HCAI5DS02 – Data Analytics and Visualization.

Lecture – 09

Introduction to Data Visualizations.

Why is Data Visualization Important?

What is Important in Data Visualization?

Siman Giri

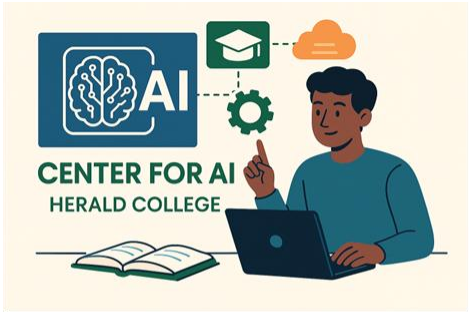


image generated via copilot.

Section-I
Fundamentals of Data Visualizations.

{1. The Whats and Whys of Data Visualization.}

1.1 Data Visualization: Definition.

- **Visualization** is the **process** of :
  - **transforming data and information** to **graphical representations** {**interactive or otherwise**}.
    - to **understand** the **complex data**
    - **explore** the **pattern** and **observe** the **trend** present in **data**
      - for **confirmation** / **presentation.**

visualization noun

vi·su·al·i·za·tion ˌvi-zhə-wə-lə-ˈzā-shən ˌvi-zhə-lə-, ˌvizh-wə-lə-

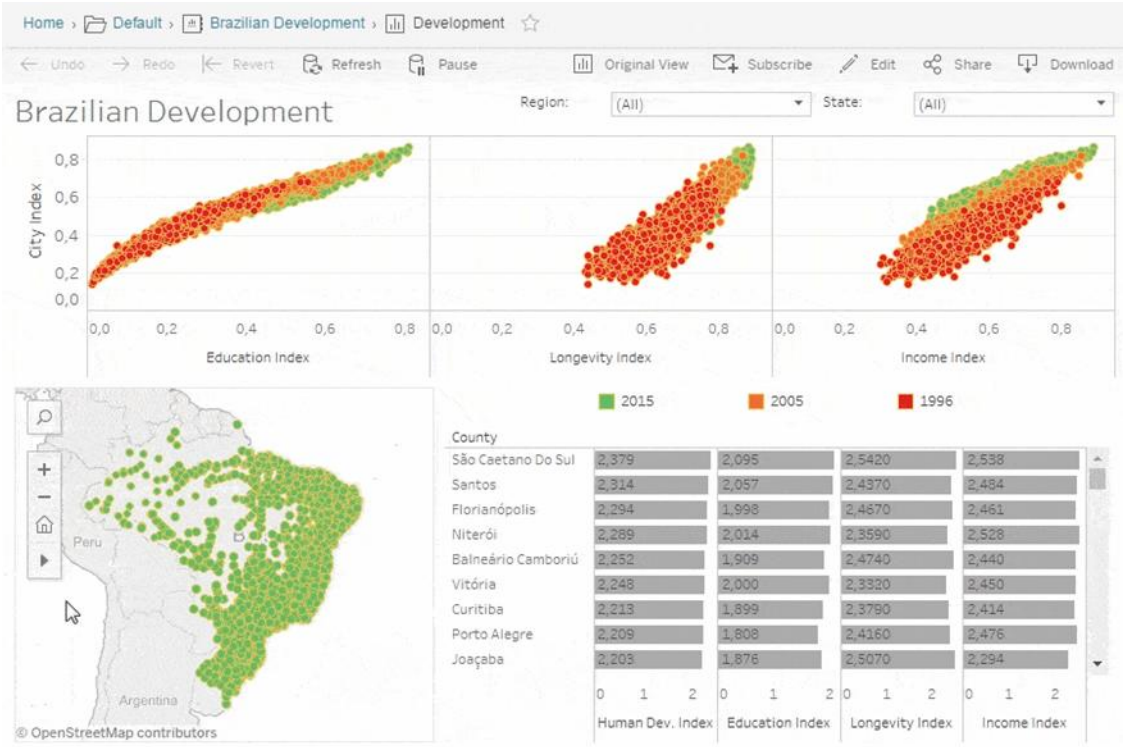
Synonyms of *visualization* ›

1 : formation of mental visual images

2 : the act or process of interpreting in visual terms or of putting into visible form

3 : the process of making an internal organ or part visible by the introduction (as by swallowing) of a radiopaque substance followed by radiography

As per → Merriam Webster Dictionary.



1.2 Evolution of Data Visualization.

- **Disclaimer!!!: Visualization does not always mean the use of computer:**

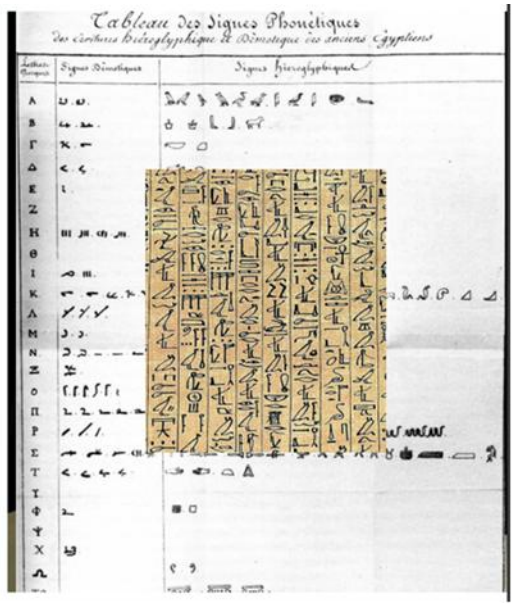
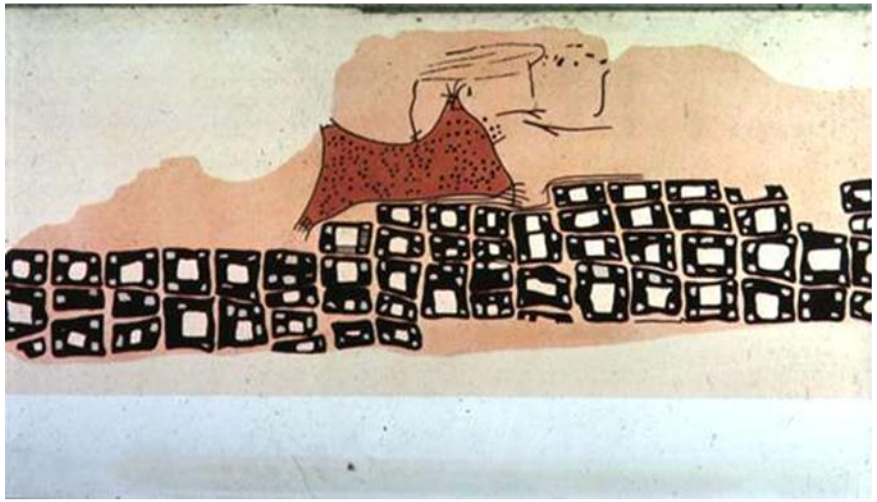


Fig: Logographic System

Why it started and Where are we?

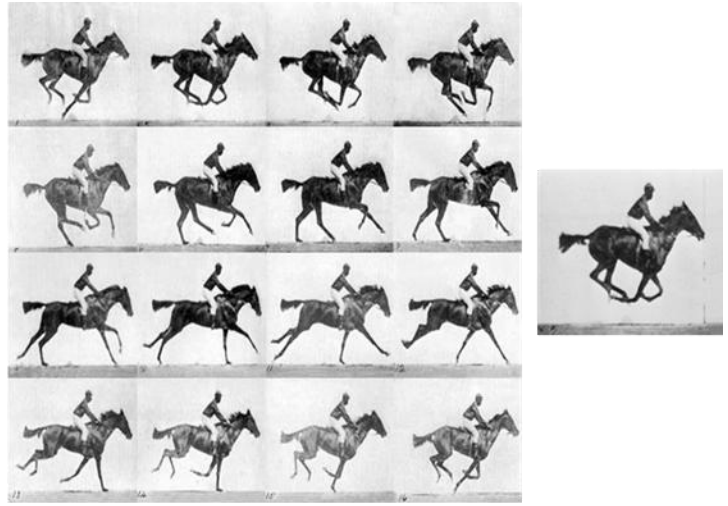
1.2.1 Evolution → Record an Event.

- To keep a **record** of an **important event**.



Konya town map, Turkey, c. 6200 BC

Leonardo Da Vinci, ca. 1500



E.J. Muybridge, 1878

1.2.2 Evolution → Analyze a Data.

First time to show changing values graphically

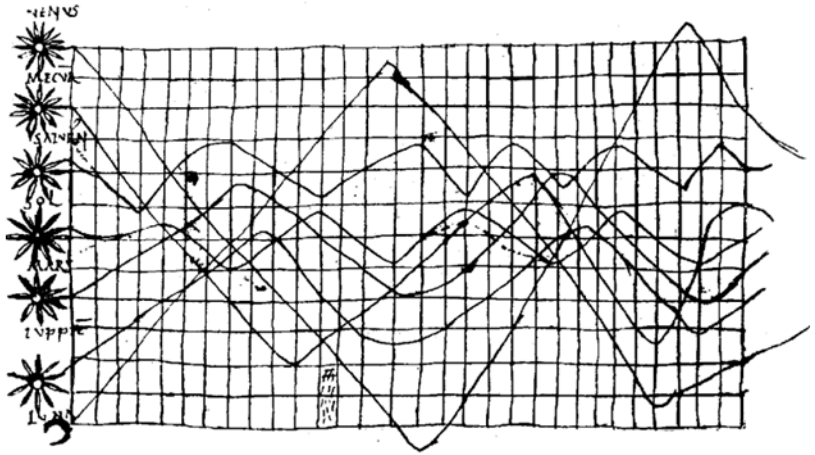


Fig: Planetary Movement Diagram, c. 950

First weather map

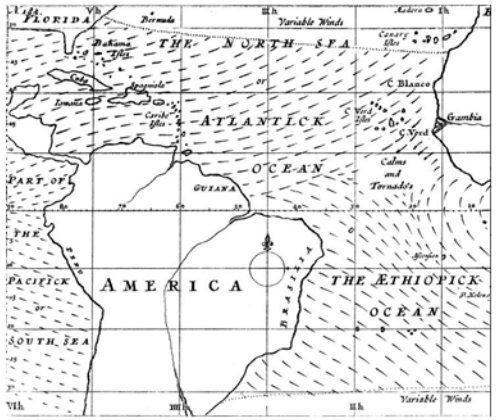


Fig: Halley's Wind Map, 1686

Import and Export of Scotland in 1781

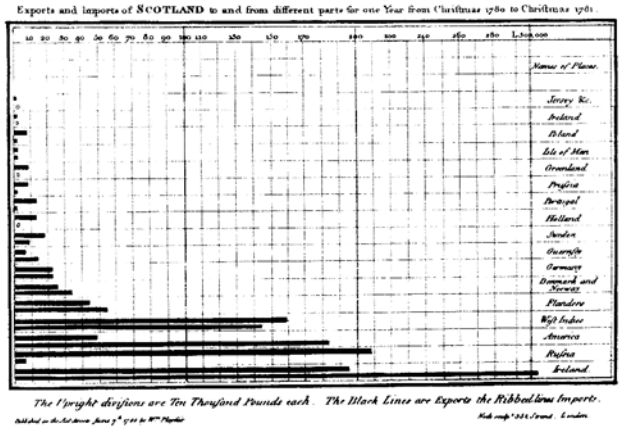
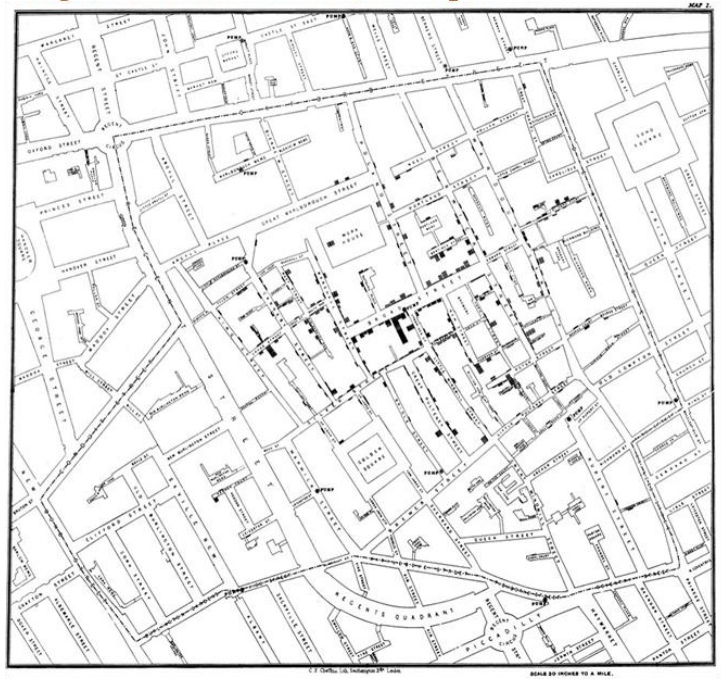


Fig: bar graph 1781

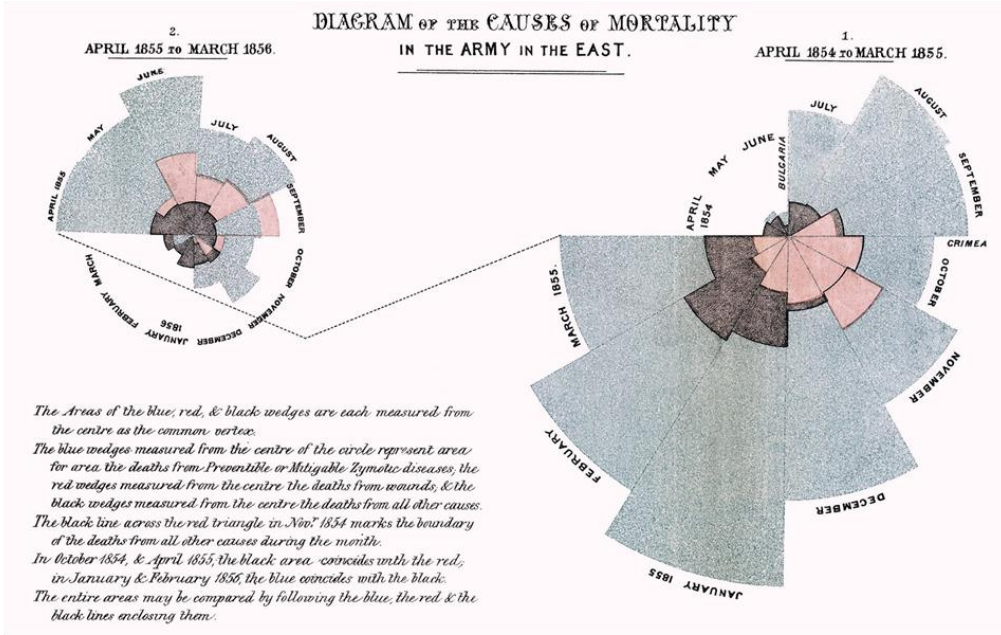
1.2.3 Evolution → Policy Making.

Fig: John Snow's Cholera Map 1854



Snow's map of Broad Street Pump map did not only represent statistical data; but helped to transformed the public health policy showcasing the connection between the water source and cholera outbreak.

