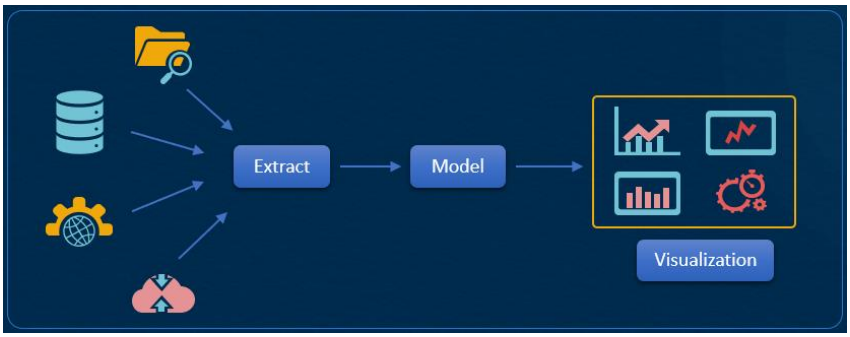
Data Visualization and Storytelling

Content Overview: -

- Data Visualization
  - Definition
  - Significance
  - Data Visualization for decision making
- Chart
  - Definition
  - Elements of Charts
  - Types of Charts with best practices
    - Bar Chart
    - Column Chart
    - Line Chart
    - Dual Axis Chart
    - Area Chart
    - Stacked Bar Chart
    - Pie Chart
    - Scatter Plot Chart
    - Waterfall Chart
    - Funnel Chart
    - Heat Map
- Dashboards
  - Definition
  - Types of Dashboards
- Storytelling and Data Visualization: Best practices

Data Visualization

Data Visualization is a process of converting data in the form of chart, diagram, picture etc. to help the decision making in mean time.
Data visualization is the graphical representation of information and data. By using visual elements like charts, graphs, and maps, data visualization tools provide an accessible way to see and understand trends, outliers, and patterns in data.
Additionally, it provides an excellent way for employees or business owners to present data to non-technical audiences without confusion.



So as per there are 3 steps involved till data visualization step.

- Extraction
- Model
- Visualization

As we know extraction is extracting the data from data source. We store the extracted data in local system to ensure the data availability all the time. Once the data is being stored in local systems. It is passed to model. So, what is model. In model step basically data cleaning is done by manipulating the data. In data cleaning operations like removing null or duplicate values, creating links between table etc. has done. Through data cleaning, data is transformed into most compatible form to make visualization. So, model part is very essential because an inefficient model can give performance issues in visualization stage.

The outcome of the model step may be a single table or multiple tables which are interlinked together. Then this data is used for creating visualization like charts, graphs etc. For creating charts, we need dimensions and metric. These visualizations help in generating useful insights for strategic business decisions.

Data Visualization Significance

- Data visualization allows users to see several different perspectives of the data.
- Data visualization makes it possible to interpret vast amounts of data.
- Data visualization offers the ability to note exceptions in the data.
- Data visualization allows the user to analyze visual patterns in the data.
- Exploring trends within a database through visualization by letting analysts navigate through data and visually orient themselves to the Patterns in the data.

Our eyes are drawn to colors and patterns. We can quickly identify red from blue, and squares from circles. Our culture is visual, including everything from art and advertisements to TV and movies. Data visualization is another form of visual art that grabs our interest and keeps our eyes on the message. When we see a chart, we quickly see trends and outliers. If we can see something, we internalize it quickly. It's storytelling with a purpose. If you have ever stared at a massive spreadsheet of data and couldn't see a trend, you know how much more effective a visualization can be.

Data visualization data allows users of to see several different perspectives of the data. If I give you some millions of records and ask you to analyze, it's not that easy for anyone to analyze the data in a couple of minutes. But the same data is going to be represented in the form of a bar graph or a pie chart or just seeing picture itself you can easily analyze which area or which customer or where the sales has been high or when the sales has been low. So, data visualization allows the user to see the data in a very different perspective. Data visualization makes it possible to interpret vast amounts of data.

Data visualization also offers the ability note exceptions in data. It allows a user to analyze visual patterns in the data, because you are going to convert your data into various formats. Data visualization also leads to data analytics where you can analyze the trend within your data. We can also navigate from higher level of hierarchy to the lower level of hierarchy using different types of graphs in the data visualization.

Data Visualization for Decision Making

To make informed decisions, you need accurate data. Data visualization can help you get more of the right data and derive valuable insights from it. For Example

Improved Speed and Efficiency: - Data visualization saves time by summarizing a large amount of information into an easy-to-understand format that can be understood within seconds. Data analytics can also be done in real-time with a graphical element. This allows you to always monitor the performance of your business and take necessary action right away. Even more so, it helps you see all your data at once in a way you understand, also in real time.

