for an analyst will fail. But when we design visualizations by aligning the details and format with the audience's role and context, we transform a simple graph into an effective decision-making tool.

1.3.1.2 Storytelling: Guiding the Audience Through the Data

This naturally brings us to the importance of storytelling. Storytelling with data is the process of turning raw numerical information into a meaningful narrative that guides the viewer from context to insight. Without storytelling, a visualization becomes merely a static display of numbers. With storytelling, it becomes an engaging explanation that clarifies patterns, emphasizes key comparisons, and supports strategic decisions.

A well-structured story leads the audience gently through the data. The visuals are arranged in a logical order, beginning with background context, moving toward exploration, and finally arriving at insights. Annotations play the role of a teacher's commentary, gently pointing out what matters and clarifying what the audience should notice. These annotations act like quiet instructions that help the viewer understand deeper meaning without feeling overwhelmed.

Balancing simplicity and detail is essential. The story must remain simple enough for the audience to follow without confusion, but sufficiently detailed to retain credibility and analytical strength. An effective story maintains this balance by highlighting only the most meaningful patterns and trends, ensuring that key messages stand out clearly. When audience analysis and storytelling are combined, visualizations evolve from simple charts into powerful communication instruments that are memorable, engaging, and deeply informative.

1.3.2 Tailoring Visualizations to Different Stakeholders

To communicate data effectively, we must think strategically. Accuracy alone is not sufficient; the visualization must be adapted to the needs and expectations of the intended audience. Different stakeholders view data through different lenses, and their decisions depend entirely on how the information is presented. Before choosing a single chart or color, we must analyze each audience as a unique case. The amount of detail, choice of visual encoding, and level of interactivity must be planned with their responsibilities and expertise in mind.

Failing to adjust visualizations appropriately results in confusion, information overload, or disengagement. Providing an executive with excessive detail slows decision-making, while offering an analyst only a high-level chart removes the ability to investigate trends. Therefore, the process of tailoring visualizations is critical for transforming raw data into meaningful communication.

Below is the expanded explanation written in sentences of how different stakeholder groups should be approached.

Executives require quick insight and fast decision-making. Their primary goal is to understand overall performance at a glance. The visual strategy for executives emphasizes summary metrics, key comparisons over time, and clear signals of opportunity or risk. Because executives have limited time and prefer direct messages, conciseness becomes the most important aspect. Tools such as high-level dashboards, gauges, and KPI cards serve this purpose effectively.

Managers and team leads focus on operational monitoring and interventions. Their role demands a balance between detail and clarity. They must see aggregated data along with the option to explore deeper when needed. The visual strategy for managers supports this by providing operational dashboards that combine high-level summaries with drill-down paths. Visual tools such as small multiples and annotated bar or line charts help them identify issues and take corrective action.

Analysts and technical users concentrate on in-depth exploration and anomaly detection. Their work requires detailed, granular visualizations that encourage investigation. The visual strategy therefore offers interactive dashboards, scatter plots, multi-dimensional charts, and flexible parameter controls. These tools allow analysts to investigate relationships, explore patterns, and validate models.

Clients and external stakeholders look for clear explanations of trends, outcomes, or business impact. Since they may not be experts in the subject, the visual strategy centers on simplicity, clarity, and persuasion. They benefit from simplified dashboards,

infographics, and annotations that highlight key findings while minimizing technical complexity.

Perfect alignment between the stakeholder's goals and the visualization's design ensures that insights are not only accessible but also actionable. When the visualization reflects the audience's priorities and knowledge level, communication becomes efficient and impactful. Here is our dossier detailing the communication strategy for key audience segments: Table 1.3.1

Table 1.3.1 Stakeholder Visual Adaptation Strategy

Stakeholder Group	Primary Goal	Visual Strategy	Key Tools
Executives	Strategic insight & quick decision-making.	Focus on summary metrics, high-level comparisons over time, and clear signals of performance or risk. Conciseness is paramount.	High-level Dashboards, Gauges, KPI Cards
Managers & Team Leads	Operational monitoring & intervention.	Balance detail and clarity, providing aggregate views alongside the ability to drill down into specific components.	Operational Dashboards, Small Multiples, Annotated Bar/Line Charts
Analysts & Technical Users	In-depth exploration & anomaly identification.	Must provide detailed, granular visualizations that allow for self-directed investigation and interpretation.	Interactive Dashboards, Scatter Plots, Multi-dimensional Charts, Parameters/Filters
Clients & External Stakeholders	Understanding trends, outcomes, or business impact.	Must simplify complexity, emphasize clarity, and tell a persuasive story.	Simplified Dashboards, Infographics, Annotated Narrative Charts

The Power of Alignment

By meticulously understanding the stakeholder's goals, their expertise level, and their decision-making requirements, designers can make informed choices regarding visualization types, levels of detail, and presentation styles.

This deliberate alignment ensures that the hard-won insights buried within the data

are not only accessible but are genuinely actionable, directly contributing to the priorities of each audience and significantly enhancing both comprehension and ultimate impact.

1.3.3 Storytelling with Data: Crafting Narratives Through Data Visualization

Now that we understand clarity, audience needs, and message adaptation, we can move into the final key element of effective visualization: storytelling. Storytelling transforms static information into a dynamic journey that leads the viewer from discovery to understanding.

1.3.3.1 The Narrative Engine: Data Storytelling

Storytelling is what gives structure to data. While data contains truth, it does not automatically communicate it. Storytelling brings coherence and meaning by arranging information in a way that naturally explains what is happening and why it matters. Without a narrative, data feels fragmented; with narrative, data becomes memorable, persuasive, and actionable.

A narrative guides the viewer through insights step by step. It removes noise, highlights what is important, and ensures that the viewer stays focused on the intended message. In this way, storytelling acts as the engine that powers effective visualization.

1.3.3.2 Tools of Emphasis: Annotations and Balance

Annotations, labels, and highlights serve as the voice of the narrator within the visualization. They gently guide the viewer's focus toward important trends, turning points, or exceptions. Without annotations, the viewer may overlook crucial insights. With annotations, the story becomes clearer and more accessible.

Balancing simplicity and detail ensures that the story remains digestible yet credible. An overly complex visualization overwhelms the viewer, while an oversimplified visualization lacks depth. A good story emphasizes patterns and comparisons that matter most while minimizing distractions.

1.3.3.3 The Impact of a Well-Told Story

A well-told data story adapts to the audience, ensuring that each group receives the right level of explanation and detail. Executives may need only a summary narrative, while analysts may need a comprehensive exploration. Regardless of the audience, a strong story elevates visualizations from simple charts into tools that support reasoning, influence decisions, and communicate insights in a memorable way.

Recap

- ◆ Effective data visualization depends on understanding who the audience is and what insights they seek.
- ◆ Different stakeholders executives, analysts, clients, or the public require tailored visualizations that match their goals, expertise, and decision-making needs.
- ◆ Audience analysis helps determine the appropriate level of detail, type of chart, and communication style.
- ◆ Contextual design ensures that visuals remain relevant and focused on the specific questions each audience wants answered.
- ◆ Storytelling with data transforms visualizations from static representations into engaging narratives that highlight patterns, trends, and implications.
- ◆ A clear narrative structure introduction, insight exploration, and conclusion guides viewers through data logically and persuasively.
- ◆ Annotations, color emphasis, and sequencing are essential techniques for maintaining clarity and focus within a story.
- ◆ Well-designed stories with data promote better understanding, retention, and decision-making.
- ◆ The ultimate goal of data storytelling is not just to present information but to inspire action and insight.

Objective Questions

1. Which of the following best describes the purpose of tailoring visualizations for different audiences?
2. Why is audience analysis important in data visualization?
3. When presenting to top-level executives, visualizations should primarily:
4. Provide an example of how a narrative can be structured in a data story to improve audience understanding.
5. Analyze two dashboards: one tailored for executives, one for technical teams. Which elements differ and why?
6. Evaluate a data story with excessive text and minimal visuals. What is the main drawback?

7. Compare two visualizations of the same dataset: one uses interactive filters, the other is static. Which approach better engages the audience and why?
8. Design a short data story for a quarterly sales report targeting the marketing team. What key elements would you include?
9. Create a visual dashboard for both executives and analysts. How would you structure it to meet the needs of both audiences?
10. Justify why designing audience-specific visualizations is critical for decision-making in organizations.

Answers

1. To ensure data aligns with audience needs and understanding
2. It ensures the message is clear, relevant, and actionable for the viewer
3. Focus on key metrics, trends, and summaries
4. Lead the audience from context to insight and conclusion
5. Executive dashboards focus on summaries; technical dashboards include detailed metrics and comparisons
6. It overwhelms the audience and reduces comprehension
7. Interactive charts, because they allow exploration and deeper insights
8. Key insights, context, visualizations, and clear conclusions
9. Include high-level summaries for executives and detailed metrics for analysts, using filters and drill-downs
10. It maximizes engagement, understanding, and informed decision-making

Assignments

1. Audience Analysis in Data Visualization

Scenario:

Your team has two dashboards displaying the same dataset. One dashboard is cluttered with excessive tables and colors, while the other uses high-level summaries and simple charts.

Assignment Task:

- ◆ Compare the two dashboards and analyze which one is more effective for different audiences (executives vs. technical teams).
- ◆ Identify the elements that contribute to clarity, comprehension, and cognitive load.
- ◆ Justify your recommendations for improving audience-specific communication.

2. Evaluating: Storytelling with Data

Scenario:

A marketing dashboard has been designed, but the team feels it does not communicate the key insights effectively. You are asked to review it and suggest improvements.

Assignment Task:

- ◆ Evaluate the current dashboard's narrative and explain where the story is unclear or misleading.
- ◆ Recommend ways to structure the data story to lead the audience from context to insight and conclusion.
- ◆ Discuss how annotations, visual cues, and contextual explanations can enhance comprehension.

3. Designing Audience-Specific Visualizations

Scenario:

You have been assigned a quarterly sales dataset and need to create a data story for both executives and analysts.

Assignment Task:

- ◆ Design a small infographic or dashboard that meets the needs of both audiences.

- ◆ Explain how you would include high-level summaries for executives and detailed metrics for analysts using charts, filters, or drill-down features.
- ◆ Ensure your design applies principles of clarity, efficiency, and visual appeal to maximize understanding and engagement.

Reference

- ◆ <https://www.storytellingwithdata.com/blog>
- ◆ <https://public.tableau.com/app/learn/how-to-videos>

Suggested Reading

1. Knaflic, Cole Nussbaumer. *Storytelling with Data: A Data Visualization Guide for Business Professionals*. Wiley, 2015.
2. Dykes, Brent. *Effective Data Storytelling: How to Drive Change with Data, Narrative and Visuals*. Wiley, 2019.

UNIT 4

Visualization Best Practices and Trends

Learning Outcomes

At the end of this unit, the learner will be able to :

- ◆ explain the importance of visualization best practices in creating clear and meaningful data insights
- ◆ describe the key principles involved in designing effective dashboards for decision-making
- ◆ identify common pitfalls in data visualization and explain how they affect interpretation
- ◆ discuss the role of interactive elements in enhancing user engagement and data exploration
- ◆ summarize current and emerging trends in data visualization and their relevance in modern analytics

Prerequisite

In today's fast paced data-driven world, organizations rely heavily on dashboards and visual reports to monitor performance, track progress, and support real time decision-making. Imagine a company that manages its sales, customer feedback, and logistics data across multiple departments. Each team creates its own reports some using complex charts, others presenting dense tables or colorful visuals with no consistent format. As a result, decision makers struggle to interpret key insights quickly, often missing critical trends hidden behind poor design choices.

Without proper visualization standards, dashboards can become cluttered, confusing, or misleading, leading to errors in judgment and delayed actions. On the other hand, a well structured, intuitive, and interactive visualization can transform complex datasets into clear, actionable intelligence.

Understanding the principles of effective visual design, common pitfalls to avoid, and emerging visualization trends is therefore essential for creating meaningful, accurate, and user friendly dashboards that empower better business outcomes.



Keywords

Data Clarity and Accuracy, Visualization Pitfalls, Interactive Visualizations

Discussion

1.4.1 Visualization Best Practices and Trends

Effective data visualization goes beyond presenting information; it is about creating visuals that are clear, accurate, and actionable. In modern organizations, dashboards and interactive reports are essential tools for monitoring performance, identifying trends, and supporting decision-making. However, designing impactful visualizations requires understanding best practices, common pitfalls, and emerging trends to ensure that the visuals communicate insights effectively without causing confusion or misinterpretation.

Dashboard Design Principles focus on clarity, simplicity, and user-centric layout. Key aspects include logical arrangement of charts, consistent use of colors and fonts, prioritizing critical information, and maintaining a balance between detail and overview. A well designed dashboard enables users to interpret information quickly, compare metrics, and make informed decisions.

Common Pitfalls in visualization include cluttered layouts, misleading scales, excessive colors, or unnecessary decoration that distracts from the core message. Recognizing and avoiding these mistakes enhances accuracy, comprehension, and credibility.

Interactive Visualizations add another layer of engagement, allowing users to explore data dynamically, filter views, and gain insights tailored to their needs. Interactive dashboards empower users to drill down into details, test hypotheses, and uncover patterns that static visuals might miss.

Emerging Trends in visualization include real time analytics, AI-assisted charting, storytelling dashboards, and the integration of multimedia elements to communicate insights more effectively. Staying updated with these trends ensures that visualizations remain relevant, modern, and impactful in rapidly evolving data environments.

By mastering best practices and keeping pace with trends, learners can create visualizations that are not only visually appealing but also functionally powerful, supporting better analysis, decision-making, and storytelling with data.

1.4.2 Dashboard Design Principles

Dashboards are powerful tools that provide a consolidated view of key metrics and insights, allowing users to monitor performance, detect trends, and make informed decisions quickly. Designing an effective dashboard requires careful attention to layout, clarity, and user needs. The following principles are essential for creating dashboards that are both functional and visually appealing:

1. **Clarity and Simplicity :** The primary goal of a dashboard is to communicate information clearly. Avoid cluttering the interface with excessive charts, text, or decorative elements. Use only relevant metrics and visuals that support the main objectives of the dashboard. A simple design helps users focus on the most critical insights without distraction.
2. **Logical Layout and Visual Hierarchy :** Organize information in a way that mirrors how users interpret data. Place the most important or high-level metrics at the top or top-left, as these areas draw the viewer's attention first. Group related information together, and maintain consistent alignment and spacing to enhance readability. Visual hierarchy ensures that users can easily distinguish key insights from supporting data.
3. **Consistency in Design :** Maintain consistency in colors, fonts, labels, and chart types across the dashboard. Consistent design reduces cognitive load, prevents confusion, and creates a cohesive visual experience. For example, using a consistent color for positive versus negative trends allows viewers to interpret data quickly without reorienting themselves for each chart.
4. **Appropriate Chart Selection :** Choose chart types that best represent the underlying data and the insights you want to convey. Use line charts for trends over time, bar charts for comparisons, pie charts for simple proportion analysis, and scatter plots for relationships between variables. Selecting the right chart type enhances comprehension and prevents misinterpretation.
5. **Emphasis on Key Metrics :** Highlight the most important metrics using size, color, or placement, so they stand out. KPIs (Key Performance Indicators) should be prominent and easily distinguishable from secondary data. This ensures that decision-makers can quickly identify performance gaps or opportunities without searching through detailed data.
6. **Interactivity and Drill-Down Options :** Incorporate interactive elements such as filters, drop-down menus, and drill-down capabilities. Interactivity allows users to explore data at multiple levels, focus on specific segments, and derive insights tailored to their needs. It also helps maintain a clean dashboard layout by avoiding excessive static information.
7. **Context and Comparisons :** Provide context for metrics by including historical data, benchmarks, or targets. Users interpret values more accurately when they understand what is typical or expected. Comparative visuals such as trend lines or performance against targets help users grasp significance at a glance.
8. **Accessibility and Responsiveness :** Ensure the dashboard is accessible to all intended users, including considerations for color blindness, font size, and screen resolutions. A responsive design ensures the dashboard can be viewed and interacted with effectively across different devices such as desktops, tablets, and mobile phones.

By adhering to these dashboard design principles, visualizations become not just informative but actionable. Well-designed dashboards guide users effortlessly through

data, support timely decisions, and enhance overall data-driven performance. The fig 1.4.1 shows example of an Effective Business Dashboard.

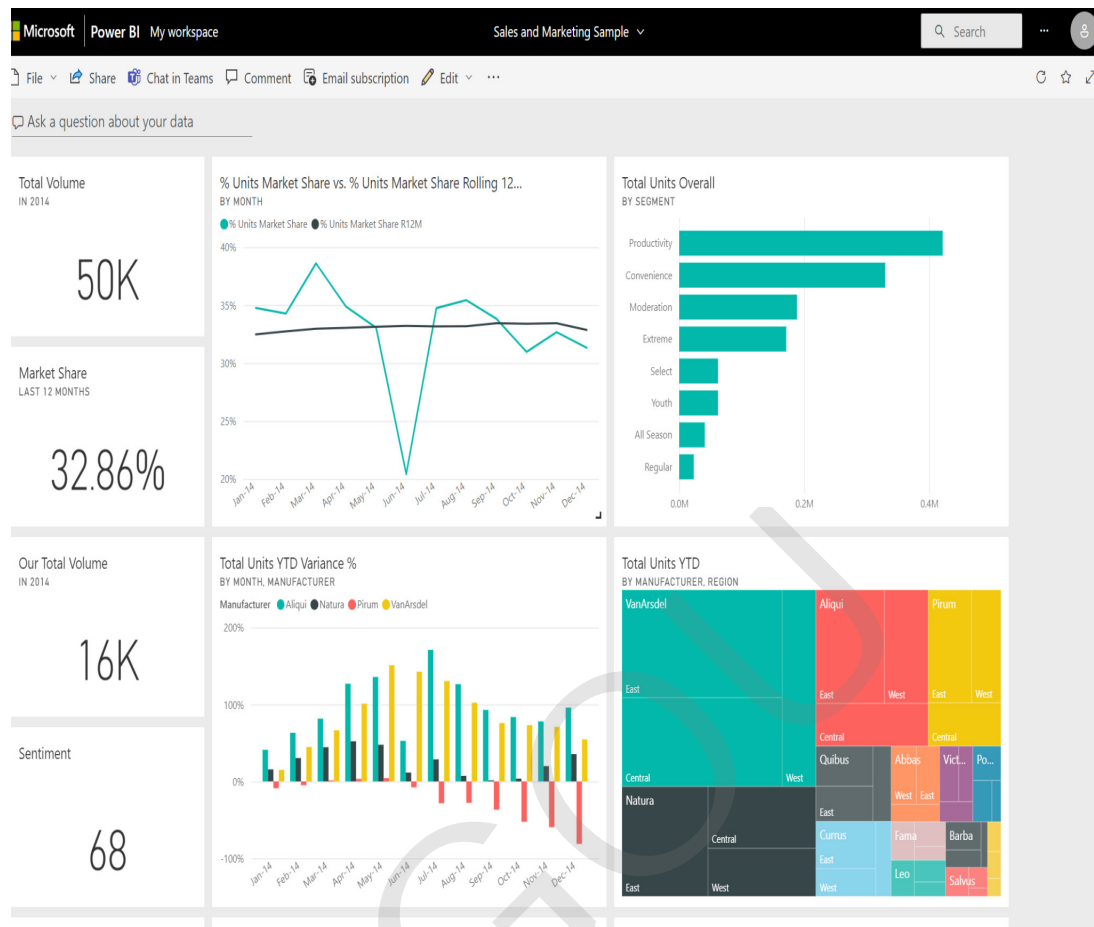


Fig 1.4.1 Example of an Effective Business Dashboard

1.4.3 Common Pitfalls in Visualization and How to Avoid Them

Even well-intentioned visualizations can fail if common design pitfalls are not avoided. Understanding these mistakes and applying corrective strategies ensures that data is communicated accurately, clearly, and effectively.

1. **Cluttered Dashboards and Overcrowded Visuals** : Including too many charts, tables, or decorative elements on a single dashboard overwhelms users and obscures key insights.

How to avoid : Prioritize essential metrics, remove unnecessary elements, and group related data logically. Use white space strategically to improve readability.