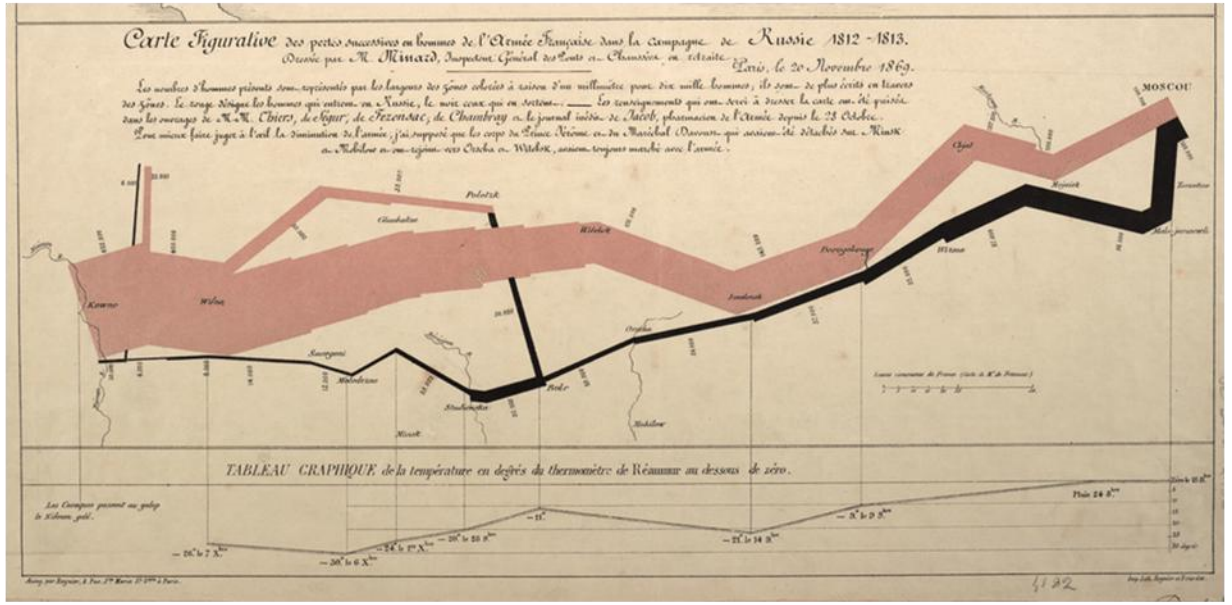
Fig: Florence Nightingale and Rose Diagram 1856



Nightangle's "polar Area Chart" depicted the mortality rate in military hospitals caused by preventable disease conditions mostly caused due to sanitary conditions rather than battle related injury.

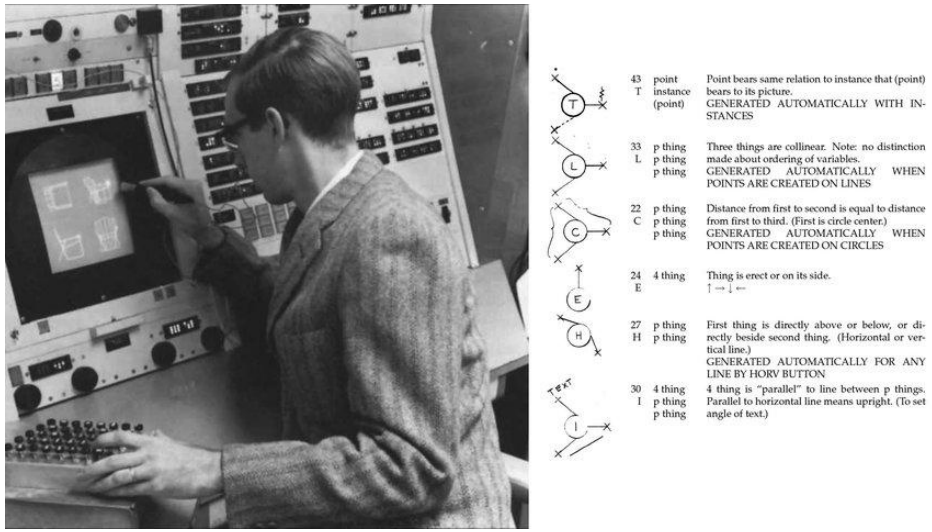
1.2.4 Evolution → Strategy and Planning.

Fig: Charles Minard's Flow Map of Napoleon's Russian Campaign



This visualization was a masterpiece as it integrates multiple data types such as troop numbers, distance, temperature , geography all in one graphic. Minard's ability to **visualize data** from various sources into a cohesive visual narrative showcased the evolution of data visualization during this period.

1.2.5 Evolution → Modern Technology.



43 point T instance (point) — Point bears same relation to instance that (point) bears to its picture. GENERATED AUTOMATICALLY WITH INSTANCES

33 p thing L p thing p thing — Three things are collinear. Note: no distinction made about ordering of variables. GENERATED AUTOMATICALLY WHEN POINTS ARE CREATED ON LINES

22 p thing C p thing p thing — Distance from first to second is equal to distance from first to third. (First is circle center.) GENERATED AUTOMATICALLY WHEN POINTS ARE CREATED ON CIRCLES

24 4 thing E — Thing is erect or on its side. ↑ → ↓ ←

27 p thing H p thing — First thing is directly above or below, or directly beside second thing. (Horizontal or vertical line.) GENERATED AUTOMATICALLY FOR ANY LINE BY HORV BUTTON

30 4 thing I p thing p thing — 4 thing is "parallel" to line between p things. Parallel to horizontal line means upright. (To set angle of text.)

Fig: Ivan Sutherland, Sketchpad, 1963

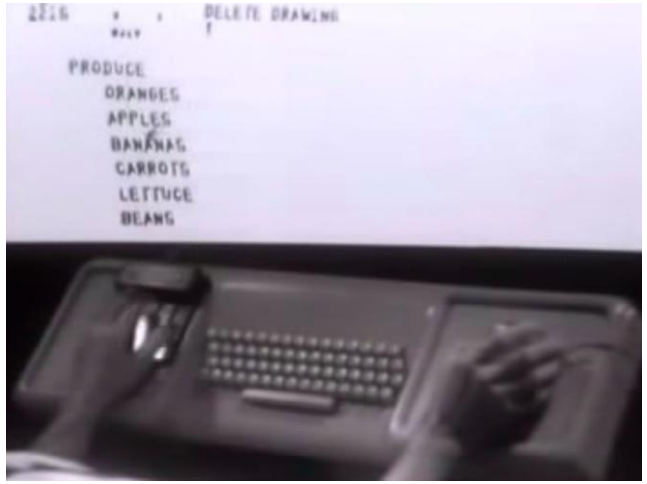


Fig: Doug Engelbart "oN-Line System" 1968

- The early days of data visualization were marked by pioneers who recognized the power of visual representation to convey complex data and drive change.
  - Over time, software developers and designers recognized the potential of computer graphics for data visualization.
  - This led to the development of specialized visualization tools that could handle complex data, offer interactive features, and present data in a more visually appealing manner.

1.2.6 Evolution → Computer Based Data Visualizations.

- **Computer-based visualization** systems provide
  - **visual representations of datasets designed to help people carry out tasks more effectively.**
- With the **rise of web-based platforms**, **data visualization** became **more accessible to the masses**.
  - No longer confined to **the realm of experts**, anyone with an internet connection could now
    - **create, share, and interact with visual data representations.**
  - This democratization has led to a surge in
    - **innovative visualization techniques** and
      - a **broader appreciation** for the **power of visual data**.

1.3 Why Visualizations?

- **Descriptive Statistics** may be **misleading** sometimes.
- For example: **Anscombe's quartet**.

Dataset I		Dataset II		Dataset III		Dataset IV	
x	*y*	*x*	*y*	*x*	*y*	*x*	*y*
10.0	8.04	10.0	9.14	10.0	7.46	8.0	6.58
8.0	6.95	8.0	8.14	8.0	6.77	8.0	5.76
13.0	7.58	13.0	8.74	13.0	12.74	8.0	7.71
9.0	8.81	9.0	8.77	9.0	7.11	8.0	8.84
11.0	8.33	11.0	9.26	11.0	7.81	8.0	8.47
14.0	9.96	14.0	8.10	14.0	8.84	8.0	7.04
6.0	7.24	6.0	6.13	6.0	6.08	8.0	5.25
4.0	4.26	4.0	3.10	4.0	5.39	19.0	12.50
12.0	10.84	12.0	9.13	12.0	8.15	8.0	5.56
7.0	4.82	7.0	7.26	7.0	6.42	8.0	7.91
5.0	5.68	5.0	4.74	5.0	5.73	8.0	6.89

Descriptive Statistics.

Property	Value	Accuracy
Mean of *x*	9	exact
Sample variance of *x*: s_x^2	11	exact
Mean of *y*	7.50	to 2 decimal places
Sample variance of *y*: s_y^2	4.125	±0.003
Correlation between *x* and *y*	0.816	to 3 decimal places
Linear regression line	$y = 3.00 + 0.500x$	to 2 and 3 decimal places, respectively
Coefficient of determination of the linear regression: R^2	0.67	to 2 decimal places

img source: "Wikipedia"

1.3.1 Why Visualizations?

- **Descriptive Statistics** may be **misleading** sometimes.
  - But **if you Visualize:**

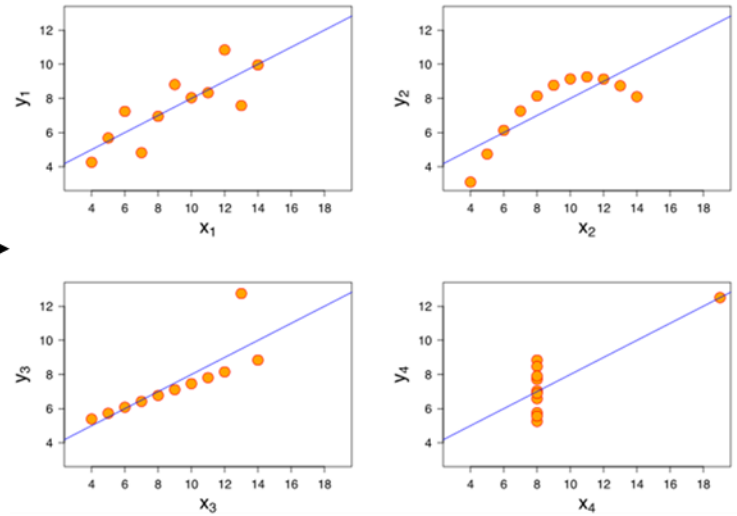
Dataset I		Dataset II		Dataset III		Dataset IV	
x	*y*	*x*	*y*	*x*	*y*	*x*	*y*
10.0	8.04	10.0	9.14	10.0	7.46	8.0	6.58
8.0	6.95	8.0	8.14	8.0	6.77	8.0	5.76
13.0	7.58	13.0	8.74	13.0	12.74	8.0	7.71
9.0	8.81	9.0	8.77	9.0	7.11	8.0	8.84
11.0	8.33	11.0	9.26	11.0	7.81	8.0	8.47
14.0	9.96	14.0	8.10	14.0	8.84	8.0	7.04
6.0	7.24	6.0	6.13	6.0	6.08	8.0	5.25
4.0	4.26	4.0	3.10	4.0	5.39	19.0	12.50
12.0	10.84	12.0	9.13	12.0	8.15	8.0	5.56
7.0	4.82	7.0	7.26	7.0	6.42	8.0	7.91
5.0	5.68	5.0	4.74	5.0	5.73	8.0	6.89

Descriptive Statistics.

Property	Value	Accuracy
Mean of *x*	9	exact
Sample variance of *x*: s_x^2	11	exact
Mean of *y*	7.50	to 2 decimal places
Sample variance of *y*: s_y^2	4.125	±0.003
Correlation between *x* and *y*	0.816	to 3 decimal places
Linear regression line	$y = 3.00 + 0.500x$	to 2 and 3 decimal places, respectively
Coefficient of determination of the linear regression: R^2	0.67	to 2 decimal places

img source: "Wikipedia"

Fig: Anscombe's quartet



img source: "Wikipedia"

1.3.2 Why Visualizations?

1. **To communicate:**
   - When data attributes are simplified into a visual language, patterns and trends can reveal themselves for easy comprehension.
     - "*The **goal of visualization** is **to aid our understanding of data** by leveraging **the human visual system's** highly tuned ability to see **patterns, spot trends, and identify outliers**.*"
       - Heer, Bostock, & Ogievetsky et. al "A tour through Visualization Zoo"
2. **Transform data into information:**
   - In this era of big data, visualization is a powerful way to make sense of the data.
   - "*Now through **Advanced Data Visualization**, potential exists for nontraditional and more **visually rich approaches**, especially in regard to **more complex** (i.e. thousands of dimensions or attributes) or larger (i.e., billion of rows) data sets, to **reveal insights** **not possible** through **conventional means**.*"
     - Evelson : Trends 2011 And Beyond: Business Intelligence.
3. **To show evidence:**
   - Data graphics are used to findings, new insights or results.
   - The data graphic makes the evidence clear when it shows an interpretable result such as trend or pattern.
   - ***Cautions!!!** Using data graphics as evidence are only as good as the insight or message communicated.*

Section-I
Fundamentals of Data Visualizations.

{2. The Context → Who, What, and How.}

2.1 Communicating with Data: The Who, The What and The How?

1. **The Who (Who is your audience?)**
   - Every communication begins **with the audience**.
     - Ask: ***Who am I speaking to?***
   - Their **background**, **expertise**, and **priorities** will **shape your approach**.
   - A technical audience (**like analysts or data scientists**) may expect
     - details about **methodology, assumptions, and models.**
   - A business audience (like executives or clients) may care more about
     - **insights, impact, and decisions**.
   - Knowing ***who*** ensures your communication is relevant and engaging.
2. **The What (What should your audience know or do?)**
   - This is **the action or takeaway**.
     - Decide clearly: ***What do I want my audience to know, believe, or act upon after this?***
     - A data story without a clear action risks being forgotten.
   - Whether it's **recommending a strategy**, **highlighting a risk**, or **showing an opportunity**,
     - "***the what***" keeps your **communication purposeful and outcome-driven.**

2.1.1 Communicating with Data: The Who, The What and The How?

3. **The How (How will you communicate and support your point?)**
   - Once you know your audience and desired action, ask: *How will I make my case?*
   - This includes both
     - **the mechanism** (**presentation, dashboard, report, or conversation**) and
     - **the evidence** (**the data itself**).
   - Choose **visuals and formats** that **fit the audience**, and use **the most relevant data to support your point**.
   - Too **much detail can overwhelm**, while **the right balance strengthens credibility and clarity**.

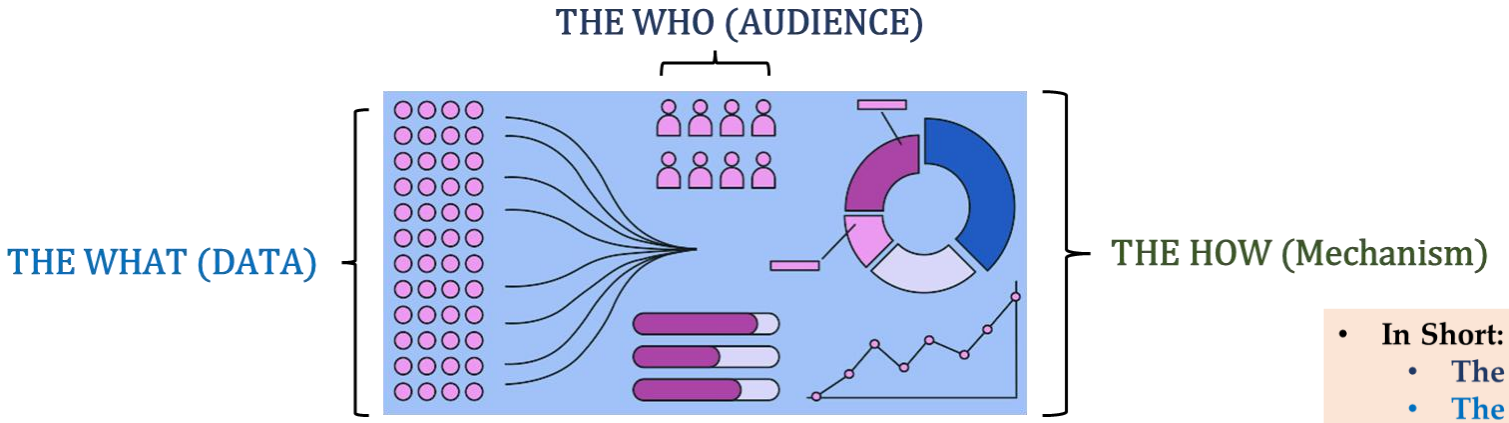


Fig: Communicating with DATA.

