

HOW A SEMANTIC LAYER CAN HELP YOU TELL BETTER STORIES WITH DATA

Ganes Kesari is an entrepreneur, AI thought leader, author, and TEDx speaker

Ganes Kesari is an entrepreneur, AI thought leader, author, and TEDx speaker. He co-founded Gramener, where he heads Data Science Advisory and Innovation. He advises executives on decision-making with data.



Why Data Storytelling?

Perhaps it's cliché to say that we [generate](#) 2.5 quintillion bytes of data every day. It's anybody's guess how many reports we generate with all this data every minute. Business intelligence (BI) reports and dashboards were once viewed as a savior for business users drowning in data.

Today, users are drowning in dashboards.

Dashboards have turned into glorified tabular reports. Yes, many have fancy visuals with striking colors and great interactivity, but most dashboards don't make it easy to consume the data. They are not insightful or actionable, so they score very low on memorability.

Can you recall a powerful dashboard you saw at work last month? On the contrary, I bet that you can remember a dozen stories from your childhood – stuff that you read, watched on the screen, or just heard from others.

Stories are powerful. Data stories bring insights to life. “Numbers have an important story to tell,” said Stephen Few, a data visualization expert. However, “they rely on you to give them a clear and convincing voice,” he added.

Just as good children's stories have a moral for the readers to take away, good data stories have an insight that the audience can act upon. Data stories pack insights from data and deliver them in a way that engages the audience and moves them to action. The intent of data storytelling is to address an organization problem by recommending a prescriptive action.

There's no better way to understand the impact of a data story than to look at a real-world example.

A Data Story on Innovation, Technology, and Entrepreneurship

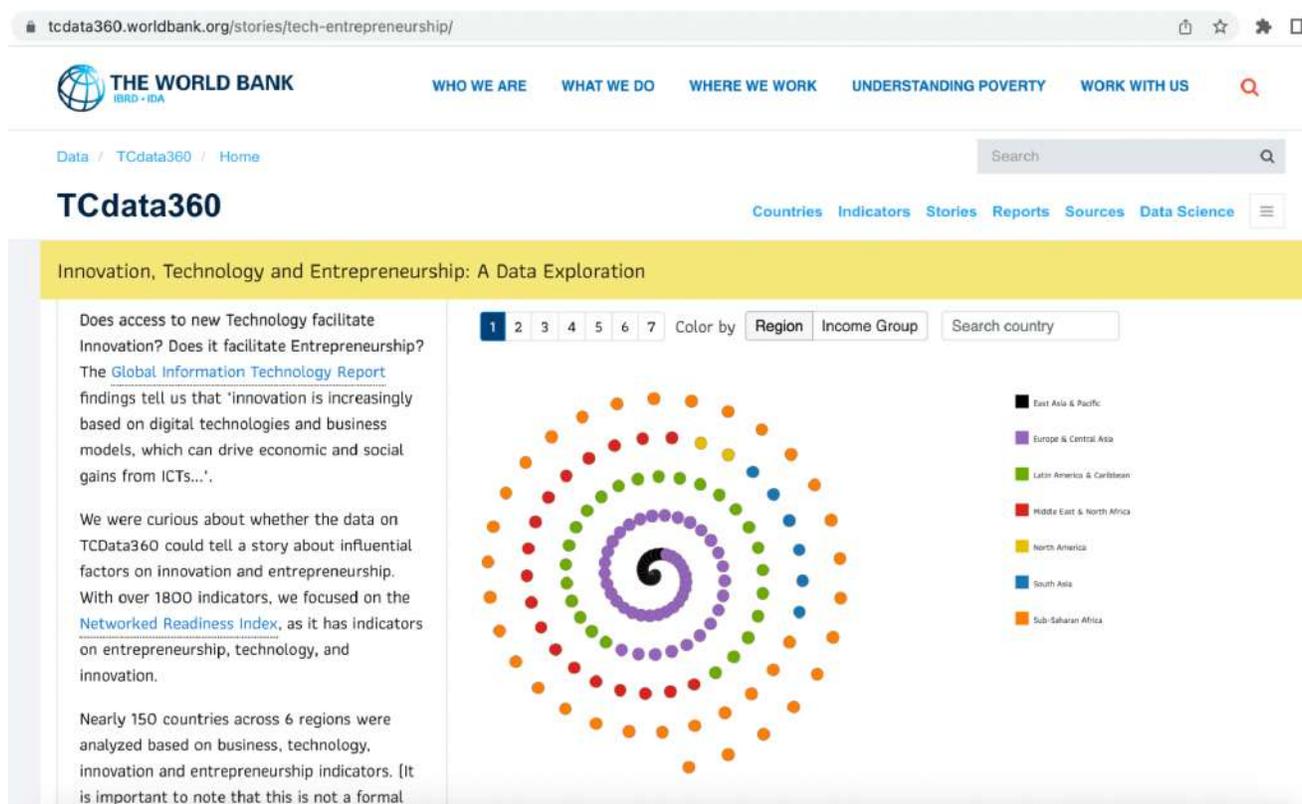
Which countries have an environment that fosters innovation and entrepreneurship? How does the presence or absence of technology infrastructure impact innovation?

