GROUPON

INTRODUCTION TO DATA VISUALIZATION & TABLEAU

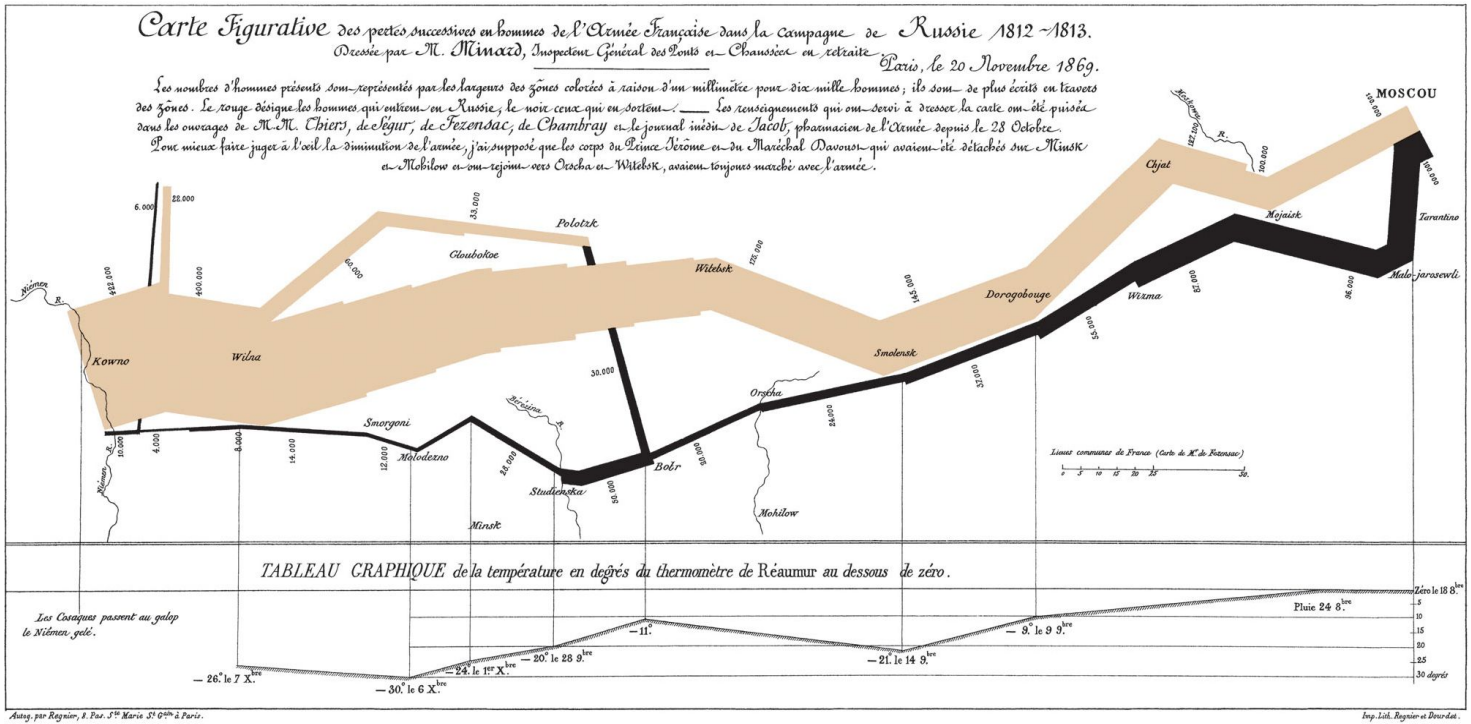
Overview of **Data Visualization** and **Tableau**

Sandeep Anand Pandey
Senior Data Analyst
Groupon

What is Data Visualization ?