2. **Misleading Scales and Axes** : Manipulating axis scales, truncating data, or using inconsistent units can distort trends and comparisons, leading to incorrect interpretations.

How to avoid : Always use consistent scales, start axes at zero when appropriate, and clearly label units to provide accurate context for the data.

3. **Excessive Use of Colors and Visual Effects :** Too many colors, gradients, or 3D effects can distract from the data, create confusion, and reduce readability.

How to avoid : Use a limited, purposeful color palette to highlight key metrics. Avoid decorative effects that do not enhance comprehension.

4. **Ignoring Audience Needs :** A visualization designed without considering the audience's expertise, objectives, or decision-making needs can be irrelevant or confusing.

How to avoid : Conduct audience analysis to tailor visuals, level of detail, and type of charts to meet the expectations of different stakeholders.

5. **Overcomplicating Charts :** Complex visualizations with multiple dimensions, overlapping data points, or intricate legends may confuse users instead of informing them.

How to avoid : Keep charts simple, focus on the story you want to tell, and break complex data into smaller, digestible visuals if necessary.

6. **Lack of Context or Comparisons :** Presenting raw numbers without benchmarks, targets, or historical trends makes it difficult for viewers to understand significance.

How to avoid: Provide context through comparative visuals, trend lines, or performance against targets, enabling viewers to interpret the data meaningfully.

7. **Inconsistent or Ambiguous Labeling :** Vague titles, missing labels, or inconsistent terminology can lead to misinterpretation of visualized data.

How to avoid : Ensure that all axes, legends, and data points are clearly labeled and that terminology is consistent across the dashboard.

8. **Neglecting Accessibility :** Visualizations that ignore accessibility considerations, such as color blindness or mobile responsiveness, can exclude significant portions of the audience.

How to avoid : Use accessible color palettes, readable fonts, and responsive layouts to ensure inclusivity and usability.

By recognizing and avoiding these common pitfalls, designers can create visualizations that are accurate, clear, and actionable. Applying these best practices enhances decision-making, reduces cognitive load, and ensures that dashboards effectively communicate the intended insights. The fig 1.4.2 shows Common Data Visualization Pitfalls and their solutions.

Common Pitfalls in Data Visualization & How to Avoid Them

Communicating Data Accurately, Clearly, and Effectively

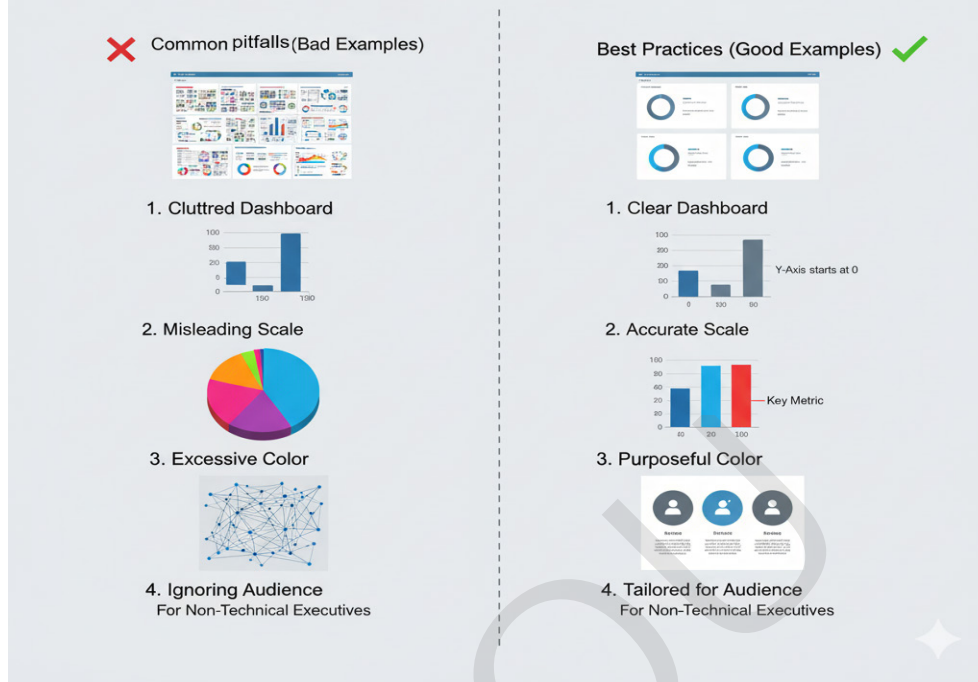


Fig 1.4.2 Common Data Visualization Pitfalls and their Solutions

1.4.4 Interactive Visualizations

Interactive visualizations allow users to actively engage with data rather than passively viewing static charts. By providing dynamic controls such as filters, drop-down menus, sliders, and drill-down options, interactive dashboards empower users to explore information at multiple levels, uncover patterns, and answer their specific questions.

The main advantage of interactive visualizations is that they accommodate diverse user needs. For example, an executive may use high-level filters to view overall performance trends, while an analyst can drill down into granular data for deeper investigation. This flexibility ensures that a single dashboard can serve multiple stakeholders with varying objectives, reducing the need for multiple separate reports.

Interactivity also enhances comprehension and retention. By allowing users to manipulate the data, focus on areas of interest, and visualize outcomes in real time, interactive dashboards make complex datasets more understandable. Techniques such as highlighting, conditional formatting, and linked visuals further guide the user's attention to key insights without overwhelming them.

However, interactivity should be applied thoughtfully. Overloading a dashboard with too many interactive elements can increase cognitive load, confuse users, or slow performance. The design should prioritize simplicity and relevance, ensuring that every interactive feature serves a clear purpose.

With the increasing availability of modern tools like Tableau, Power BI, and Python dashboards (Dash, Plotly), interactive visualizations have become a standard for data-driven decision making. They provide a richer user experience, foster exploration, and support timely insights that static visuals cannot deliver.

1.4.5 Current and Emerging Trends in Visualization

The field of data visualization is continuously evolving, driven by technological advances, increasing data volumes, and changing user expectations. Staying updated with current and emerging trends is essential for creating visuals that are not only relevant but also impactful and engaging.

1. **Real-Time Analytics and Dashboards :** Organizations increasingly require dashboards that display data in real time, enabling immediate decision-making and rapid response to changing conditions. Real-time visualizations are particularly valuable in operations, finance, marketing, and logistics, where timely insights can drive performance and mitigate risks.
2. **Storytelling Dashboards :** Integrating narrative elements into dashboards helps transform raw data into meaningful stories. Storytelling dashboards guide users through insights, highlight critical trends, and provide context that aids interpretation. These dashboards combine visualization best practices with narrative structure to create engaging and actionable experiences.
3. **AI-Assisted Visualization :** Artificial intelligence and machine learning are being incorporated into visualization tools to automate chart recommendations, detect anomalies, and highlight patterns that may not be immediately obvious. AI-assisted visualizations reduce manual effort, enhance accuracy, and support deeper insights for both technical and non-technical users.
4. **Interactive and Immersive Experiences :** Beyond basic interactivity, modern visualization tools now offer immersive experiences such as virtual reality (VR), augmented reality (AR), and 3D interactive visualizations. These technologies allow users to explore complex datasets in more intuitive and engaging ways, providing new perspectives on multidimensional information.
5. **Mobile and Responsive Visualizations :** As users increasingly access dashboards on mobile devices, responsive design has become essential. Dashboards that adapt seamlessly to different screen sizes and devices ensure accessibility, usability, and consistent experience for all users, regardless of platform.
6. **Data Democratization and Self-Service Tools :** Emerging trends emphasize empowering all users to analyze data independently. Self-service BI platforms enable employees across departments to generate visualizations, explore datasets, and gain insights without relying heavily on IT or data teams. This democratization encourages data-driven culture and faster decision-making.
7. **Integration of Multimedia and Contextual Data :** Incorporating images, videos, maps, or external contextual information into dashboards enhances storytelling and helps users understand complex phenomena. For example,

combining geospatial data with traditional charts allows for richer insights in fields such as logistics, marketing, or urban planning.

By adopting these trends, organizations can design visualizations that are not only functional and accurate but also engaging, intuitive, and forward-looking. Keeping pace with technological advancements ensures that visualizations remain effective communication tools in an increasingly data-driven world.

Recap

- ◆ Effective dashboards and visualizations communicate insights clearly, accurately, and efficiently.
- ◆ Dashboard Design Principles include clarity, simplicity, logical layout, visual hierarchy, appropriate chart selection, consistency, interactivity, and accessibility.
- ◆ Common Pitfalls in visualization include cluttered dashboards, misleading scales, excessive colors, ignoring audience needs, overcomplicating charts, lack of context, inconsistent labeling, and poor accessibility.
- ◆ Interactive Visualizations empower users to explore data dynamically, filter views, and drill down for detailed insights.
- ◆ Emerging Trends include real time analytics, storytelling dashboards, AI-assisted visualizations, immersive experiences (VR/AR), responsive dashboards for mobile devices, self service BI tools, and integration of multimedia/contextual data.
- ◆ Applying best practices and leveraging trends ensures that dashboards and visualizations are actionable, engaging, and aligned with user goals.

Objective Questions

1. What is the main focus of visualization best practices?
2. Which tool provides a consolidated view of key metrics?
3. Which element helps users explore data dynamically?
4. Using too many colors in a chart results in what?
5. Which type of visualization helps users drill down into data?
6. What should be maintained throughout a dashboard for better readability?
7. What is distorted when misleading scales are used?

8. Which trend involves showing data updates instantly?
9. What type of dashboards tell stories through visuals?
10. What technology is used to automate chart recommendations?
11. Which feature makes dashboards usable on mobile devices?
12. Using 3D effects in charts often reduces what?
13. What is essential to include for better understanding—context or raw data?
14. What helps users distinguish important metrics quickly?

Answers

1. Clarity
2. Dashboard
3. Interactivity
4. Confusion
5. Interactive
6. Consistency
7. Interpretation
8. Real-time
9. Storytelling
10. AI
11. Responsive
12. Readability
13. Context
14. Hierarchy

Assignments

1. Explain the importance of following visualization best practices when designing dashboards. Give any two examples.
2. Describe the principles of dashboard design and explain why clarity and simplicity are essential for effective data communication.
3. Identify any three common pitfalls in data visualization and explain how each of them can be avoided with simple examples.
4. Write a short note on interactive visualizations. Explain how features like filters and drill down options help users explore data.
5. Discuss any two emerging trends in data visualization and explain how they are transforming modern dashboards and reporting.

Reference

- ♦ <https://powerbi.microsoft.com/en-us/blog/>
- ♦ <https://www.interaction-design.org/literature/topics/data-visualization>

Suggested Reading

1. Few, Stephen. *Information Dashboard Design: The Effective Visual Communication of Data*. Analytics Press, 2013.
2. Knaflic, Cole Nussbaumer. *Storytelling with Data: Let's Practice!* Wiley, 2020.
3. Yau, Nathan. *Data Points: Visualization That Means Something*. Wiley, 2013.

BLOCK-2

Data Visualisation Using Power BI

UNIT 1

Data Preparation and Integration in Power BI

Learning Outcomes

At the end of this unit, the learner will be able to :

- ◆ identify Power BI Desktop, Service, and Mobile, and understand their main uses
- ◆ connect Power BI to different data sources and explain Import, DirectQuery, and Live Connection modes
- ◆ load data efficiently and understand how Power BI stores and manages data
- ◆ clean and prepare data in Power Query Editor
- ◆ apply common data transformations like merge, append, pivot/unpivot, and data type changes

Prerequisite

In a data-driven environment, organizations rely on accurate, timely, and well-structured information to make informed decisions. Whether it is sales performance, customer behavior, financial metrics, or operational efficiency, every business process generates data that must be collected, connected, and transformed before it can be meaningfully analyzed. Yet, raw datasets often scattered across Excel files, databases, cloud platforms, and third-party applications offer little value until they are properly integrated and prepared.

Modern analytics tools like Power BI bridge this gap by enabling users to bring data from diverse sources into a unified platform. To fully benefit from these capabilities, learners should have a foundational understanding of where data comes from, how it is stored, and why it requires cleaning before analysis. Familiarity with spreadsheets, basic data types, and common issues such as duplicates, missing values, or inconsistent formats helps learners better appreciate the importance of data preparation.

Understanding these basics sets the stage for exploring Power BI versions, connecting to various data sources, loading data efficiently, and applying essential cleaning and transformation techniques. This foundational knowledge ensures learners are ready to build reliable, accurate, and meaningful analytical models as they progress through the unit.

Keywords

Power BI Desktop, Data Connectors, Power Query Editor, Data Cleaning, Data Types, Import Mode, DirectQuery, Data Transformation, Duplicate Removal, Data Profiling.

Discussion

In today's data-driven world, organizations rely heavily on analytical tools to transform raw data into meaningful insights. Power BI is one of the most widely adopted business intelligence platforms designed to help businesses visualize their data, generate reports, and support informed decision making. Before learning how to prepare and clean data for Power BI, it is essential for learners to understand what Power BI is and how it compares with other visualization tools commonly used in the industry.

2.1.1 What is Power BI?

Power BI is a business analytics and data visualization tool developed by Microsoft. It allows users to connect to a wide range of data sources, transform and model data, build interactive reports, and share dashboards across an organization. Power BI provides a seamless environment for converting raw data into visually appealing and interactive insights that support strategic decision making.

Power BI consists of three main components:

- ◆ Power BI Desktop, a Windows-based application for creating reports and data models
- ◆ Power BI Service, a cloud based platform for sharing, publishing, and collaborating on reports
- ◆ Power BI Mobile, an app for viewing dashboards on smartphones and tablets

Power BI is widely recognized for its user friendly interface, flexibility, and strong integration with Microsoft products such as Excel, SQL Server, Azure, and Microsoft 365. Its ability to handle large datasets and support advanced analytics through DAX (Data Analysis Expressions) makes it a preferred tool for data analysts, business intelligence professionals, and decision makers.

2.1.2 Why Power BI is important

Modern organizations handle massive volumes of data from multiple sources such as sales systems, financial databases, marketing platforms, operational records, and more. Power BI simplifies the process of connecting to this data, shaping it into a structured format, and creating intuitive visualizations. Its automation features, cloud based sharing, and scheduled refresh options make it an essential tool for maintaining up-to-date business insights.

Furthermore, Power BI empowers both technical and non technical users to explore data independently. With features like natural language queries (“Q&A”) and one-click insights, business users can analyze data without requiring programming knowledge.

2.1.3 Comparison of Power BI with other Data Visualization Tools

Although several tools exist for data visualization and business intelligence, Power BI has gained significant popularity due to its affordability, scalability, and strong ecosystem. The following comparison highlights how Power BI differs from other industry tools:

2.1.3.1 Power BI vs Tableau

Tableau is known for its advanced visualization capabilities and ability to handle large datasets efficiently. It provides deep customization for visual storytelling and is highly suitable for enterprise level BI.

Power BI, while offering robust visualization features, is generally more user friendly and integrates better with Microsoft products. It is more cost effective and easier for beginners, especially those familiar with Excel.

Key differences:

- ◆ Tableau is preferred for complex and highly customized visuals.
- ◆ Power BI offers better integration with familiar tools like Excel and Azure.
- ◆ Power BI is more budget friendly for organizations of all sizes.

2.1.3.2 Power BI vs Qlik Sense

Qlik Sense uses an associative data model that allows users to explore data relationships more dynamically. It supports advanced analytics, but learning Qlik requires more technical understanding.

Power BI follows a more straightforward data model (star/snowflake schemas) and provides a more intuitive interface.

Key differences:

- ◆ Qlik Sense is strong in associative data exploration.
- ◆ Power BI offers a smoother learning curve and a richer ecosystem.

- ◆ Qlik is suitable for complex enterprise BI, while Power BI fits both small and large organizations.

2.1.3.3 Power BI vs Excel

Excel is a widely used spreadsheet application with powerful formulas, pivot tables, and charts. However, it is not designed for automated refresh, large scale dashboards, or collaborative online reporting.

Power BI extends Excel's capabilities with interactive reports, scalable dashboards, and cloud publishing.

Key differences:

- ◆ Excel is ideal for ad-hoc analysis; Power BI is better for automated, interactive dashboards.
- ◆ Power BI handles larger datasets and supports scheduled refresh.
- ◆ Excel's visuals are limited compared to Power BI's interactive capabilities.

2.1.3.4 Power BI vs Google Data Studio (Looker Studio)

Looker Studio (formerly Google Data Studio) is a free visualization tool integrated with Google Analytics and Google Ads. It is effective for digital marketing dashboards.

Power BI, however, offers more advanced modelling capabilities, better data transformation, and broader connectivity.

Key differences:

- ◆ Looker Studio is excellent for marketing and web analytics.
- ◆ Power BI is better for enterprise reporting, transformations, and security.
- ◆ Power BI supports complex DAX calculations; Looker Studio does not.

2.1.4 Introduction to Data Preparation in Power BI

Power BI is one of the most widely used Business Intelligence (BI) tools in the world, known for its ability to connect, clean, prepare, transform, and visualize data seamlessly. Data preparation forms the foundation of every analytical process because the accuracy of insights and visualizations entirely depends on the quality and structure of input data.

In any BI or data analytics project, it is estimated that nearly 70–80% of the total project time is spent in data cleaning, transformation, and preparation, rather than visualization or modeling. Hence, mastering the data preparation process in Power BI is essential for every data analyst or BI professional.

2.1.4.1 What is Data Preparation?

Data Preparation refers to the process of converting raw, unstructured, or inconsistent data into a clean and structured format suitable for analysis and visualization. It involves several tasks such as:



- ◆ Extracting data from multiple sources (files, databases, APIs, cloud, etc.)
- ◆ Transforming the data into a consistent and standardized format
- ◆ Cleaning the data by removing errors, duplicates, and irrelevant records
- ◆ Loading the prepared data into Power BI for modeling and visualization

This entire process is often described as ETL – Extract, Transform, and Load.

Power BI offers a complete ETL system within its Power Query Editor, eliminating the need to rely on external tools or write complex SQL or Python scripts for data cleaning.

2.1.4.2 Why Data Preparation is Important

Real-world datasets are rarely clean or analysis-ready. They often come from multiple departments, systems, or manual entry sources resulting in inconsistencies, missing values, and formatting errors.

Some common problems found in raw data include:

1. **Missing or Null Values** – e.g., missing customer names or product prices that can lead to incorrect totals or averages.
2. **Inconsistent Date Formats** – for instance, one dataset may use “DD-MM-YYYY” while another uses “MM/DD/YYYY”.
3. **Different Naming Conventions** – variations like “South Region”, “south”, or “SOUTH” can lead to duplicate categories.
4. **Duplicate Records** – repeated entries due to data import or entry errors.
5. **Incorrect Data Types** – numerical fields stored as text or date fields not recognized as valid dates.
6. **Outliers and Invalid Values** – e.g., negative quantities or unrealistic sales amounts.

If these issues are not detected and corrected before visualization, they can distort results and lead to wrong business decisions.

Real-World Impact

According to a Gartner research report, poor data quality costs enterprises an average of USD 12.9 million annually in rework, missed opportunities, and poor decision-making. This highlights that data quality management is not just a technical requirement but a business necessity.

2.1.5 Power BI’s Approach to Data Preparation

Power BI simplifies and automates much of the data preparation process through its Power Query Editor.

This interactive interface allows users to:

- ◆ Connect to 160+ data sources
- ◆ Apply cleaning and transformation operations with simple point-and-click actions
- ◆ Automatically record each step as a script in M Language, ensuring repeatability and consistency

Each action such as “Remove Duplicates”, “Split Column”, or “Change Data Type” is stored as a transformation step, and these steps can be refreshed automatically when new data is loaded.

Example Scenario

Consider a retail company that receives monthly sales data from three regional branches. Each branch maintains its files in slightly different formats one in Excel, another in CSV, and the third through a database export. Before analysis, the data analyst must:

- ◆ Standardize date formats
- ◆ Align product categories and region names
- ◆ Merge all three datasets into one consolidated table
- ◆ Remove duplicates and ensure sales figures are numeric

This cleaning and preparation process happens within Power BI’s Power Query Editor before building the final sales dashboard.