The World Economic Forum's (WEF) Global Information Technology [report](#) collects data from around 150 countries to help answer such questions. Similarly, reports from the World Bank's Doing Business project and WEF's global competitiveness index publish hundreds of valuable indicators.

However, insights from such reports don't reach enough people in the public domain.

The World Bank decided to apply the power of data storytelling to change that by launching [TCdata360](#) in 2017, an initiative to enable data consumption by the general public. The team partnered with Gramener to create the following interactive [data story](#) as part of this project.

The story sets the context by sharing how it will answer some critical questions by analyzing 1800+ indicators across 150 countries.



(Picture - World Bank Data Story: Panel 1 - Setting the context)

For example, it shows how Malaysia is the only outlier in the group of countries that are 'Most Favorable' to technology innovation. None of the upper-middle income group countries feature here.

Innovation, Technology and Entrepreneurship: A Data Exploration

The Peer Groups

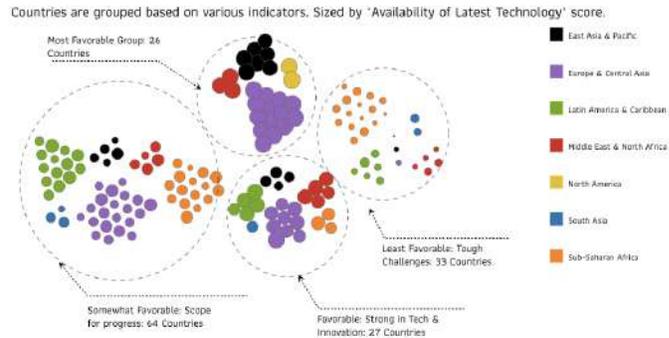
We analyzed 150 Countries based on scores generated by Business **1**, Technology **1** & Innovation **1** indicators, then used K-means clustering to group countries into four peer groups.

- Most Favorable: Leaders in Business, Technology and Innovation
- Favorable: Strong in Technology & Innovation; Average in Business
- Somewhat Favorable: Room for improvement in Business, Technology and Innovation
- Least Favorable: Tough challenges across indicators

Malaysia is the only upper middle income country **1** in the 'Most Favorable' group. Rwanda, is the only low income Sub Saharan country in the 'Favorable' group. Technology & Innovation is impacted by weak business environment in 'Somewhat Favorable' group. Kyrgyz Republic is the only Europe & Central Asian country in 'Least Favourable' group.

[Read More ...](#)

1 2 3 4 5 6 7 Color by **Region** Income Group Search country



(Picture - World Bank Data Story: Panel 2 - Grouping countries based on performance)

What drives innovation and makes the favorable and most favorable groups stand out? An evaluation of relationships across parameters suggests that countries with better availability of latest technologies, investment in research & development (R&D), and higher collaboration between industry and universities stand out on innovation.

Innovation, Technology and Entrepreneurship: A Data Exploration

Innovation is enabled by access to latest technology and reliance on professional management

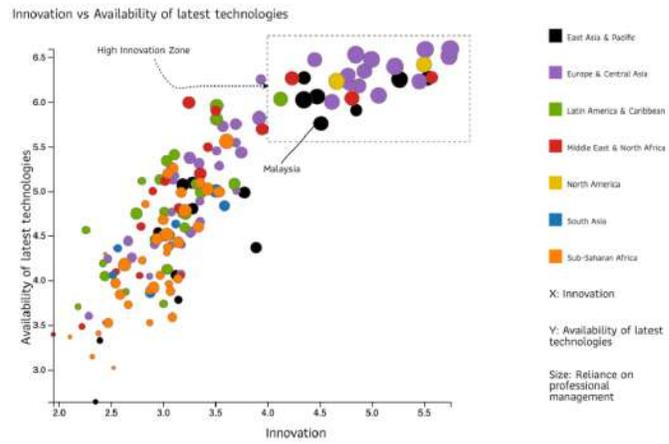
The high innovation zone features high investment in research and development (R&D) with higher collaboration with industry and universities.

Malaysia is the only upper middle income country in this zone.

High innovation countries are characterised by:

- High Income
- High Availability of Technology
- High Reliance on Professional Management
- High Innovation

1 2 3 4 5 6 7 Color by Region Income Group Search country



(Picture - World Bank Data Story: Panel 3 - Relationship between Innovation and Availability of latest technologies)

Further, these countries have fostered a business environment that enables innovation. This includes conditions such as an ease of doing business (minimal red tape), presence of a skilled work force, and availability of venture capital for financing projects. The countries that fall into the zone are highlighted below.