Greater Accuracy:

Data visualization may be objective and easy to understand, which results in increased accuracy. However, it's important that you use the right visualization type for your data to be able to draw correct conclusions from the analysis. For example, pie charts are not great at showing numerical data with many categories because they imply equal value among slices even if that isn't the case.

Simplified Communication: -

Easy-to-understand reports also impact how you share information with other people in your company. For example, you might use a flowchart or bubble chart to show different stages of a process or how certain operations are linked to overall performance. Visualization can thus prove to be an important tool for effective internal communication.

In turn, this fosters a great level of collaboration across your team and across departments which previously may have never even communicated.

Empowered Collaboration: -

Better data collaboration within your team allows everyone in your company to see the complete picture and make informed decisions based on this information rather than having a limited view of the data.

Improve collaboration, communication, and decision-making in your company simply by visually representing the data your company is dealing with.

Chart

A chart is a graphical representation of worksheet data. Charts can make data interesting, attractive, and easy to read and evaluate. They can also help you to analyze and compare data. A chart combines text, symbols and/or graphics to show the relationship between multiple data sets. They are a popular method for sharing information in different settings.

Charts enable you to visually compare multiple sets of data. It can help people better understand and remember information. Many people understand a picture more quickly than blocks of text.

A compelling chart can help you make your point more convincingly and lend credibility to your presentation.

Charts are most valuable when you need to quickly convey the relationship between two or more sets of data, and you'd like your audience to retain this information. They have countless uses, as modern businesses lean on data to shape their strategies and priorities.

For instance, a chart can quickly show how sales and/or expenses have changed for a business unit over a set period of time. Or it could display something more specific, like an increase or decline in the number of orders fulfilled per hour at a particular warehouse.

Elements of Chart

While the various types of charts show information in different ways, most share a few common elements. These include the following:

Title

The title should let viewers quickly grasp the data presented in the chart. An example: "Sales of Products A, B and C Over the Past Five Years."

The Horizontal and Vertical Axes

Apart from pie charts, most charts have vertical and horizontal axes. Each is used to present at least one set of data. Both also incorporate scales that indicate the starting and ending point of the information presented.

Horizontal Axis

The horizontal axis, also known as the x-axis, typically represents the value that's known — often the time. Say a chart shows ice cream sales over time. The years typically would go on the horizontal axis.

Vertical Axis

The vertical or y-axis usually shows the "unknown" value. In the above example of ice cream sales, the volume of ice cream sales for each of the relevant years would be shown on the vertical access.

Data Series

The data series refers to the information set presented in the chart. All charts are based on some type of data series. Returning to the example of ice cream sales, it would be the amount of ice cream sales each year.

Key

A key or legend helps explain the information in the chart. Say the chart breaks out ice cream sales by region and uses differently colored bars for each region. The key would indicate which color correlated with which region.

There are myriad different types of charts, graphs and other visualization techniques that can help analysts represent and relay important data.

Types of Charts & Graphs

Bar Chart

- The bar chart or bar graph is one of the most common data visualizations. They're sometimes also referred to as column charts. Bar charts are used to compare data along two axes.
- One of the axes is numerical, while the other visualizes the categories or topics being measured. Bar graphs can help you compare data between different groups or to track changes over time.
- Bar graphs are most useful when there are big changes or to show how one group compares against other groups.
- You can use a bar chart with vertical bars or horizontal bars. On vertical bar graphs, numerical values are on the y axis (vertical axis); on horizontal bars, they are on the x-axis (horizontal axis).
- A bar graph should be used to avoid clutter when one data label is long or if you have more than 10 items to compare.

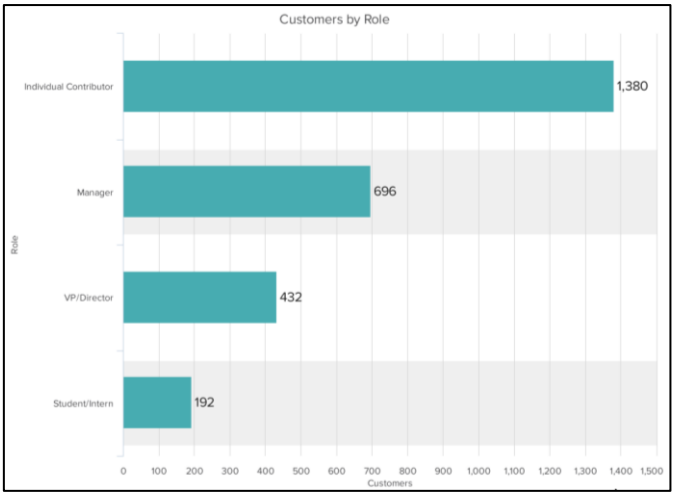
Use Case

The example below compares the number of customers by business role.