2.1.5.1 Key Benefits of Data Preparation in Power BI

Table 2.1.1 Key Benefits of Data Preparation in Power BI

Benefit	Description
Accuracy	Ensures reports and visuals are based on reliable and consistent data.
Automation	Transformations are stored as reusable steps that can refresh automatically.
Efficiency	Eliminates repetitive manual cleaning using Excel formulas or scripts.
Scalability	Handles large datasets from multiple sources effectively.
Transparency	Each cleaning step is documented and can be reviewed later for auditing.

Data preparation is the backbone of Power BI analytics. Without a well cleaned and structured dataset, even the most advanced visualizations or DAX calculations can produce misleading insights. By leveraging the Power Query Editor, analysts can efficiently implement ETL workflows, ensuring the data entering the model is accurate, standardized, and trustworthy.

2.1.6 Power BI Versions

Table 2.1.2 Power BI Versions

Version	Usage	Audience
Power BI Desktop (Free)	Development of reports	Students, Data analysts
Power BI Pro (Paid)	Sharing, Collaboration	Business Teams
Power BI Premium	Enterprise level workloads, large data	Enterprise

Power BI Desktop is used 90% of the time by data analysts for report creation and data preparation.

2.1.7 Data Connectors in Power BI

Power BI supports more than 160+ connectors, including:

- ◆ Excel, CSV, Text
- ◆ SQL Server, MySQL, Oracle
- ◆ Web connectors (API)
- ◆ SharePoint folders
- ◆ Azure SQL DB
- ◆ Snowflake
- ◆ Google Big Query

Power BI makes connecting to data as simple as selecting a connector, entering authentication details and then selecting tables (Fig 2.1.1 Power BI Data Source Ecosystem).

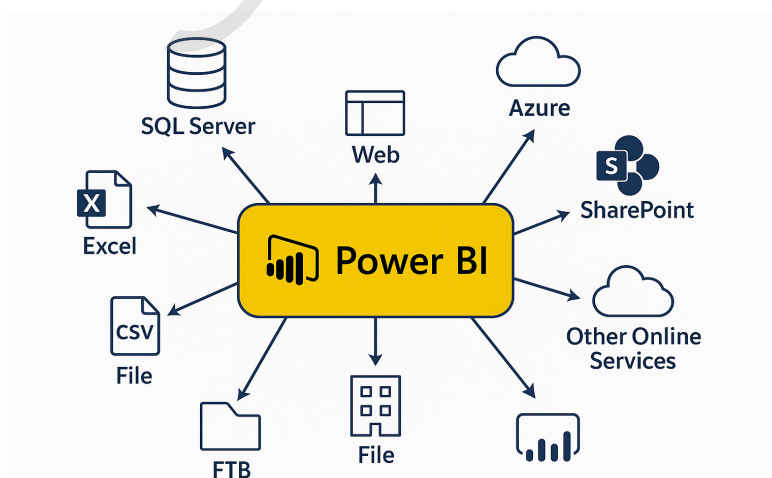


Fig 2.1.1 Power BI Data Source Ecosystem

2.1.8 Data Loading Methods

Table 2.1.3 Data Loading Methods

Method	Description	When to use
Import	Stores entire data inside .pbix	For small to medium datasets
DirectQuery	Connects in real-time, no import	Live Streaming, Real Time dashboards
Live Connection	Analysis Services connection	Enterprise cubes

2.1.9 Data Cleaning using Power Query Editor

The Power Query Editor is a GUI based data cleaning and transformation interface.

Common cleaning tasks include:

- ◆ Remove Duplicates
- ◆ Replace Values
- ◆ Split Columns
- ◆ Merge Columns
- ◆ Trim & Clean functions
- ◆ Remove Null rows
- ◆ Change Data Types

2.1.10 Data Transformations

Transformation is converting raw data to useful data. Some major transformations:

- ◆ Group By
- ◆ Pivot / Unpivot
- ◆ Conditional Columns
- ◆ Creating Custom Columns (M language)
- ◆ Filtering rows
- ◆ Date / Time transformations

Transformation steps are recorded automatically and are applied in sequence each time the dataset refreshes.

Every action done in Power Query Editor is converted into M Language Script automatically in the background.

2.1.11 Data Integration

Power BI can combine data from multiple sources using:

- ◆ Append Queries (Union)
- ◆ Merge Queries (Join)

This enables combining data from various departments like sales + HR + finance into a single unified model.

Common Challenges in Data Preparation

- ◆ Data coming from multiple systems
- ◆ Missing business definitions
- ◆ Date dimension issues
- ◆ Duplicate data entries
- ◆ Human manual excel errors

Power BI reduces manual coding efforts drastically.

Recap

- ◆ Power BI transforms raw data into meaningful insights using interactive dashboards and reports.
- ◆ It has a user-friendly interface suitable for beginners and advanced analysts.
- ◆ Strong integration with Microsoft tools (Excel, Azure, Office 365).
- ◆ Connects easily to multiple data sources using built-in connectors.
- ◆ Allows combining and analyzing data without advanced technical skills.
- ◆ More affordable and easier to learn than tools like Tableau and Qlik Sense.
- ◆ Offers strong data modelling and enterprise-level deployment options.
- ◆ Tight integration with Excel enhances usability in organizations.
- ◆ Supports Import mode for fast in-memory performance.
- ◆ Supports DirectQuery for real-time data access.
- ◆ Supports Live Connection for direct connectivity to enterprise models.
- ◆ Power Query Editor is the central tool for data cleaning and shaping.
- ◆ Provides transformations like merge, append, pivot, and unpivot.

- ◆ Supports data type conversions, grouping, splitting, and aggregation.
- ◆ Data profiling helps identify missing values, errors, and inconsistencies.
- ◆ Ensures datasets are clean and reliable before modelling or visualization.
- ◆ Clean, well-prepared data leads to accurate and meaningful reports.

Objective Questions

1. Power BI Desktop is primarily used for:
2. Which feature in Power Query Editor helps identify missing values, errors, and inconsistencies?
3. Import Mode in Power BI is best suited for:
4. Which connector would you use to bring data from an Excel file into Power BI?
5. Which transformation should be used to combine two tables with the same columns but different rows?
6. Changing a column from “Text” to “Date” is part of which activity in Power BI?
7. Which view in Power BI Desktop allows users to see loaded tables in tabular form?
8. DirectQuery Mode is used when:
9. The Unpivot Columns transformation is useful when:
10. Which of the following is not a Power BI data connectivity option?
11. Power BI is considered better than Excel for dashboards because:
12. Tableau is often preferred over Power BI when:
13. Which Power BI feature removes duplicate rows automatically?
14. Looker Studio (Google Data Studio) is best suited for:
15. When preparing data for analysis, the first recommended step is to:

Answers

1. Creating reports and performing data transformations
2. Data Profiling
3. In-memory analysis with faster performance
4. Excel Connector
5. Append Queries
6. Data Cleaning
7. Data View
8. Real-time access to the underlying source is required
9. Convert columns into rows
10. HTML Export
11. It allows interactive visuals and automated refresh
12. Users require deeper visualization customization
13. Remove Duplicates
14. Marketing dashboards integrated with Google platforms
15. Clean and transform the data

Assignments

1. Explain the steps involved in loading data from an Excel file into Power BI. Describe how Power BI identifies column data types and how the user can modify them.
2. Compare Import Mode and DirectQuery Mode in Power BI. Discuss at least four major differences and provide examples of situations where each mode is preferred.
3. Using the Power Query Editor, perform data cleaning on a sample dataset by removing duplicates, handling missing values, and renaming fields. Document each transformation step clearly.

4. Describe the process of merging two tables in Power Query Editor. Provide an example of a real-time business scenario where merging is necessary.
5. Data profiling plays an important role in identifying issues in a dataset. Explain the types of profiling information available in Power BI and how each helps in preparing clean data.

Reference

- ◆ <https://learn.microsoft.com/power-bi/fundamentals/>
- ◆ <https://learn.microsoft.com/power-query/>
- ◆ <https://www.sqlbi.com/>

Suggested Reading

1. Microsoft Documentation – Power BI Basics:
<https://learn.microsoft.com/power-bi/fundamentals>
2. Power Query Overview – Microsoft Learn:
<https://learn.microsoft.com/power-query>
3. Book: Introducing Microsoft Power BI – Alberto Ferrari & Marco Russo
4. Tutorial: Power BI Data Loading and Transformation – SQLBI Articles <https://www.sqlbi.com>

UNIT 2

Data Modelling and Dax Fundamentals

Learning Outcomes

At the end of this unit, the learner will able to :

- ◆ explain the role of Power Query in data preparation and perform essential data cleaning and transformation tasks
- ◆ build structured data models by establishing relationships between tables and applying best practices for model design
- ◆ differentiate between calculated columns and measures, and apply them appropriately within a data model
- ◆ construct and evaluate DAX expressions to perform advanced calculations, aggregations, and analytical logic
- ◆ optimize data models and DAX computations to improve report performance and accuracy
- ◆ apply Power Query, data modelling principles, and DAX functions collaboratively to create meaningful insights and analytical outputs in Power BI

Prerequisite

Modern organizations rely on data-driven decision-making, where information from multiple sources, sales records, customer interactions, financial systems, and operational logs must be combined, cleaned, and analyzed to uncover meaningful insights. However, raw datasets often come with inconsistencies such as missing values, duplicate entries, or mismatched formats, making analysis difficult and unreliable. Before insights can be generated, this scattered information must be systematically prepared and structured.

In a typical business scenario, a sales manager might receive data from different branches, each using its own format. Without proper transformation and modelling, combining these files could lead to incorrect totals, broken relationships, or misleading

reports. Power BI solves these challenges through its powerful data preparation layer (Power Query), robust data modelling capabilities, and calculation engine powered by DAX. These components work together to convert raw, unorganized data into a reliable analytical model.

To successfully engage with this unit, learners should have a basic understanding of spreadsheets, familiarity with tabular data, and awareness of common data issues such as duplicates or missing entries. A foundational knowledge of simple formulas or logical expressions is helpful, as it prepares learners to explore Power Query transformations, build structured data models, and write DAX expressions for deeper analysis. This background ensures that learners are ready to move from raw data handling to creating meaningful, accurate, and dynamic insights in Power BI.

Keywords

Data Modelling, Relationships, Cardinality, Star Schema, Snowflake Schema, Fact Table, Dimension Table, Calculated Column, DAX, Measures, Row Context, Filter Context.

Discussion

2.2.1 Introduction to Data Modeling and DAX Fundamentals

After preparing and cleaning the data in Power BI using Power Query Editor, the next major step in the analytics process is Data Modeling. This step is where the cleaned data is structured, related, and enriched so that it can support efficient analysis, calculations, and visualization.

In simpler terms, data modeling is the “logical brain” of Power BI, connecting different data tables and defining how they relate to each other through relationships, hierarchies, and business logic.

2.2.2 What is Data Modeling?

Data Modeling is the process of organizing, structuring, and connecting multiple tables in a way that mirrors real-world business processes. It defines how data tables interact and how Power BI understands their relationships.

Power BI uses a relational data model, meaning it connects multiple tables through keys (unique identifiers like CustomerID or ProductID).

The goal of data modeling is to:

- ◆ Reduce redundancy (no repeated data)
- ◆ Improve query performance
- ◆ Make calculations using DAX more accurate and efficient
- ◆ Create a scalable structure that can handle growing data volumes

2.2.2.1 Why is Data Modeling Important?

A properly designed data model ensures:

- ◆ Data Consistency: All visuals and calculations use the same underlying relationships.
- ◆ Performance Optimization: Reports load faster because Power BI can query data efficiently.
- ◆ Reusability: Measures and relationships can be reused across multiple visuals and reports.
- ◆ Ease of Maintenance: Changes in one table automatically propagate correctly through relationships.

Without proper data modeling, even a clean dataset can lead to:

- ◆ Incorrect results in measures and KPIs
- ◆ Duplicated or missing values in reports
- ◆ Slow report performance
- ◆ Confusing relationships or filter propagation issues

Real-World Analogy

Think of your Power BI model like a map of a shopping mall:

- ◆ Each shop is like a table.
- ◆ Corridors connecting the shops are like relationships between tables.
- ◆ Store categories (Food, Fashion, Electronics) represent dimension attributes.
- ◆ Sales receipts or transactions represent fact tables.

Without a clear map (data model), you cannot navigate from one shop (table) to another effectively.

2.2.2.2 Power BI Model Types

Power BI supports different model architectures depending on the dataset and performance needs:

Star Schema vs Snowflake Schema

In Power BI and most data modeling tools, you will often hear two terms , Star Schema and Snowflake Schema. Both are ways of organizing data tables (Fact & Dimension tables) so that Power BI can understand relationships between them efficiently.

1. What is Star Schema?

A Star Schema is the simplest and most popular data model used in Power BI.

It looks like a star — a central Fact table surrounded by multiple Dimension tables.

Each Dimension table connects directly to the Fact table using a Primary Key–Foreign Key relationship.

Example : A Sales Fact Table connects to Customer, Product, Date, and Store Dimension tables.

Key Idea

All dimensions are directly linked to the Fact table — no sub-tables or hierarchies between dimensions.

2. What is a Snowflake Schema?

A Snowflake Schema is an extended version of the Star Schema, where one or more Dimension tables are normalized (split into sub-tables).

It looks like a snowflake shape — more complex and branched.

Example: In a Snowflake model, the Product Dimension might be further split into Category and Subcategory tables.

Key Idea

Some dimension tables have their own related sub-dimensions, creating multiple levels of relationships.

Table 2.2.1 Comparison Table

Feature	Star Schema	Snowflake Schema
Structure	Fact table in the center, dimensions around it	Dimensions are normalized into sub-tables
Complexity	Simple, flat structure	Complex, hierarchical structure
Performance	Fast — fewer joins needed	Slightly slower — more joins required
Storage	Uses more storage (redundancy)	Uses less storage (eliminates redundancy)

Ease of Use in Power BI	Easy to design and visualize	Harder to manage relationships
Best For	Power BI, Tableau, Excel reports	Data warehouses needing high normalization
Recommended For Beginners	* Yes	No (too complex)

Real-World Analogy

Think of your data as a library system.

- ◆ Star Schema = One large catalog that directly lists books, authors, and genres together simple and quick to search.
- ◆ Snowflake Schema = A detailed catalog where “Genres” and “Authors” are stored in separate sub-sections more organized, but requires more steps to find what you need.

Example from Retail Business

Table 2.2.2 Star Schema Example

Table	Description
Sales	Fact table with SalesAmount, Quantity, Profit
Customer	Name, Region, Segment
Product	ProductID, ProductName, Category, Subcategory
Date	Date, Month, Year

Table 2.2.3 Snowflake Schema Example

Table	Description
Sales	Fact table (same as above)
Customer	CustomerID, Name, RegionID
Region	RegionID, RegionName
Product	ProductID, SubCategoryID
SubCategory	SubCategoryID, CategoryID
Category	CategoryID, CategoryName
Date	DateID, Month, Year

Here, Product and Customer dimensions are split into multiple related sub-tables, forming a “snowflake-like” shape.

Which Should You Use in Power BI?

Use of Star Schema

- ◆ Power BI, Tableau, Excel, or visualization-focused projects
- ◆ Fast performance and easy maintenance
- ◆ Business users and students learning BI

Avoid Snowflake Schema unless:

- ◆ You're working on a large enterprise data warehouse
- ◆ Storage optimization or normalization is a strict requirement
- ◆ Star Schema = Simplicity, Speed, Clarity
- ◆ Snowflake Schema = Complexity, Normalization, Slightly slower queries
- ◆ Star Schema is the recommended approach for Power BI and most self-service BI tools.
- ◆ Snowflake Schema may be useful when dealing with very large and complex datasets, but not ideal for beginners.
- ◆ Always start with a Star Schema, and only move to a Snowflake design if performance or data integrity issues demand it.

Understanding the structural difference between these two models is fundamental for efficient data loading in Power BI (See Fig 2.2.1 Star schema vs Snowflake Schema).

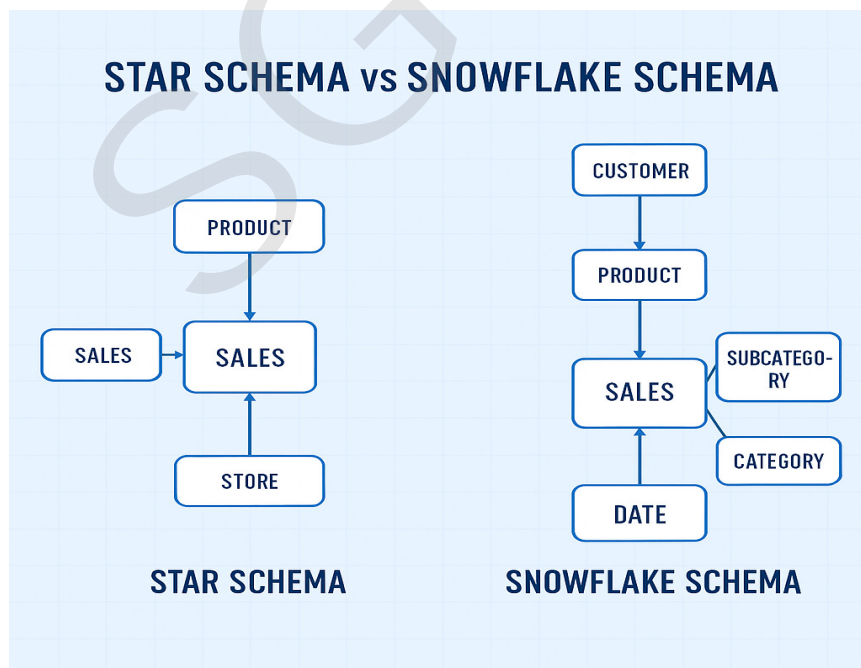


Fig 2.2.1 Star schema vs Snowflake Schema

Table 2.2.4 Components of a Data Model

Component	Description
Fact Tables	Contain measurable, quantitative data such as sales, profit, or quantity.
Dimension Tables	Contain descriptive information such as customer details, product categories, or regions.
Relationships	Logical connections between fact and dimension tables using key fields.
Calculated Columns	Custom columns created using DAX for row-wise computations.
Measures	Aggregated values (like SUM, COUNT, AVERAGE) computed using DAX dynamically.

Example: Retail Business Model

Let's consider a Retail Sales Dataset.

Table 2.2.5 Example: Retail Business Model

Table	Example Fields	Type
Sales	OrderID, ProductID, CustomerID, Date, SalesAmount, Profit	Fact
Products	ProductID, ProductName, Category, SubCategory	Dimension
Customers	CustomerID, CustomerName, Region, Segment	Dimension
Calendar	Date, Month, Quarter, Year	Dimension

In Power BI:

- ◆ The Sales table sits in the center (Fact table).
- ◆ All other tables (Products, Customers, Calendar) connect to it via relationships.

This structure forms a Star Schema, which makes DAX calculations and visuals easier, faster, and logically correct.

2.2.2.3 Why NOT Flat Tables / Single Excel Approach?

- ◆ Too many repeated customer names
- ◆ Too many repeated product details in every row
- ◆ Performance degradation (slow visuals + heavy file)
- ◆ Hard to calculate DAX across entities

Hence Star Schema is the BI standard (Fig 2.2.2 Star Schema Example for PowerBI).