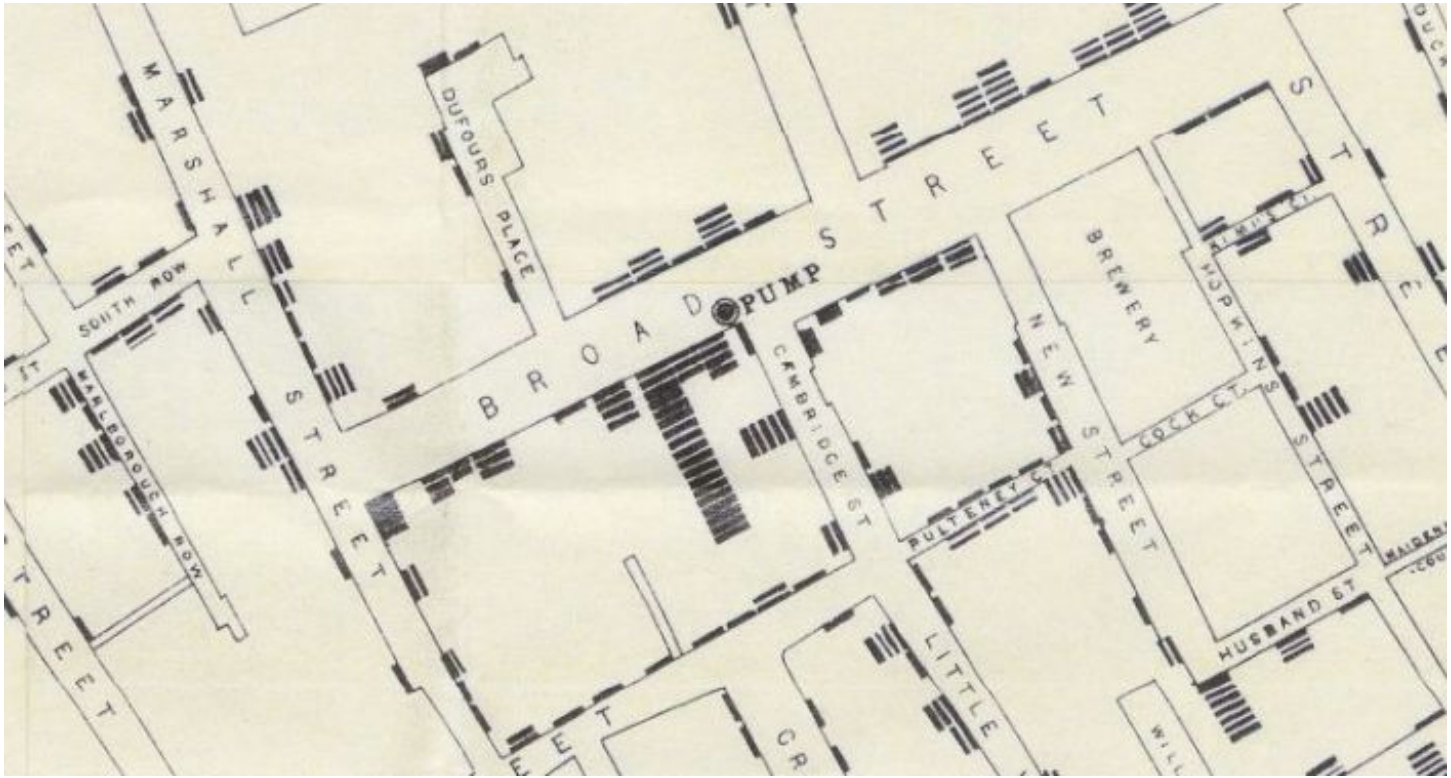
Data visualization is the **graphical representation of data** to help people understand the patterns, trends, and insights contained within the data

Minard's graphic of Napoleon invasion on Russia (1812).



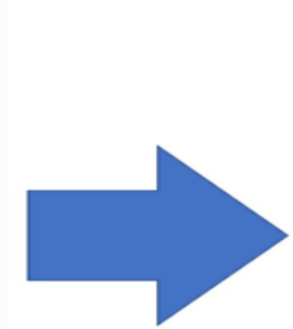
Some more example of great data visualization