It makes it easy to see that there is more than twice the number of customers per role for individual contributors than any other group.

A bar graph also makes it easy to see which group of data is highest or most common.

For example, at the start of the pandemic, online businesses saw a big jump in traffic. So, if you want to look at monthly traffic for an online business, a bar graph would make it easy to see that jump.



Other use cases for bar graphs include:

- Product comparisons
- Product usage
- Category comparisons
- Marketing traffic by month or year
- Marketing conversions

Column Chart

This is one of the most common types of data visualization tools. There's a reason we learn how to make column charts in elementary school. They're a simple, time-honored way to show a comparison among different sets of data. You can also use a column chart to track data sets over time.

A column chart will include data labels along the horizontal (X) axis with measured metrics or values presented on the vertical (Y) axis, also known as the left side of the chart. The Y-axis will normally start at 0 and go as high as the largest measurement you're tracking.

You can use column charts to track monthly sales figures, revenue per landing page, or similar measurements. Consistent colors help keep the focus on the data itself, though you can introduce accent colors to emphasize important data points or to track changes over time.

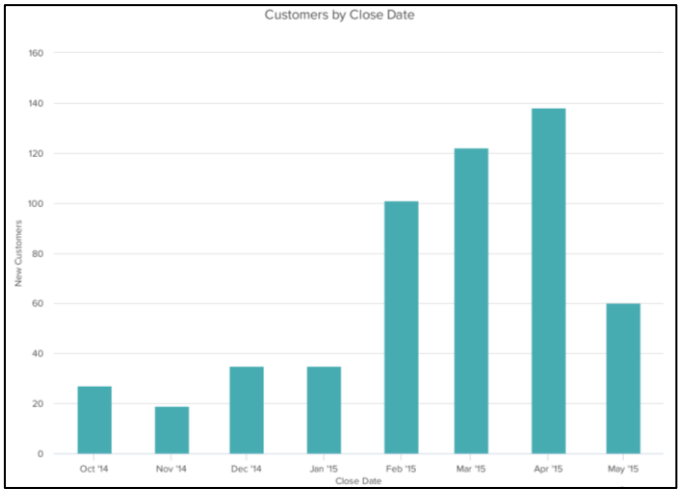
Design Best Practices for Column Charts:

- Use consistent colors throughout the chart, selecting accent colors to highlight meaningful data points or changes over time.
- Use horizontal labels to improve readability.
- Start the y-axis at 0 to appropriately reflect the values in your graph.

Use Case

While column charts show information vertically, and bar graphs show data horizontally. While you can use both to display changes in data, column charts are best for negative data.

For example, warehouses often track the number of accidents that happen on the shop floor. When the number of incidents falls below the monthly average, a column chart can make that change easier to see in a presentation.



In the example above, this column chart measures the number of customers by close date. Column charts make it easy to see data changes over a period of time. This means that they have many use cases, including:

- Customer survey data, like showing how many customers prefer a specific product or how much a customer uses a product each day.
- Sales volume, like showing which services are the top sellers each month or the number of sales per week.
- Profit and loss, showing where business investments are growing or falling.

Line Graph

A line graph reveals trends or progress over time, and you can use it to show many different categories of data. You should use it when you chart a continuous data set.

A line chart or line graph is a data visualization type that showcases changing data over time. Like a bar graph, the line chart has an x and y-axis. The difference is that both axes contain numerical values representative of the data.

A line graph can have one line or several. In the case of a chart with several lines, each one represents a category. Every category has a color, and the description is detailed in the legend. For an effective line graph, use not more than four or five lines and make sure the colors are different enough to be differentiated visually.

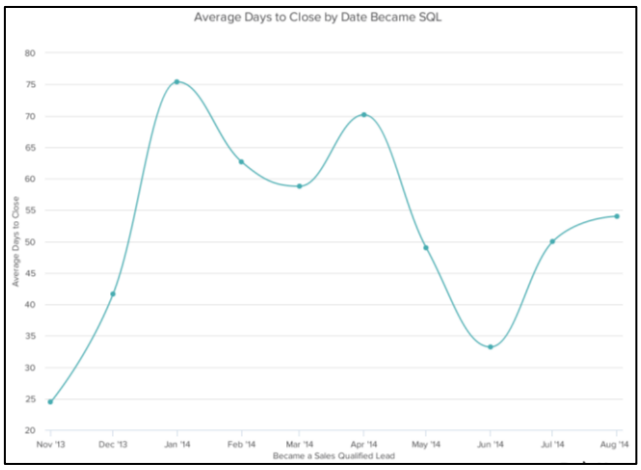
Design Best Practices for Line Graphs:

- Use solid lines only.
- Don't plot more than four lines to avoid visual distractions.
- Use the right height so the lines take up roughly 2/3 of the y-axis' height.