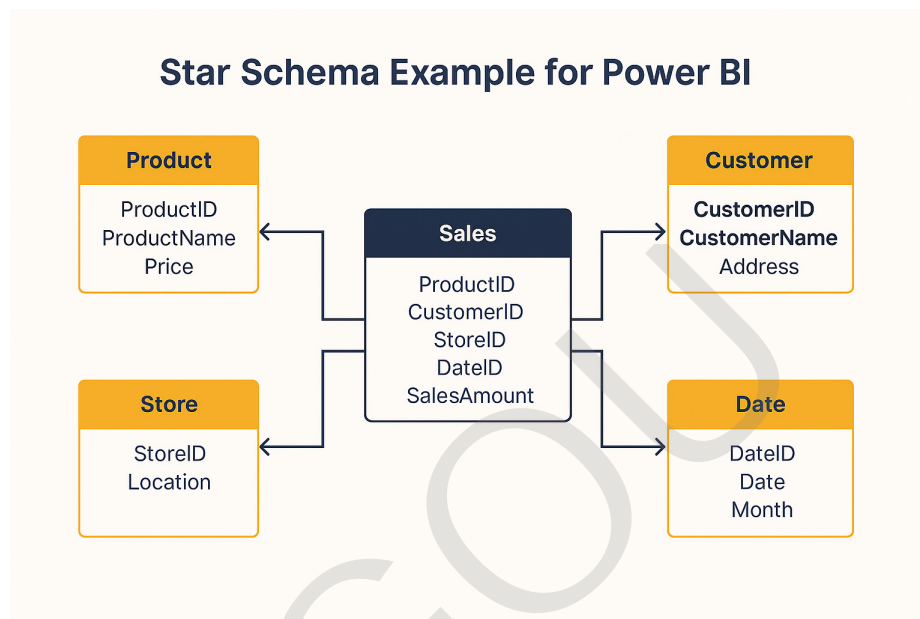


Fig 2.2.2 Star Schema Example for PowerBI

2.2.3 Introduction to DAX (Data Analysis Expressions)

After the model is created, we use DAX (Data Analysis Expressions) to build business logic and analytical calculations on top of it. DAX allows you to:

- ◆ Define Calculated Columns (row-level computations)
- ◆ Create Measures (aggregated calculations)
- ◆ Apply Filters dynamically in calculations
- ◆ Perform Time Intelligence (YTD, MTD, YOY comparisons)

DAX is:

- ◆ Formula Language (similar to Excel formulas)
- ◆ Functional Language, not procedural
- ◆ Case-insensitive
- ◆ Used both for Calculated Columns and Measures

Example :

Total Sales = SUM(Sales[SalesAmount])

Profit Margin = DIVIDE(SUM(Sales[Profit]), SUM(Sales[SalesAmount]))

These expressions are stored in the model and can be reused across multiple visuals.

Table 2.2.6 Importance of DAX in Data Modeling

Advantage	Description
Dynamic Calculations	Measures update automatically when filters are applied.
Custom Business Logic	Allows implementing logic specific to business rules.
Interactivity	Enables dynamic visuals that respond to slicers.
Reusability	A single DAX measure can be used in multiple reports.

Common Mistakes to Avoid in Modeling

1. Using a single flat table instead of building a relational model.
2. Connecting tables without defining proper key columns.
3. Ignoring data types (text vs numeric mismatches).
4. Creating unnecessary bidirectional relationships.
5. Overusing calculated columns instead of measures.

2.2.3.1 Role of Power Query Output in Modeling

Unit 1 transformations produce clean structured tables. Unit 2 starts after the data is loaded inside Power BI Model View.

- ◆ Data Preparation → Power Query
- ◆ Data Modeling → Model View
- ◆ Calculation Logic → DAX

All three together form BI pipeline.

A poor model design leads to incorrect DAX results even if your formulas are right.

Data Modeling is the core engine of Power BI. It connects the technical (data preparation) and analytical (visualization) stages together. When combined with DAX, it enables Power BI to perform complex calculations, business metrics, and real-time data analysis efficiently.

Mastering Data Modeling and DAX Fundamentals empowers us to:

- ◆ Create logical and efficient BI models
- ◆ Build accurate KPI dashboards
- ◆ Answer complex business questions dynamically
- ◆ Lay a strong foundation for advanced DAX and data visualization topics in later units

2.2.3.2 Relationships in Power BI

Once fact and dimension tables are identified, we need to create proper relationships between them. A relationship is a connection established between two tables based on a matching column.

Example:

Sales(ProductID) → Products(ProductID)

Sales(CustomerID) → Customers(CustomerID)

Sales(Date) → Calendar(Date)

When relationships are defined correctly:

- ◆ Filters propagate correctly
- ◆ Slicers work correctly
- ◆ Measures give correct results

Without proper relationships, DAX will give wrong numbers even if the formula is correct.

Table 2.2.7 Types of Relationships (Cardinality)

Cardinality	Meaning	Example
One-to-Many (1:*)	One unique value in dimension connects to many rows in fact	One Product → Many sales transactions
One-to-One (1:1)	Both tables have single unique matching records	Rare in retail
Many-to-Many (:)	Both tables have repeated values	Avoid unless required

Table 2.2.8 Cross Filter Direction

Cross Filter Direction	Meaning	Usage
Single Direction	Filter flows only one way (recommended)	Standard Star Schema
Both Direction	Filter flows both ways	Only when needed for specific DAX models

In star schema → ALWAYS prefer Single Direction filtering from Dimension → Fact.

Common Mistake

Beginners often set all relationships to “Both Direction” thinking it solves filtering problems, but it creates ambiguity, circular filtering & incorrect totals. Always first try Single Direction.

Active vs Inactive Relationships

Power BI allows multiple relationships between the same tables ... but only one can be active at a time.

Example: Sales Table can connect to Calendar table using:

- ◆ Order Date
- ◆ Ship Date

Only one date field can be an active relationship(solid line). The other one becomes inactive (dotted line).

Inactive relationships are still usable using DAX function: USERELATIONSHIP()

This is common in retail analysis for Shipping performance vs Order performance comparison.

Table 2.2.9 Relationship View Best Practices

Best Practice	Reason
Keep model as Star Schema	Widely accepted BI standard
Avoid Snowflake / Many-to-Many	Hard to control filtering
Make dimension tables low row count	Faster slicing
Always have a Calendar table	Time Intelligence depends on it
Use single direction filtering	Most predictable behavior

2.2.3.3 Calculated Columns vs Measures

When working with DAX in Power BI, you can create two types of calculations — Calculated Columns and Measures. Although both use the same DAX language, they are used in different contexts, have different evaluation behaviors, and serve different purposes in the data model. This is one of the most fundamental DAX concepts every Power BI learner must understand clearly.

a) What is a Calculated Column?

A Calculated Column is a new column added to an existing table. It performs calculations row by row, just like formulas in Excel.

Once created, the result of each row is stored in the model, increasing the model size.

Characteristics of Calculated Columns:

- ◆ Evaluated for every row in a table (row context)
- ◆ Physically stored in the dataset (increases memory usage)
- ◆ Recalculated only when the data refreshes
- ◆ Can be used like any other field in visuals, slicers, or filters

Example : Calculated Column

Let's say your Sales table has the following columns:

Table 2.2.10 Example: Calculated Column

Product	Quantity	Unit Price
Laptop	3	60,000
Mouse	10	500
Keyboard	5	1,000

Now, you want to create a column showing the Total Sale per row:

- ◆ Total Line Amount = Sales[Quantity] * Sales[Unit Price]

This creates a Calculated Column named Total Line Amount which stores:

Table 2.2.11 Example: Calculated Column

Product	Quantity	Unit Price	Total Line Amount
Laptop	3	60,000	180,000
Mouse	10	500	5,000
Keyboard	5	1,000	5,000

Each row is evaluated individually — this is called Row Context.

b) What is a Measure?

A Measure performs aggregated calculations on your data — such as SUM, AVERAGE, COUNT, or complex business KPIs.

Unlike Calculated Columns, Measures:

- ◆ Do not create a new column in the table
- ◆ Do not store results permanently
- ◆ Are calculated only when needed (based on filters in visuals)

Characteristics of Measures:

- ◆ Evaluated dynamically using Filter Context
- ◆ Calculated on the fly
- ◆ Do not increase data model size
- ◆ Cannot be used in slicers or filters (because they are not static fields)
- ◆ Used mainly in visuals, cards, KPIs, and reports

Example : Measure

If you want to calculate Total Sales for the entire table, you can write:

- ◆ Total Sales = SUM(Sales[SalesAmount])

Now, this measure dynamically calculates Total Sales based on the filters or slicers in the report.

Table 2.2.12 Key Differences Between Calculated Columns and Measures

Feature	Calculated Column	Measure
Evaluation Context	Row Context	Filter Context
Storage	Stored in model (consumes memory)	Not stored (calculated dynamically)
Performance	Slower for large datasets	Faster and optimized
Visibility	Appears in the data table	Appears only in the Fields pane (for visuals)
Use in Filters/Slicers	Yes	No
Use Case	Derived attributes (e.g., Profit Category)	Aggregations and KPIs (e.g., Total Profit, Sales Growth)

When to Use Which?

Table 2.2.13 When to Use Which

Situation	Recommended Choice
You need to create a category or label (e.g., "High Sales" or "Low Sales")	Calculated Column
You need to create totals, averages, or KPIs	Measure
You want a field to appear in slicers or filters	Calculated Column
You want to optimize report speed	Measure
You need a value that depends on filters (like region, category)	Measure

Real-World Example (Retail Case)

Dataset Fields:

- ◆ SalesAmount
- ◆ Profit
- ◆ ProductCategory
- ◆ Region

You want to:

- ◆ Classify each sale as “Profit” or “Loss” →

ProfitStatus = IF(Sales[Profit] > 0, “Profit”, “Loss”)

→ Calculated Column (because it’s per row)

- ◆ Calculate total profit by region →

Total Profit = SUM(Sales[Profit])

→ Measure (because it’s aggregated by filter context)

2.2.3.4 Fundamental Study of DAX

DAX (Data Analysis Expressions) is the formula language used in Power BI, Excel Power Pivot, and Analysis Services. It helps you create custom calculations, measures, and time-based analysis such as year-to-date (YTD) or month-over-month growth.

In simple words:

DAX is to Power BI what formulas are to Excel — but much more powerful!

Why DAX is Needed

After you have cleaned and modeled your data, you often need to:

- ◆ Calculate totals and averages
- ◆ Create business KPIs (like Profit %, YoY Growth, etc.)
- ◆ Compare current vs. previous periods
- ◆ Build interactive reports where numbers change with filters

Just like Excel functions — but DAX works with tables and relationships too.

Table 2.2.14 Basic DAX Examples

Purpose	DAX Formula	Explanation
Sum of Sales	SUM(Sales[SalesAmount])	Adds all Sales Amount values
Count of Transactions	COUNTROWS(Sales)	Counts number of rows in the Sales table
Distinct Customers	DISTINCTCOUNT(Sales[CustomerID])	Counts unique customers
Average Quantity Sold	AVERAGE(Sales[Quantity])	Finds average quantity sold

* Tip: Use DISTINCTCOUNT() instead of COUNT() when you want unique counts, like the number of unique customers or invoices.

Understanding DAX Evaluation Contexts

This is the most important DAX concept in Power BI. The result of any DAX formula depends on the context in which it is calculated.

There are two main types of contexts:

Table 2.2.15 two main types of contexts

Type of Context	Meaning	Used In
Row Context	DAX works row by row — one record at a time	Calculated Columns
Filter Context	DAX works on the data that is visible after filters or slicers are applied	Measures

Example to Understand Context

Imagine you have a Power BI report with a slicer for Region.

Table 2.2.16 Example

Filter Applied	Measure Result
No filter	Shows total sales of all regions
Region = “South”	Shows only South region’s sales
Category = “Furniture”	Shows sales of only Furniture items

This happens because Measures automatically respond to the filter context — they always calculate based on the current filters in the report.

CALCULATE Function — The Heart of DAX

Definition:

CALCULATE() changes or modifies the filter context before evaluating a calculation. It is used whenever you want to apply specific conditions in your measure.

Syntax:

CALCULATE(<expression>, <filter1>, <filter2>, ...)

Retail Example:

Sales (Furniture) =

```
CALCULATE(
    SUM(Sales[SalesAmount]),
    Products[Category] = “Furniture”
)
```

Explanation :

This measure calculates the total sales only for the “Furniture” category, even if other filters (like region or date) are applied.

Table 2.2.17 Additional DAX Examples

Measure Name	Formula	Description
Total Sales	SUM(Sales[SalesAmount])	Adds up all sales amounts
Total Quantity	SUM(Sales[Quantity])	Adds total units sold

Total Profit	SUM(Sales[Profit])	Adds total profit
Average Unit Price	DIVIDE([Total Sales], [Total Quantity])	Computes selling price per unit
Distinct Customers	DISTINCTCOUNT(Sales[CustomerID])	Counts unique customers
Profit %	DIVIDE([Total Profit], [Total Sales])	Profit margin as a percentage

2.2.3.5 Introduction to Time Intelligence

Time Intelligence in Power BI

Time Intelligence means calculations performed over time based contexts like:

- ◆ What is the Sales Year till Date?
- ◆ How much Sales this Month compared to the previous Month?
- ◆ Quarter wise trend?
- ◆ What was the sales in the same period last year?

Power BI DAX has huge Time Intelligence support.

You MUST have a proper Date Table and it must be marked as Date Table in Power BI.

To create a Date Table (Calendar Year):

DateTable = CALENDAR (DATE(2019,1,1), DATE(2025,12,31))

Then mark as Date Table → Model View → right click table → Mark as Date Table.

Base Measure We Use

Total Sales = SUM(Sales[SalesAmount])

This is the measure used everywhere now.

Table 2.2.18 Standard International Calendar (Jan 1 Year Start)

Calculation	DAX
YearToDate Sales	Sales YTD = TOTALYTD([Total Sales], 'DateTable'[Date])
Month To Date Sales	Sales MTD = TOTALMTD([Total Sales], 'DateTable'[Date])
Quarter To Date Sales	Sales QTD = TOTALQTD([Total Sales], 'DateTable'[Date])

Previous Year Sales	Sales PY = CALCULATE([Total Sales], SAMEPERIODLASTYEAR('DateTable'[Date]))
---------------------	---

Indian Financial Year (April 1 – March 31 Model)

Same formula... BUT add the FY end date inside the functions:

Table 2.2.19 FY end date inside the functions

Calculation	DAX
YTD (Indian FY)	Sales YTD FYIN = TOTALYTD([Total Sales], 'DateTable'[Date], "03/31")
MTD (same)	no change needed, MTD is month based.
QTD (same)	Quarter changes depending on Fiscal Quarter mapping.
Previous FY same period	Sales PY FYIN = CALCULATE([Total Sales], SAMEPERIODLASTYEAR('DateTable'[Date])) (this also automatically respects Fiscal if Fiscal applied in Date Table)

Why Time Intelligence matters in Retail?

- ◆ Retail sales are extremely seasonal
- ◆ Offers like Onam, Diwali, Christmas change business cycle
- ◆ Year comparison tells business if performance improving
- ◆ Q1, Q2, Q3, Q4 performance helps category managers & pricing strategy teams

Real Business Implementation Tip

Time intelligence is ONE of the most asked interview topics for Power BI roles. 95% industry KPI dashboards contain YoY % change measures. Time Intelligence requires a proper Calendar table.

Recap

- ◆ A data model organizes tables and relationships to support efficient analysis.
- ◆ Relationships connect tables using keys and are defined by cardinality and filter direction.
- ◆ Star schemas are preferred for analytical models due to simplicity and performance.



- ◆ Fact tables store measurable data, while dimension tables store descriptive attributes.
- ◆ DAX (Data Analysis Expressions) is used to create calculations in Power BI.
- ◆ Calculated columns operate row-by-row, while measures compute results dynamically based on filters.
- ◆ Understanding row context and filter context is essential for accurate DAX calculations.
- ◆ A well-structured data model improves report performance and ensures meaningful insights.

Objective Questions

1. Which table type typically contains numeric and aggregatable values?
2. A relationship where multiple rows in one table match a single row in another table is called:
3. Which view in Power BI is used to create and manage relationships between tables?
4. DAX is primarily used for:
5. A calculated column is evaluated during:
6. Which DAX function is used for aggregating numeric fields?
7. Filter context affects:
8. A star schema is preferred because it:
9. The RELATED() function is used to:
10. Which relationship cardinality is most commonly used in fact–dimension models?
11. A data model contains duplicate product names in the dimension table. What should the developer check first?
12. A measure returns unexpected values when filtered. This is most likely caused by an incorrect understanding of:
13. A manager wants to calculate “Sales per Customer.” Which DAX function should be used to count customers?

14. Which scenario requires a many-to-many relationship?
15. A data model includes several unnecessary columns causing performance issues. The best optimization technique is to:

Answers

1. Fact table
2. Many-to-one
3. Model view
4. Creating calculations such as measures and columns
5. Data refresh
6. SUM
7. Only measures
8. Simplifies calculations and improves performance
9. Bring column values from a related table
10. One-to-many
11. Key fields used to create relationships
12. Filter context
13. DISTINCTCOUNT
14. Each order contains multiple products, and each product appears in many orders
15. Remove unused columns

Assignments

1. Explain the differences between star schema and snowflake schema with suitable diagrams. Discuss which schema is better suited for analytical reporting and why.
2. Create a sample data model consisting of one fact table and at least three dimension tables. Describe the cardinality and relationship types used between them.
3. Write DAX expressions to calculate Total Sales, Average Sales, and Year-over-Year Growth. Explain how filter context affects the result of each measure.
4. Explain the difference between calculated columns and measures in Power BI. Provide three examples where calculated columns are preferred and three where measures are essential.
5. Consider a dataset with sales transactions and product details. Design a relationship model and explain how slicers and filters work across fact and dimension tables due to cross-filter directions

Reference

- ◆ <https://learn.microsoft.com/power-bi/transform-model/>
- ◆ <https://learn.microsoft.com/power-bi/dax/>
- ◆ <https://www.sqlbi.com/>

Suggested Reading

1. *The Definitive Guide to DAX* – Marco Russo and Alberto Ferrari
SQLBI Articles on Data Modelling and DAX
<https://www.sqlbi.com>
2. Microsoft Documentation – Data Modelling in Power BI
<https://learn.microsoft.com/power-bi/transform-model>
3. Microsoft Documentation – DAX Basics
<https://learn.microsoft.com/power-bi/dax>

UNIT 3

Creating and Publishing Visual Reports

Learning Outcomes

At the end of this unit, the learner will be able to :

- ◆ choose the right visualization type such as charts, graphs, maps, or KPIs based on the data and purpose
- ◆ create and customize Power BI visuals so they are clear and easy to understand
- ◆ use interactive features like slicers, filters, drill-down, and drill-through to explore data better
- ◆ add custom visuals from the Power BI marketplace to improve reports
- ◆ format visuals with good colors, labels, tooltips, and layout for better presentation
- ◆ arrange pages and visuals neatly to prepare a complete Power BI report
- ◆ publish reports to the Power BI Service and know the basics of sharing and using workspaces

Prerequisite

As organizations move toward data-driven decision-making, the ability to translate raw data into clear, interactive visuals has become essential. Decision-makers often rely on dashboards and reports to identify trends, monitor performance, and explore underlying patterns. However, without well-structured data and an understanding of basic analytical concepts, visualizations may fail to convey insights accurately or may even mislead users.

Before building reports in Power BI, learners must be comfortable working with datasets understanding rows, columns, and different data types. A familiarity with foundational cleaning and modelling concepts ensures that the visuals they create

are based on reliable and organized data. For example, a sales dashboard built on inconsistent date formats or incorrect relationships may show misleading totals, defeating the purpose of visualization.

Learners entering this unit should also have a basic grasp of how Power BI handles data, including data loading and transformations. This prepares them to focus on choosing the right visualization type, applying interactivity features such as drill-downs and slicers, and designing user-friendly dashboards. With these fundamentals in place, learners can confidently move into creating compelling visuals, customizing reports, and publishing dashboards that support meaningful, real-time decision-making.

Keywords

Visualizations, Slicers, Filters, Drill-down, Drill-through, KPIs, Story telling, Power BI, Publishing, Sharing Reports

Discussion

2.3.1 Introduction to Visual Reporting

After data preparation and modeling, the next and most visible stage in Power BI is Visual Reporting the process of converting your data into insights that can be understood quickly and acted upon. Data visualization is not just about creating charts; it's about telling a meaningful story that helps decision-makers see trends, spot problems, and take informed actions.

Definition:

Data visualization is the graphical representation of data using visual elements like charts, graphs, and maps to communicate insights clearly and effectively.

Each step builds towards one goal : "Make data easy to see, understand and act upon"

Why Visuals Matter ?

Humans process visuals 60,000 times faster than text. Good visuals turn numbers into patterns, comparisons, and decisions. Table 2.3.1 shows how the visualization helps in particular purpose.

Table 2.3.1: Visualization Purpose

Purpose	How Visualization Helps
Identify trends	Line charts show upward or downward movement
Compare categories	Bar charts show which category performs better
Show composition	Pie or Tree-map shows contribution to total
Track performance	KPI cards summarize key metrics instantly
Discover relationships	Scatter plots show correlation between variables

2.3.1.1 The Visualization Workflow in Power BI

Creating an effective Power BI report involves moving through a series of clear steps that transform raw data into meaningful insights. Each step helps build visuals that are accurate, interactive, and easy to understand. Figure 2.3.1 depicts the visualization workflow in Power BI.