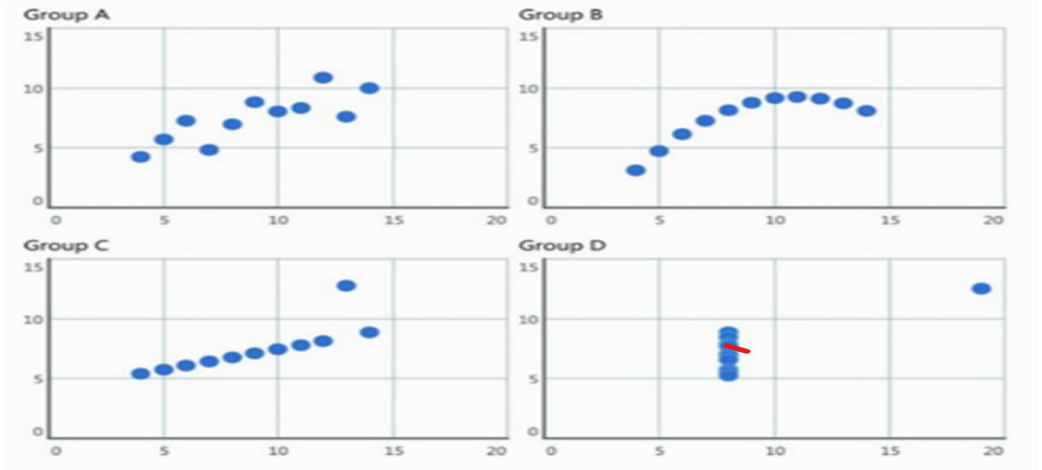
In 1854, Dr. John Snow was convinced that cholera was spreading via tainted water and decided to display neighborhood mortality data directly on a map. This method clearly revealed a cluster of cases around a specific pump.



Why Data Visualization is Required?

Group A		Group B		Group C		Group D	
x	y	x	y	x	y	x	y
10.00	8.04	10.00	9.14	10.00	7.46	8.00	6.58
8.00	6.95	8.00	8.14	8.00	6.77	8.00	5.76
13.00	7.58	13.00	8.74	13.00	12.74	8.00	7.71
9.00	8.81	9.00	8.77	9.00	7.11	8.00	8.84
11.00	8.33	11.00	9.26	11.00	7.81	8.00	8.47
14.00	9.96	14.00	8.10	14.00	8.84	8.00	7.04
6.00	7.24	6.00	6.13	6.00	6.08	8.00	5.25
4.00	4.26	4.00	3.10	4.00	5.39	19.00	12.50
12.00	10.84	12.00	9.13	12.00	8.15	8.00	5.56
7.00	4.82	7.00	7.26	7.00	6.42	8.00	7.91
5.00	5.68	5.00	4.74	5.00	5.73	8.00	6.89





Why Data Visualization is Required?

Data visualization can be used in many contexts in nearly every field, like public policy, finance, marketing, retail, education, sports, history, and more. Here are the benefits of data visualization:

Storytelling: People are drawn to colors and patterns in clothing, arts and culture, architecture, and more. Data is no different—colors and patterns allow us to visualize the story within the data.

Accessibility: Information is shared in an accessible, easy-to-understand manner for a variety of audiences.

Visualize relationships: It's easier to spot the relationships and hidden patterns within a data set when the information is presented in a graph or chart.

Exploration: More accessible data means more opportunities to explore, collaborate, and inform actionable decisions.

Reduce Cognitive Load :- Human mind is poor when it comes to number and calculation but its visual capabilities are at par and

Better business decisions :- With insight the process to taking decisions is not based on hunch but its backed by data

Revenue by Order Date

	Order Date		
Month of Order Date	2017	2018	2019
January	3,462,955	5,073,964	4,489,201
February	4,085,613	4,151,193	3,462,768
March	5,538,079	4,515,525	4,350,290
April	4,084,146	3,882,583	4,350,370
May	4,072,340	4,903,053	4,353,285
June	3,806,565	4,114,336	4,166,717
July	3,923,299	4,502,293	4,032,060
August	5,154,886	4,020,643	4,409,062
September	5,152,675	4,488,484	4,173,283
October	4,992,063	4,016,114	4,304,134
November	3,486,640	4,674,851	4,529,394
December	4,821,273	5,120,622	1,908,381

Revenue by Order Date

	Order Date		
Month of Order Date	2017	2018	2019
January	3,462,955	5,073,964	4,489,201
February	4,085,613	4,151,193	3,462,768
March	5,538,079	4,515,525	4,350,290
April	4,084,146	3,882,583	4,350,370
May	4,072,340	4,903,053	4,353,285
June	3,806,565	4,114,336	4,166,717
July	3,923,299	4,502,293	4,032,060
August	5,154,886	4,020,643	4,409,062
September	5,152,675	4,488,484	4,173,283
October	4,992,063	4,016,114	4,304,134
November	3,486,640	4,674,851	4,529,394
December	4,821,273	5,120,622	1,908,381

Data Visualization Process

Understand the **context**

1. Who is your audience
2. What decision they will take
3. How to enable them to take that decision

Choosing and **effective visual** (Type of Chart, Graph, Map)

Declutter (Clutter is Your Enemy!!)