Use Case

Line graphs help users track changes over short and long periods of time. Because of this, these types of graphs are good for seeing small changes.

Line graphs can help you compare changes for more than one group over the same period. They're also helpful for measuring how different groups relate to each other.



A business might use this type of graph to compare sales rates for different products or services over time.

These charts are also helpful for measuring service channel performance. For example, a line graph that tracks how many chats or emails your team responds to per month.

Dual Axis Chart

A dual-axis chart allows you to plot data using two y-axes and a shared x-axis. It has three data sets. One is a continuous set of data and the other is better suited to grouping by category. Use this chart to visualize a correlation or the lack thereof between these three data sets.

Best Practices for Dual Axis Chart:

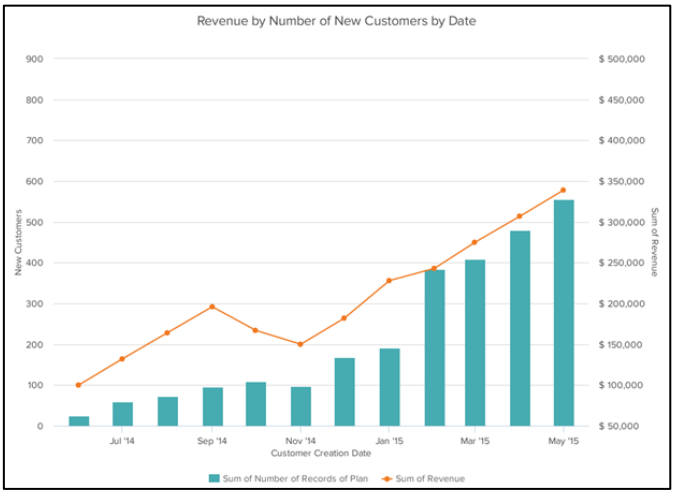
- Use the y-axis on the left side for the primary variable because brains are naturally inclined to look left first.
- Use different graphing styles to illustrate the two data sets, as illustrated above.
- Choose contrasting colors for the two data sets.

Use Case

A dual-axis chart makes it easy to see relationships between different data sets. They can also help with comparing trends.

For example, the chart above shows how many new customers this company brings in each month. It also shows how much revenue those customers are bringing the company.

This makes it simple to see the connection between the number of customers and increased revenue.



You can use dual-axis charts to compare:

- Price and volume of your products
- Revenue and units sold
- Sales and profit margin
- Individual sales performance

Area Chart

An area chart is basically a line chart, but the space between the x-axis and the line is filled with a color or pattern.

It is useful for showing part-to-whole relations, like showing individual sales reps' contributions to total sales for a year. It helps you analyze both overall and individual trend information.

Best Practices for Area Chart:

- Use transparent colors so information isn't obscured in the background.
- Don't display more than four categories to avoid clutter.
- Organize highly variable data at the top of the chart to make it easy to read.

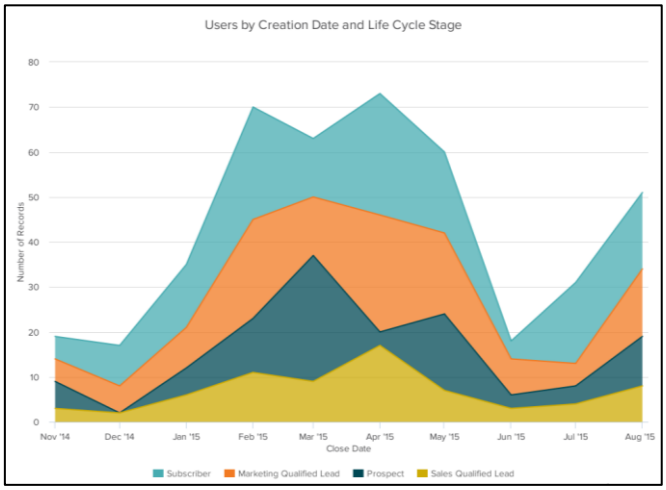
Use Case

Area charts help show changes over time. They work best for big differences between data sets and help visualize big trends.

For example, the chart below shows users by creation date and life cycle stage.

A line chart could show that there are more subscribers than marketing qualified leads. But this area chart emphasizes how much bigger the number of subscribers is than any other group.

These types of charts and graphs make the size of a group and how groups relate to each other more visually important than data changes over time.