Step 1 : Data Model

The process begins with preparing and organizing the data. This includes cleaning the dataset, creating relationships between tables, and arranging fields properly. A well-designed data model ensures accuracy and smooth functioning of all visuals in the report.

Step 2 : Select Visualization

Once the data model is ready, the next step is choosing the appropriate visualization. Depending on the type of data and the objective, you may select charts, graphs, maps, or KPIs. The right visual helps communicate information clearly and effectively.

Step 3 : Add Fields

After selecting a visualization, fields are added to it by dragging and dropping relevant data elements. These fields may include categories, values, dates, or measures. Adding fields brings the visual to life by filling it with actual data.

Step 4 : Format and Customize

The visual is then formatted to enhance clarity and appearance. This includes adjusting colors, labels, titles, fonts, and layout. Customizing tooltips or applying conditional formatting can also help make the visual more readable and engaging.

Step 5: Add Interactions

To make the report more dynamic, interactive elements such as slicers, filters, drill-down, and drill-through are added. These interactions allow users to explore the data in different ways and gain deeper insights.

Step 6 : Tell the Story

Finally, visuals are arranged in a meaningful order to present a clear story. Consistent design, logical flow, and focused insights help convey the main message of the report. This step ensures that the report not only looks good but also communicates the intended information effectively.

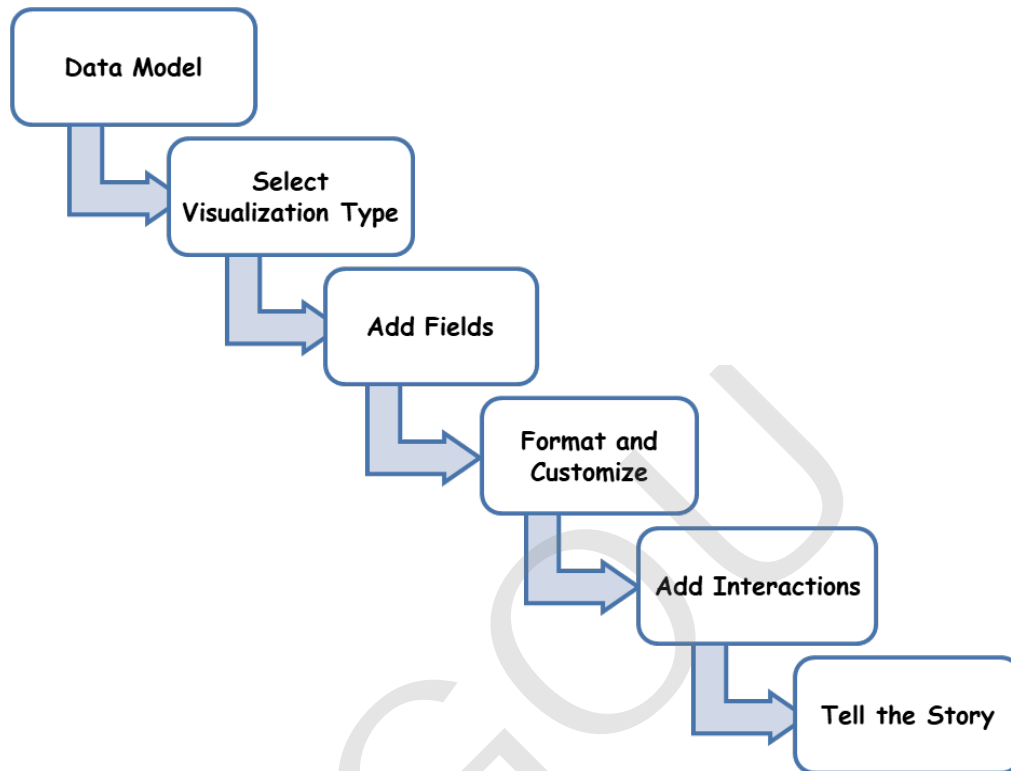


Fig. 2.3.1: The Visualization Workflow in Power BI

2.3.1.2 Choosing the Right Chart Type

Selecting the correct chart is the first step to telling an effective story. Every chart answers a specific kind of business question. Common chart types in Power BI and their use cases are described in Table 2.3.2.

Table 2.3.2 Common Chart Types in Power BI and Their Use Cases

Chart Type	Purpose	Best Used For
Column / Bar Chart	Compare values across categories	Sales by Region, Profit by Product
Line Chart	Show trends over time	Monthly or Yearly Sales Trends
Area Chart	Show cumulative values or volume over time	Overall Growth Trends

Pie / Donut Chart	Show percentage contribution	Share of Sales by Segment (up to 3–4 items only)
Treemap	Display part-to-whole relationships	Category contribution when many small items exist
Scatter Chart	Show relationship between two variables	Discount vs Profit, Sales vs Quantity
Funnel Chart	Visualize process stages or drop-offs	Lead → Conversion → Purchase
Card / KPI	Highlight a key performance metric	Total Sales, Profit %, Active Customers
Matrix / Table	Display detailed numeric data	Product-wise or Region-wise breakdown
Map (Filled / Bubble)	Show geographic distribution	Regional or Country-level Sales
Gauge Chart	Compare performance to target	Actual vs Target Sales Progress

2.3.2 Designing KPIs (Key Performance Indicators)

KPIs are the heart of every dashboard. They summarize performance in a single glance. A KPI (Key Performance Indicator) is a measurable value that shows how effectively a business is achieving its key objectives. KPI examples for retail business explained in Table 2.3.3.

Table 2.3.3: KPI Examples for Retail Business

KPI	Formula	Description
Total Sales	SUM(Sales[SalesAmount])	Total revenue generated
Total Profit	SUM(Sales[Profit])	Profit after costs
Profit Margin %	DIVIDE([Total Profit],[Total Sales])	Profitability measure
Sales Growth %	([Sales YTD] - [Sales PY]) / [Sales PY]	Growth rate over previous year
Top Product Category	TOPN(1, VALUES(Product[Category]), [Total Sales])	Category contributing highest sales

Given that a **KPI (Key Performance Indicator)** is a measurable value showing how

effectively a business achieves its key objectives, the layout for prioritizing these critical metrics is shown in fig 2.3.2.

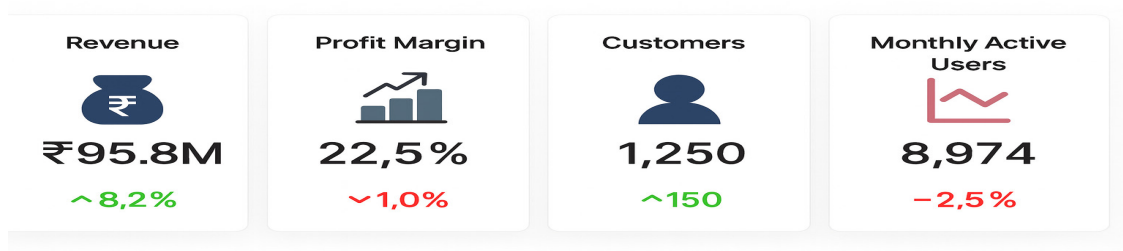


Fig 2.3.2 Layout of KPI Section

Best Practices for KPI Design

- ◆ Keep KPIs minimal (3–5 per dashboard).
- ◆ Use consistent colors (green = good, red = below target).
- ◆ Always compare with a target (use arrows or icons).

Place KPIs at the top of dashboard for immediate visibility.

2.3.2.1 Filters, Slicers, and Interactivity

Power BI's power lies in interactivity, the ability for users to explore data from different perspectives without coding.

A. Filters

At its core, a filter is a mechanism used to limit the data displayed in your visuals. Filters allow you to strip away the noise and focus on specific data points such as a particular time period, product category, or geographic region. By controlling what is shown and what is hidden, filters enable you to curate the data story, ensuring the user sees exactly what is relevant to their current analysis.

The Three Levels of Filter Scope

One of the most important concepts to master is that filters in the “Filters Pane” work hierarchically. They can be applied at three distinct levels, each serving a different purpose:

- **Visual-Level Filters:** These are the most granular filters. They affect only a single specific chart or visual selected on the canvas. For example, if you have a bar chart showing “Total Sales,” you could apply a visual-level filter to show only “Top 5 Products.” This filter will not change the numbers on any other card, pie chart, or table surrounding it. Use this when one specific visual needs to display a restricted subset of data while the rest of the page remains broad.

- **Page-Level Filters:** These filters impact every visual on the current page (or tab) of the report. If you have a multi-page report where one page is dedicated entirely to “European Operations,” you would apply a Page-Level filter for *Region = Europe*. This ensures that every chart on that specific page reflects only European data, but once the user clicks to the next page, that filter no longer applies.
- **Report-Level Filters:** These are global filters that affect every visual, on every page, across the entire report. These are typically used for high-level constraints. for example, if you are generating a report specifically for the “2024 Fiscal Year,” you would apply a Report-Level filter for *Year = 2024*. No matter where the user navigates within the report, they will only ever see data belonging to that year.

Example : Show only data for Region = “South”.

B. Slicers

Slicers are visual filters users can interact with directly on the report page. Table 2.3.4

Table 2.3.4 Slicers

Type of Slicer	Use Case
Dropdown / List	Filter data by Category or Region
Date Range	View data for specific period
Numeric Range	Filter by Profit or Quantity range
Hierarchical	Filter by Category → Subcategory

Example : A slicer for “Year” lets users toggle between 2023, 2024, etc., and automatically updates all visuals.

C. Drill Down and Drill Through

- ◆ **Drill Down :** View data at a more detailed level (e.g., Year → Quarter → Month → Day). Activated by clicking the down arrow on a visual.
- ◆ **Drill Through :** Navigate to another page to view detailed information for a selected item.

Example : Clicking on Electronics in a bar chart opens a detailed “Product Performance” page showing brand-wise sales and profit. Benefits of interactivity is shown in Table 2.3.5

Table 2.3.5 Benefits of Interactivity

Benefit	Description
Flexibility	Users can explore without changing source data
Personalization	Each user sees data relevant to them
Discover Insights	Identify anomalies or patterns easily
Engagement	Keeps dashboards dynamic and exploratory

2.3.3 Storytelling with Data

Data storytelling goes beyond numbers, it combines visuals + narrative + context to guide the audience toward understanding. Data Storytelling is the structured presentation of data-driven insights using visuals, text, and context to explain what happened, why it happened, and what actions to take. Elements of an effective story is depicted in Table 2.3.6.

Table 2.3.6 Elements of an Effective Story

Element	Description
Narrative (What)	A logical sequence of insights introduction, conflict, resolution
Data (Why)	Facts and metrics that support the story
Visuals (How)	Charts that make insights clear and engaging

2.3.3.1 Storytelling Framework in Power BI

Storytelling Framework in Power BI refers to the structured approach used to convert raw data into meaningful insights through clear, interactive, and visually engaging reports. It guides users through each step starting from organizing the data to presenting the final message so that the report not only displays information but also communicates a compelling story that supports decision-making.

- 1. Set the Scene:** Start with KPIs or summary cards to establish context.
Example: “Overall sales reached ₹1.2 Cr this quarter.”
- 2. Show the Trend:** Use line charts to show how performance evolved.
Example: “Sales dropped 15% in August due to a supply issue.”
- 3. Highlight the Key Insight:** Use color or annotation to emphasize what matters
Example: “South Region contributed 45% of the loss.”
- 4. Conclude with Action:** Provide takeaways or next steps.
Example: “Re-evaluate pricing in South region for Q4.”

2.3.3.2 Dashboard Design Principles

Creating a dashboard in Power BI is as much an art as it is a science. While the DAX formulas provide the logic, the design provides the clarity. A poorly designed dashboard can hide important insights behind clutter, while a well-designed one guides the user to the answer immediately. Below are the core best practices for creating effective dashboards.

1. **Use a Clean and Consistent Layout:** The layout is the skeleton of your dashboard. Humans instinctively look for patterns, so your design should rely on a consistent grid structure. Align your charts perfectly to create clean lines and use white space (negative space) effectively to give the eyes a place to rest. A chaotic layout forces the user to expend mental energy figuring out where to look, rather than understanding what they are seeing. A good rule of thumb is to place the most critical KPIs (Key Performance Indicators) in the top-left corner, as that is where the eye naturally lands first in most cultures.
2. **Group Related Visuals Together:** Apply the principle of Proximity, items that are placed near each other are perceived as being related. If you have a slicer for “Year,” place it near the trend line it controls. If you have three charts related to “Sales Performance,” group them in one section, perhaps separated from “HR Metrics” by a subtle background border or white space. This logical grouping helps users create a mental map of the data, allowing them to navigate complex information intuitively.
3. **Use Color with Purpose (Not Decoration):** Color is a powerful communication tool, not just a way to make the report look “pretty.” Avoid the “fruit salad” effect where every bar in a chart is a different color for no reason (Table 2.3.7).
 - ◆ *Use semantic colors* : Stick to standard associations (e.g., Green for positive/growth, Red for negative/decline).
 - ◆ *Use contrast to highlight* : If you want to draw attention to one specific bar in a chart, make that bar a bold color (like blue) and turn all other bars gray. This directs the user’s attention instantly to the insight you want to highlight.
4. **Keep Fonts and Sizes Consistent:** Typography establishes hierarchy. Your report should have a strict set of rules for fonts:
 - ◆ *Dashboard Title:* Largest font (e.g., 24pt).
 - ◆ *Section Headers:* Medium font (e.g., 14-16pt).
 - ◆ *Data Labels/Axis Text:* Smaller font (e.g., 10-12pt). Don’t mix different font families; stick to one standard font (like Segoe UI or Arial) throughout the report. Inconsistency in font size or style makes the report look unprofessional and distracts from the data.
5. **Avoid Too Many Visuals (The “Less is More” Rule):** One of the most common mistakes is trying to answer every possible question on a single page. This creates “Cognitive Overload.” A dashboard should not be an encyclopedia, it should be a summary.

- ◆ *The 6-8 Rule:* Aim for a maximum of 6 to 8 visuals (including card visuals) per page.
 - ◆ If you have more data to show, use Drill-throughs or Tooltips to hide the detail until the user specifically asks for it. If everything is important, then nothing is important.
6. **Always Test with End-Users for Readability:** You are not designing the dashboard for yourself; you are designing it for an audience. What is obvious to you (the creator) might be confusing to them. Before finalizing a report, perform a “squint test” or ask a colleague to look at it for 5 seconds. Ask them: “What is the first thing you noticed?” and “Is the main message clear?” Also, consider accessibility to ensure your color choices are distinguishable for users with color blindness (avoiding red/green reliance) and that the text is large enough to be read on the devices your stakeholders use (e.g., laptops vs. projector screens).

Table 2.3.7 Colors and Themes

Type	Recommendation
Corporate Dashboards	Use company brand palette
Analytical Dashboards	Use muted colors (grays, blues)
Highlighting	Use bright accent color (orange/red) for exception

Example : Dashboard Layout for Financial Report shown below in fig 2.3.3

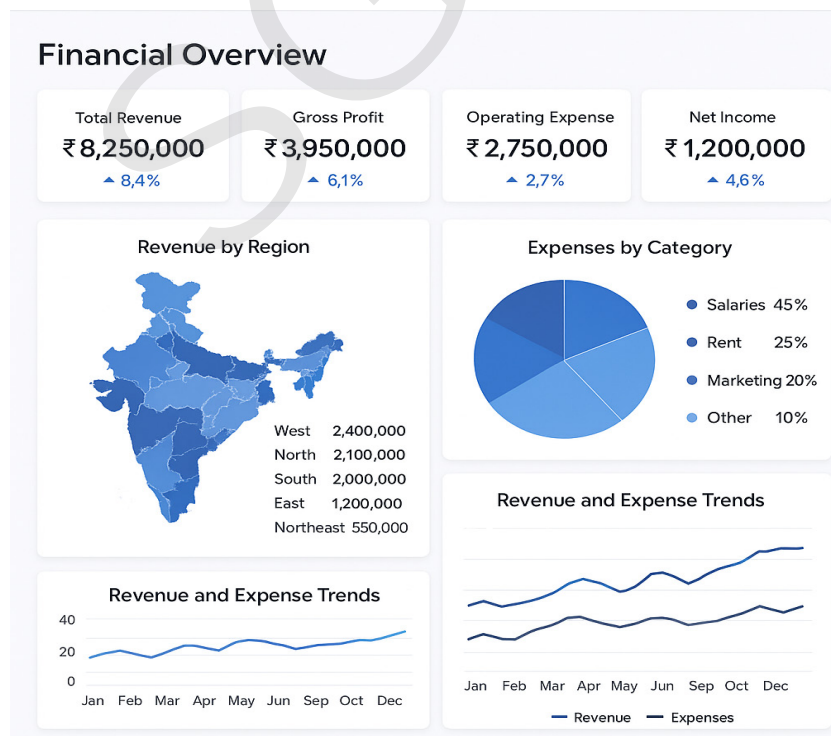


Fig 2.3.3 Dashboard Layout for Financial Report

Example : Dashboard layout for retail report shown in Fig 2.3.4.

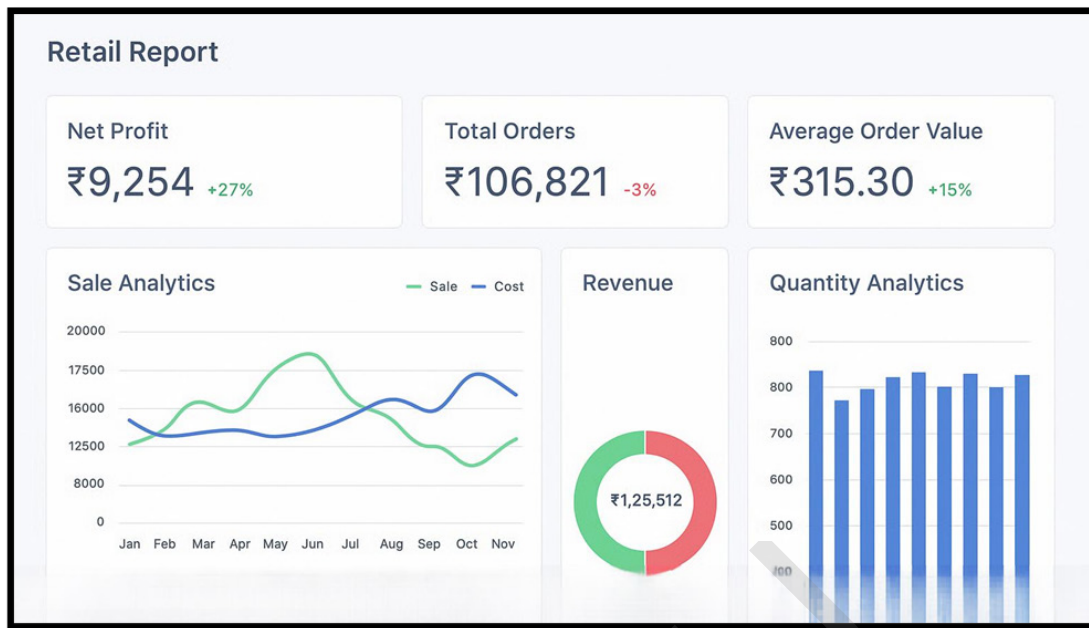


Fig 2.3.4 Dashboard Layout for Retail Report

Tip: Always maintain good contrast (light background + dark text or vice versa).

Choosing visuals based on question type is described in table 2.3.8.

Table 2.3.8 Choosing Visuals based on Question Type

Question in Retail	Best Visual
How did sales change over the period?	Line Chart
Which category contributes more revenue?	Bar / Column Chart
Which region performs best?	Bar Chart
Compare Current Year vs Previous Year sales	Multi-line chart / Clustered Column
KPI summary for management	Card visual

2.3.3.3 Chart types :

To ensure the visualization effectively supports decision-making in a specific field like retail, the practical application of different chart types and their business interpretations are detailed in Fig 2.3.5.

