BEST PRACTICES IN DATA VISUALIZATION

Best Practices in Data Visualization

Data visualization is the representation of data in graphical form. It aims to simplify data for users to understand easily. From a technical perspective, it is important to choose the right tool, type of chart, style, and layout, which are all essential elements of modern data visualization. But maybe even more important is to understand the purpose of a visual or dashboard (a combination of related visual elements). What requirements business users have, what questions they expect to find answers for, what the level of detail should be and so on – all such questions need to be discussed to create meaningful and insightful visuals, rather than a good-looking but not helpful chart.

Main Principles

Data visualization is a form of communication that portrays raw data in a neat, appealing graphical form that makes it easier to extract meaningful information. To create a good visual, the main principles need to be followed. Some of them are:

1 **Accuracy**
A visual must represent correct and full information. Misleading visuals are not acceptable.

2 **Convenience**
A visual must be user-friendly in terms of style, layout, readability, etc.

3 **Scalability**
A visual must have the ability to accommodate increasing volumes of data and data sources so that reports can be kept up to date.

4 **Simplicity**
A visual must include only those elements that are required for the report and that bring value to it.

5 **Standardization**
Often, visuals are not a single item, but a set of visuals or dashboards, and they can be a part of boards, workbooks, stories, etc. All these should have a standard form, if possible, in terms of design, data structures, layouts, etc.

6 **Interactivity**
In most cases, interactivity allows the user to achieve more insights from the visual, and look at it from different perspectives.

Most of these principles must not be compromised, but some can be omitted for specific use cases. For example, if a visual will be used only as a static view (maybe as part of a presentation), the “interactivity” principle can be omitted.

Dashboards

Even if we can refer to visualization as a single chart, in most real cases, visual reports are not just a chart or few separate charts on different pages, but, rather, a logical combination of visuals.

A dashboard can consist of various data sets that help users gather different types of data in one place. The user interface is designed in a way that can help combine information for:

- Analytics
- Presentation
- Process and Operations

The art of dashboarding is a separate topic that needs to be carefully considered before providing the final visuals. Sometimes we can have great visuals as separate items, but in the case of wrong "dashboarding," the readability, insightfulness, and general feel of a dashboard can be spoiled.

Dashboards need to be constructed properly and not overloaded with data, using standardized visuals to achieve the perception of a dashboard as a whole, rather than as a collection of different items.

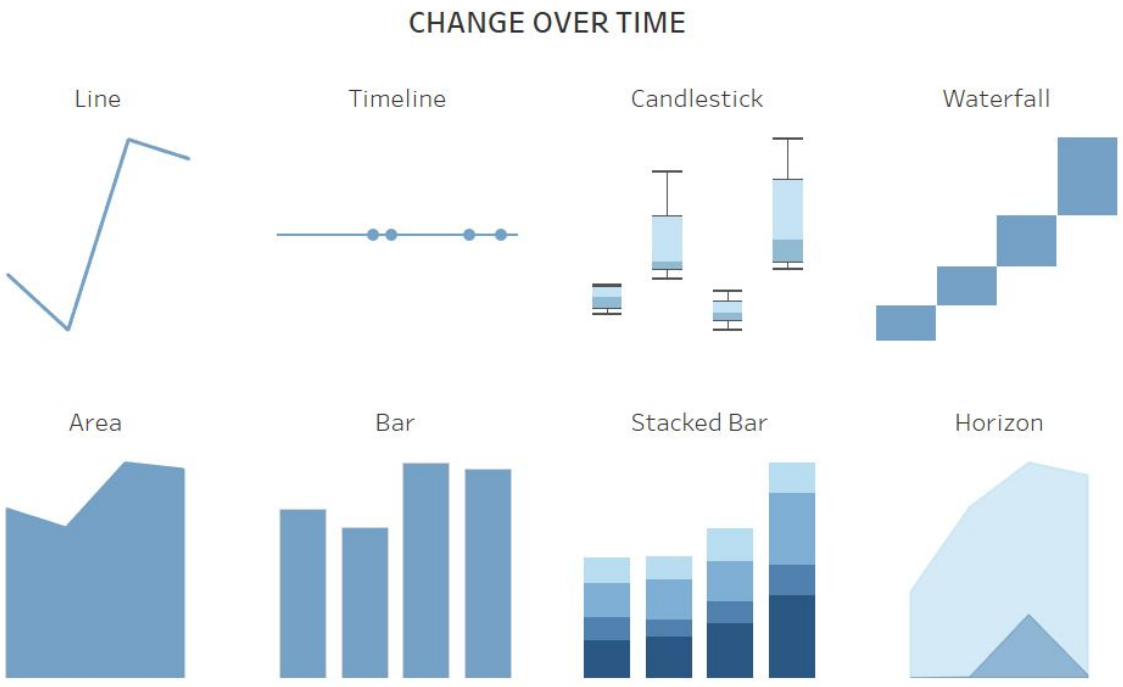
Sometimes, dashboards are grouped into a set that combines reports from homogeneous areas, like marketing and finance reports, or other logical groupings common to the business (loan-related reports, sales channel reports, etc.).