Innovation, Technology and Entrepreneurship: A Data Exploration

Innovation is aligned to the strength of business and innovation environment in most High Income Countries

The highlighted zone features higher scores in Business & Innovation Environment to support entrepreneurship.

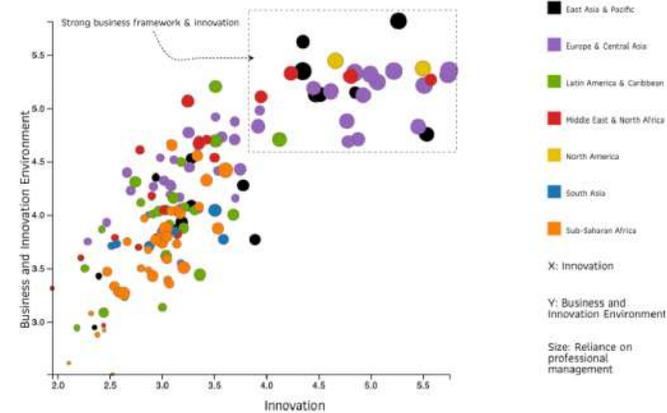
Developed **East Asia & Pacific countries** - Hongkong, Malaysia, Singapore, Japan & New Zealand have supporting business grounds for Innovation. Puerto Rico is the only economy in **Latin America and Caribbean** with stronger business framework.

This group is characterized by:

- High Income
- High Innovation
- High Business & Innovation Environment

1 2 3 4 5 6 7 Color by **Region** Income Group Search country

Innovation vs Business and Innovation Environment

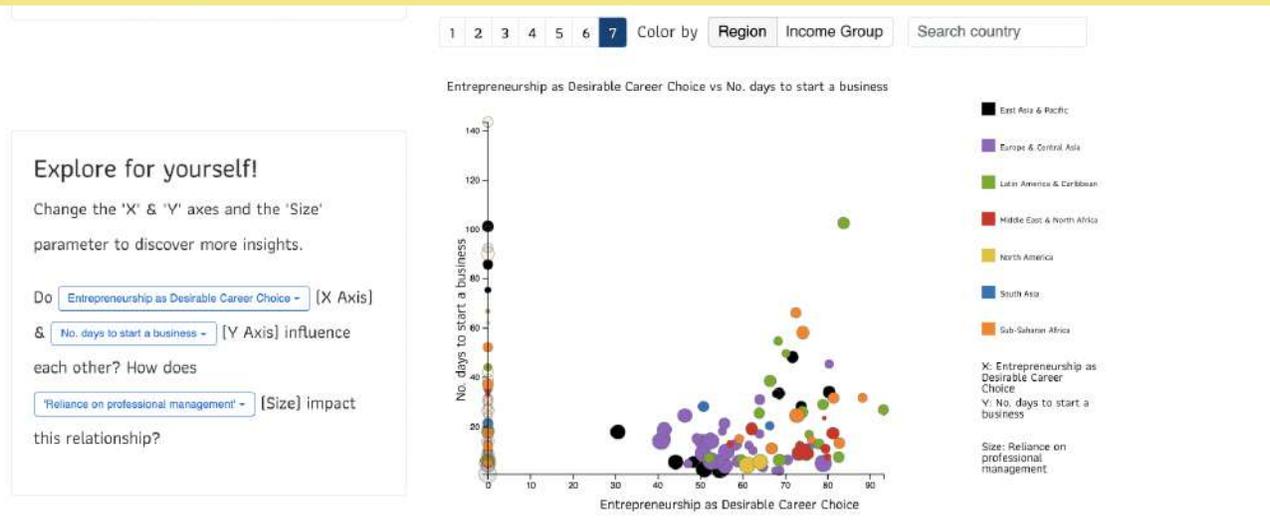


(Picture - World Bank Data Story: Panel 4 - How a favorable Business and Innovation environment drives Innovation)

In summary, the countries that score high on innovation have good availability of the latest technologies, better investment in R&D, higher collaboration between industries and universities, and an environment that makes it easy to do business. This suggests clear actions for the developing countries that aspire to foster innovation and entrepreneurship.

We can see how this interactive visual story establishes a clear context, uses engaging visuals to highlight data insights, and drives home the takeaways irrespective of user background. Such insights, narrated in a scrollytelling format (scrolling + storytelling), attract the audience's attention. Once users gain familiarity with the framework to understand insights, the final story panel encourages them to discover their own stories.

Innovation, Technology and Entrepreneurship: A Data Exploration



(Picture - World Bank Data Story: Panel 7 - Enabling self-discovery by users)

Demystifying Data Stories

We've seen the power of data storytelling with public data. But what's the biggest roadblock to creating them at enterprises? The first challenge is the lack of understanding of a data story.