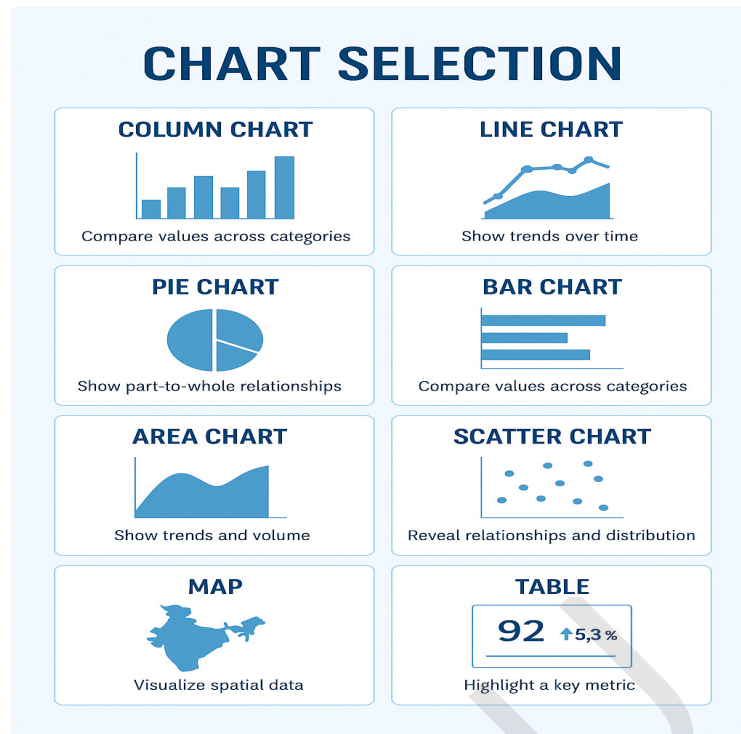


Fig 2.3.5: Chart Types + Business Interpretation (Retail Context)

1. Column & Bar Charts

These are the workhorses of data visualization, designed for comparing values across different categories. They are the most accurate visuals for human perception because our eyes compare lengths very easily.

Retail Example: Use these to compare Sales by Product Category, analyze Sales by Region, or see the split between Online vs. Store channels.

2. Line Chart

The Line Chart is the standard for visualizing trends over continuous intervals, usually time. It is the best tool to answer the question, “How is the data changing?” or “Is the trend moving up or down?”

Retail Example: Plotting Sales Trends on a monthly, weekly, or quarterly basis to spot seasonality.

3. Area Chart

Similar to a line chart, an Area Chart fills the space below the line with color. While the line highlights the trend, the filled area emphasizes the volume or magnitude of the change. It is particularly useful for showing cumulative totals.

Retail Example: Visualizing Cumulative Sales build-up throughout the year to see total volume growth.

4. Pie & Donut Charts

These visuals show part-to-whole proportions. However, they should be used strictly for datasets with very few categories (ideally 2 to 4). If there are too many slices, the chart becomes unreadable and academically unsound.

Retail Example: Perfect for a simple binary comparison, such as the Share of Online Sales vs. Physical Store Sales.

5. Scatter Plot

A Scatter Plot is used to investigate the relationship (correlation) between two continuous numerical variables. It helps you see patterns, clusters, or outliers that aren't obvious in standard tables.

Retail Example : Plotting Sales vs. Discount % to identify if offering higher discounts is actually hurting overall profitability.

6. Bubble Chart

This is an advanced variation of the scatter plot. It allows you to visualize three dimensions of data simultaneously: the X-axis and Y-axis show relationships, while the size of the bubble represents a measure of magnitude.

Retail Example: Each bubble represents a product; the axes show Sales vs. Profit, and the Bubble Size represents Quantity Sold.

7. Tree Map

A Tree Map displays hierarchical data as a set of nested rectangles. The size of each rectangle corresponds to its value, allowing users to spot patterns and dominant categories in large datasets instantly.

Retail Example: A breakdown of Product Categories by Sales, where the top-performing categories appear as larger, more prominent blocks.

8. Funnel Chart

This visual is specifically designed for analyzing stages in a linear process. It visually depicts how data reduces (or “funnels down”) as it moves through sequential steps, making it easy to spot drop-off points.

Retail Example: Analyzing the E-commerce conversion flow: Product Awareness → View Product → Add to Cart → Checkout → Purchase.

9. KPI Card Visuals

KPI (Key Performance Indicator) Cards are single-number visuals used to communicate the most critical metrics immediately. They act as the “headlines” of your dashboard, ensuring the user sees the most important numbers first.

◆ *Retail Example:* Displaying distinct, large numbers for Total Revenue, Total Profit, or Net Margin at the top of the report.

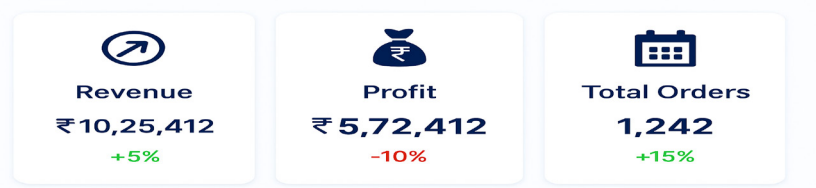


Fig 2.3.6: KPI Card Visuals

2.3.4 How to Select Visuals in Power BI?

Selecting visuals in Power BI involves choosing the chart or graphic that best represents your data and communicates your message effectively. Start by understanding the type of data you are working with numerical, categorical, time-based, or geographical. Then identify the purpose of your analysis, such as showing comparisons, trends, relationships, distribution, or performance. Match this purpose to an appropriate visual, like bar and column charts for comparisons, line charts for trends, scatter plots for relationships, maps for geographical data, and KPIs or cards for key performance indicators. Always keep visuals simple, easy to read, and suitable for your audience. Finally, test different visuals and select the one that clearly highlights the insight you want to communicate. Choosing visuals based on question type shown in Table 2.3.9.

Table 2.3.9 Choosing Visuals based on Question Type

Type of Question	Preferred Visual
Comparison between items	Bar / Column
Trend over time	Line Chart
Distribution	Histogram / Box Plot
Relationship	Scatter Plot
Part-to-Whole	Treemap / Stacked Column (NOT Pie)

Golden rule: If the question is time-based → almost always Line Chart.

2.3.4.1 KPI Design Principles

A KPI is NOT just a number. It must tell the meaning. A good KPI in Power BI should always answer 3 things:

- ◆ Where am I today? (Current Value)
- ◆ Compared to what? (Benchmark / last year / target)
- ◆ So what? (Direction → is this good or bad?)

Storytelling Focus : Questions → Visuals → Narrative

Instead of directly creating visuals → always start with QUESTIONS.

2.3.4.2 Retail Core Business Questions

In the retail industry, understanding key business questions is essential for improving

performance, attracting customers, and increasing profitability. These core questions help retailers analyze sales trends, customer behavior, product performance, and operational efficiency. By focusing on these areas, businesses can make better decisions, optimize strategies, and respond effectively to market demands. Measures used based on question type described in Table 2.3.10.

Table 2.3.10 Measures used based on Question Type

Business Question	Visual	Measure Used
Is our overall Sales increasing month by month?	Line Chart (Monthly Trend)	Sales
Which Product Category drives most revenue?	Bar Chart	Sales by Category
Which Region underperforms consistently?	Bar Chart	Sales by Region
Did festival season (Onam/ Diwali) boost sales?	Line Chart with period highlights	Sales vs Date
Is discount actually converting to higher sales?	Scatter	Sales vs Discount

Example Story Flow (Retail):

1. Start with overall sales trend (Line chart) → highlight drop in Aug.
2. Click Region Slicer → South region shows sharp drop.
3. Drill into Products → Accessories category specifically dropped.
4. Final Story Statement → “Accessory category discount strategy needs revision in South region. Drop started in mid-August.”

This is how “data → story → decision” happens.

2.3.4.3 Visualization Best Practices

The DO's

- ◆ **Use simple visuals first:** Stick to the classics like Bar and Line charts. They are universally understood and require zero learning curve for your audience, unlike complex custom visuals.
- ◆ **Keep consistent formatting:** Maintain uniformity in fonts, alignment, and titles across all report pages. This ensures users focus on the changing data, not the changing design.

- ◆ **Use minimal but meaningful colors:** Color should highlight insights (e.g., Red for decline, Green for growth), not decorate the page. Avoid the “rainbow effect” where every bar is a different random color.
- ◆ **Always label axes clearly:** Never leave your user guessing. Ensure every chart has readable titles, axis labels, and legends so the metric being measured is obvious.
- ◆ **Show KPIs on top:** Follow the natural reading hierarchy. Place critical “headline” numbers (KPIs) at the top left, followed by trends in the middle, and detailed tables at the bottom.
- ◆ **Sort visuals based on business logic:** Don’t default to alphabetical sorting. Arrange data meaningfully (e.g., Highest Sales to Lowest) to instantly highlight top performers and issues.
- ◆ **Test mobile layout:** If your stakeholders check data on the go, use Power BI’s mobile layout view to ensure the dashboard is readable on a phone screen, rather than just shrinking the desktop view.

The DON'Ts

- ◆ **Avoid Pie charts for many categories:** Never use a pie chart for more than 3-4 categories. Human eyes struggle to compare small angles; a bar chart is almost always a better alternative.
- ◆ **Avoid clutter:** Don’t cram too much onto one canvas. If you have too many visuals, the main message gets lost. Split the content into multiple pages or use drill-throughs.
- ◆ **Don’t prioritize “Art” over Data:** The goal is business insight, not artistic expression. Avoid distracting background images, 3D effects, or shadows that clutter the view.
- ◆ **Don’t mix font types:** Stick to one standard font family throughout the report. Using multiple font styles makes the dashboard look messy and unprofessional.
- ◆ **Don’t force people to guess the context:** A number like “Sales: \$5M” is meaningless on its own. Always add context such as “vs. Last Year” or “vs. Target” so the user immediately knows if the number is good or bad.

2.3.5 Publishing a Report in Power BI

After designing visuals, KPIs, and dashboards in Power BI Desktop, the next major step is to publish and share your report with others. Power BI provides a seamless way to move your report from your local computer (Power BI Desktop) to the Power BI Service (Cloud), where it can be viewed, shared, and refreshed automatically.

What Does “Publishing” Mean?

Publishing is the critical step that bridges the gap between individual creation and organizational consumption. It involves uploading your local Power BI Desktop file (.pbix) to the Power BI Service, a cloud-based SaaS (Software as a Service) platform

accessible at *app.powerbi.com*. This action transforms a static file on your computer into a live, shared asset in the cloud.

Key Capabilities After Publishing

1. **Universal Access (Web & Mobile)** Once published, the report is no longer tied to your specific laptop. Users with the right permissions can access the dashboard securely through any standard web browser or via the Power BI Mobile app on iOS and Android. This ensures decision-makers have access to insights anytime, anywhere.
2. **Full Interactivity** Moving to the cloud does not sacrifice functionality. Viewers can fully interact with the report clicking slicers, applying filters, and using drill-down features to explore the data. They get the full analytical experience without needing to install the Power BI Desktop software or understand how to build the report.
3. **Automated Data Refreshes** One of the biggest advantages of the Service is automation. You can configure Scheduled Refreshes (e.g., every morning at 9:00 AM) so the dashboard always displays the latest data. This eliminates the manual drudgery of updating and emailing Excel files or PDFs every day.
4. **Secure Collaboration** The Service provides a managed environment for sharing. You can control exactly who sees the report through email sharing, distributing via “Power BI Apps,” or embedding reports directly into Microsoft Teams. Access is governed by your organization’s security protocols, ensuring sensitive data remains protected.

2.3.5.1 Steps to Publish a Power BI Report

Below are the standard steps for publishing a report from Power BI Desktop:

Step 1: Save Your Work

Before publishing, save your .pbix file locally.

Step 2: Sign In to Power BI Service

In Power BI Desktop, click “Sign in” (top right corner) using your Microsoft or organizational account.

Step 3: Click Publish

Go to: Home Tab → Publish → Select Destination Workspace

Step 4: Choose a Workspace

Select the workspaces shown in table 2.3.11 (for example, My Workspace or Sales Reports) in the Power BI Service.



Step 5: Open in Browser

After a successful upload, click “Open in Power BI” to view the report live in the cloud.

A. Understanding Workspaces

In the Power BI Service, a Workspace acts as a dedicated container for your content. Think of it as a specific project folder or a “hub” where you organize your datasets, reports, and dashboards. The type of workspace you choose determines the level of collaboration, security, and who can access the content (Table 2.3.11).

Table 2.3.11 Understanding Workspaces

Type of Workspace	Description
My Workspace	Personal workspace visible only to the creator
Shared / App Workspace	Collaborative area for teams and departments
Premium Workspace	High-performance environment for enterprise-level sharing

B. Components Published to the Power BI Service

In the Power BI Service, your content is not just a single file; it is organized into three distinct layers that build upon one another. It begins with the raw data logic, moves into detailed multi-page analysis, and culminates in a high-level overview for quick monitoring. Understanding the specific role of each component is essential for effectively managing your workspace (Table 2.3.12).

Table 2.3.12 Understanding Component

Component	Description
Dataset	The data model and queries used by the report
Report	Pages and visuals created in Power BI Desktop
Dashboard	Custom layout created in the Power BI Service using pinned visuals

C. Refreshing Data After Publishing

After publishing, you can configure data refresh schedules so that your dashboard always shows the latest information.

- ◆ Automatic Refresh : Set up in Power BI Service → Datasets → Schedule Refresh
- ◆ Manual Refresh: Click the “Refresh Now” option.

- ◆ Gateway Connection: Required for on-premise databases (explained in Unit 4).

Example: A retail sales dashboard can be refreshed daily at 6 AM to show previous day's transactions.

D. Sharing Reports

Power BI provides multiple sharing options (Table 2.3.13).

Table 2.3.13 Multiple Sharing Options

Method	Description
Share Dashboard	Directly share dashboard URL with team members
Publish to Web	Generates a public URL (not secure, used for learning/demo only)
Embed in App / Website	Add interactive Power BI visuals to applications
Export	Export to PDF, PowerPoint, or Excel snapshot

Note: "Publish to Web" should not be used for confidential or internal company data.

2.3.5.2 Security Considerations When Publishing

Security is paramount when moving data to the cloud. Power BI provides robust protection, but it requires the report creator to use the features correctly. Power BI publishing process described below in Fig.2.3.7

- ◆ **Only authorized users can view shared dashboards:** Publishing does not mean "public." Access is strictly controlled via Microsoft 365 logins. Users must be explicitly granted permission (via email or group) and typically require a Power BI Pro license to view secure reports.
- ◆ **Data stored in the Power BI Service is encrypted:** You do not need to manage your own servers to be secure. Microsoft automatically encrypts your data both **at rest** (while it is stored in their data centers) and in transit (while it is moving between the server and your browser), meeting enterprise compliance standards.
- ◆ **Sensitive data should never be shared via "Publish to Web": Critical Warning:** The "Publish to Web" feature creates a public link that can be accessed by *anyone* on the internet and indexed by Google. Use this *only* for public data (like government census stats). Never use it for internal company data, as there is no login required to view it.
- ◆ **Use Row-Level Security (RLS) to restrict access:** RLS allows you to publish a single report that adapts to the user. Instead of creating 50 different reports

for 50 sales managers, you create one. When “Manager A” logs in, the data-set filters automatically to show *only* their region’s data, while “Manager B” sees only theirs.

Power BI Publishing Process described below in fig 2.3.7

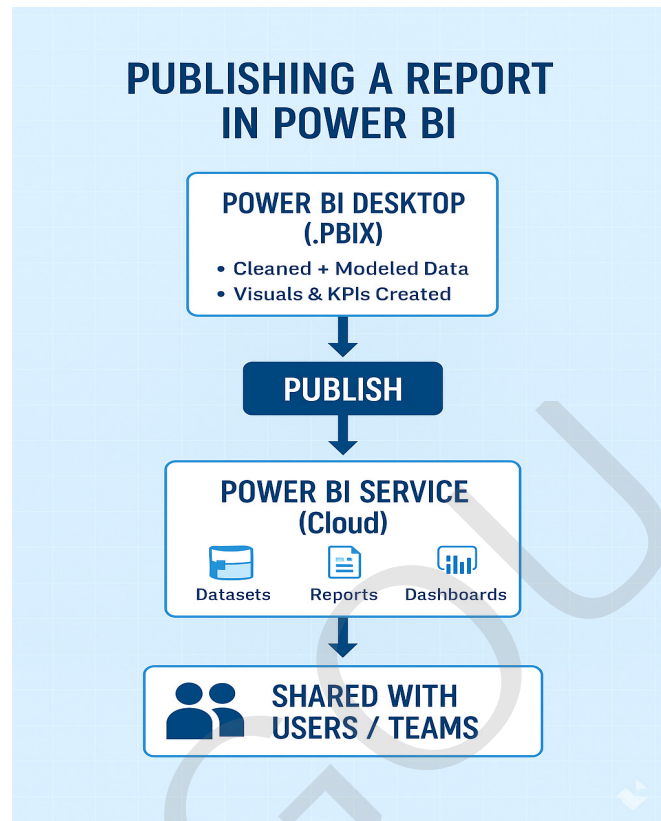


Fig 2.3.7 Publishing a report

Best Practices for Publishing Reports

Before you hit that “Publish” button, run through these checks to ensure your report is professional, performant, and ready for business use.

- ◆ **Always rename visuals and measures with meaningful names:** A user should never see technical database names like “Sum of Sales_Amount_TBL.” Rename titles and axis labels to plain business language, such as “Total Revenue.” This makes the report look polished and reduces confusion.
- ◆ **Keep reports under 1 GB for smooth cloud performance:** For Power BI Pro users, the dataset size limit is 1 GB. Even if you are under the limit, smaller models render faster in the browser. Remove unused columns and hidden tables to keep the file lean and the user experience snappy.
- ◆ **Publish to a specific workspace (not default “My Workspace”):** “My Workspace” is tied to your individual account. If you leave the company, that content becomes difficult to manage. Always publish to a Standard or App Workspace so the team retains ownership and access to the report regardless of personnel changes.

- ◆ **Test filters and slicers before sharing:** Conduct a “Quality Assurance” (QA) pass. Click every button and change every slicer to ensure visuals interact correctly. A broken filter or a blank chart destroys user confidence in the data immediately.
- ◆ **Use descriptive report names:** Avoid generic names like “Report1” or “Final_Final_v2.” Use a consistent naming convention that includes the topic and the time period (e.g., “**Retail_Sales_2024_Q1.pbix**”). This helps users find the right file when searching the Service.
- ◆ **Use the Power BI Service to monitor dataset refresh status:** Publishing is not the end of the road. Regularly check the dataset settings in the Service to ensure Scheduled Refreshes are not failing. If the data is stale (e.g., last updated two weeks ago), the report becomes useless for decision-making.

Recap

- ◆ Power BI provides various visualization types such as bar charts, maps, KPIs, tables, and matrices.
- ◆ Filters and slicers allow users to interact with the report and explore data deeply.
- ◆ Drill-down and drill-through enhance navigation within hierarchical data.
- ◆ Bookmarks enable users to save specific report views for presentations and storytelling.
- ◆ Custom visuals from AppSource can extend the analytical capabilities of Power BI.
- ◆ Proper formatting, layout consistency, and theme usage improve clarity and professionalism.
- ◆ Reports can include multiple pages organized logically for better user experience.
- ◆ Publishing to Power BI Service enables collaboration, sharing, and scheduled refreshes.
- ◆ Understanding workspace settings is essential for managing shared reports securely.

Objective Questions

1. Which visualization is most appropriate for showing monthly sales trends?
2. What component allows users to filter data directly from the report canvas?
3. Drill-through is used to.....
4. Which view in Power BI Desktop is used to design visual reports?
5. Conditional formatting is mainly used to
6. What feature helps users explore data in a hierarchy (e.g., Year → Quarter → Month)?
7. To save the current configuration of a page and use it for navigation, you should use:
8. Which option allows users to sort visual data by specific fields?
9. A retail report needs a visual to show the contribution of each category to total sales. Which chart is best?
10. Publishing a report to Power BI Service allows users to
11. Which feature supports navigation between multiple pages in a report?
12. To show the top 5 and bottom 5 products separately, a developer should use
13. What is the purpose of tooltips in Power BI visuals?
14. Which file format is used to publish reports from Power BI Desktop?
15. A user wants to interactively switch between different views of the same page (e.g., light mode and dark mode). Which feature supports this?