- **In Short:**
  - **The Who → Shapes your framing.**
  - **The What → defines your purpose.**
  - **The How → connects purpose to data and delivery.**

2.2 The Who, The What and The How – Example.

- **Scenario:**
  - Imagine you are a **Product Manager** at a retail company. You just ran an **experimental pilot loyalty program** during the summer aimed at increasing repeat purchases from younger customers (18–25 years old). You surveyed customers at the start and end of the pilot to understand whether and how perceptions towards your brand changed. You believe the data shows a **great success story** — increased engagement, higher purchase frequency, and improved brand sentiment. You would like to continue and expand the program going forward.
- **Distinguish the The Who, The What and The How for the presentations.**

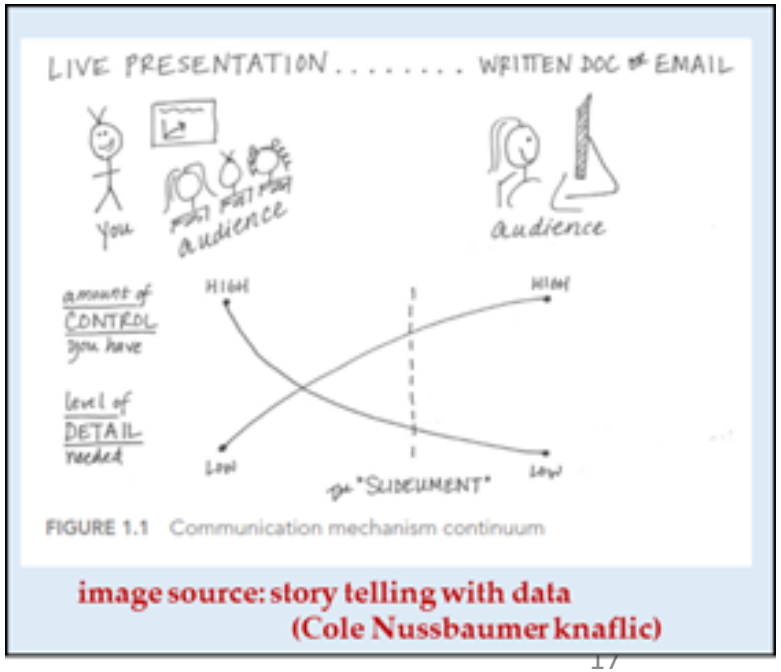


image source: story telling with data (Cole Nussbaumer knaflic)

2.2.1 The Who, The What and The How – Example.

1. **The Who (Audience):**
   - **Executives / Leadership Team** → Interested in overall business impact and ROI.
   - **Marketing Department** → Wants to see how customer sentiment shifted and what messaging resonated.
   - **Finance Team** → Focused on costs versus long-term value.
   - ***Why this matters:***
     - Each group needs a slightly different emphasis, but all are stakeholders in deciding whether to scale the program.
2. **The What (Message / Takeaway):**
   - Pilot program was **successful** → measurable uplift in repeat purchases (+18%), stronger customer satisfaction, and higher engagement.
   - Program should be **continued and scaled** because benefits outweigh costs.
   - The company risks **losing momentum** if it does not act now.
3. **The How (Communication & Evidence):**
   - Use **before-and-after comparisons** (surveys, purchase frequency, sentiment analysis).
   - Present **clear visuals** (bar charts for purchase frequency, trend lines for engagement, word clouds for sentiment).
   - Highlight **ROI metrics** (e.g., $X increase in revenue vs. $Y spent on the pilot).
   - Tell a **story with data**: start with the business problem (low retention), show the intervention (pilot program), present results (data), and end with recommendation (scale the program).

Section-II
The Graphics.

{3.Visual Encodings: Creating an Image/graphs from Data.}

3.1 Visual Encoding: The Core of Data Visualization

- **Definition:**
  - **Visual encoding** is the process of **mapping data variables to visual elements** in a graphic.
  - It's the **essential step that turns raw numbers** into **a visual form** that humans can **interpret** quickly.

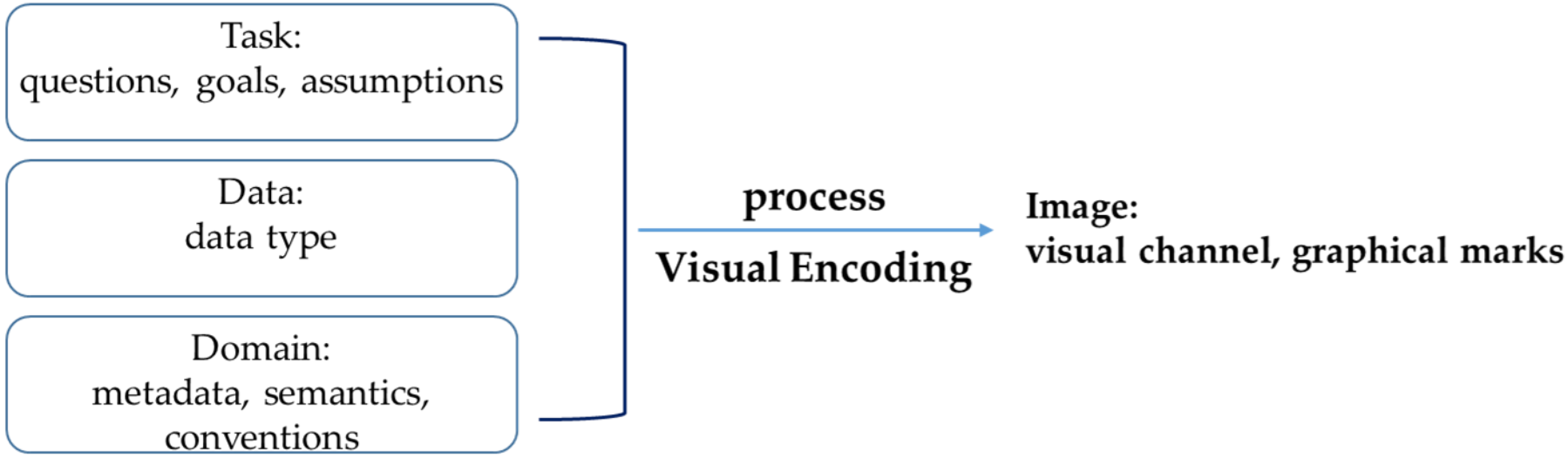


Fig: Visual Encodings: The Big Picture.

3.2 The Components of Visual Encodings.

- When we talk about **graphics in data visualization**, every graphic (chart/plot/graph) has
  - **two foundational components**, plus **supporting variables**:
    1. **Coordinate System:**
       - This defines the **space** in which your **data is plotted**.
       - It provides the **axes, grid, scale, and orientation** for mapping data.
       - Examples:
         - **Cartesian (x–y plane, most common)**
         - **Polar (radar charts, pie charts)**
         - **Geographic (maps, longitude/latitude)**
       - Think of it as the **"stage"** where your **data will be displayed**.
    - Example:
      - The most common coordinate system is **2D Cartesian**,
        - where each data point is specified by **x and y values**.
      - The axes are continuous scales that can represent positive or negative values, with equidistant grid spacing.
        - Typically, **x and y represent different units**.

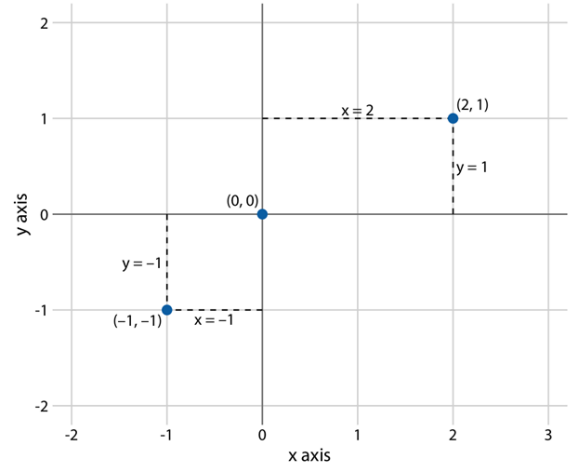


Fig: Anatomy of a 2d co-ordinate system

3.2.1 The Components of Visual Encodings.

2. **Visual Marks (Visual Aesthetics/Channels/Variables):**
   - These are the **basic shapes** or elements used to represent data in the coordinate system.
     - **Types of Marks:**
       - **Point → scatter plot**
       - **Line → line chart, trend visualization**
       - **Area → area chart, histograms**
       - **Bar → bar/column chart**
       - **Text → labels, annotations**
     - **Visual channels (aesthetics):**
       - **ways to encode additional dimensions of data:**
         - **Position (x, y placement – the strongest channel)**
         - **Size (larger/smaller point = bigger value)**
         - **Shape (different symbols for categories)**
         - **Color (hue, brightness, saturation)**
         - **Orientation (direction of marks, like arrow plots)**
         - **Texture/pattern**
   - This is the **"actor"** on the stage.

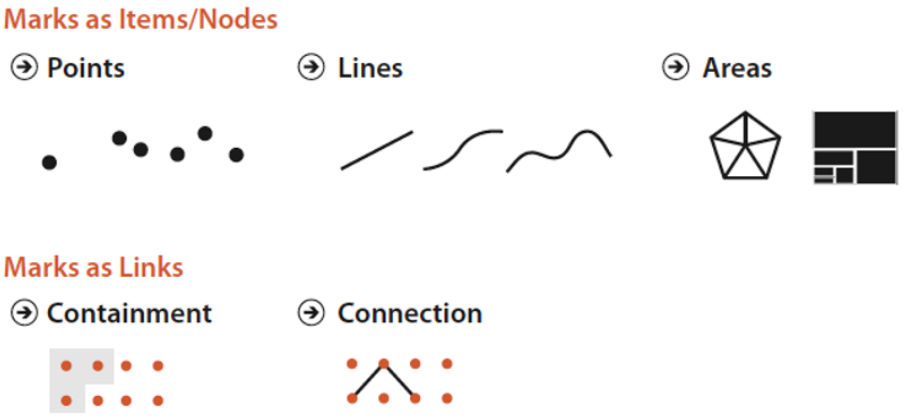
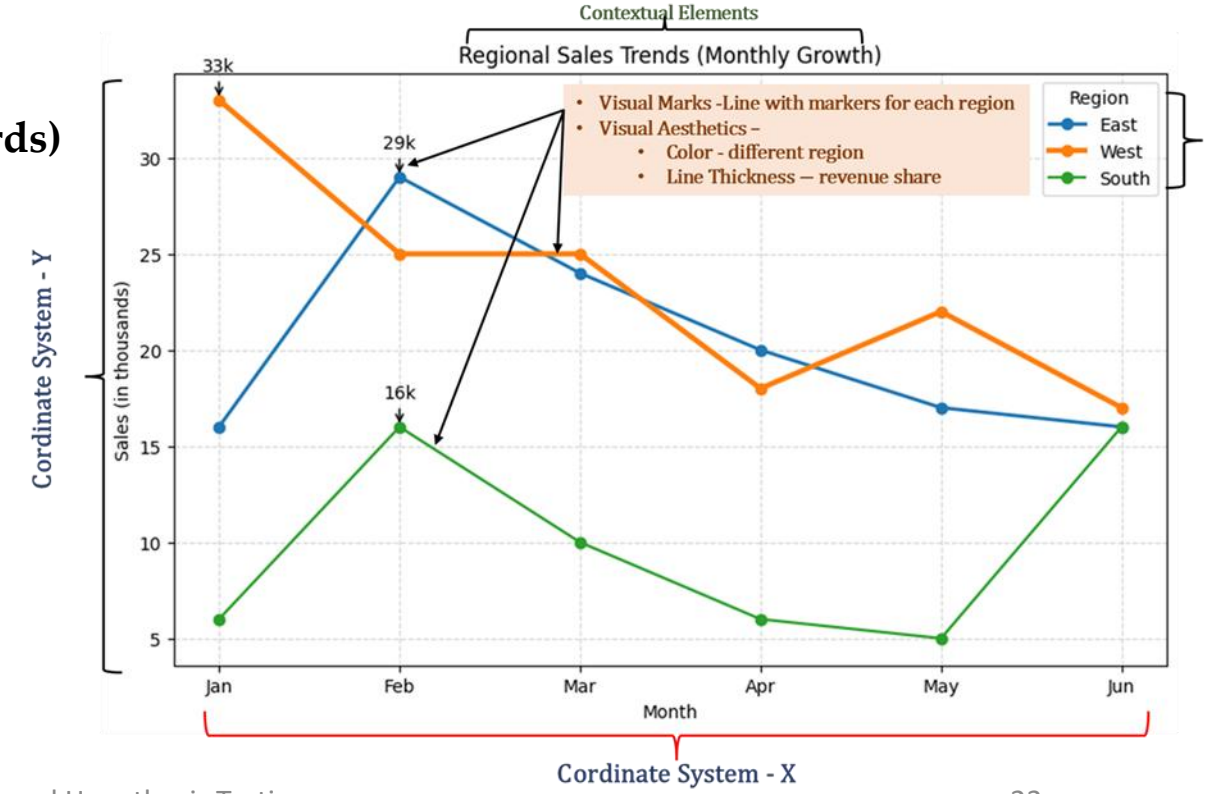


Fig: Some common visual marks in use.

3.2.2 The Components of Visual Encodings.

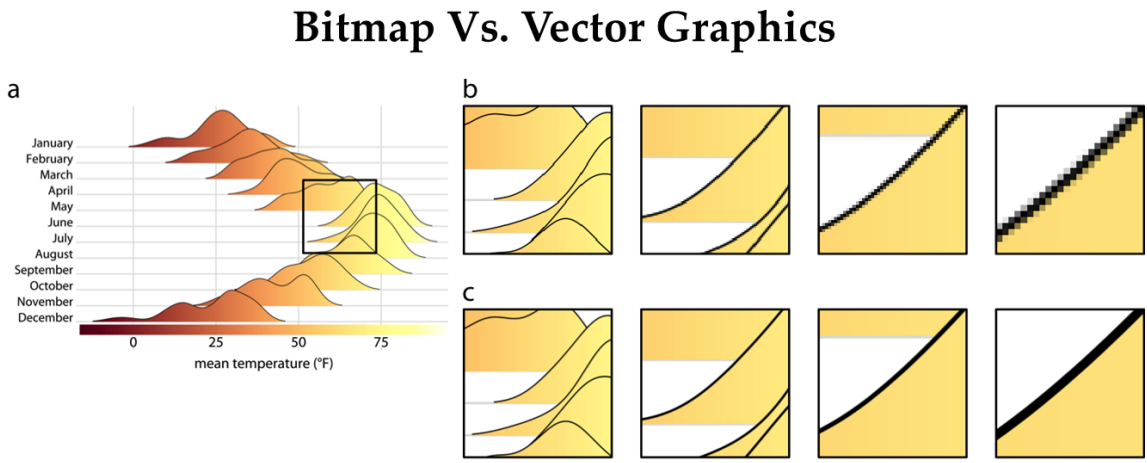
3. **Other Variables (Contextual Elements):**
   - These don't directly map data but help **interpret** it:
     - **Titles, labels, legends**
     - **Scales (linear, logarithmic, categorical)**
     - **Gridlines, reference lines**
     - **Interaction (hover tooltips, filters in dashboards)**
   - These are like the **narration and props**
     - **that help the audience understand the story**.

Fig: Components of Visual Encodings.



3.3 Visual Encoding and File Formats.

- **File format** is **not a type of visual encoding**,
  - but it determines **how the encoded visualization is stored, shared, or displayed**.
- **Key points:**
  1. **Purpose of file formats:**
     - Preserve the visual encoding (the chart you created).
     - Enable sharing, embedding, or printing.
  2. **Common file formats:**
     - **Raster formats** (pixel-based):
       - PNG, JPEG, TIFF →
         - good for reports, presentations; lose scalability.
     - **Vector formats** (geometry-based):
       - SVG, PDF, EPS →
         - scalable without losing quality; better for dashboards, posters.
  3. **Interactive formats:**
       - HTML, JSON (Plotly, Vega, Altair) →
         - preserve interactivity like hover, zoom, filters.