Given the high public interest in data stories and the marketability of this phrase, the term is overused in the industry. It is often tossed around out of context because people use it interchangeably with data visualizations, dashboards, or infographics.

Harvard Business School [defines](#) data storytelling as “the ability to effectively communicate insights from a dataset using narratives and visualizations. It can be used to put data insights into context for and inspire action from your audience.” (emphasis mine)

The outcome targeted with data stories is the effective communication of data insights to inspire user action. Stories achieve this by contextualizing the insights and communicating information visually with powerful narratives.

The Anatomy of a Data Story

What are the ingredients of a data story? There are three elements that make up any strong data story. We'll go through each one of them so that you understand their guidelines and learn how you can craft them using the data available in your organization.

The Three Key Elements of Every Powerful Data Story

Gartner analyst, James Richardson, [breaks](#) data stories into three key elements: visualization, context, and narrative.

- **a) Visualization** is the first element that helps convert insights identified from the underlying data into visual charts. This is achieved by employing information design principles to choose the right chart, colors, and interactivity.
- **b) Context** helps the audience relate to the story by adding organizational background and explaining the insights and takeaways. This calls for solid business expertise and a mastery of the domain workflows.
- **c) A narrative** is built by laying out the visuals in an engaging sequence to craft a coherent and compelling story that promotes action. This is best done with a clear understanding of the user, priorities, pain points, and preferences.

When these three elements are combined to craft a data story, actionable decisions become a reality. The audience understands the insights in the context of the organizational scenario and the recommended actions leap out of the visual narrative.

In the previous storytelling example, you would notice the use of pertinent and robust visualizations – for example, distribution charts, bubble charts, and x-y plots. Notice how background context, highlights, and takeaways are peppered throughout every story panel to aid user consumption. Finally, the visuals were stitched into a seven-panel narrative to surface recommendations and inspire action – these constituted the beginning, middle, and end of our data story.

Every memorable data story you might have engaged with in the past would have these three elements. Let's now look at some best practices to create these elements.

a) Visualization: Converting Data Insights Into Engaging Visuals

The genesis of every data story is the audience and their unmet needs. One must start by understanding the users and their organizational roles. Define the business problem to be solved based on the user's requirements. When there is sufficient clarity on the issue, you can evaluate the likely approaches to address it using technology.

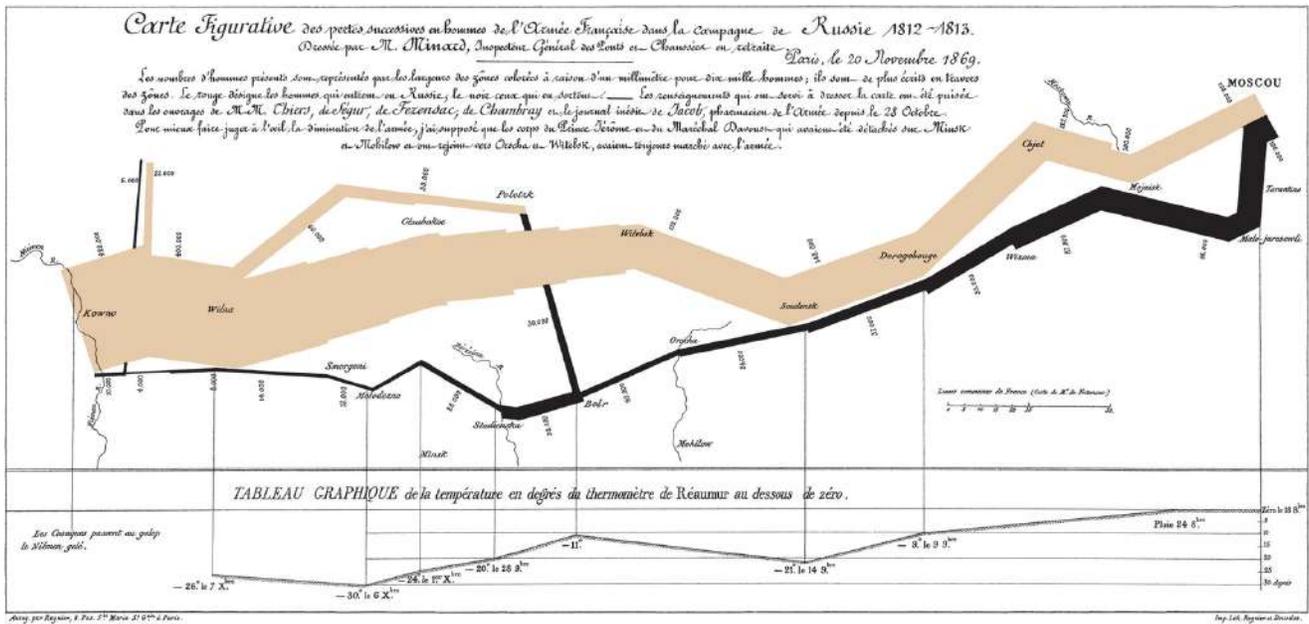
This is the stage when the need for data and analytics becomes evident. When a data-driven approach is the best way to solve the business problem, figure out the data you need. This data must then be collected, transformed, and cleaned to prepare for analytics.