Answers

1. Line Chart
2. Slicer
3. Navigate to a detailed page related to a selected data point
4. Report View
5. Highlight important values
6. Drill-down
7. Bookmarks
8. Sort by
9. Pie/Donut Chart
10. Schedule data refresh and share reports
11. Buttons
12. Visual-level filters
13. To show additional information on hover
14. .pbix
15. Bookmarks

Assignments

1. Create a multi-page Power BI report using at least five different visualization types.
2. Explain the difference between drill-down, drill-through, and report tooltips.
3. Design a Power BI dashboard using slicers and filters.
4. Apply conditional formatting to a table or matrix visualization.
5. Publish a report to Power BI Service and configure sharing settings.

Reference

1. <https://learn.microsoft.com/power-bi/visuals/>
2. <https://learn.microsoft.com/power-bi/create-reports/power-bi-report-interactions>
3. <https://www.sqlbi.com/>

Suggested Reading

1. Microsoft Documentation – Power BI Visualization Guide
<https://learn.microsoft.com/power-bi/visuals>
2. Microsoft Documentation – Interacting with Reports (Filters, Slicers, Drill-down)
<https://learn.microsoft.com/power-bi/create-reports/power-bi-report-interactions>
3. SQLBI – Visualization Best Practices
<https://www.sqlbi.com>
4. Book: Storytelling with Data – Cole Nussbaumer Knaflic
(Recommended for understanding effective visual communication)
5. Microsoft Documentation – Publish and Share Reports in Power BI Service
<https://learn.microsoft.com/power-bi/collaborate-share/service-share-reports>

UNIT 4

Power BI Deployment and Security

Learning Outcomes

At the end of this unit, the learner will be able to :

- ◆ explain core Power BI security features, including workspace roles, Row-Level Security (RLS), and governance controls
- ◆ describe the functioning of the Power BI Service, including workspaces, datasets, dashboards, and sharing models
- ◆ configure Power BI Embedded by setting up architecture components such as capacity, authentication, and embedding tokens
- ◆ implement and manage the On-Premises Data Gateway to ensure secure hybrid data connectivity and scheduled refreshes

Prerequisite

As organizations transition to cloud-based analytics, ensuring secure access to data and reports becomes a critical responsibility. Power BI's ecosystem spanning the Power BI Service, Embedded solutions, and On-Premise Gateways allows businesses to share insights across teams, applications, and even external platforms. However, without an understanding of basic data governance and reporting workflows, it becomes difficult to manage permissions, protect sensitive information, or maintain reliable data connectivity.

Before exploring security features and deployment options in Power BI, learners should have a foundational understanding of how reports and datasets are created, published, and consumed. Familiarity with Power BI Desktop and the process of building visualizations prepares learners to appreciate why controlled access, row-level security, and user roles are essential in shared environments. Likewise, a basic awareness of cloud concepts such as authentication, data refresh, and remote connectivity helps learners understand how the Power BI Service and the On-Premise Gateway function together.



Having this foundational knowledge enables learners to progress confidently into advanced topics such as securing reports, embedding dashboards in external applications, managing gateways for hybrid data scenarios, and implementing organization-wide governance models. With these prerequisites in place, learners are prepared to explore Power BI's powerful security and deployment capabilities.

Keywords

Workspaces, Deployment, Row-Level Security (RLS), On-Premise Data Gateway, Data Refresh, Data Governance.

Discussion

In modern organizations, data driven decisions depend not only on the creation of meaningful reports but also on the proper deployment, sharing, and protection of these analytical assets. Power BI provides a comprehensive platform that supports secure report distribution, scalable deployment options, governance policies, and controlled access to organizational data. Once a report is created in Power BI Desktop, the next step involves publishing it to the Power BI Service, where teams can collaborate, consume insights, and automate data refreshes. This shift from local development to cloud-based deployment ensures that information is readily accessible to decision-makers and remains up-to-date with the latest business data.

Effective deployment begins with understanding how Power BI Service organizes content through *workspaces*. A workspace serves as a central hub where datasets, reports, dashboards, and dataflows are stored and managed. Team members are assigned different roles such as Viewer, Contributor, Member, or Admin depending on the access they require. This controlled environment helps maintain governance and ensures that only authorized individuals can create, modify, or share analytics within the workspace.

Security is a critical aspect of Power BI deployment, especially when handling sensitive business information. Power BI offers *Row Level Security (RLS)*, which restricts data visibility for individual users based on predefined roles. For example, a sales manager may view performance data for the entire organization, while a regional sales representative is allowed to see only their region. RLS enforces this differentiation by applying filters when the user accesses the report, thereby preventing unauthorized access. In large enterprises, this form of granular security is essential for preserving confidentiality and compliance with internal policies.

Organizations using on-premise databases or legacy data systems often require a secure bridge between their local environment and the cloud. Power BI resolves this need through the *On-Premise Data Gateway*, a software component installed on local servers or computers. The gateway securely transfers data from on-premise sources to the Power BI Service without exposing internal databases publicly. It also supports scheduled refresh operations, ensuring that reports reflect the latest available data. The gateway has become a vital component for hybrid data environments, where both cloud based and on-premise systems need to work seamlessly together.

Beyond internal use, Power BI also supports scenarios where business reports must be embedded into external applications such as websites, customer portals, or enterprise software systems. *Power BI Embedded* enables developers to integrate Power BI visuals into custom applications, allowing users to interact with analytics without directly accessing the Power BI interface. This is particularly useful for organizations that wish to provide analytical capabilities to clients, partners, or large user groups with varying access requirements.

Reliable deployment also requires strong governance, which includes monitoring data refreshes, managing user permissions, ensuring dataset certification, and maintaining audit controls. Administrators use the Power BI Admin Portal to oversee usage metrics, enforce security settings, and define data policies. This level of centralized governance ensures that Power BI content is managed responsibly and that organizational standards for security and compliance are consistently upheld.

Thus, Unit 4 focuses on equipping learners with the skills needed to deploy, secure, and manage Power BI reports in professional environments. Understanding the roles of workspaces, security models, gateways, deployment modes, and collaborative workflows enables learners to deliver fully operational business intelligence solutions that are not only insightful but also safe, scalable, and aligned with enterprise data practices.

2.4.1 Introduction

Modern organizations rely on Power BI not only for creating dashboards but also for securely deploying and managing them across teams. Effective deployment ensures that insights are accessible, up-to-date, and governed properly. This unit focuses on the process of publishing reports, managing workspaces, configuring security, establishing gateway connectivity, and understanding enterprise deployment scenarios.

2.4.2 Power BI Service

The Power BI Service (app.powerbi.com) is the central environment where insights are published, governed, interacted with, and shared. While Power BI Desktop is the “Artist’s Studio” where reports are created, the Service is the “Gallery” where the world views them.

It is not merely a storage system; it is a living SaaS (Software as a Service) ecosystem that coordinates authoring, data refresh, team collaboration, and secure distribution.

2.4.2.1 Power BI Service Architecture — The Dual Layer Operating System

To handle millions of users simultaneously, Power BI Service is built on Microsoft Azure and operates through two intertwined clusters. You can think of this as a restaurant: one team handles the guests (Front End), and the other handles the cooking (Back End).

1. Web Front End (WFE) — The Gateway Layer

The WFE acts as the traffic controller. When you type in the URL or open the mobile app, the WFE is the first point of contact.

- ◆ **Authentication:** It immediately verifies “Who are you?” using **Azure Active Directory (Azure AD)**.
- ◆ **Routing:** Once your identity is confirmed, it determines which data center holds your specific data and routes your request there.
- ◆ **Session Management:** It manages the connection to ensure the browser loads the static UI elements (the buttons, menus, and frames) quickly.

2. Back End (BE) — The Computational Core

The BE is where the heavy lifting happens. Once the WFE approves the user, the BE takes over.

- ◆ **Data Processing:** It interacts with the actual data storage (Azure SQL Database and Blob Storage).
- ◆ **Query Execution:** It runs the DAX queries required to generate the numbers.
- ◆ **Security Enforcement:** Before showing any number, it checks **RLS (Row Level Security)** rules to ensure the user is allowed to see that specific data point.
- ◆ **Rendering:** It generates the visualization structures and sends them back to the user’s screen.

The Flow:

When a report is opened: User Login → WFE Authenticates → BE Retrieves Data & Applies Security → Report Renders.

2.4.2.2 Workspaces

Workspaces are the distinct “containers” or folders where analytic work takes place. They prevent chaos by keeping Finance reports separate from HR reports.

- ◆ **My Workspace:** A personal sandbox for individual use. Content here cannot be easily shared with the organization.

- ◆ **App Workspaces:** Collaborative environments where teams work together on shared content.

Roles in a Workspace

Managing access to these workspaces is critical for governance. Power BI provides four distinct roles:

Table 2.4.1 PowerBI roles

Role	Description & Best Practice
Admin	The Owner. Has full control over the workspace, including deleting it and managing other users' access. <i>Best Practice: Limit to 1-2 people per workspace).</i>
Member	The Manager. Can add content, publish reports, and share the "App" with the wider organization. They cannot delete the workspace or change Admin settings.
Contributor	The Developer. Can create, edit, and delete reports and datasets within the workspace. They cannot modify who else has access. <i>(Best Practice: Assign this to analysts building the reports).</i>
Viewer	The Consumer. Can strictly view and interact with the content (filter, slice). They cannot edit or change anything. <i>(Best Practice: Assign this to management or stakeholders).</i>

2.4.2.3 Deployment & Refresh Process

Once a workspace is set up, the lifecycle of a report follows a standard path:

1. **Publishing:** Moving the .pbix file from Desktop to the Service.
2. **Configuring Gateway:** If the data is on-premise, the Gateway must be mapped to the dataset.
3. **Scheduling Refresh:**
 - ◆ **Scheduled Refresh:** The standard method where data updates at specific times (e.g., 6:00 AM daily).
 - ◆ **DirectQuery:** A real time connection where the report queries the database every time the user clicks a visual (no data is stored in the cloud).

2.4.3 Security

Every analytic system requires a protective framework, an invisible shield that determines who can see what, how information is handled, and how responsibly it moves through the organization. Power BI's security model is designed to work effortlessly with the way human cognition interprets information: clarity, boundaries, and structure.

Security is not just a technical layer; it is a communication system. It tells each viewer, *"Here is exactly what you are permitted to know,"* while silently concealing everything else.

The visual below (Fig 2.4.1 Power BI Security Model) illustrates the multi-layered approach Power BI uses to protect data, from the initial user login down to the specific rows of data they are allowed to see.

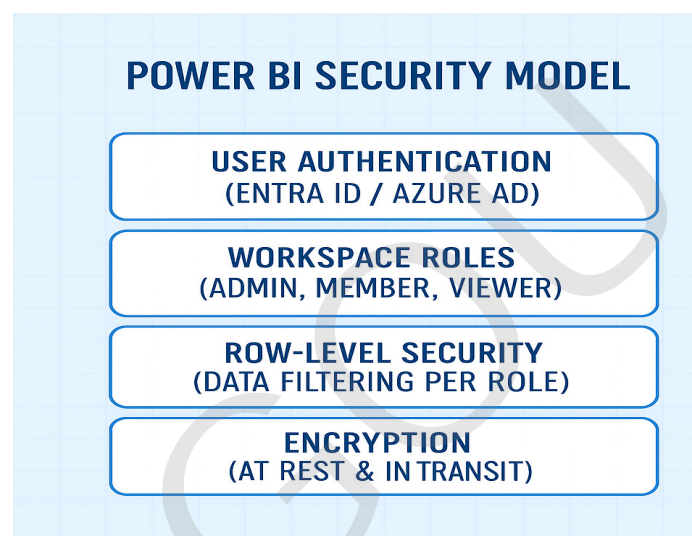


Fig 2.4.1 Power BI Security Model

2.4.3.1 Row Level Security (RLS) Adaptive Visibility

RLS behaves like a responsive filter. It ensures that two individuals viewing the same report do not necessarily see the same underlying rows of data. The system automatically applies rules that shape the dataset according to the viewer's identity.

Example structure of a rule:

[Region] = "North"

This allows the same report to communicate different, role appropriate insights with no duplication of content.

2.4.3.2 Object Level Security (OLS)

While RLS focuses on rows, OLS protects at an even deeper level entire columns or tables. If an organization must hide financial measures, confidential HR details, or sensitive metrics, OLS becomes the silent gatekeeper.

It ensures that restricted elements are not merely unreadable; they are invisible.

2.4.3.3 Governance Features

As analytics environments grow from a few files to thousands of reports, maintaining discipline becomes essential. Power BI provides governance mechanisms that act as the “organizational memory” and safety net of the system. These features ensure that data remains secure, accurate, and compliant as it scales.

1. Sensitivity Labels Sensitivity labels allow organizations to classify and protect data based on how confidential it is (e.g., “Public,” “Internal,” “Confidential,” or “Highly Restricted”). These labels are integrated with Microsoft Purview Information Protection.

- ◆ **Why it matters:** The protection is **persistent**. If a user exports a report with a “Highly Confidential” label to Excel or PDF, the file remains encrypted and restricted, even if it leaves the Power BI environment.

2. Usage Metrics Usage metrics provide built in reports that reveal exactly how content is being consumed. Admins and content creators can see total views, views per day, and a ranking of the most active users.

- ◆ **Why it matters:** This helps in measuring ROI (Return on Investment). If a dashboard took three months to build but has zero views, usage metrics highlight the adoption problem. Conversely, it helps identify “abandoned” reports that can be deleted to declutter the workspace.

3. Audit Logs Audit logs are the “black box” recorder of the Power BI tenant. They capture detailed records of every meaningful action taken by users, such as viewing a report, exporting data, publishing a dataset, or changing permissions.

- ◆ **Why it matters:** These logs are critical for regulatory compliance (GDPR, HIPAA). If sensitive data leaks, the security team can query the audit logs to trace exactly who accessed the report and when they exported the data.

4. Certified & Promoted Datasets (Endorsement) To prevent the spread of bad data, Power BI allows datasets to be “Endorsed.”

- ◆ **Promoted:** The content creator marks the dataset as ready for others to use (e.g., “I have finished checking this”).
- ◆ **Certified:** A higher standard where a central authority (like IT or a Data Governance team) verifies that the dataset is accurate, secure, and follows company standards.
- ◆ **Why it matters:** It establishes a “**Single Source of Truth.**” Users know that if they build a report using a *Certified* dataset, the numbers can be trusted.

5. Lineage & Impact Analysis Data Lineage provides a visual map showing the flow of data from the original source (e.g., SQL Server) → Dataset → Report → Dashboard.

- ◆ **Impact Analysis:** This feature allows you to see the consequences of a change *before* you make it. If you plan to delete a column in a dataset, Impact Analysis

will list every report and dashboard across the organization that will break if you do so.

6. Permission Management This involves the granular control of user access. Governance requires a strict strategy for who gets added to Workspaces (Creators) versus who receives the Power BI App (Consumers).

- ◆ **Why it matters:** Proper permission management prevents “over sharing” and ensures that unauthorized users cannot edit critical datasets or inadvertently delete content.

2.4.4 Power BI Embedded

If security is the shield, Power BI Embedded is the bridge. It brings analytics into external applications, customer portals, supplier dashboards, and SaaS platforms without requiring external users to carry Power BI licenses.

Here, the goal is simple: **deliver insights wherever users already are**. Instead of forcing an external client to log into a separate Microsoft portal, the data comes to them, embedded naturally within the software they use every day.

2.4.4.1 Why Embedded Matters

Embedding analytics transforms static applications into interactive data experiences. In this model, the “Host Application” (your website or software) acts as the container. It controls the navigation, the layout, and the user authentication, while Power BI supplies the visual intelligence engine behind the scenes. This combination allows organizations to offer sophisticated dashboards as part of their digital ecosystem without exposing their internal infrastructure or requiring users to understand Power BI.

Core Benefits of Embedding To understand why developers choose Embedded over standard sharing, we look at three key advantages:

- ◆ **Cost-Effective for Large Scale Audiences:** Licensing 10,000 external users with individual Power BI Pro licenses is often financially unviable. Power BI Embedded solves this by shifting to a **capacity-based model** (Azure A-SKUs). You pay for the computing power of the server, not the number of users. This makes it highly scalable and cost-effective for applications with large or fluctuating user bases.
- ◆ **Allows Full Branding and Customization (White Labeling):** External clients should feel like they are using *your* product, not Microsoft’s. Embedded allows for “White-Labeling,” meaning you can strip away the Power BI branding. You can customize colors, fonts, and behaviors so the charts blend seamlessly with your application’s UI, maintaining a consistent brand identity.
- ◆ **Does Not Require Each User to Have a Power BI License:** This is the “App Owns Data” scenario. The end user (e.g., your customer) does not need a Power BI account or an Office 365 login. Your application handles the authentication in the background using a Service Principal. This removes friction; the user simply logs into your portal and sees the data immediately.

2.4.4.2 Authentication Methods for Embedded Analytics

Embedding analytics into an external application requires a robust security handshake. You cannot simply “open the door” to your data; you need a system that verifies identity without exposing your internal network. Power BI supports several authentication models, each designed for specific deployment patterns.

1. Service Principal Authentication (Recommended) This is the industry standard for production environments, particularly in the “App Owns Data” scenario (where your customers view dashboards without Power BI licenses).

- ◆ **How it works:** Instead of using a human username and password, you register your application in Azure Active Directory (Azure AD). The application itself becomes the “user.”
- ◆ **Why it is superior:** It eliminates dependency on a specific employee’s account. If the developer leaves the company or changes their password, the application does not break. It allows for fully automated, headless authentication.

2. Managed Identities Managed Identities are an evolution of the Service Principal, designed specifically for applications hosted within Microsoft Azure (e.g., an Azure Web App or Function).

- ◆ **How it works:** Azure automatically manages the identity of the resource.
- ◆ **Why it is superior:** It removes the need for developers to manage or rotate secrets and credentials in the code. The security handshake happens internally between Azure services, significantly reducing the risk of credential theft.

3. Master User Authentication (Legacy) In this older model, a real Power BI Pro account (e.g., admin@company.com) is hard coded into the application to handle authentication.

- ◆ **The Problem:** This method is considered “Legacy” and is generally discouraged. It poses significant security risks because it often requires disabling Multi Factor Authentication (MFA) for that account to work automatically. Furthermore, it creates a single point of failure; if the account is locked, the entire embedded analytics platform goes down.

4. Embed Tokens While the methods above determine *who* connects to Power BI, the **Embed Token** determines *what* they can see.

- ◆ **The Concept:** When an external user logs into your portal, your application generates a unique, temporary “key” called an Embed Token.
- ◆ **Why it is critical:** This token is valid only for a short time (e.g., 1 hour) and grants access strictly to the specific report required. It ensures that even if a malicious actor intercepted the token, they could not access the wider workspace or underlying datasets. It effectively builds a firewall between the external user and your internal permissions.

2.4.5 On-Premise Data Gateway

While the Power BI Service excels in the cloud, the reality is that many organizations operate in a hybrid world. Critical data often resides in on-premises databases like SQL

Server, Oracle, or even local Excel files due to security policies or legacy infrastructure.

The **On-Premise Data Gateway** acts as the secure, encrypted bridge connecting these two worlds. It ensures that sensitive data stays safely inside the corporate network firewall while still being accessible for cloud analytics and dashboard refreshes. Without a gateway, the Power BI Service effectively has no way to “reach down” and see your local data.

2.4.5.1 Gateway Architecture - A Secure Exchange Mechanism

The gateway’s architecture is designed to prioritize security. It does not open inbound ports (which would be a security risk). Instead, it relies on outbound connections to the Azure Service Bus.

As illustrated in **Fig 2.4.2 Gateway Architecture**, the communication flow operates as a constant negotiation:

1. **The Request:** The Power BI Service (Cloud) generates a query; for example, a scheduled refresh triggers a request for the latest sales numbers.
2. **The Bridge:** The Gateway Cloud Service queues this request. The On-Premise Data Gateway (installed on your local server) constantly checks this queue.
3. **The Execution:** When the local gateway sees the pending request, it securely retrieves the information from the local source (e.g., the SQL Server shown in the diagram).
4. **The Return:** The data is compressed, encrypted, and sent back up to the cloud to update the report.