3.3.1 Some Common File Formats.

Table 27.1: Commonly used image file formats

Acronym	Name	Type	Application
pdf	Portable Document Format	vector	general purpose
eps	Encapsulated PostScript	vector	general purpose, outdated; use pdf
svg	Scalable Vector Graphics	vector	online use
png	Portable Network Graphics	bitmap	optimized for line drawings
jpeg	Joint Photographic Experts Group	bitmap	optimized for photographic images
tiff	Tagged Image File Format	bitmap	print production, accurate color reproduction
raw	Raw Image File	bitmap	digital photography, needs post-processing
gif	Graphics Interchange Format	bitmap	outdated for static figures, Ok for animations

Section-II
The Graphics.
{4. Choosing the Visuals}

4.1 Choosing the Visuals.

4.1.1 Visualizing Amounts.

- The most common charts used to visualize amounts are:
  - **bar charts and it's variants**
  - **dots chart**
  - **heatmap**

Fig: Bar and Dots chart

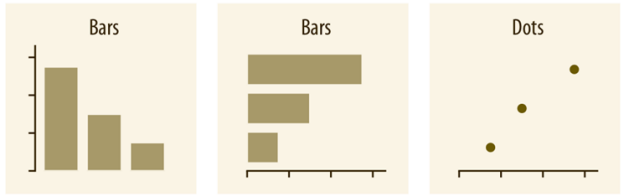
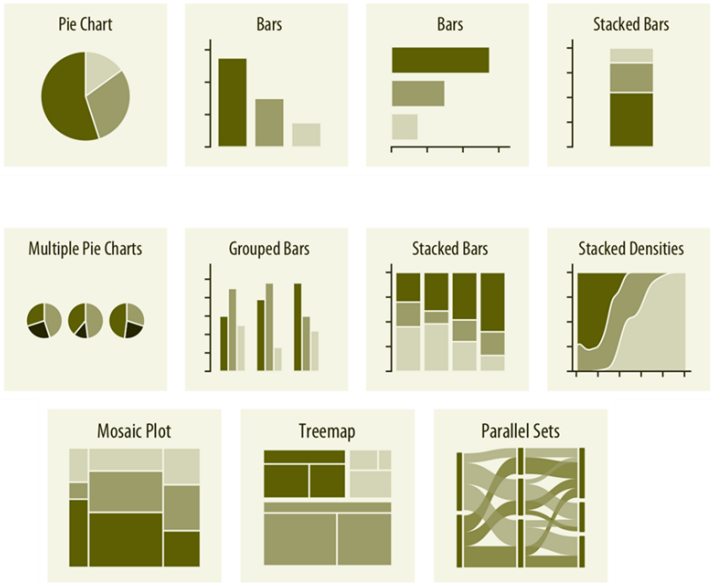


Fig: It's Variants

4.1.2 Visualizing Proportions.

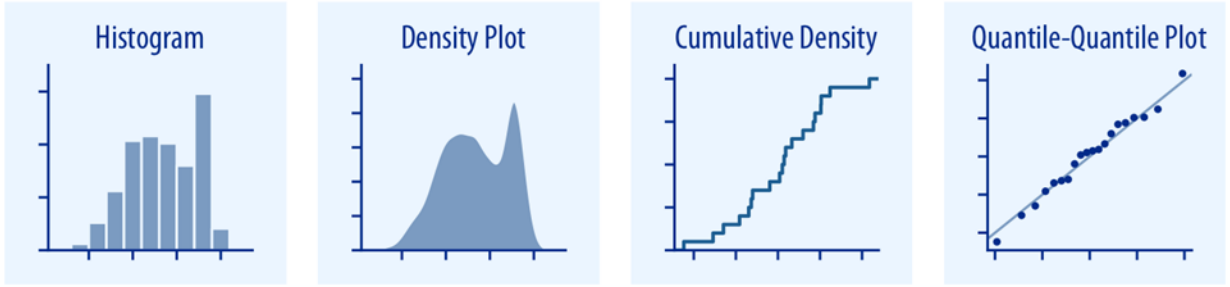
- Proportions can be visualized as
  - pie charts,
  - side-by-side bars,
  - or stacked bars,



4.1 Choosing the Visuals.

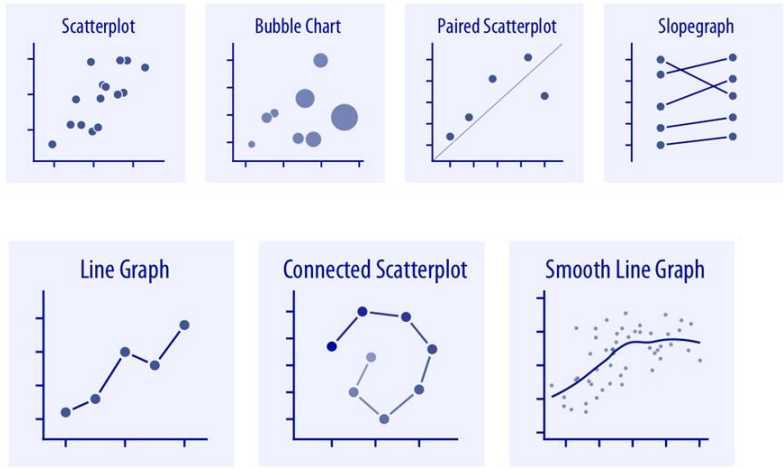
4.1.3 Visualizing Distributions.

- **Histograms and density plots** provide
  - the most intuitive visualizations of a distribution
  - but both require arbitrary parameter choices and can be misleading.



4.1.4 Visualizing Relationships

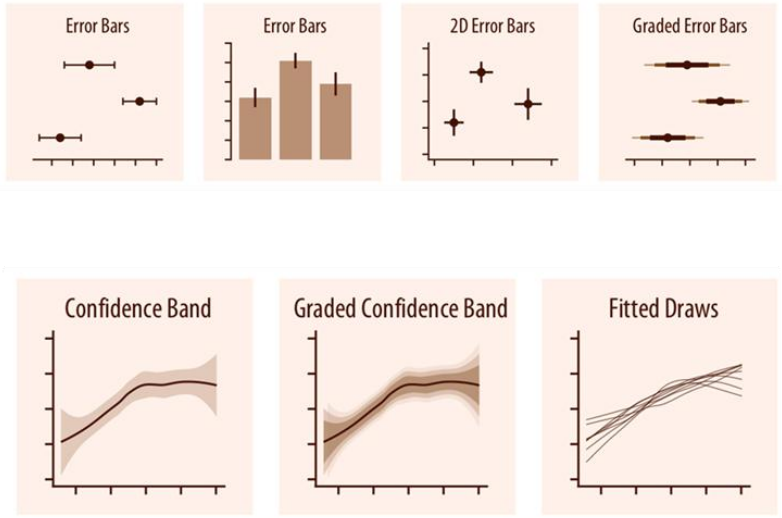
- **Scatterplots** represent the archetypical visualization when
  - we want to **show one quantitative variable relative to another**
- **Paired data** can also be shown as a **slope graph** of
  - **paired points** connected by straight lines



4.1 Choosing the Visuals.

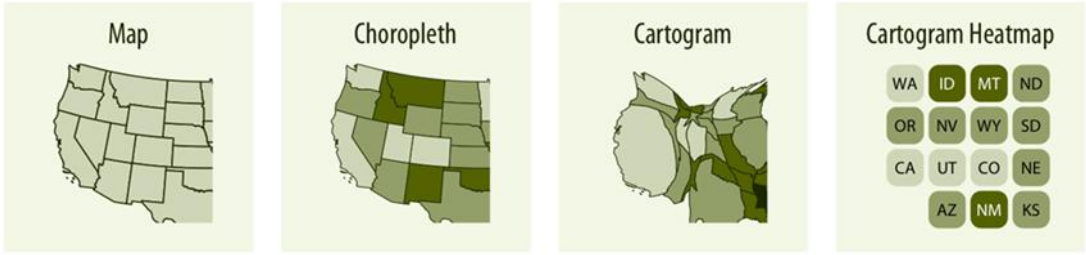
4.1.5 Visualizing Uncertainty.

- **Error bars** show **the range of likely values** for
  - an estimate, extending from the reference point horizontally or vertically.
- **For line graphs**, **confidence bands** serve the same purpose,
  - showing the range the line may pass through at a given confidence level.



4.1.6 Visualizing Geospatial Data.

- **Geospatial data** is usually shown on a **map**,
  - which projects globe coordinates onto a flat surface,
  - preserving shapes and distances approximately.
- In a **choropleth map**, regions are colored based on data values to visualize differences across areas.



4.2 Visuals to Avoid – Never use Pie Chart.

- **Pie Charts are evil**:

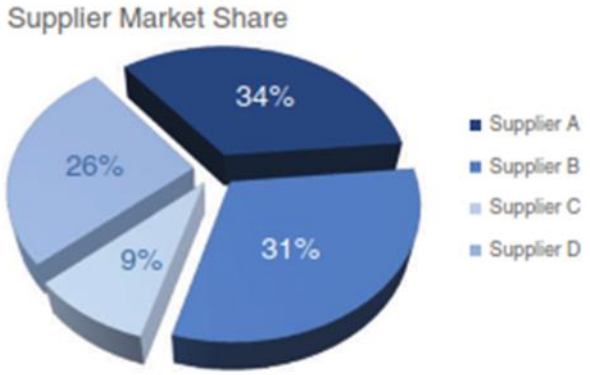


FIGURE 2.22 Pie chart with labeled segments

© "story telling with data" – Cole N. Knaflic.

- **Make an Observation:**
  - *Which supplier is the largest based on this visual?*

- **Alternate to Pie Charts:**

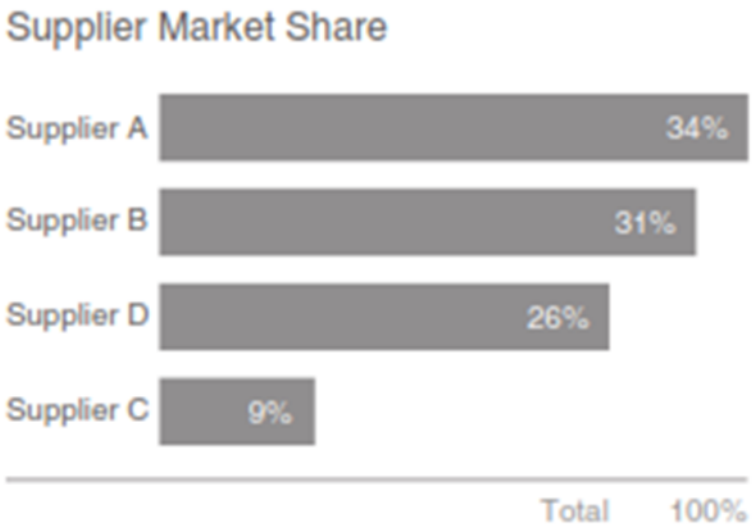


FIGURE 2.23 An alternative to the pie chart

© "story telling with data" – Cole N. Knaflic.

4.3 Visuals to Avoid – Misleading Scales.

- **Secondary y – axis** is **not a good idea** **in general**.

FIGURE 2.26 Secondary y-axis

- **Suggested Alternatives:**

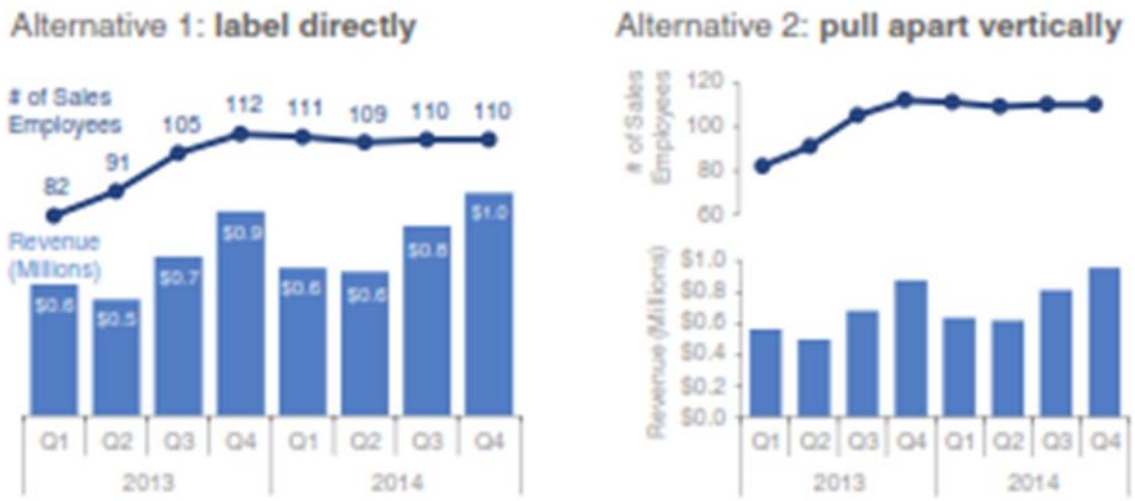
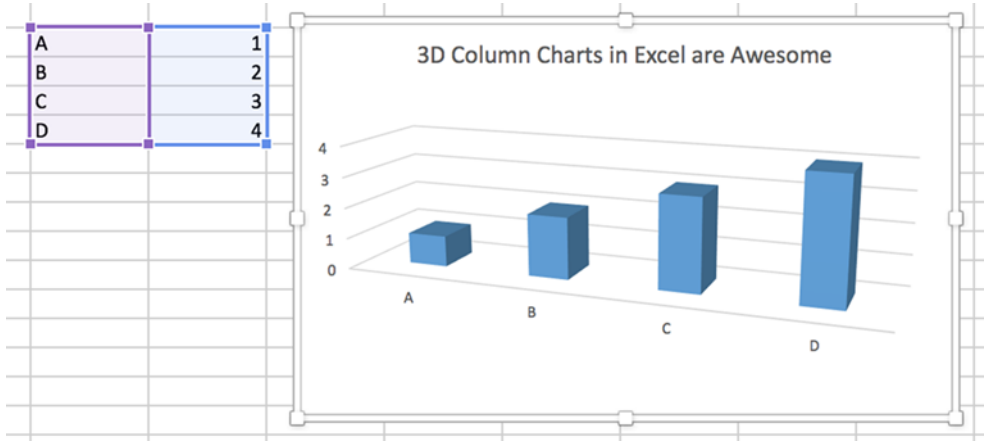


FIGURE 2.27 Strategies for avoiding a secondary y-axis

4.4 Visuals to Avoid – No 3D Plots

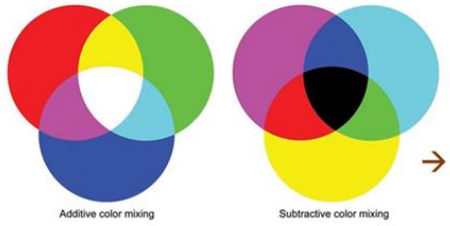
- Never use **3D Plots**:



Section-II
The Graphics.
{5. Choosing the Colors}

5.1 Understanding the Colors : HUE.

- **HUE:**
  - Hue is the **base color**.
  - **Primary hues cannot be formed** by **mixing other colors.**
  - Two families of **primary Hues**:
    1. **Additive (light-based):**
       - Colors are created by adding light of different colors together.
       - The more colors you add, the closer to white the result becomes.
         - **Example: Red + Green + Blue Light ⇒ White.**
       - Used in screens, monitors, projectors.
    2. **Subtractive (pigment – based):**
       - Colors are created by subtracting (absorbing) light using pigments or inks.
       - The more pigments you mix, the closer to black the result becomes.
         - **Example: Cyan + Magenta + Yellow ink ⇒ Black.**
       - Used in printing, painting , dyes.