Identifying insights through analytics is all about answering a series of business questions. Sharp and well-framed business questions help you discover hidden patterns in the data. For this, you'll need access to high-quality data and metrics computed using pertinent business rules.

Depending on whether you must answer questions such as "what happened?", "why did it happen?", or "what will happen?", you can choose the pertinent type of analytics – descriptive analytics, diagnostic analytics, or predictive analytics. Then convert the answers found by these approaches into actionable recommendations, also called prescriptive analytics. This helps answer the question, "what should you do next?".

Data visualization is a technique that helps present the discovered analytical insights into visual charts. But how do you pick the right visual for a given insight? [Several](#) chart picker guides can help. [The Data Visualization Catalogue](#) has put together a nifty reference that recommends visuals based on the kind of insight pattern you want to show.

One of the finest public examples of data visualization is the chart created by [Charles Joseph Minard](#), a French engineer, in 1869. He recreated Napoleon Bonaparte's disastrous invasion of Russia in 1812. Over six data types are presented as charts, such as the Sankey flow diagram and line chart. Created for consumption by the general public, descriptive insights explain what happened during the campaign. Visual design guru, Edward Tufte, considers this to be the best graphic ever created.



(Picture - Napoleon's March by Charles Joseph Minard, [Wikipedia Commons](#))

b) Context: Making Your Story Relatable to the Audience

What do your users need to know before looking at the visualization? What should they notice when they see it to place it in context? What is the takeaway and recommendation to help them make the decision?

This is where context comes into play. When the right context is added to a visualization, it helps address all the above questions. Here are three guidelines for contextualizing your data story:

1. Annotate the key takeaways and anomalies

Focus the reader's attention on key messages through in-place callouts or text boxes. Highlight any outliers or anomalies so that they are easy to spot. Explain the backstory and the top takeaways from the visual to make – not just the insights – but also the potential actions obvious.

There is an element of human judgment involved in choosing the right annotations. Compute all relevant business metrics and then evaluate which of these are most important for your story.

2. Add comparative metrics for benchmarking

Patterns in data are often contextualized only when they are compared against relevant indicators such as organizational or business unit averages, market indices, and industry benchmarks. Ensure you normalize the metrics and standardize units across the visual.

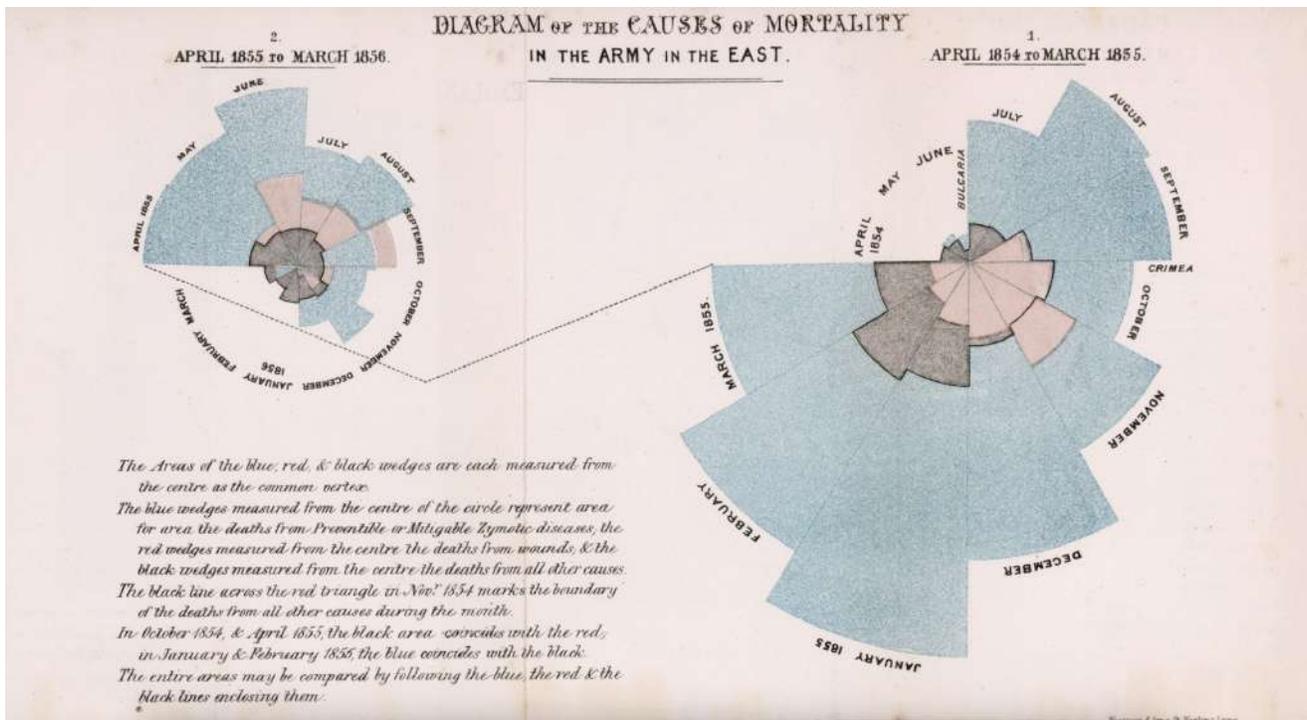
To bring in the right comparisons, you need ready access to all relevant KPIs and metadata across all of your data sources. Keep these handy because you'll reuse them throughout the story.