Area graphs can help your business to:

- Visualize which product categories or products within a category are most popular.
- Show key performance indicator (KPI) goals vs. outcomes.
- Spot and analyze industry trends.

Stacked Bar Chart

A stacked chart is a form of bar chart that shows the composition and comparison of a few variables, either relative or absolute, over time. Also called a stacked bar or column chart, they look like a series of columns or bars that are stacked on top of each other.

A Stacked Bar Graph is used to show how a larger category is divided into smaller categories and what the relationship of each part has on the total amount.

Best Practices for Stacked Bar Chart:

- Best used to illustrate part-to-whole relationships.
- Use contrasting colors for greater clarity.
- Make the chart scale large enough to view group sizes in relation to one another.

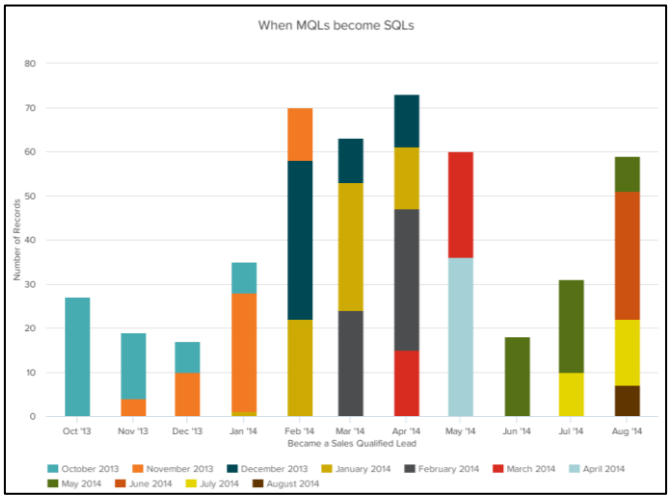
Use Case

These graphs are helpful when a group starts in one column and moves to another overtime.

For example, the difference between a marketing qualified lead (MQL) and a sales qualified lead (SQL) is sometimes hard to see. The chart above helps stakeholders see these two lead types from a single point of view– when a lead changes from MQL to SQL.

Stacked bar charts are excellent for marketing. They make it simple to add a lot of data on a single chart or to make a point with limited space.

These types of graphs can show multiple takeaways, so they're also super for quarterly meetings when you have a lot to say, but not always a lot of time to say it.



Stacked bar charts are also a smart option for planning or strategy meetings. This is because these charts can show a lot of information at once, but they also make it easy to focus on one stack at a time or move data as needed.

You can also use these charts to:

- Show the frequency of survey responses.
- Identify outliers in historical data.
- Compare a part of a strategy to its performance as a whole.

Pie Chart

A pie chart represents one static number, divided into categories that constitute its individual portions. When you use one, you'll represent numerical amounts in percentages. When you sum up all of the separate portions, they should add up to 100%.

The data in a pie chart represent parts of a whole. The entirety of the circle is the whole, and each wedge is a relevant section.

The best type of data for a pie chart has no more than five or six parts. Any more than this makes the wedges too thin at the center. If more than three values are similar to each

other, it will be difficult to discern the difference. The best pie charts use contrasting colors that fit well together, making each wedge visually different from the one next to it.

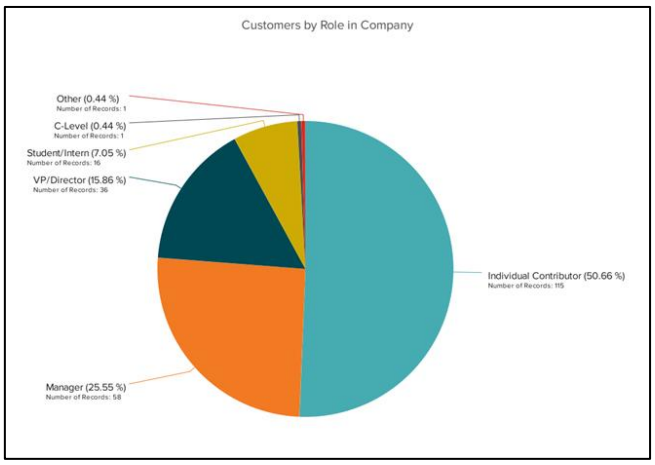
Best Practices for Pie Chart:

- Don't illustrate too many categories to ensure differentiation between slices.
- Ensure that the slice values add up to 100%.
- Order slices according to their size.

Use Case

The image below shows another example of customers by role in the company.

The bar graph example shows you that there are more individual contributors than any other role. But this pie chart makes it clear that they make up over 50% of customer roles.