→ Fig: Primary Colors (Hue)

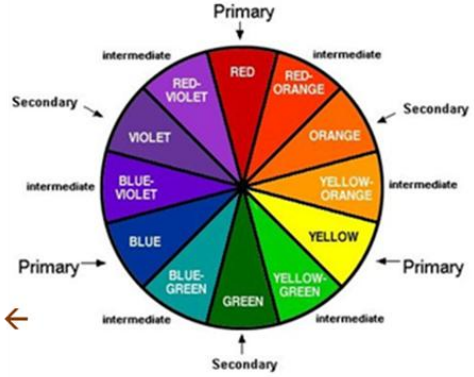


Fig: Secondary Colors ←

5.1.1 Understanding the Colors : Saturation.

- **Saturation:**
  - **Saturation** measures the **purity or intensity of a hue**:
    - **100% saturation → completely pure hue, no gray.**
    - **Lower saturation → hue becomes more grayish.**
    - **0% saturation → all colors appear as gray.**
  - Think of it as the ***"Shades of Gray → amount of gray mixed into a color".***

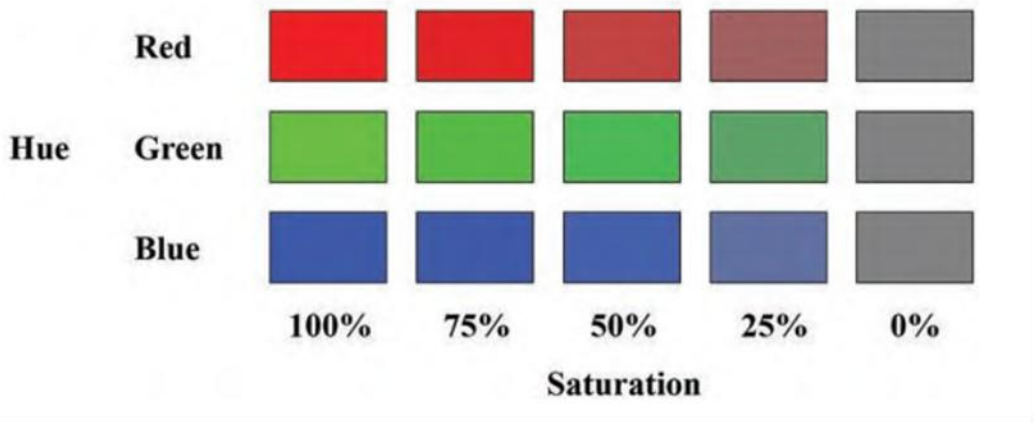


Fig: Primary Hues in the RGB primary color at 50% Luminance

5.1.2 Understanding the Colors : Luminance.

- **Luminance:**
  - **Luminance** measures the **lightness or darkness** of a color:
    - Adding **white** → lighter color
    - Adding **black** → darker color
    - Greater differences in luminance create greater contrast between them:
    - So, **luminance** is a **good way** to **indicate hierarchy or degree**.
      - **higher contrast, useful for indicating hierarchy or emphasis**
    - Human eye can discern only ~ 6–7 degrees of luminance in color.
  - 100% luminance → white; 0% luminance → black

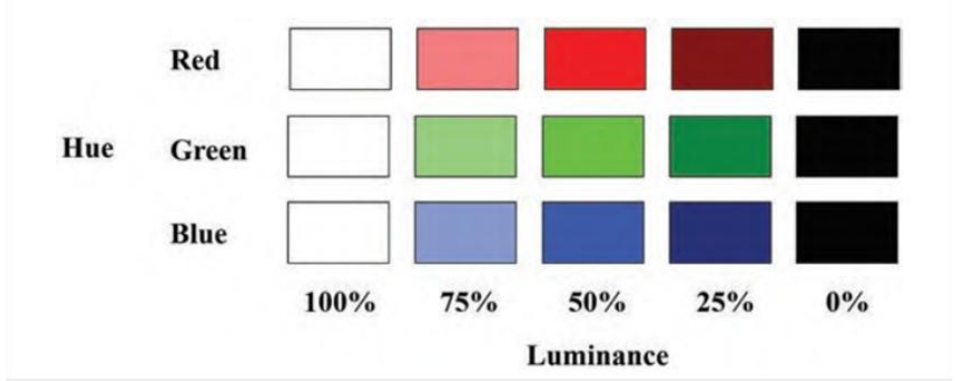


Fig: Primary Hues in the RGB primary color at different levels of luminance (at 100% Saturations)

5.2 Use of Colors !!!

- If used **wisely**, color could be the most powerful tools you have for drawing your audiences' attention.
- **How to use wisely?**
  - Resist the urge to use color for the **sake of being colorful**
  - Instead leverage color as **a strategic tool** to **highlight** the important parts of your visuals
  - **Never let your tool make this important decision for you!**

5.3 Avoid: Excessive Use of Colors .

- There is a **limit to the amount of information** that can be **communicated to the audience** using **color**.

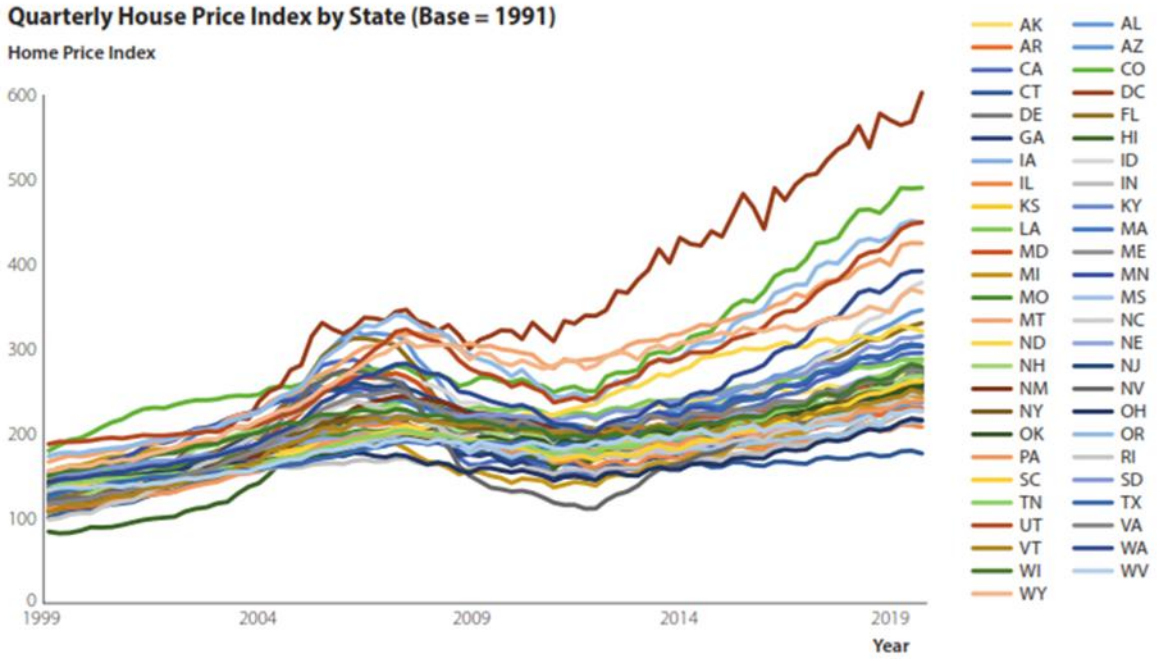


Fig: Avoid Excessive Use of Colors!!!

- **How can you make it better?**

5.3.1 Avoid: Excessive Use of Colors .

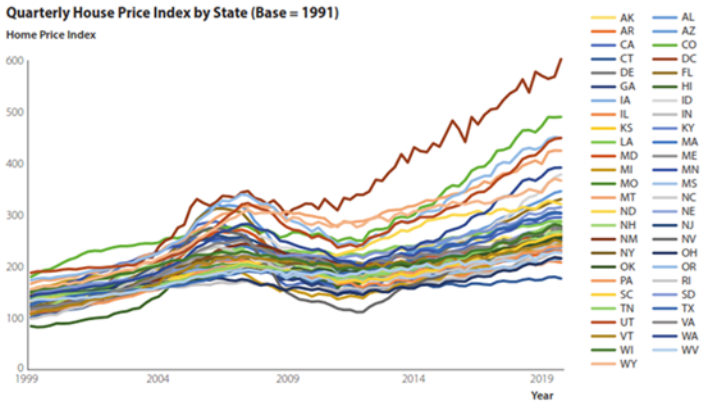
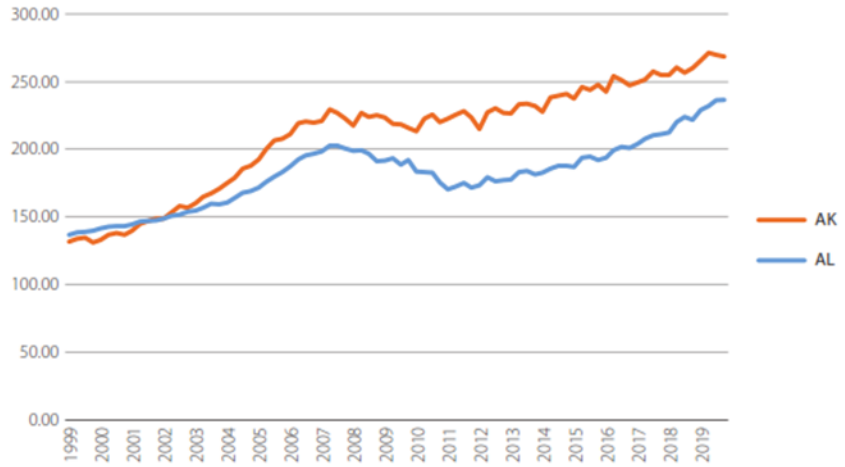
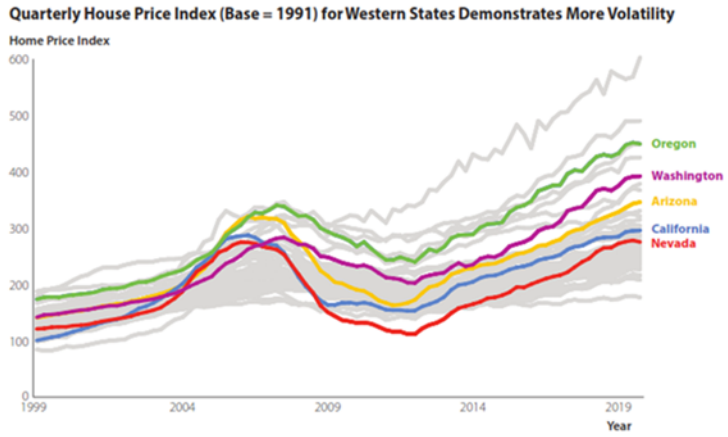


Fig: Avoid Excessive Use of Colors!!!

- How can you make it better?



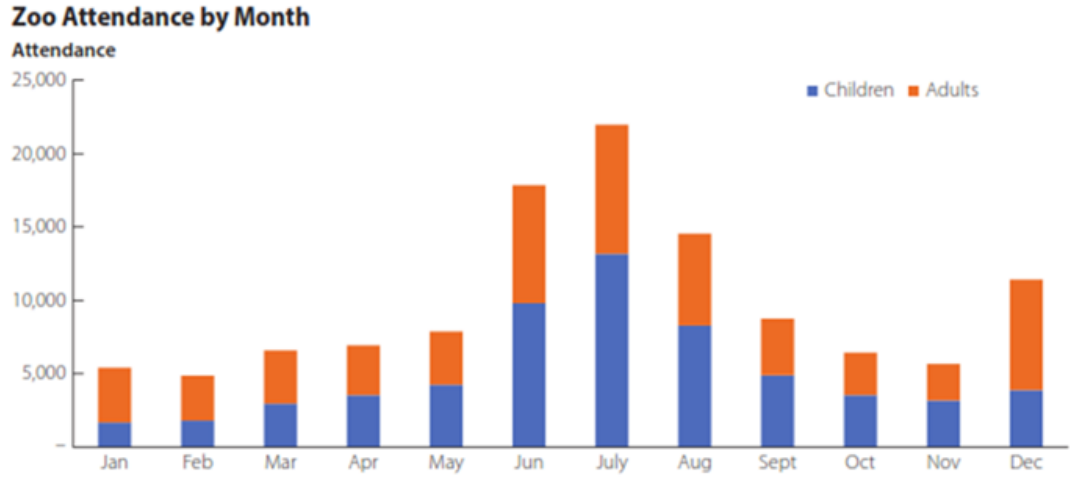
Decompose

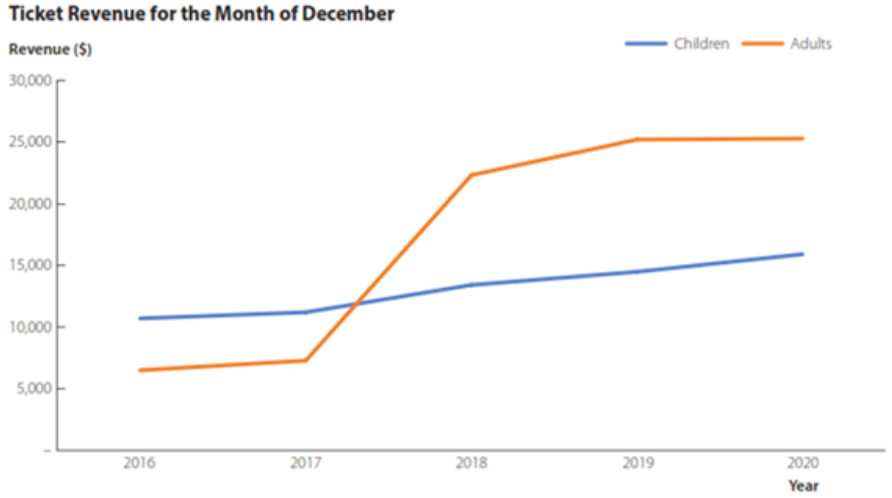


Highlight.

5.3.2 Avoid: Inconsistent Colors .

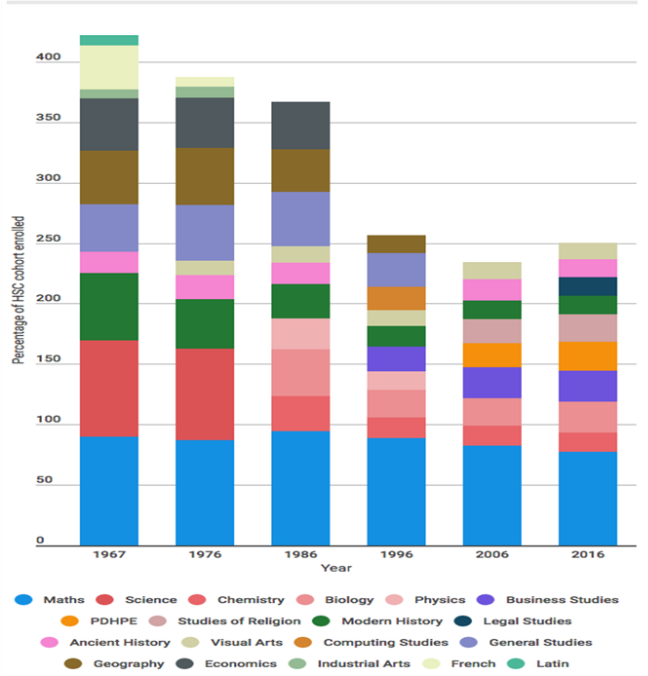
- When **creating several charts** for a **single report**, **a presentation**, **an ongoing analysis**, or a **data dashboard**, it is critical that **color is used consistently**.
- Using **different color schemes** or using **different colors** to represent categories on different charts will confuse the audience and dramatically increase its cognitive load.





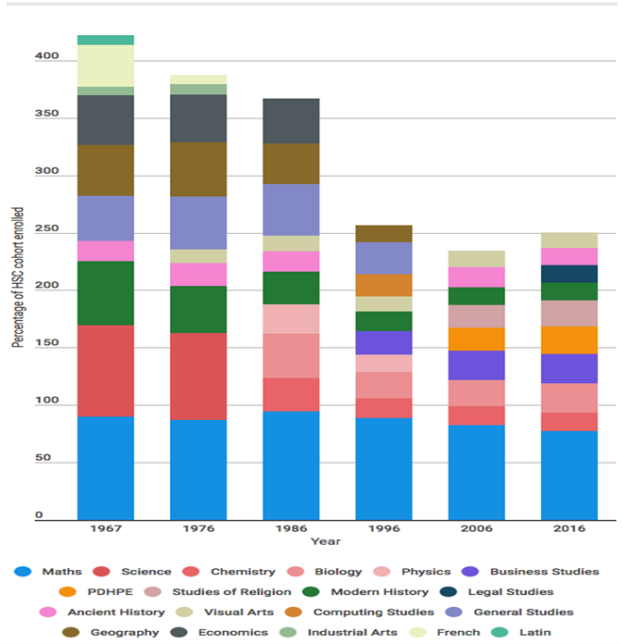
5.3.3 Avoid: Un – necessary Colors .

- **Color should** only be used **when it communicates** something that **no other aspect of a chart communicates** to the audience
- In this chart,
  - **the audience can discern** which **column corresponds to each of the models** through **the colors of the columns and the legend.**
    - although this communicates the data, we can accomplish the same communication with a chart that creates less cognitive load by avoiding the use of multiple colors.
- **How can you make it better?**



5.3.3 Avoid: Un – necessary Colors .

- **Color should** only be used **when it communicates** something that **no other aspect of a chart communicates** to the audience
- **In this chart,**
  - **the audience can discern** which **column corresponds to each of the models** through **the colors of the columns and the legend.**
    - **although** this **communicates** the data, we can accomplish the same communication with a chart that creates less cognitive load by avoiding the use of multiple colors.
- **How can you make it better?**



- **Redesign and Rethink:**
  - May be bar chart is not the right plot.