3. Add sources and footnotes

In an age of big data and real-time insights, users must understand the underlying data used. Ensure you include the data sources, aging of data, and any limitations in data collection. Often, what's excluded from the data is more important than what's included. Finally, clarify the business rules and computations done on the data.

What's often shown in a data story is a fraction of the data and metrics that analysts use. The underlying raw data and business metrics should be correct and consistent to ensure quality and trust in data stories.

For example, [Florence Nightingale](#) created a classic visual story to highlight what led to high mortality during the Crimean War. Prepared for the Members of the Parliament and civil servants with limited ability to understand statistical reports, she used a powerful visual narrated with a strong context to show that more soldiers died due to poor medical care than from battle wounds.



(Picture - Causes of Mortality by Florence Nightingale - [David Rumsey](#), Public Domain, [Wikipedia](#))

c) Narrative: Stitching Your Story Into a Logical Sequence

Visual charts are informative but seldom engaging on their own. To make the data story impactful, one must piece them into a coherent narrative. Every good story has a plot with a beginning, middle, and end.

The beginning sets up the story, explains the problem, and shows why the user must care. This is where the background is sketched out to build context. With descriptive summaries, you show the user what happened in the organization until then.

The middle part of the story details the problem and the likely contributing factors. Certain narrative styles introduce a series of issues and resolutions to build momentum. For example, you could highlight the multiple challenges that led to a drop in sales in the past quarter, the actions taken, and whether they were successful.

The end portion introduces a big disconnect and sets up the climax with a significant decision point. The novel, hidden insights identified are then presented as the resolution with a suggested action. For example, you could show how the entry of a large competitor jeopardized the achievement of the annual revenue targets and what potential measures to take based on the insights.

An excellent public example of a story narrative is the presentation by Hans Rosling on how 200 countries around the world fared over 200 years, using 120,000 numbers. [This video](#) produced by BBC brings together impressive visuals, contextualized with a strong narrative.

Composing a strong narrative with multiple visuals and business metrics calls for the data to be extremely well organized. The KPIs need to be easily accessible and the business metrics should be consistent across the entire multi-page story. For data stories that must be periodically refreshed with new data, the backend data engineering infrastructure must be top-notch. Otherwise, the implementation of data stories can turn into a nightmare.

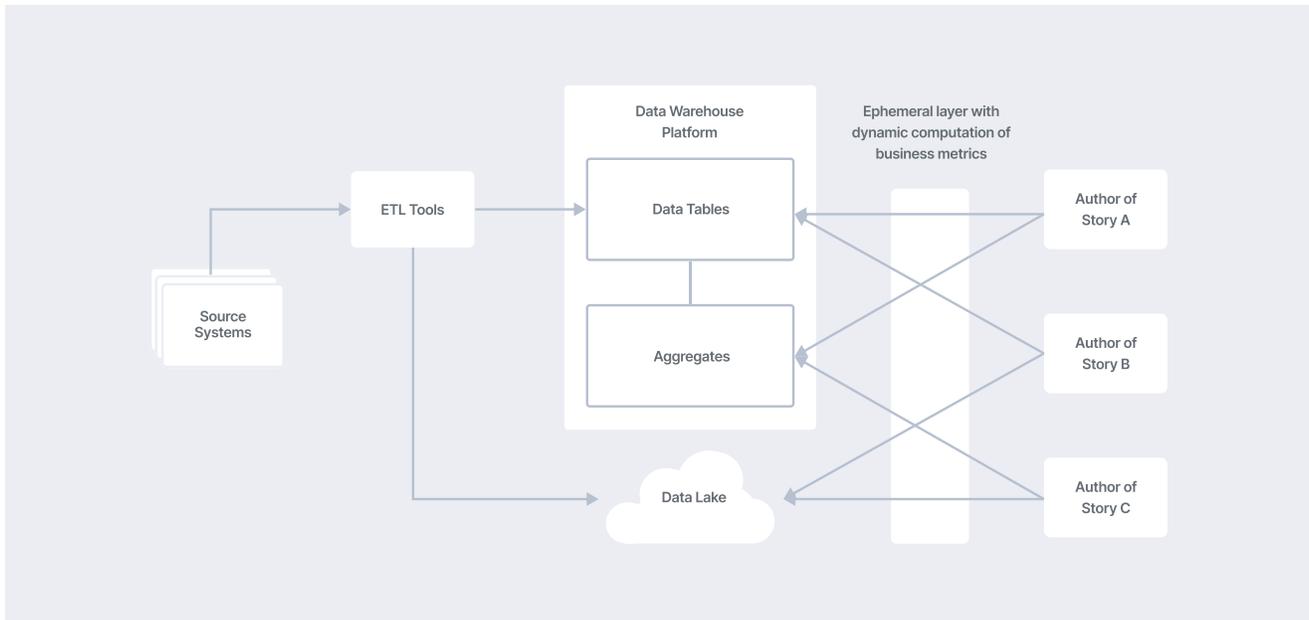
We've seen the ingredients of a data story and how to put them together. Let's go behind the scenes to understand how a semantic layer can help make this process of data story creation efficient, repeatable, and trustworthy.