Types of Visuals

There are many different graphical ways in which we can depict data. The kind of message that we want to convey to the user influences how we make this representation. The most common charts by data representation type are:

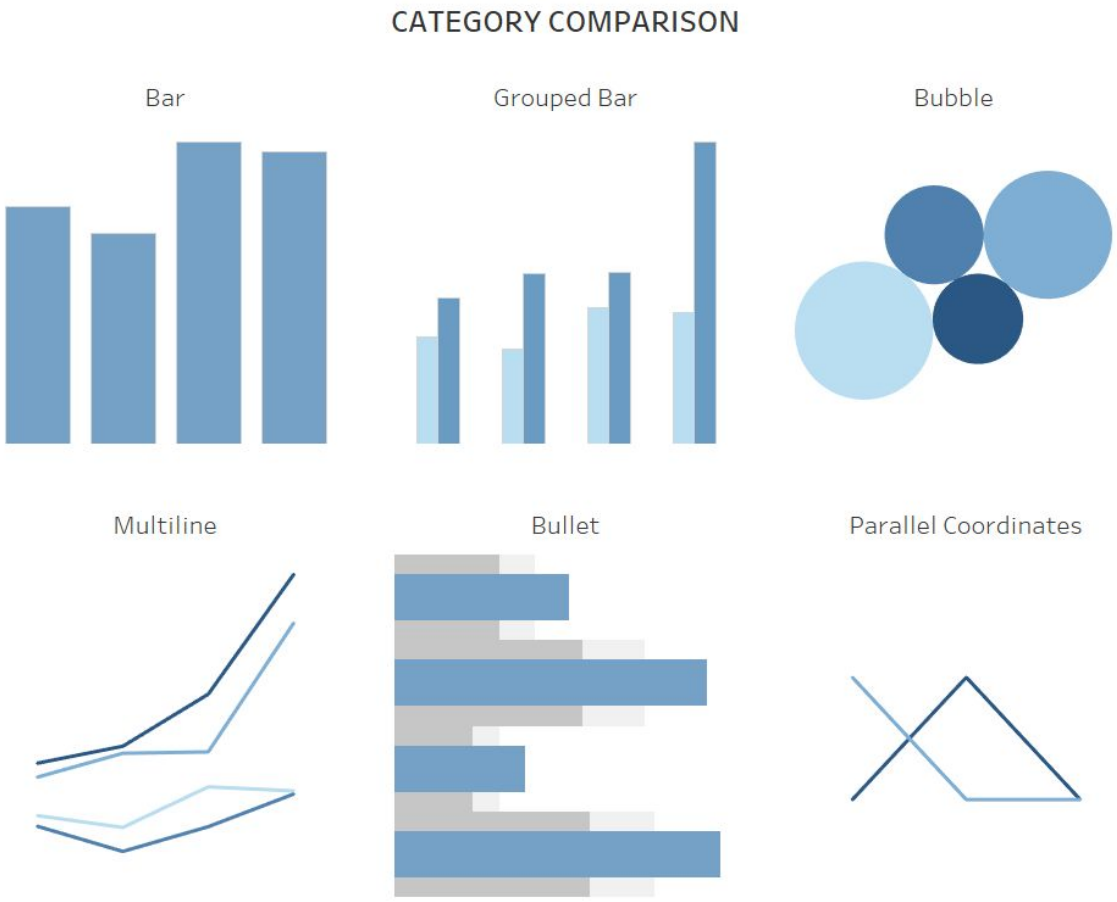
1. Change over Time

This type of chart shows how a particular subject has changed over time.