5.3.4 Do not neglect: Color Blindness.

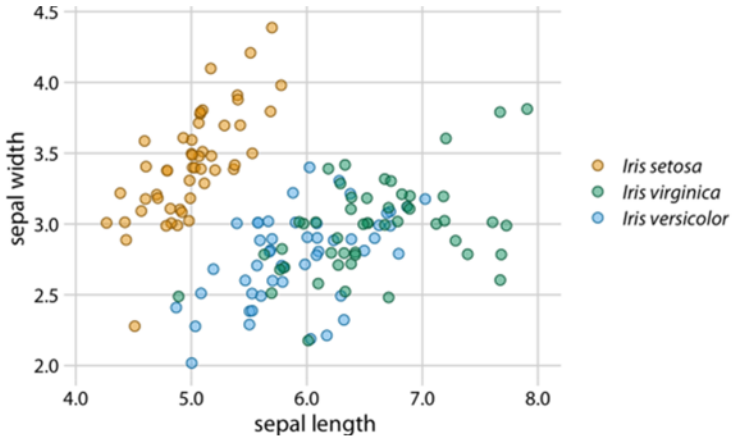


Fig: Your Graph.

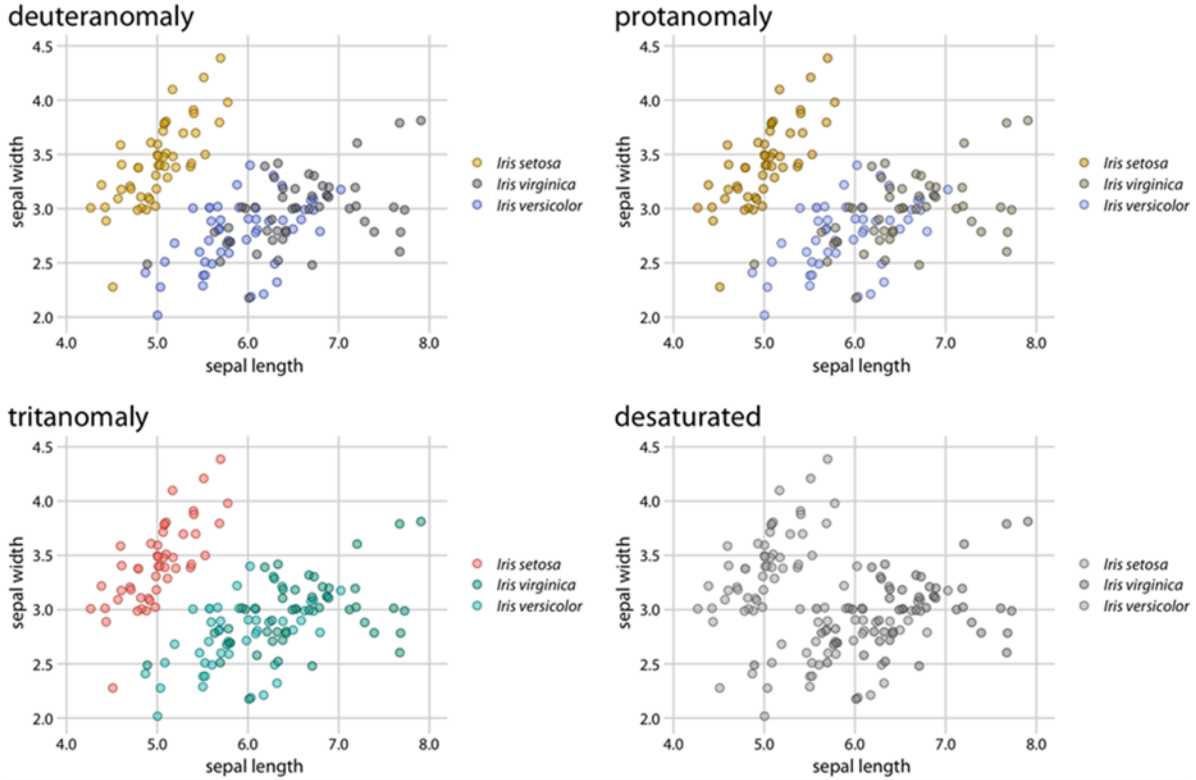


Fig: What Colorblind people sees

5.3.5 Do not neglect: Color Blindness.

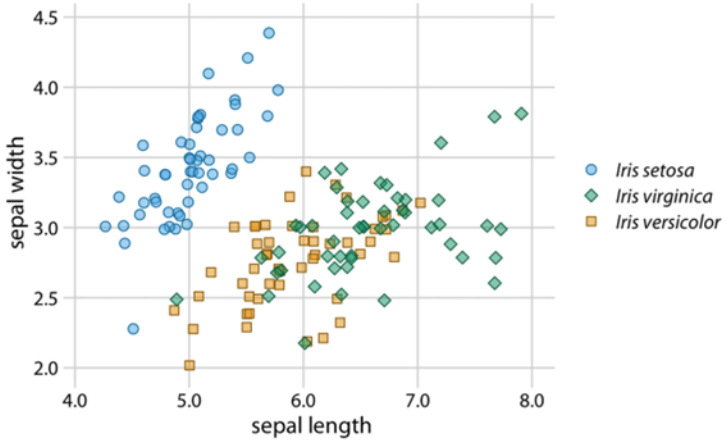


Fig: Your Graph

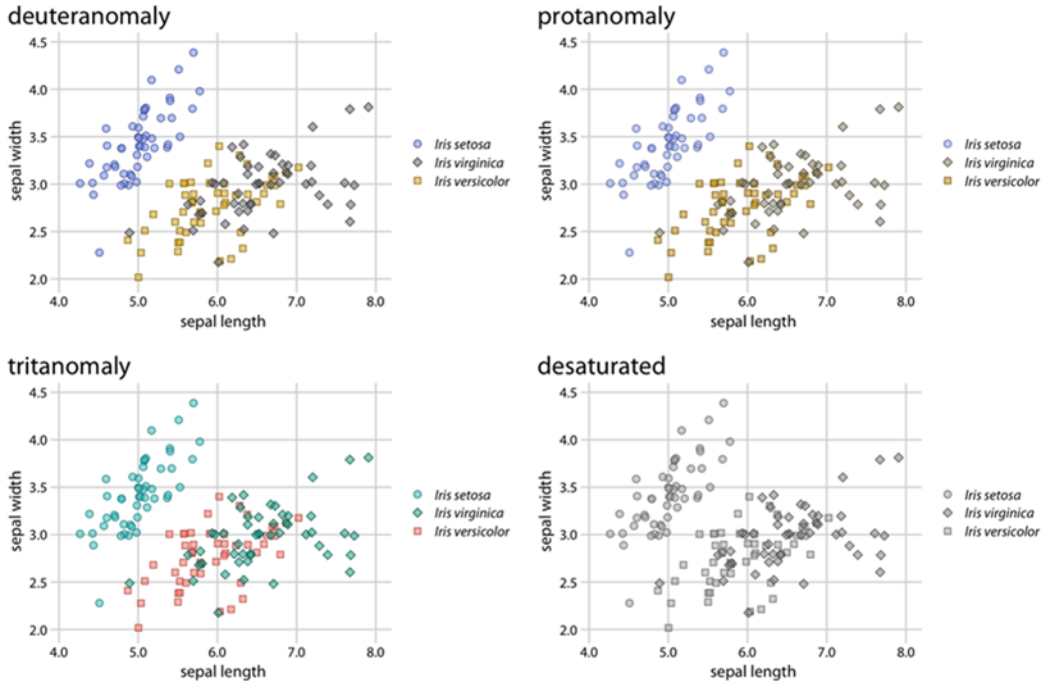


Fig: What Colorblind people sees

Section-III.

{6. What makes it Garbage!!}

6.1 Focus on Your Audience

- **You see with your brain:**
  - We do not fully see with our eyes, there is some processing that happens in eye, but majority of stuff happens in our brain that we think of as visual perception.

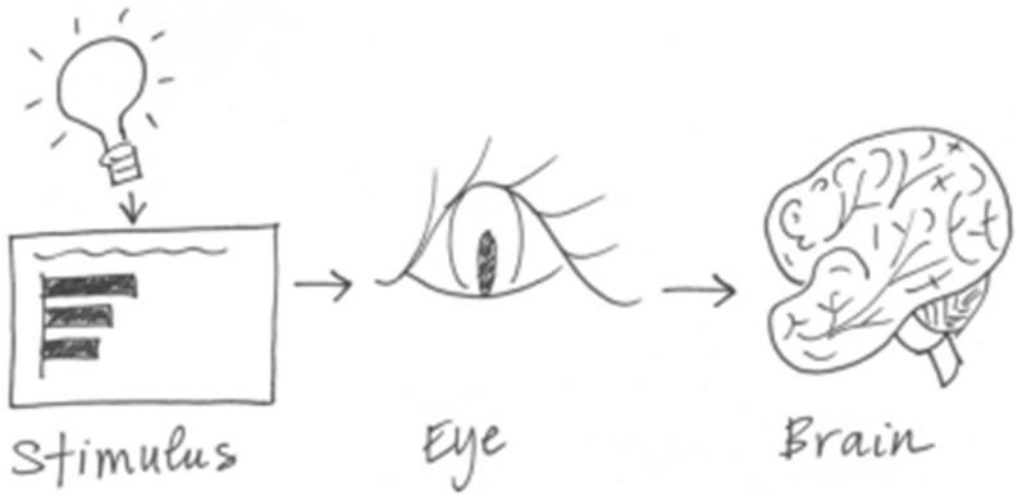


FIGURE 4.1 A simplified picture of how you see

6.2 Where to look?

- Add **some preattentive attributes** signal to guide your audience where to look.
- **Example: Count the three.**

FIGURE 4.2 Count the 3s example

6.2.1 Where to look?

- Add some preattentive attributes signal to guide your audience where to look.
- **Example: Count the three.**

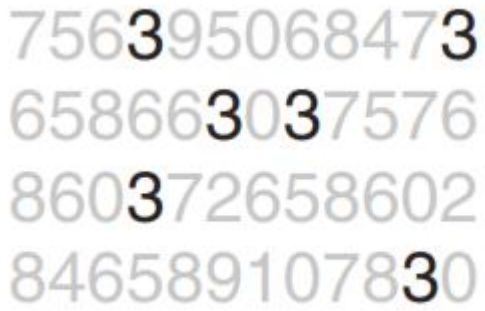


FIGURE 4.3 Count the 3s example with preattentive attributes

© "story telling with data" – Cole N. Knaflic.

- If we **use preattentive attributes** strategically, they can help us enable our audience to see what we want them to see before they even know they're seeing it!

6.3 Avoid Cognitive Overload

- Cognitive load can be thought of as the mental effort that's required to learn new information.
- **When designing a Visualization,**
  - we must try to reduce the Cognitive Load i.e. avoid the components in your graphics
  - *"that takes up large number of mental resources but does not help the audience understand the information."*

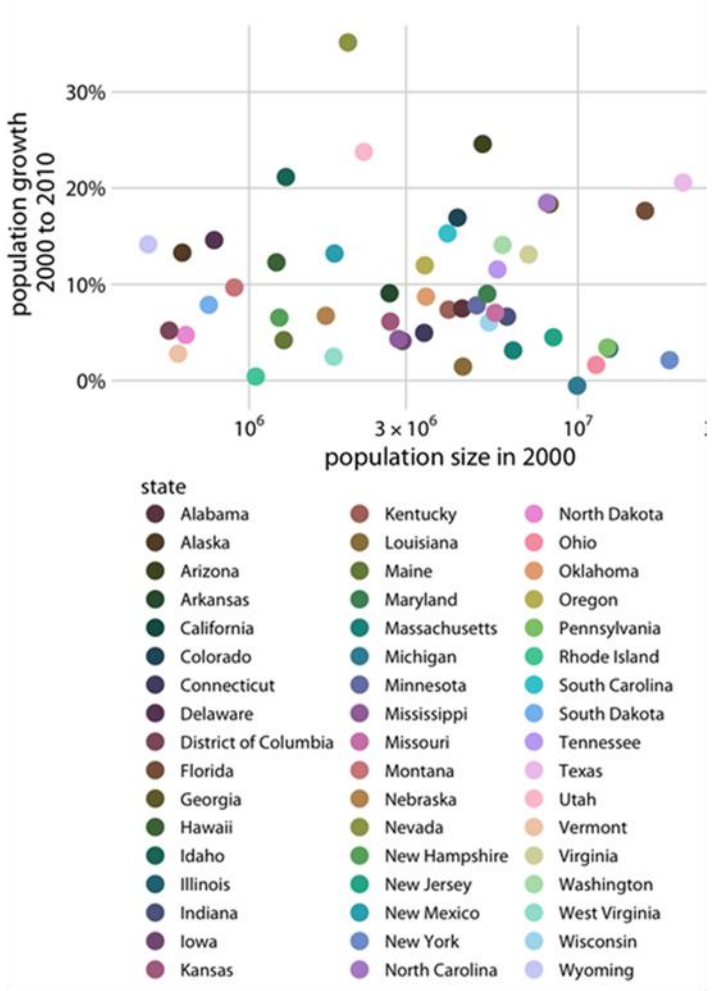


Fig: Visualization with High Cognitive Overload …

6.4 Avoid Clutter!!

- **One culprit** that **contribute to excessive or extraneous cognitive load** can be **referred as clutter**.
- These are visual elements that take up space but do not increase understanding.
- How to identify what elements are clutter?

5.4 Avoid Clutter!!

- **One culprit** that **contribute to excessive or extraneous cognitive load** can be **referred as clutter**.
- These are visual elements that take up space but do not increase understanding.
- **How to identify what elements are clutter?**
  - **Gestalt principles of visual Perception:**
  - **Gestalt psychology** explains **how humans naturally organize and perceive visual information**.
  - Using these principles helps you **reduce clutter** by making visuals more intuitive:
    - **Proximity:**
      - Elements close together are seen as related. (Group related data, avoid random spacing.)
    - **Similarity:**
      - Similar shapes, colors, or sizes are perceived as belonging together.
    - **Continuity:**
      - The eye follows continuous lines or curves, not abrupt breaks.
    - **Closure:**
      - The brain fills in missing shapes, so you don't need to over – outline everything.
    - **Figure/Ground:**
      - We distinguish objects (figure) from background (ground). Use contrast effectively.
    - **Common Fate:**
      - Elements moving or oriented in the same direction are grouped mentally.