The Role of a **Semantic Layer** in bringing a Data Story to Life

When companies create data stories, what kind of a backend setup do they usually use to prepare the data, perform analytics, and craft a visual story?

Data in most organizations is found across multiple sources – data warehouses, data lakes, and multiple siloed locations. Analysts tap into data from across these sources ad-hoc, and then use custom analytical techniques to identify insights. The business metrics are computed based on rules and formulae that the analysts manually collect from the various business users they serve.

There is a lot of manual effort and duplication of activities when analysts pull the aggregated data (ex., daily sales transactions), compute the business metrics (ex. monthly revenue), and perform other derived computations (ex. month-over-month sales growth). The illustration below shows such a conventional setup that is non-optimal for story creation.



(Picture - Story creation in a traditional organizational setup)

Three Challenges Analysts Face While Creating Data Stories

With the conventional approach, companies typically run into three common challenges while creating data stories:

a) Difficulty getting the right data from a distributed ecosystem:

With data distributed throughout the enterprise across multiple sources, it's tough to pull all the data required for answering the business questions. Finding novel, actionable insights can remain a pipedream without data of the right granularity, type, and quality. Modern visualizations can elegantly display patterns across thousands of data points. However, they need good volumes of clean data and this is tough to achieve with siloed data.

b) Duplication of efforts in the computation of business metrics:

Even when the data engineering challenge is solved, the trouble at the next layer is the computation of business metrics. There are a few dozen key metrics in organizations that all business units need to compute. However, with hundreds of story authors spread across the organization, they compute and re-compute these metrics for every project. Imagine the magnitude of duplication underway across the business.

c) Inconsistency in insights leading to a lack of trust in data stories:

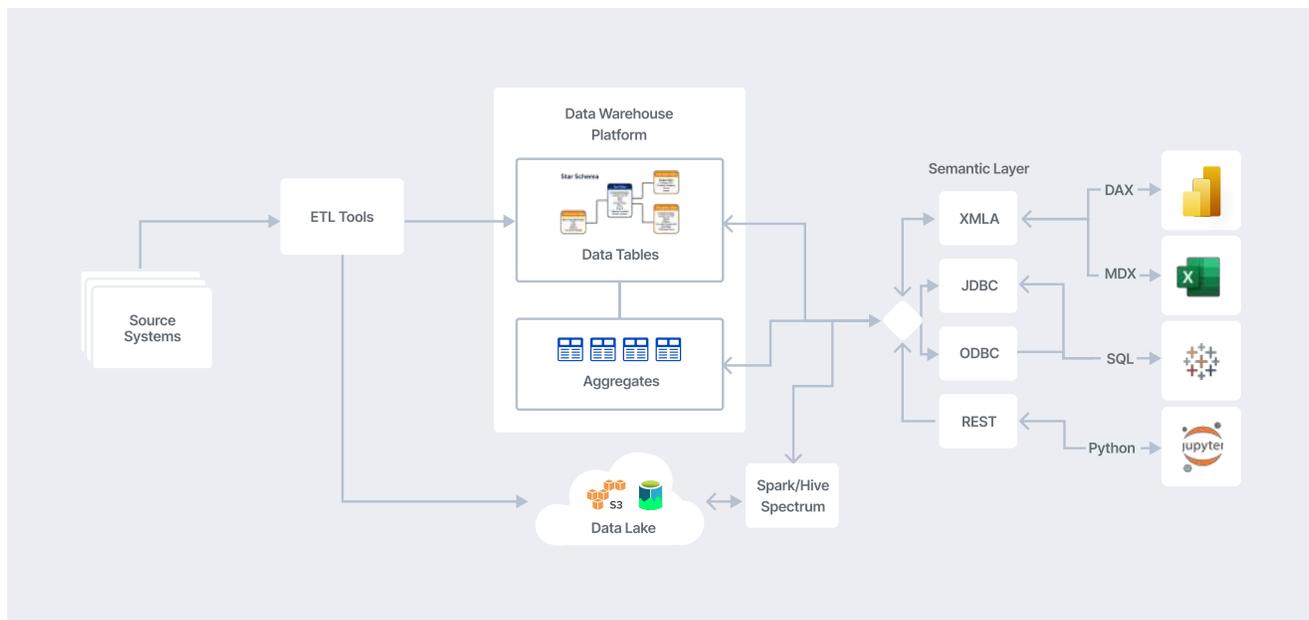
There is ample awareness about the cost of poor data quality in organizations. What's less talked about is the cost of poor quality metrics. While the duplication of business metrics is a drain on organizational effort, it's far more dangerous if all these siloed efforts produce inconsistent views of business performance. This is a pain point in most organizations today. While stories have far more visibility than tabular reports, the inconsistencies in metrics are more glaring in data stories than in traditional reports