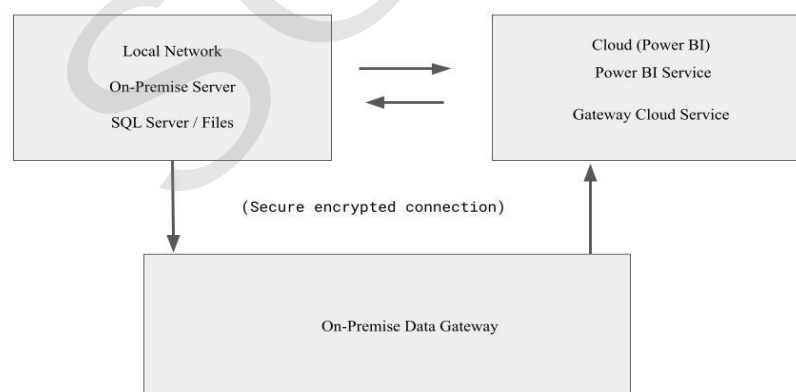


Fig 2.4.2 Gateway Architecture

This architecture ensures that the “Local Network” remains isolated from the public internet, connected only via this secure, encrypted channel.

2.4.5.2 Types of Gateways

When deploying a gateway, you must choose between two modes depending on your operational needs:

◆ **Standard Mode (Enterprise Gateway):**

- **Target Audience:** Enterprise IT and Teams.
- **Capabilities:** It supports multiple users and connects to multiple services beyond just Power BI (including PowerApps, Power Automate, and Azure Logic Apps).
- **Recommendation:** This is the standard for production environments. It should be installed on a dedicated server that runs 24/7 to ensure reports never fail to refresh.

◆ **Personal Mode:**

- **Target Audience:** Individual analysts.
- **Capabilities:** It runs as an application on a single user's laptop and supports only Power BI.
- **Limitation:** If the user's laptop is turned off or falls asleep, the gateway goes offline, and data refreshes will fail. It is strictly for testing or personal analysis.

2.4.5.3 Gateway Configuration Steps

Setting up a Standard Gateway is a critical administrative task. The process transforms the cloud-to-local connection into a reliable component of the analytics pipeline:

1. **Download and Install:** Download the Standard Gateway installer and run it on a dedicated on-premise server (avoid installing on personal laptops for production use).
2. **Sign In:** Authenticate using your **Azure Active Directory (Office 365)** credentials. This links the physical server to your organization's cloud tenant.
3. **Register the Gateway:** Assign a unique, descriptive name (e.g., "NY-Finance-Server") and create a **Recovery Key**.
 - *Note:* The Recovery Key is crucial; if the server fails, you cannot restore the gateway connection on a new machine without it.
4. **Add Data Sources:** Navigate to the Power BI Service online ([Manage Connections and Gateways](#)). Add specific data sources (e.g., "SQL Server Production DB") to the newly created gateway cluster.
5. **Map Credentials:** Enter the username and password that the gateway will use to log into the local database. This ensures the gateway has permission to read the data.

Enable Scheduled Refresh: Go to the specific dataset settings in the Power BI Service and toggle on “Gateway Connection.” You can now set up automated refresh times (e.g., Daily at 9:00 AM).

2.4.6 Deployment Scenarios

Choosing the right deployment strategy is as critical as building the report itself. The decision dictates cost, performance, and how users access the data. Power BI offers four primary deployment models to suit different organizational needs.

1. Power BI Pro

The Standard for Self-Service BI Power BI Pro is the entry-level license for professional sharing. It operates on a “Per User” model, meaning every person who creates a report *and* every person who views that report must have a paid Pro license.

- ◆ Suitable for : Small to medium-sized teams where the number of users is manageable.
- ◆ Key Constraint : It creates a “paywall.” You cannot share a report with a free user; both the sender and the receiver need a license.
- ◆ Performance : Hosted on shared capacity (resources are shared with other Microsoft customers).

2. Power BI Premium

The Enterprise Solution Power BI Premium shifts the model from “Per User” to “Capacity Based.” Instead of buying licenses for individuals, the organization purchases dedicated hardware resources (Capacity) in the Microsoft Cloud.

- ◆ Unlimited Distribution : The defining feature of Premium is that it allows “Free” users to view reports. If you have 5,000 employees who only need to *read* reports, you don’t need 5,000 licenses, you just need Premium capacity.
- ◆ Advanced Features : It unlocks AI capabilities, Paginated Reports, and faster refresh rates (up to 48 times a day).
- ◆ Suitable for : Large enterprises with a massive user base or strict performance requirements.

3. Power BI Embedded

Designed specifically for ISVs (Independent Software Vendors) and developers, this model allows you to embed Power BI visuals directly into your own custom applications, websites, or portals.

- ◆ White-Labeling : The end user might not even know they are looking at Power BI. The branding, colors, and navigation are controlled entirely by your application.
- ◆ Cost Efficiency : It uses Azure A-SKUs (pay-as-you-go). You pay for the service uptime, not for the number of users accessing the app.

4. Power BI Report Server

The On-Premises Alternative For organizations restricted by strict regulatory compliance or data sovereignty laws that prevent them from using the public cloud, Power BI Report Server offers a solution.

- ◆ How it works: It is a localized server installed on the company's own hardware.
- ◆ Trade-offs: While it keeps data entirely within the firewall, it lacks many cloud-exclusive features (like Dashboards, Q&A, and frequent monthly updates).

The Embedded Architecture Flow

Understanding how these components interact in an embedded scenario is crucial for developers. The integration of analytics into custom applications is not a simple drag-and-drop; it requires a specific architectural flow involving the Master Account (or Service Principal), the Azure AD token, and the rendering engine.

This entire end-to-end process from the initial .pbix file creation to the final secure view within a custom app is visually detailed in Fig 2.4.3.

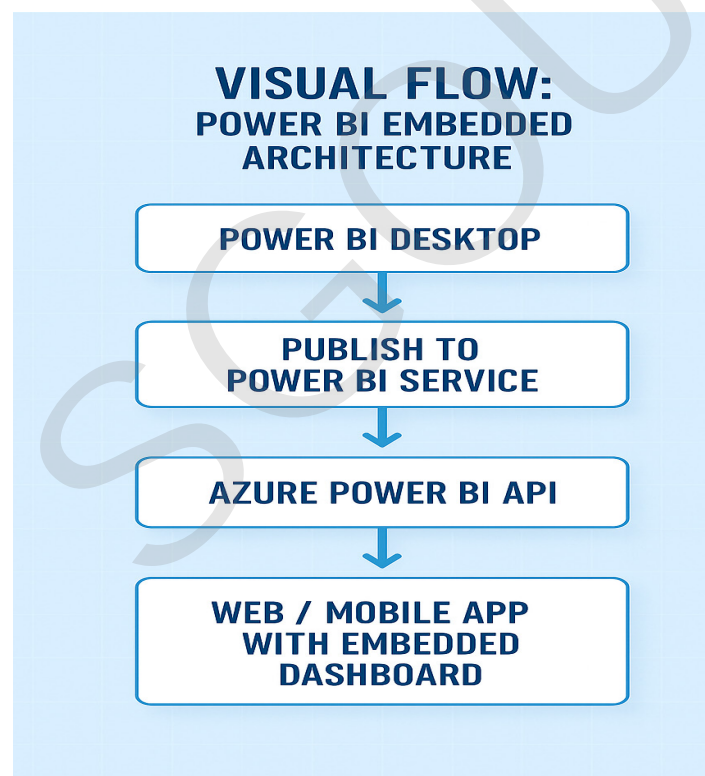


Fig 2.4.3 Power BI Embedded Architecture Visual Flow

Recap

- ◆ Power BI deployment involves publishing and managing reports within Power BI Service.
- ◆ Workspaces are organizational containers where datasets, reports, and dashboards are stored and managed.
- ◆ Row-Level Security (RLS) ensures that users can only access the portion of data relevant to their role.
- ◆ The On-Premise Data Gateway enables secure connectivity between cloud-based Power BI Service and on-premise databases.
- ◆ Power BI Embedded allows integration of Power BI visuals into external applications.
- ◆ Data refresh schedules ensure that dashboards display up-to-date information.
- ◆ Permissions and roles determine who can access, edit, or publish content within a workspace.
- ◆ Secure deployment ensures proper governance and protects sensitive organizational data.

Objective Questions

1. Power BI Service is mainly used for:
2. Which component securely connects on-premise data sources to Power BI Service?
3. Row-Level Security (RLS) is used to:
4. A workspace role that allows users to view but not edit content is:
5. What is the purpose of Scheduled Refresh in Power BI Service?
6. Power BI Embedded is used mainly for:
7. Object Level Security (OLS) can restrict access to:
8. Which workspace role has full administrative control?
9. What does the Gateway require to function properly?
10. A user needs to share a dashboard but avoid allowing edits. Which sharing method is appropriate?

11. When using DirectQuery, data is:
12. The Power BI Admin Portal is mainly used for:
13. A company needs to control who can publish content inside a workspace. What should they configure?
14. Which scenario requires using Power BI Report Server?
15. A developer wants to give customers access to reports inside a web portal without requiring Power BI licenses. Which deployment mode is ideal?

Answers

1. Publishing, sharing, and managing reports
2. Gateway
3. Restrict data visibility for specific users
4. Viewer
5. To update visuals automatically at specified times
6. Embedding analytics inside external applications
7. Entire tables or columns
8. Admin
9. Installation on an on-premise machine
10. Share with “View only” permission
11. Queried in real-time from the source
12. Managing organizational settings and governance
13. Workspace roles
14. When the organization cannot use cloud services
15. Power BI Embedded

Assignments

1. Explain the concepts of Row-Level Security (RLS) and Object-Level Security (OLS) in Power BI.

Provide examples of how each security method can be applied in an organizational environment to protect sensitive data.

2. Install and configure an On-Premise Data Gateway.

Document the steps involved, explain the role of the gateway in connecting local data sources to the Power BI Service, and discuss common issues faced during configuration.

3. Publish a Power BI report to the Power BI Service and configure data refresh schedules.

Describe how scheduled refresh works, the limitations based on licensing, and how refresh failures can be identified and resolved.

4. Compare different Power BI deployment scenarios: Power BI Pro, Power BI Premium, Power BI Embedded, and Power BI Report Server.

Discuss the features, advantages, and suitable use cases for each deployment mode in an academic or enterprise context.

5. Create a workspace in Power BI Service and assign roles to different users.

Explain the permissions associated with Viewer, Contributor, Member, and Admin roles, and discuss how proper role assignment ensures governance and security.

Reference

- ◆ <https://learn.microsoft.com/power-bi/collaborate-share/>
- ◆ <https://learn.microsoft.com/power-bi/admin/service-admin-rls/>
- ◆ <https://learn.microsoft.com/power-bi/connect-data/service-gateway-onprem/>

Suggested Reading

1. Microsoft Documentation – Power BI Service Overview
<https://learn.microsoft.com/power-bi/service-overview>
2. Microsoft Documentation – Row-Level Security (RLS)
<https://learn.microsoft.com/power-bi/admin/service-admin-rls>
3. Microsoft Documentation – Power BI Gateway
<https://learn.microsoft.com/power-bi/connect-data/service-gateway-onprem>
4. SQLBI – Articles on Data Security and Power BI Governance
<https://www.sqlbi.com>
5. Book : Mastering Microsoft Power BI – Brett Powell
(Excellent for understanding deployment, governance, and enterprise-grade security)



SREENARAYANAGURU OPEN UNIVERSITY

QP CODE:

Reg. No. :

Name:

THIRD SEMESTER DEGREE EXAMINATION
BSc DATA SCIENCE AND ANALYTICS
SKILL ENHANCEMENT COURSE
B24DS01SE - DATA VISUALIZATION (CBCS - UG)
2024-25 - Admission Onwards
MODEL QUESTION PAPER- SET A

Time: 3 Hour

Max Marks: 70

SECTION A

*Answer any **ten** questions of the following. Each question carries **one** mark.*

(10 × 1 = 10 Marks)

1. Who is considered the father of statistical graphics?
2. Which Gestalt principle denotes items seen as related because they are close together?
3. When presenting to top-level executives, visualizations should primarily _____.
4. What is the primary goal of a dashboard?
5. Write any two key differences between Power BI and Google Data Studio.
6. Which feature in Power Query Editor helps identify missing values, errors, and inconsistencies?
7. Which feature supports navigation between multiple pages in a report?
8. Drill-through is used to.....
9. Write about On-Premise Data Gateway.
10. What is the main advantage of interactive dashboards?
11. Which table type typically contains numeric and aggregatable values?
12. Which chart is commonly used to visualize the relationship between two numerical variables?

13. Which view in Power BI Desktop allows users to see loaded tables in tabular form?
14. What are real-time dashboards mainly used for?
15. What is the term for the mental effort required to process a visualization?

SECTION B

*Answer any **five** questions of the following. Each question carries **two** marks.*

(5×2 =10 Marks)

16. What is the primary purpose of data visualization?
17. What is the significance of pre-attentive attributes in data visualization?
18. Define Power of Alignment.
19. Mention any two dashboard design principles and explain briefly.
20. Write about the importance of data preparation.
21. Why is Data Modeling Important?
22. What is the difference between *Drill Down* and *Drill Through* in Power BI?
23. What is Visual Reporting?
24. Draw the gateway architecture.
25. What is a cluttered dashboard? Why is it considered a major visualization pitfall?

SECTION C

*Write a short note on any **five** questions of the following.
Each question carries **four** marks.*

(5×4 = 20 Marks)

26. Describe how data visualization transforms raw data into meaningful insights.
27. Explain Tufte's data-ink ratio and its importance in avoiding unnecessary visual clutter.
28. Describe the stakeholder visual adaptation strategies and its supporting tools.
29. Describe any four common pitfalls in data visualization and explain how they can be avoided.
30. What is Power Query Editor?. Write about Data Cleaning using Power Query Editor.

31. Explain the different types of Power BI Model.
32. Discuss the storytelling framework in Power BI
33. Explain the Power BI Service Deployment Process.
34. Explain the importance of choosing appropriate data types and scales in visualization.
35. Explain the concept of storytelling with data.

SECTION D

*Answer any **two** questions of the following. Each question carries **fifteen** marks.*

(2×15 =30 Marks)

36. Discuss the fundamentals of Data Visualization in detail
37. “Effective data visualization depends on understanding the audience and using storytelling techniques.” Explain this statement with reference to different stakeholders.
38. Discuss in detail the various types of data used in visualization (such as categorical, numerical, ordinal, nominal, discrete, and continuous).
39. Explain the process of creating effective visualizations in Power BI using charts, graphs, maps, KPIs, slicers, filters, drill-through, drill-down, and custom visuals. Discuss how each visualization type enhances data analysis and supports decision-making, using suitable examples.



SREENARAYANAGURU OPEN UNIVERSITY

QP CODE:

Reg. No. :

Name:

THIRD SEMESTER DEGREE EXAMINATION
BSc DATA SCIENCE AND ANALYTICS
SKILL ENHANCEMENT COURSE
B24DS01SE - DATA VISUALIZATION (CBCS - UG)
2024-25 - Admission Onwards
MODEL QUESTION PAPER- SET B

Time: 3 Hour

Max Marks: 70

SECTION A

*Answer any **ten** questions of the following. Each question carries **one** mark.*

(10 × 1 = 10 Marks)

1. What is the role of aesthetics in data visualization?
2. Which perceptual pillar uses size to represent magnitude?
3. A method of presenting data as a meaningful story is known as _____.
4. What type of dashboards allow users to interact and explore data dynamically?
5. Write any two components of Power BI.
6. Which connector would you use to bring data from an Excel file into Power BI?
7. Which option allows users to sort visual data by specific fields?
8. Which workspace role has full administrative control?
9. Write any two points which highlights the importance of data modelling.
10. What is the purpose of AI-assisted charting?
11. The language used to create formulas and expressions in Power BI is called _____.
12. Which chart is used to show changes in data over continuous time?
13. Charles Minard is famous for which historical visualization?
14. write one emerging trend in data visualization.
15. Which Gestalt principle distinguishes the object of focus from the background?

SECTION B

*Answer any **five** questions of the following. Each question carries **two** marks.*

(5×2 =10 Marks)

16. Why are aesthetics considered an important component of data visualization?
17. Explain the term chartjunk as described in visual design principles.
18. What is meant by “tailoring visualizations to different stakeholders”? Explain with an example.
19. What is visual hierarchy, and how does it help users interpret dashboards?
20. Write any two differences between Power BI and Tableau.
21. Write about DAX.
22. Why is visual reporting important in Power BI?
23. What is Storytelling with Data
24. Write about Drill Down and Drill Through.
25. Write any two types of Power BI Gateways.

SECTION C

*Write a short note on any **five** questions of the following.
Each question carries **four** marks.*

(5×4 = 20 Marks)

26. Explain the real-world applications of data visualization?
27. Discuss the concept of cognitive load and outline strategies to reduce it in visualization design.
28. Discuss the major components of data storytelling. How do narrative structure, visuals, and insights work together to convey a message?
29. What is meant by a “cluttered dashboard”? Explain its effects and how it can be avoided.
30. What is Data Preparation?. Write the importance of data preparation.
31. Discuss the major relationships in Power BI.
32. Explain the importance of choosing the right chart type in Power BI?
33. What is Power BI Gateway?Write its advantages.

34. Illustrate how visual perception principles influence the interpretation of graphs and dashboards, with suitable examples.
35. Describe the process of tailoring data visualizations for technical and non-technical audiences. How does the style and detail of visuals change?

SECTION D

*Answer any **two** questions of the following. Each question carries **fifteen** marks.*

(2×15 =30 Marks)

36. Discuss in detail the cognitive and perceptual foundations of effective data visualization
37. Describe the principles of dashboard design and explain why clarity and simplicity are essential for effective data communication.
38. Explain Power BI and its components. Describe the process of data preparation in Power BI, including data cleaning, transformation, and integration. Also, discuss the benefits of proper data preparation.
39. Discuss in detail the various security features available in Power BI, including data protection, access control, row-level security, workspace roles, and data governance. Explain how these features ensure secure deployment and usage of Power BI reports and dashboards.

സർവ്വകലാശാലാഗീതം

വിദ്യാതൽ സ്വതന്ത്രരാകണം
വിശ്വപൗരരായി മാറണം
ഗ്രഹപ്രസാദമായ് വിളങ്ങണം
ഗുരുപ്രകാശമേ നയിക്കണം

കുരിശുട്ടിൽ നിന്നു ഞങ്ങളെ
സൂര്യവീഥിയിൽ തെളിക്കണം
സ്നേഹദീപ്തിയായ് വിളങ്ങണം
നീതിവൈജയന്തി പാറണം

ശാസ്ത്രവ്യാപ്തിയെന്നുമേകണം
ജാതിഭേദമാകെ മാറണം
ബോധരശ്മിയിൽ തിളങ്ങുവാൻ
ജ്ഞാനകേന്ദ്രമേ ജ്വലിക്കണം

കുരീപ്പുഴ ശ്രീകുമാർ

SREENARAYANAGURU OPEN UNIVERSITY

Regional Centres

Kozhikode

Govt. Arts and Science College
Meenchantha, Kozhikode,
Kerala, Pin : 673002
Ph: 04952920228
email:rckdirector@sgou.ac.in

Thalassery

Govt. Brennen College
Dharmadam, Thalassery
Kannur, Pin : 670106
Ph: 04952990494
email:rctdirector@sgou.ac.in

Tripunithura

Govt. College
Tripunithura, Ernakulam
Kerala, Pin : 682301
Ph: 04842927436
email:rcdirector@sgou.ac.in

Pattambi

Sree Neelakanta Govt. Sanskrit College
Pattambi, Palakkad
Kerala, Pin : 679303
Ph: 04662912009
email:rcddirector@sgou.ac.in

**DON'T LET IT
BE TOO LATE**

**SAY
NO
TO
DRUGS**

**LOVE YOURSELF
AND ALWAYS BE
HEALTHY**



SREENARAYANAGURU OPEN UNIVERSITY

The State University for Education, Training and Research in Blended Format, Kerala

Data Visualization

COURSE CODE :B24DS01SE



YouTube



Sreenarayanaguru Open University

Kollam, Kerala Pin- 691601, email: info@sgou.ac.in, www.sgou.ac.in Ph: +91 474 2966841