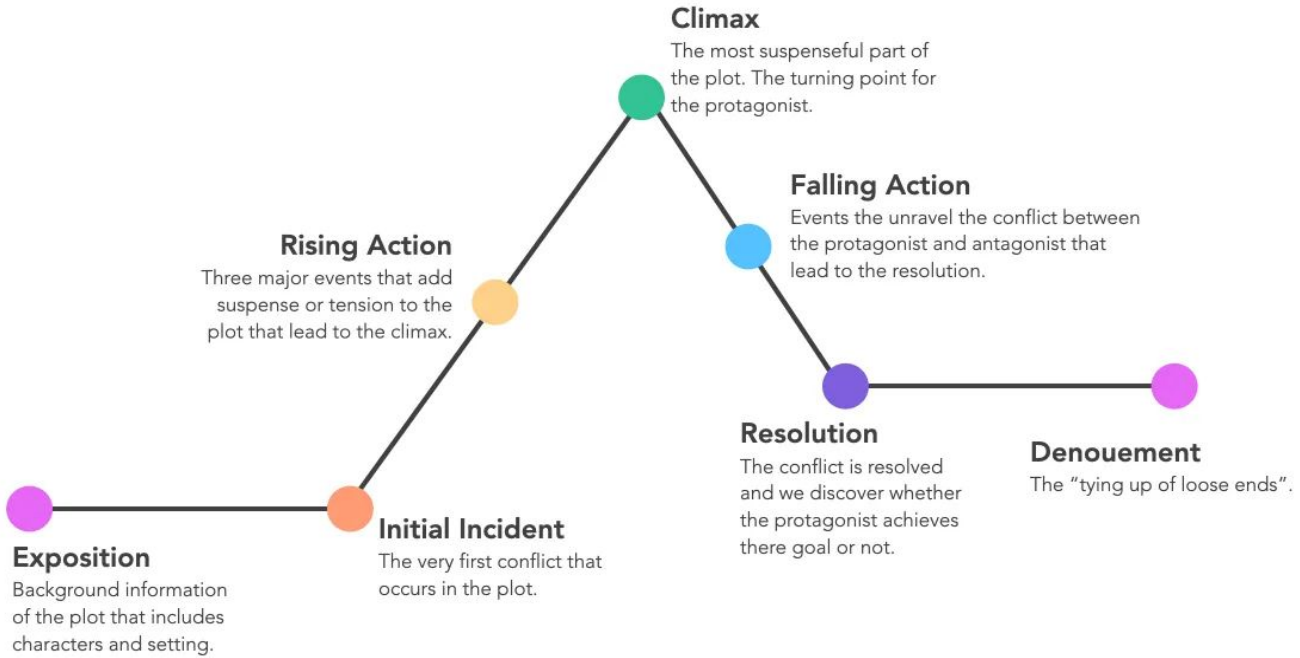
Focus your audience attention on the main issue.Use preattentive attribute to achieve it

Think like a **designer**. Reiterate the process and take inspiration from others

Data Storytelling + Narrative Arch

Data storytelling is the process of transforming data analyses into an understandable storyline for a wider audience in order to influence the decisions of business users and other stakeholders. Data storytelling presents data analysis findings in layman's terms.

A **Narrative or Story Arc** is the chronological construction of a plot in a novel or story. It is a great tool which can be leverage in modern day data storytelling .



Preattentive Attributes

How many number 3s are there?

Count the number of 3s and time how long it takes you. It can take a long time to count the numbers when nothing is highlighted.