How a Semantic Layer Helps Address These Challenges

By [definition](#), "a semantic layer is a business representation of corporate data that helps end-users access data autonomously using common business terms. A semantic layer maps complex data into familiar business terms such as product, customer, or revenue to offer a unified, consolidated view of data across the organization."

It can be empowering if analysts could work with a common business layer of data to create their stories rather than having to dip into the raw data for manual computations. A semantic layer can be the perfect middleware to enable insight spotting and story creation.

Here's how a semantic layer fits into the story creation workflow:



(Picture - Organizational setup with a good semantic layer: Source - AtScale)

There are [three features](#) of a robust semantic layer that make it a key enabler for data storytelling. They directly address the three challenges discussed above:

1. Data Integration: A semantic layer supports data blending across multiple data platforms and data sources, thereby minimizing the need for data movement. By abstracting the format, location, and complexity of data, a semantic layer truly democratizes data access.

2. Semantic Modeling: A semantic model organizes the data to reflect the business meaning of data items and their relationships. Imagine if we could directly access the business metrics and KPIs from a standardized central layer rather than having to compute them for every report or application repeatedly. A semantic layer makes this a reality.

3. Consumption Integration: A good semantic layer enables a live connection to all data sources within an organization. It provides a business-friendly interface to data for all consumer personas who might author a story. This ensures consistency and standardization across the organization.

Today, organizations are moving towards refreshing and automating their data stories every time the underlying source data changes. A semantic layer becomes increasingly invaluable for orchestrating and maintaining data stories within organizations.

How to Create a Workplace Where **Data Stories Thrive**

Knowing what a data story is and how to create one is a big step forward. However, institutionalizing data storytelling within an organization takes an almost equal effort.

Getting business users to adopt data stories isn't easy. Often, leaders mistakenly assume that hiring a team to create data stories, equipping them with the tools, and training them on the techniques will lead to successful business outcomes. Nothing is further from the truth.

Here are three recommendations to help you bridge this gap:

a) Educate your users on data literacy and the power of stories:

Business stakeholders are often used to traditional modes of data consumption – BI reports and dashboards. Educate them on what's possible with data. Help them get fluent in data literacy, including reading, writing, and communicating with data. Inspire them with examples of data storytelling and have executives sponsor initiatives to infuse storytelling across the organization.

b) Don't limit data storytelling to analytics applications:

People often restrict storyifying of data to dashboards or applications. Data storytelling can be applied to any mode of communication that presents information, from business presentations and emails to employee communications or investor reports. Encourage your teams to experiment with creative outputs irrespective of the communication channel.

c) Institutionalize processes to help users embrace data stories:

Leaders play a key role in change management and shifting the mindset of users to adopt data stories. Have them run business reviews using outputs presented as data stories. Encourage teams to present business plans and performance summaries as stories. Make data storytelling the de facto standard when it comes to communicating insights within the organization.

As famed psychologist Daniel Kahnemann said, “No one ever made a decision because of a number. They need a story.” It is high time every leader leverages this powerful technique to supercharge decision-making within their organization.



Ganes Kesari is an entrepreneur,
AI thought leader, author, and
TEDx speaker

Ganes Kesari is an entrepreneur, AI thought leader, author, and TEDx speaker. He co-founded Gramener, where he heads Data Science Advisory and Innovation. He advises executives on decision-making with data. He helps apply data science to solve organizational challenges, tell stories with data, and build winning analytics teams.

Ganes contributes articles to leading magazines such as Forbes and Entrepreneur. He teaches guest lectures on data science in schools such as Princeton University and runs corporate training on transforming organizations with data & analytics.