Example – Removing Clutter …

1. How to Declutter?

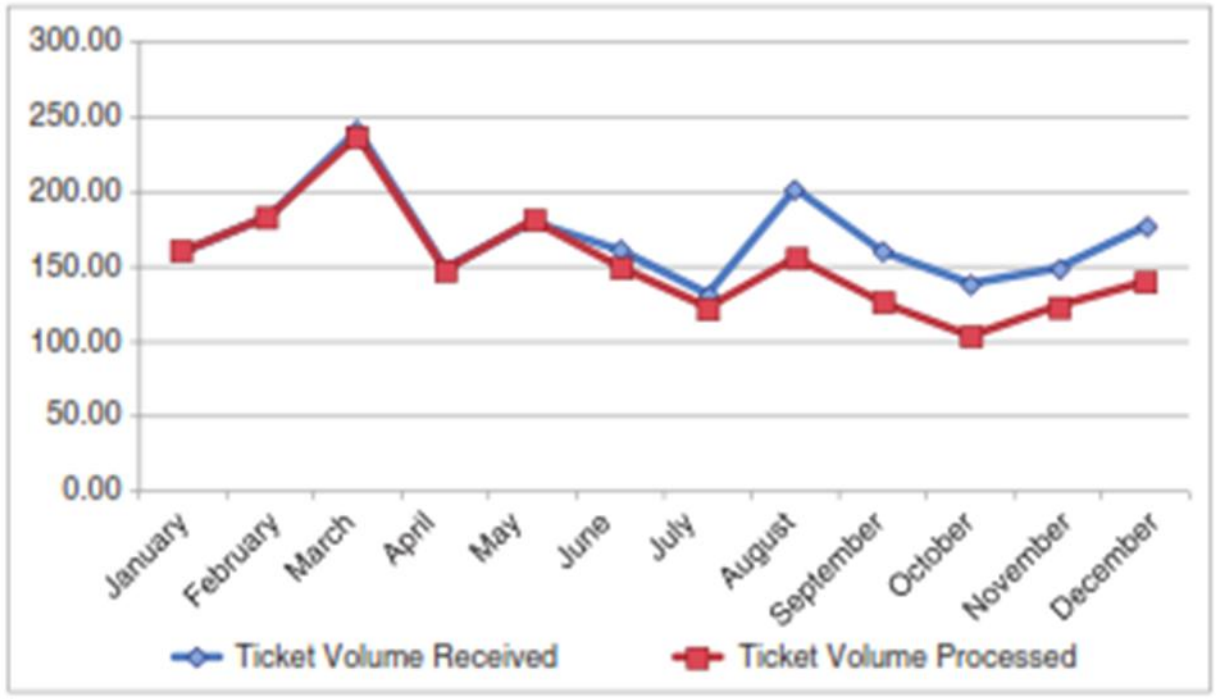


Fig: Original Graph

2. Remove Chart Border!

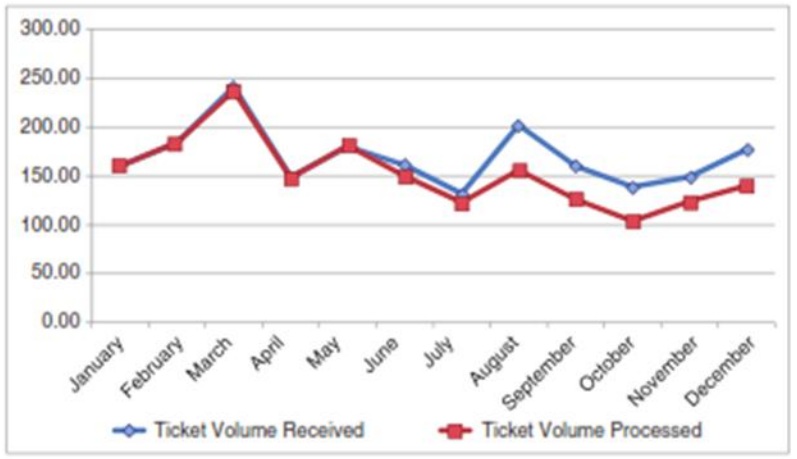
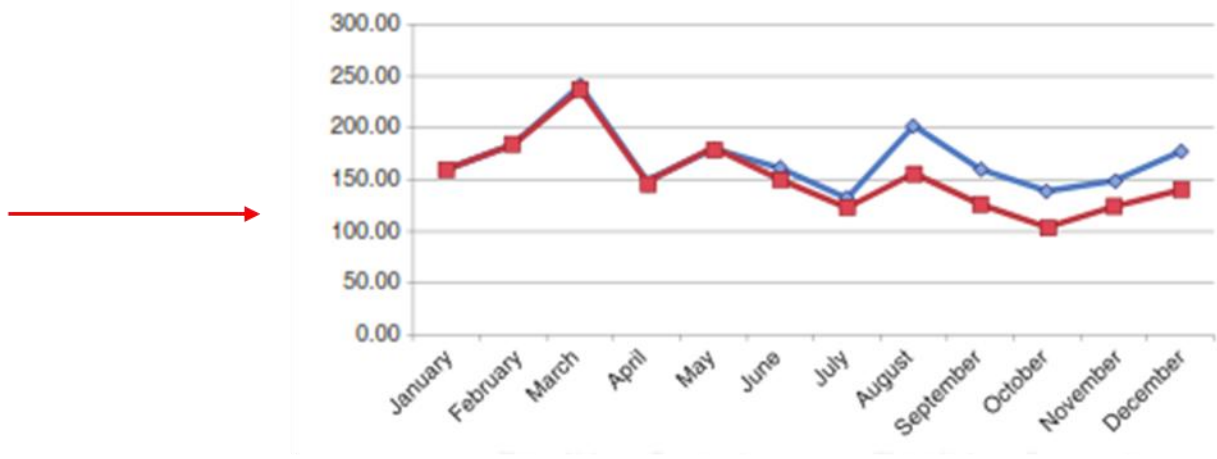


Fig: Original Graph



3. Remove Gridlines!!

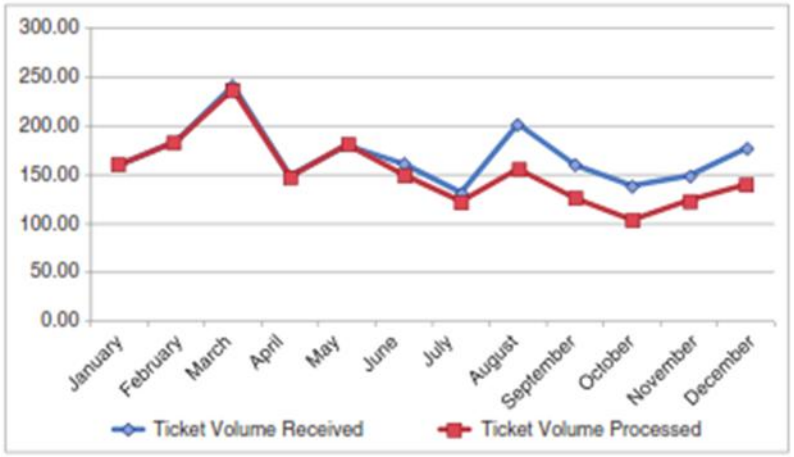
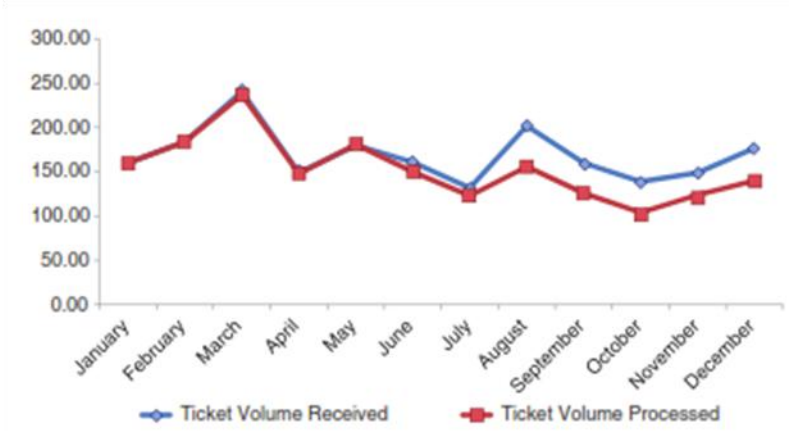


Fig: Original Graph



4. Remove Data Markers!!

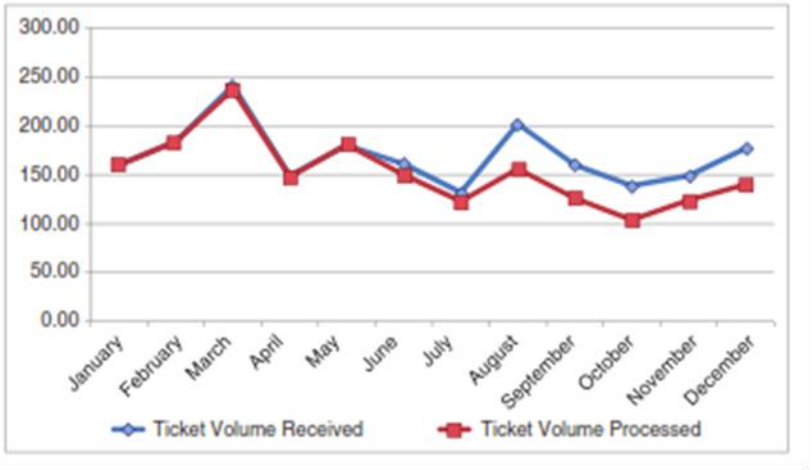
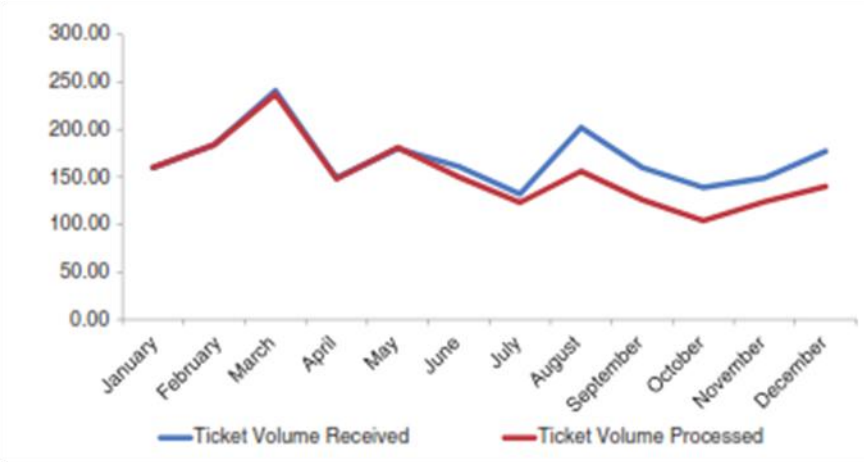


Fig: Original Graph



5. Clean up axis labels!!

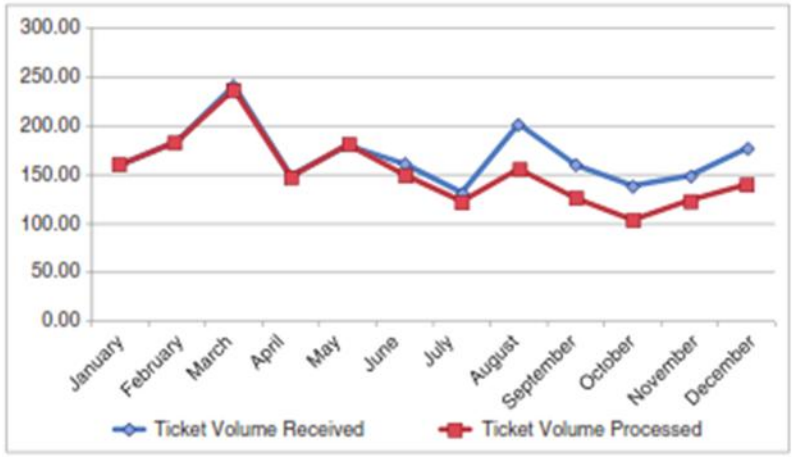
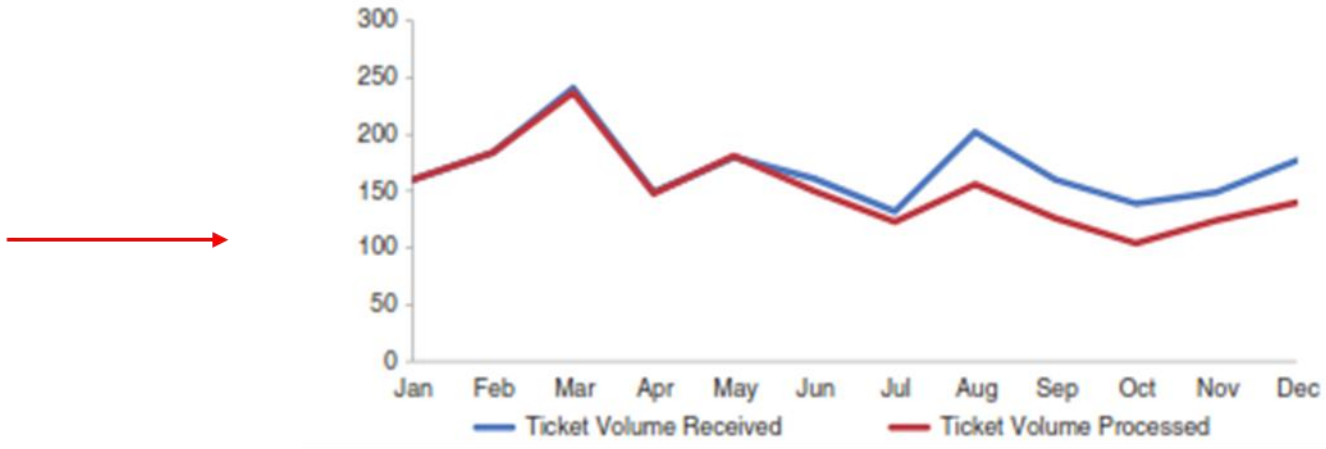


Fig: Original Graph



6. Label data directly

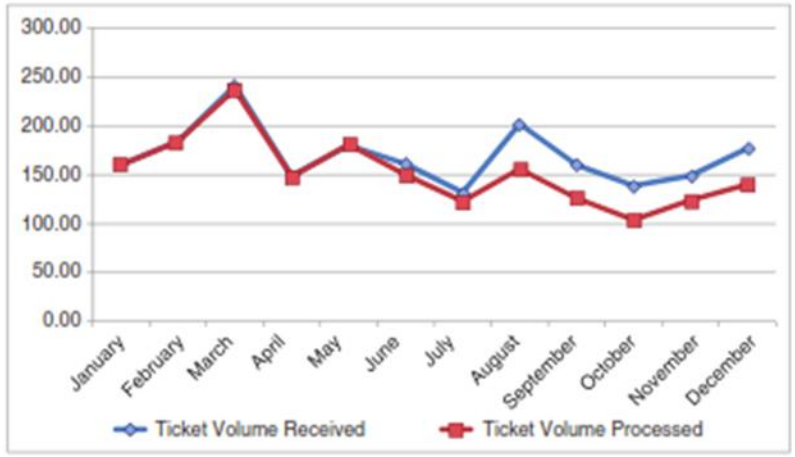
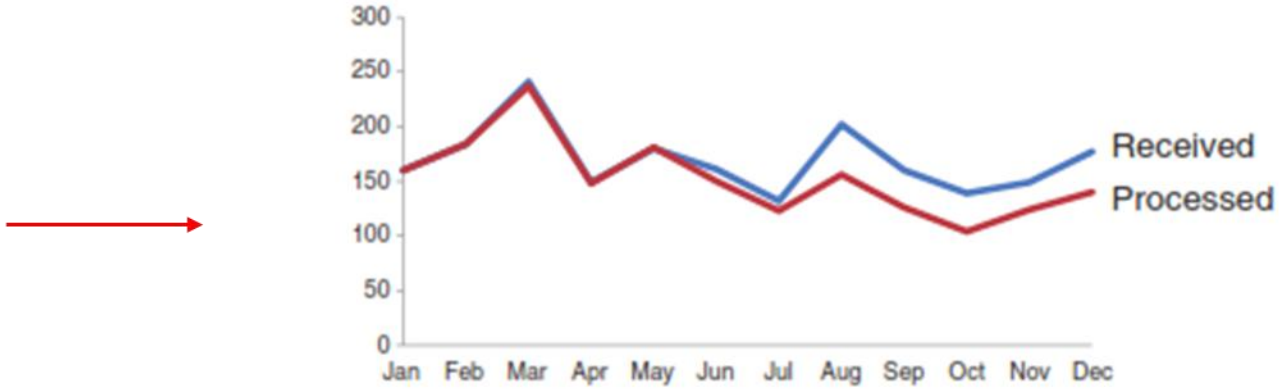