5	7	3	4	3	7	1	1	5	3
6	3	9	4	5	5	9	3	3	6
4	8	9	5	8	8	10	10	5	8
3	8	3	3	9	1	8	9	6	9
7	6	1	7	2	10	4	3	2	8

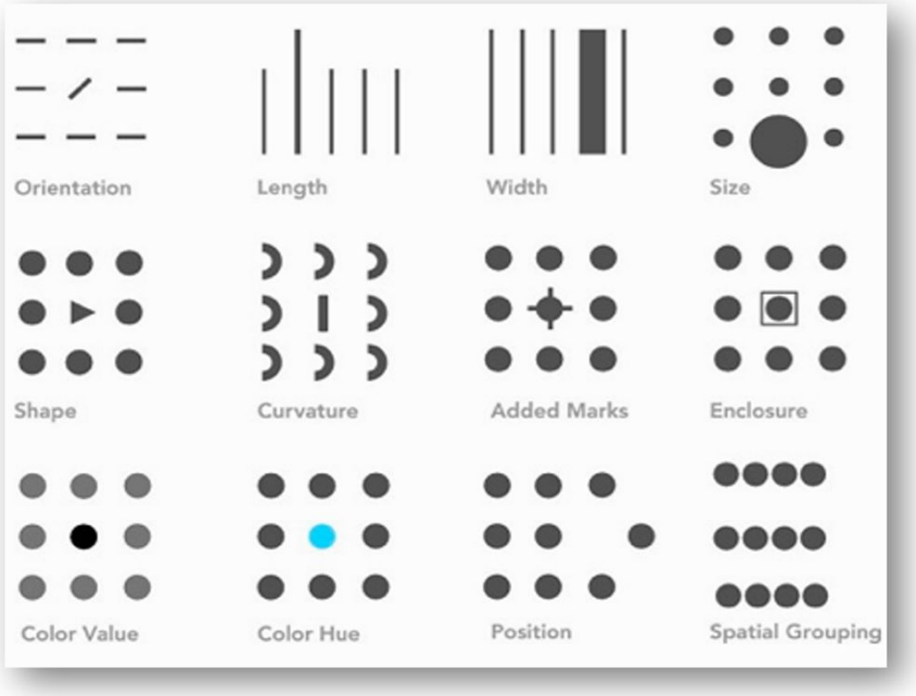
Viz inspiration: Knaflic, C (2015), Storytelling with Data. WILEY

Design: @datawoj

Preattentive Attributes

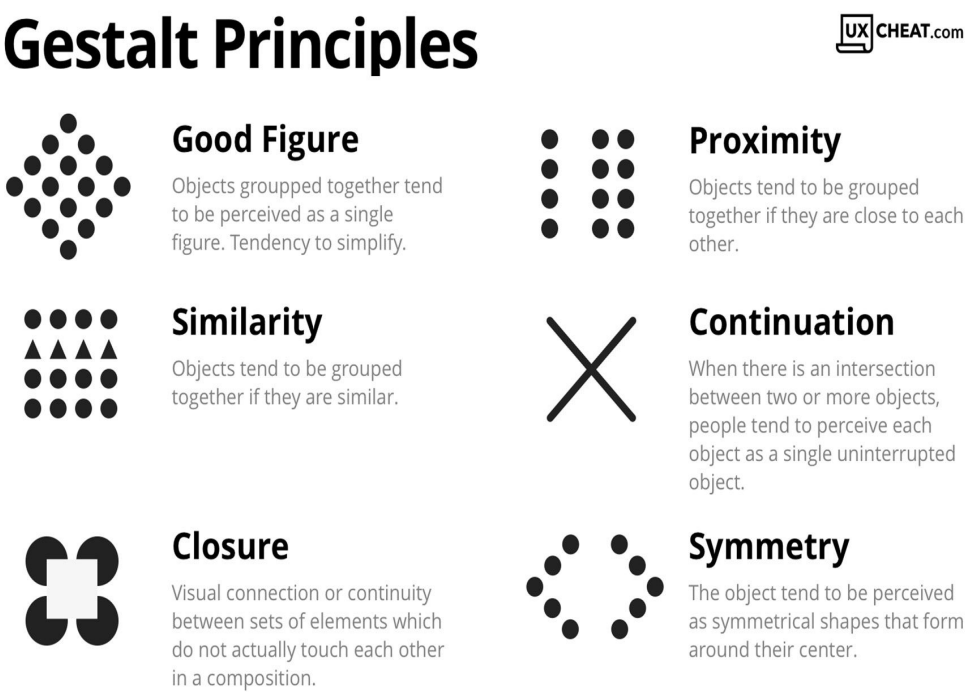
Preattentive Attributes

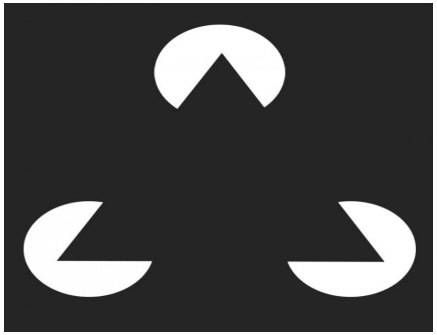
Preattentive attributes are visual properties that we notice without using conscious effort to do so and determine what information catches our attention, occurs before the conscious mind starts to pay attention and is the immediate cognitive experience of processing the visual world.