Pie charts make it easy to see a section in relation to the whole, so they are good for showing:

- Customer personas in relation to all customers
- Revenue from your most popular products or product types in relation to all product sales
- Percent of total profit from different store locations

Scatter Plot Chart

A scatter plot or scattergram chart will show the relationship between two different variables or reveals distribution trends. Use this chart when there are many different data points, and you want to highlight similarities in the data set. This is useful when looking for outliers or for understanding the distribution of your data.

Best Practices for Scatter Plots:

- Include more variables, like different sizes, to incorporate more data.
- Start the y-axis at 0 to represent data accurately.
- If you use trend lines, only use a maximum of two to make your plot easy to understand.

Use Case

Scatter plots are helpful in situations where you have too much data to quickly see a pattern. They are best when you use them to show relationships between two large data sets.

In the example below, this chart shows how customer happiness relates to the time it takes for them to get a response.

Great use cases for this type of graph make it easy to see the comparison of two data sets. This might include:

- Employment and manufacturing output
- Retail sales and inflation
- Visitor numbers and outdoor temperature
- Sales growth and tax laws

Try to choose two data sets that already have a positive or negative relationship. This type of graph can also make it easier to see data that falls outside of normal patterns.

Waterfall Chart

A waterfall chart is used to show how an initial value changes with intermediate values, either positive or negative, and results in a final value.

Use this chart to reveal the composition of a number. An example of this would be to showcase how overall company revenue is influenced by different departments and leads to a specific profit number.

These types of charts and graphs make it easier to understand how internal and external factors impact a product or campaign as a whole.

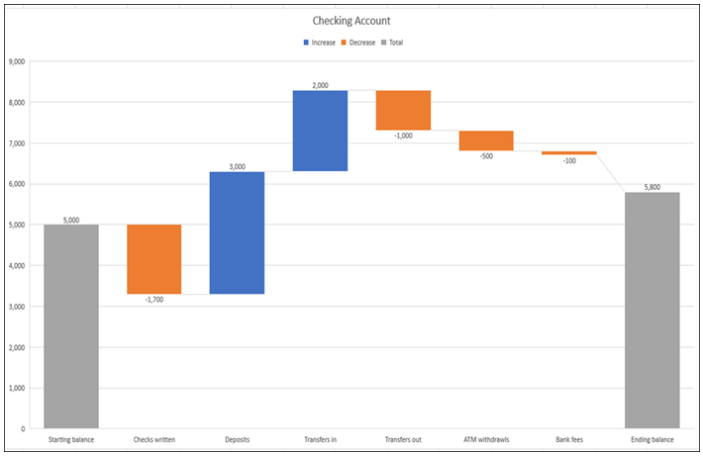
Best Practices for Waterfall Charts:

- Use contrasting colors to highlight differences in data sets.
- Choose warm colors to indicate increases and cool colors to indicate decreases.

Use Case

In the example below the chart moves from the starting balance on the far left to the ending balance on the far right. Factors in the center include deposits, transfers in and out, and bank fees.

A waterfall chart offers a quick visual that makes complex processes and outcomes easier to see and troubleshoot. For example, SaaS companies often measure customer churn.



This format can help visualize changes in new, current, and free trial users, or changes by user segment.

You may also want to try a waterfall chart to show:

- Changes in revenue or profit over time.
- Inventory audits.
- Employee staffing reviews.

Funnel Chart

A funnel chart is a specialized chart type that demonstrates the flow of users through a business or sales process. The chart takes its name from its shape, which starts from a broad head and ends in a narrow neck. The number of users at each stage of the process are indicated from the funnel's width as it narrows.

A funnel chart shows a series of steps and the completion rate for each step. Use this type of chart to track the sales process or the conversion rate across a series of pages or steps.

Best Practices for Funnel Charts:

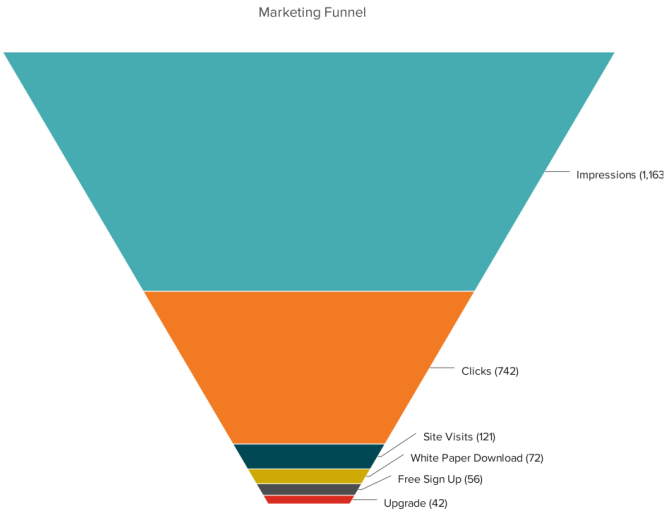
- Scale the size of each section to accurately reflect the size of the data set.
- Use contrasting colors or one color in gradated hues, from darkest to lightest as the size of the funnel decreases

Use Case

The most common use case for a funnel chart is the marketing or sales funnel. But there are many other ways to use this versatile chart.