Fig: Original Graph



7. Leverage Consistent Color!!

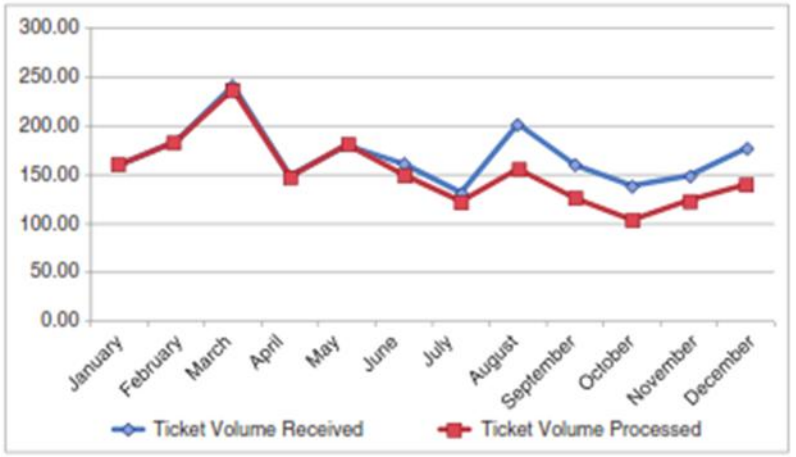
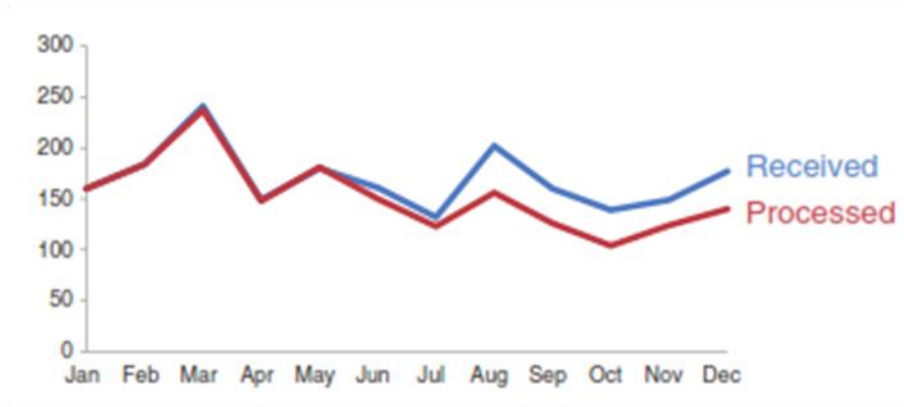
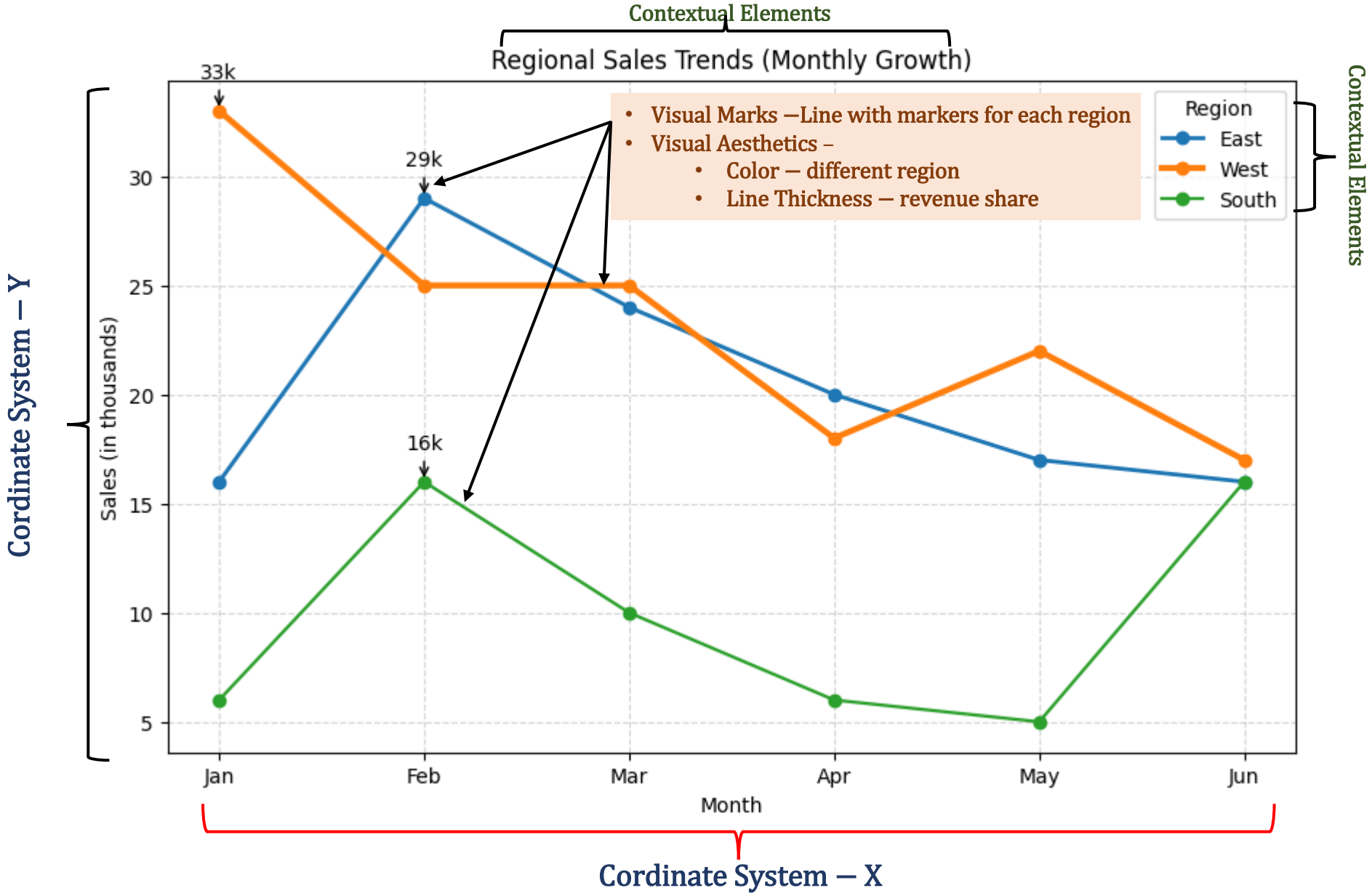
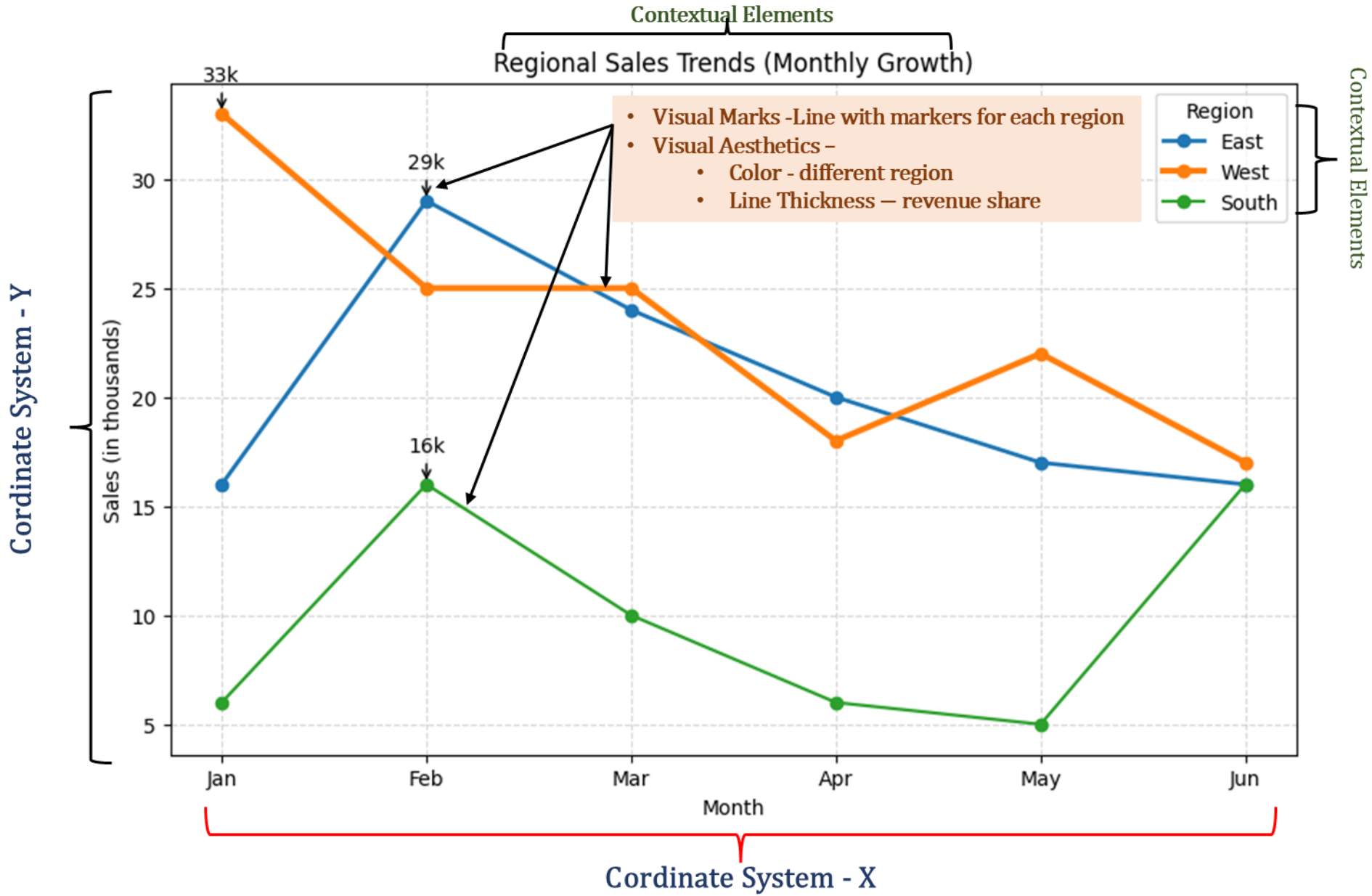


Fig: Original Graph



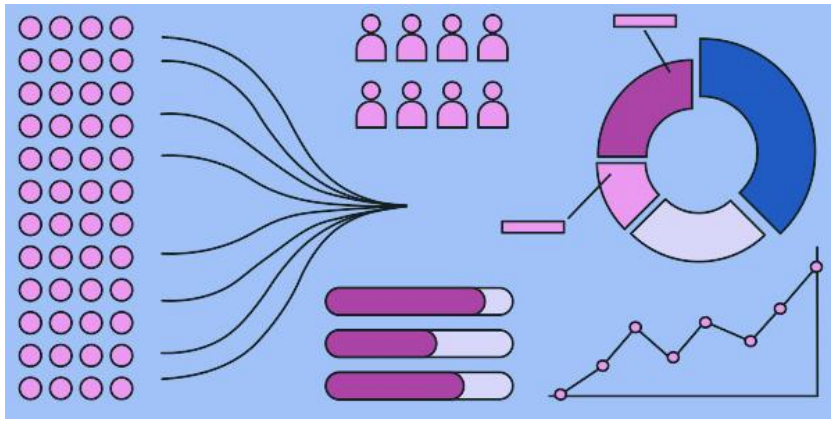
Thank You


Contextual Elements
Regional Sales Trends (Monthly Growth)
33k
29k
16k
• Visual Marks –Line with markers for each region
• Visual Aesthetics –
• Color – different region
• Line Thickness – revenue share
Region
East
West
South
Contextual Elements
Cordinate System – Y
Sales (in thousands)
30
25
20
15
10
5
Jan
Feb
Mar
Apr
May
Jun
Month
Cordinate System – X


Contextual Elements
Regional Sales Trends (Monthly Growth)
33k
29k
16k
• Visual Marks -Line with markers for each region
• Visual Aesthetics –
• Color - different region
• Line Thickness – revenue share
Region
East
West
South
Contextual Elements
Cordinate System - Y
Sales (in thousands)
30
25
20
15
10
5
Jan
Feb
Mar
Apr
May
Jun
Month
Cordinate System - X

THE WHO (AUDIENCE)

THE WHAT (DATA)



THE HOW (Mechanism)

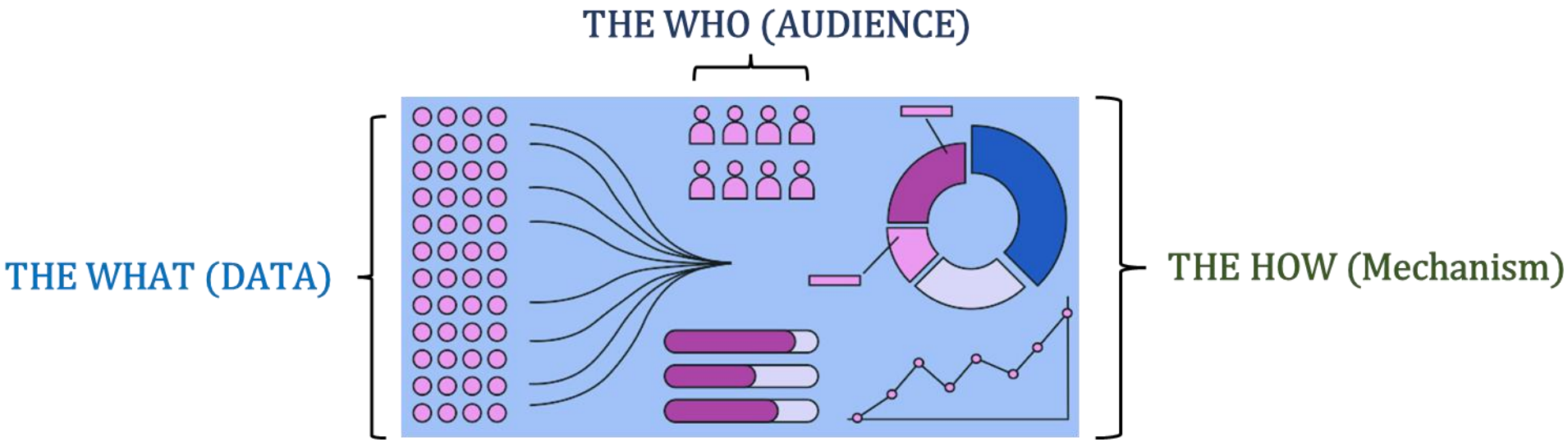


Fig: Communicating with DATA.

1.3.2 Evolution → Analyze a Data.

First time to show changing values graphically

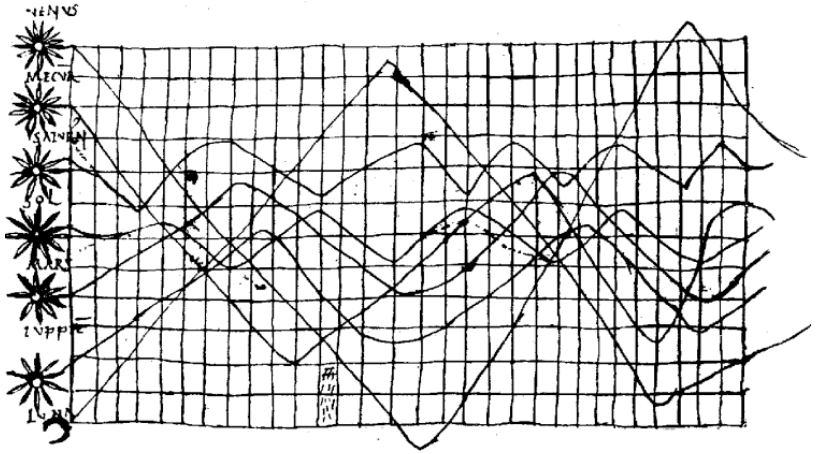


Fig: Planetary Movement Diagram, c. 950

First weather map

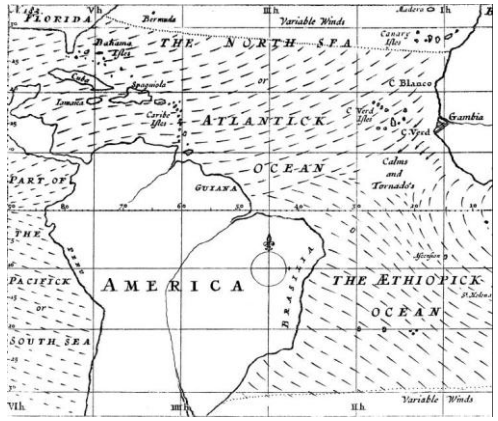


Fig: Halley's Wind Map, 1686

Import and Export of Scotland in 1781

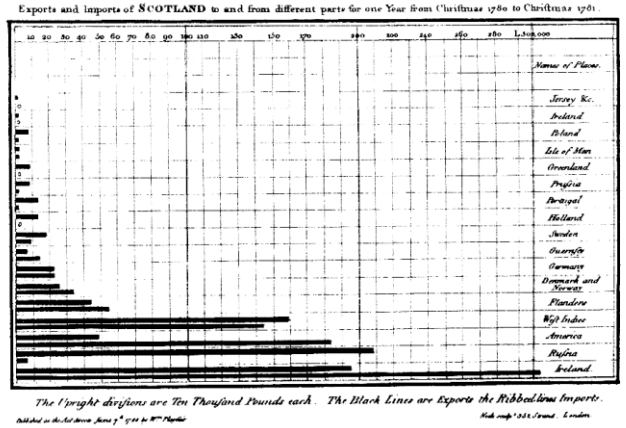


Fig: bar graph 1781