Gestalt Principle

The human brain is wired to see structure, logic, and patterns. It helps us make sense of the world. In the 1920s, a group of German psychologists developed theories around how people perceive the world around them, called **Gestalt principles.**





Introduction of Tableau

Tableau is a visual analytics platform transforming the way we use data to solve problems. It empowers people and organizations to make the most of their data.

The company's founders, **Christian Chabot**, **Pat Hanrahan** and **Chris Stolte**, were researchers at the Department of Computer Science at **Stanford University**. They specialized in visualization techniques for exploring and analyzing relational databases and data cubes, and started the company as a commercial outlet for research at Stanford from 1999 to 2002.

Tableau was founded in 2003 in Mountain View, California, and is currently headquartered in Seattle, Washington. In 2019 the company was **acquired by Salesforce for $15.7 billion.**

Figure 1: Magic Quadrant for Analytics and Business Intelligence Platforms

Source: Gartner March (2023)

Tableau Product Landscape

Tableau Cloud - The fast, flexible, easy-to-use self-service platform. Prepare your data, author, analyze, collaborate, publish and share in the cloud.

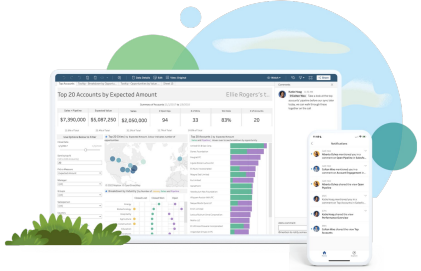


Tableau Desktop -Access, visualize, and analyze your data. With an intuitive drag and drop interface, you can uncover hidden insights and make smarter decisions faster.

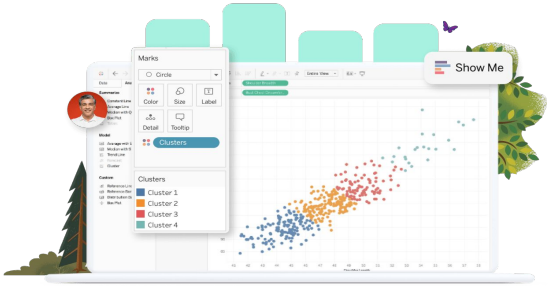


Tableau Prep -The modern approach to data preparation makes it easier and faster to combine, shape, and clean data for analysis.

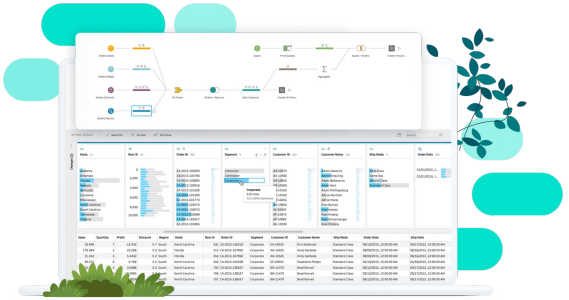


Tableau Data Management :- Combine the power of data flows, preparation, and cataloging so you can manage, connect, and trust your analytics data.

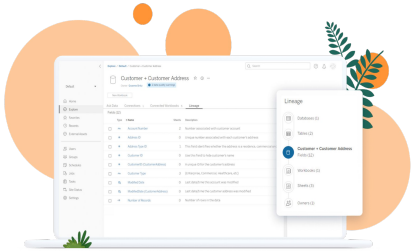


Tableau Server - Empower your organization with the freedom to explore data with governed, trusted, self-service analytics at scale.