If you have at least four stages of sequential data, this chart can help you easily see what inputs or outputs impact the final results.

For example, a funnel chart can help you see how to improve your buyer journey or shopping cart workflow. This is because it can help pinpoint major drop-off points.



Other stellar options for these types of charts include:

- Deal pipelines.
- Conversion and retention analysis.
- Bottlenecks in manufacturing and other multi-step processes.
- Marketing campaign performance.
- Website conversion tracking.

Heat Map

A heat map shows the relationship between two items and provides rating information, such as high to low or poor to excellent. This chart displays the rating information using varying colors or saturation. A heat map is a two-dimensional representation of data in which values are represented by colors. A simple heat map provides an immediate visual summary of information. More elaborate heat maps allow the viewer to understand complex data sets.

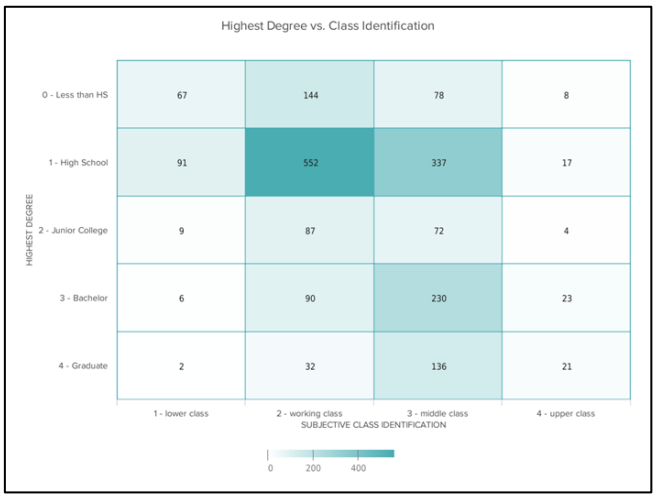
There can be many ways to display heat maps, but they all share one thing in common -- they use color to communicate relationships between data values that would be much harder to understand if presented numerically in a spreadsheet.

Best Practices for Heat Map:

- Use a basic and clear map outline to avoid distracting from the data.
- Use a single color in varying shades to show changes in data.
- Avoid using multiple patterns.

Use Case

In the example below, the darker the shade of green shows where the majority of people agree.



With enough data, heat maps can make a viewpoint that might seem subjective more concrete. This makes it easier for a business to act on customer sentiment.

There are many uses for these types of charts and graphs. In fact, many tech companies use heat map tools to gauge user experience for apps, online tools, and website design.

Another common use for heat map graphs is location assessment. If you're trying to find the right location for your new store, these maps can give you an idea of what the area is like in ways that a visit can't communicate.

Heat maps can also help with spotting patterns, so they're good for analyzing trends that change quickly, like ad conversions. They can also help with:

- Competitor research
- Customer sentiment
- Sales outreach

Dashboard

A dashboard is a visual display which contains multiple indicators on it and displayed at a time. these indicators can be called as a report or graph or maybe a filter output as well which can be showcased all together in a single page and that's what you call it as a Dashboard.

Now, dashboard is collection of multiple matrices where matrices can be referred to as some graph presenting progress reports or current statistics about any project or maybe health of a sprint and definitely be presented.

It can have multiple graphs being displayed together. Generally, we do see one graph at a time but if it comes to the presentations on a particular slide, you can add multiple graphs and that page is also called as a dashboard. In fact, many tools which are generally used for project management has a dashboard option enabled in it by default where, dashboard allows you to add any number of indicators on a particular page. When it comes to creating a dashboard, you need to be curious that what type of report should be added because not everything can be just displayed.

As you have a display, to display multiple graphs so you need to shortlist the right set of indicators to be displayed on a dashboard page which can be displayed to many people in the organization and can give you a current prospect and progress update from the monitoring point of view of a particular project. In fact, it tells you the ongoing progress as well as the health of the project.

Adding a quick example to tell you that a dashboard can be even seen in a car and why do we call it as a dashboard because it has many numbers of indicators at one place. For example, your speedometer, odometer, field indicator, the heat of the engine and not many other things. So, that can definitely tell you that what's the current status of your car, what inputs are on, what inputs are off, are you going short on fuel and many other things. This is why we call the front panel of the car when you sit inside the car is a dashboard because it has multiple indicators at one place.