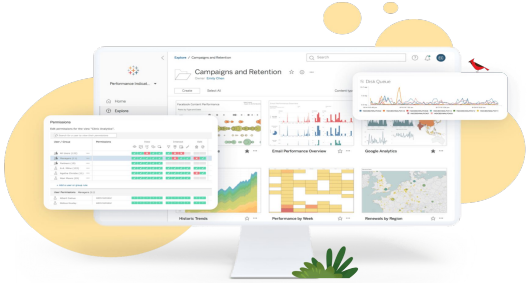


Tableau Architecture

Application Server
It processes & hold browsing and permissions for the Tableau Server web and mobile interfaces.

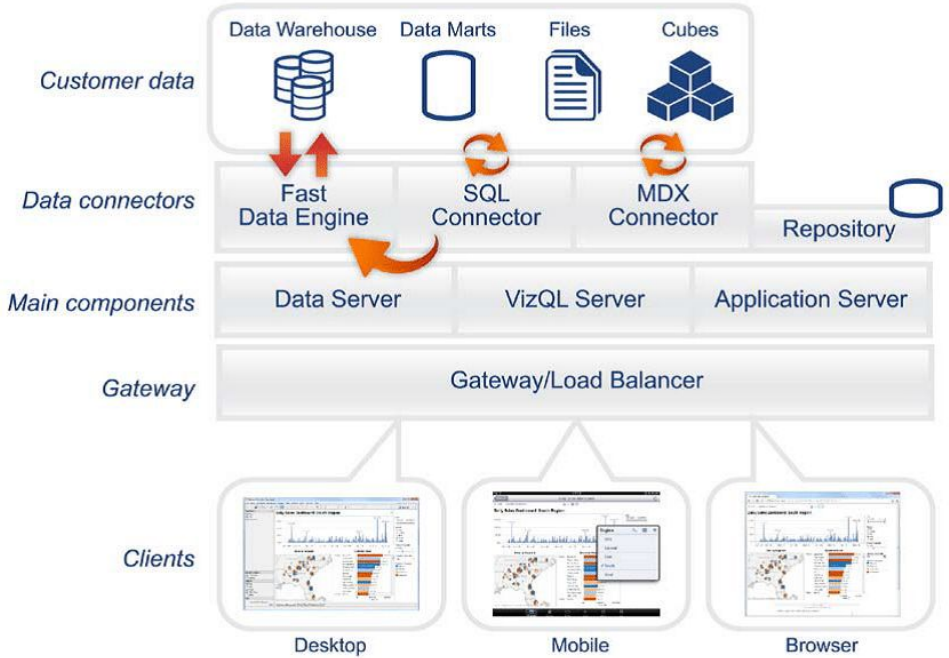
VizQL Server
If a view is released, the client sends a request to the VizQL process (vizqlserver.exe). The VizQL process then sends queries directly to the data source, returning a result set that is rendered as images and presented to the user. Each VizQL Server has its own cache that can be shared across multiple users.

Data Server
Tableau Data Server allows you centrally control and store Tableau data sources. It also manages metadata from Tableau Desktop, like calculations, definitions, and groups. The published data source can be founded on.

Backgrounder
The backgrounder refreshes scheduled extracts and manages other background tasks.

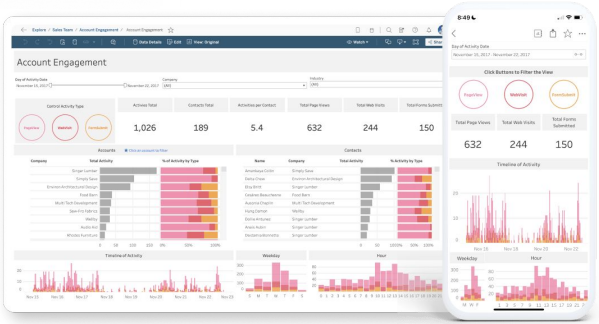
Gateway/ Load Balancer
It is the primary Tableau Server that trails requests to other components.

Clients (Web Browsers and Mobile Apps)



Why Tableau is a great data visualization tool.

- User-Friendly Interface
- Wide Range of Data Sources
- Powerful Data Preparation
- Interactive Visualizations
- Rich Library of Visualizations
- Seamless Integration
- Advance Analytics
- Collaboration and Sharing
- Scalability
- Community and Support
- Security and Governance



Questions ?

Tableau Additional Resources

- Cholera outbreak in London- Link
- Napoleon invasion on Russia - Link
- Tableau Product :- Link
- Tableau Community :- Link
- Agenda :- Link
- Tableau :- Link