Types of Dashboards

There are 4 general subtypes of dashboards:

- Strategic Dashboard
- Operational Dashboard
- Analytical Dashboard
- Tactical Dashboard

Strategic Dashboard

A strategic dashboard is a reporting tool for monitoring the long-term company strategy with the help of critical success factors. They're usually complex in their creation, provide an enterprise-wide impact to a business, and are mainly used by senior-level management.

Strategic dashboards are commonly used in a wide range of business types while aligning a company's strategic goals. They track performance metrics against enterprise-wide strategic goals. As a result, these dashboards tend to summarize performance over set time frames: past month, quarter, or year.

Operational Dashboard

An operational dashboard is one of the types of dashboards used for monitoring and managing operations that have a shorter time horizon. Since they focus on tracking operational processes, they're usually administrated by junior levels of management.

These kinds of dashboards are arguably the most common ones. They are mostly used for monitoring and analyzing a company's activities in a given business area. These dashboards are usually focused on alerting about business exceptions and are based on real-time data.

Analytical Dashboard

An analytical dashboard is a type of dashboard that contains a vast amount of data created and used by analysts to provide support to executives. They supply a business with a comprehensive overview of data, with middle management being a crucial part of its usage.

An analytical dashboard lies within its impact on historical data usage, where analysts can identify trends, compare them with multiple variables and create predictions, and targets, which can be implemented in the business intelligence strategy of a company.

Tactical Dashboard

A tactical dashboard is utilized in the analysis and monitoring of processes conducted by mid-level management, emphasizing the analysis. Then an organization effectively tracks the performance of a company's goal and delivers analytic recommendations for future strategies.

Tactical dashboards are often the most analytical dashboards. They are great for monitoring the processes that support the organization's strategic initiatives. Tactical dashboards help

guide users through the decision process. They capitalize on the interactive nature of dashboards by providing users the ability to explore the data.

Dashboard Creation and Story Telling: Best practices

As we mentioned, data visualization is all about making data, numbers, and statistics easy to understand by formatting them into a visually digestible graphic.

But in order to ensure that your data actually is as easy-to-understand as you aim to make it, there are certain guidelines that you need to follow. Some of them are as follows:

Identify your target audience.

The first step in effective data visualization and communication is identifying who you're visualizing the data for, i.e., the target audience. This can help you tailor the techniques and methods you use to a specific audience.

Some examples of your potential target audience can be marketers, business executives, social media managers, entrepreneurs, educators, students, and non-designers.

You can create a visual to help identify and represent your target audience so everyone is on the same page. The way you visualize data needs to be clear and understandable to the target audience, enabling them to process information efficiently.

Select the Right Chart.

Once your data is clean, you can select the specific type of graph or chart to visualize your data most effectively and efficiently convey the essential information in the dataset.

There are many different types of information that can be represented by various charts. You can use bar chart, line chart, pie chart, and others as per data visualization requirements.

Label Your Chart Effectively

Graphs and charts can allow us to identify patterns in our data quickly. However, for specific values that may be of significance, labels are the best way to visually represent this data.

Whether describing an experimental setup, introducing a new model, or presenting new results, you cannot explain everything using just the figure itself.

This is why a caption should always accompany your figure. The caption explains how to read the figure and provides additional precision for what cannot be graphically represented. If the numerical values are significant, they must be given somewhere in the article or written very simply on the figure. In the same way, where there is a point of significance in the model (critical domain, unique point, etc.), make sure it’s visually distinct but don’t fail to point it out again in the caption.

Emphasize the Important Points.

In data visualization, the audience must follow the story you're trying to convey by looking at your chart. This is why it's so imperative to direct the reader's attention through specific visual cues, such as reference lines or highlighted trends.

Humans can absorb more significant amounts of information visually. Our eyes are drawn to symbols that send us valuable details at a glimpse.

We usually search for patterns, because if patterns are chaotic or make no sense, it's too complicated to grasp what the image conveys.

To draw on these human behaviors, make sure that the order or style in which you show the data makes sense to audiences. The data you use can be numeric, alphabetical or sequential.

Choose the Best suitable Dashboard

To make your data visualization most effective, you need to recognize what type of dashboard you're offering to the audience. You can use Strategic, analytical, or operational which is suitable according to the audience.

Make Use of Colour.

Colors can be a useful tool in data visualization, as you can effectively communicate important information about your data through the use of different color combinations.

For example, categorical data is best represented by a distinct color for each category, while sequential data can be ordered through different shades of a single color.