2. Category Comparison

This chart reflects data from different categories. Not only does it show the data of these categories, but it compares them with one another.

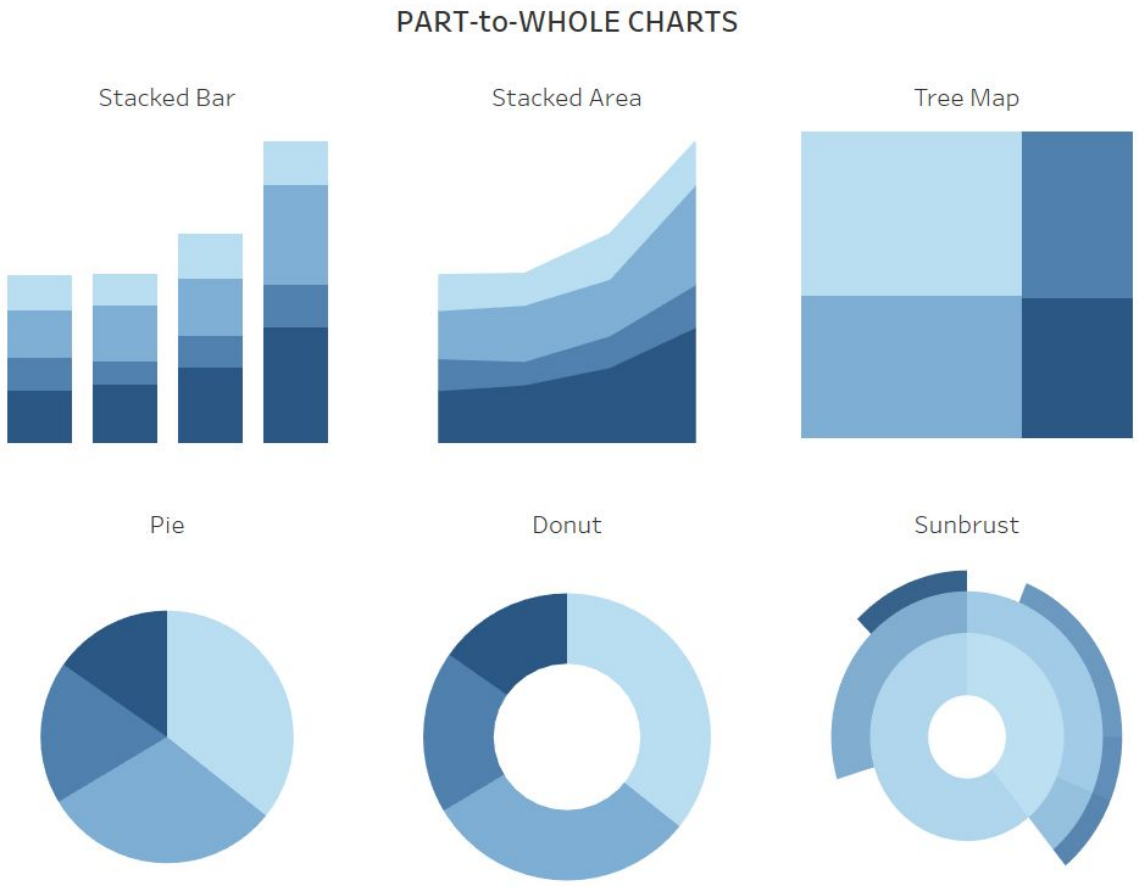


3. Ranking

This chart shows various subjects in order of rank.

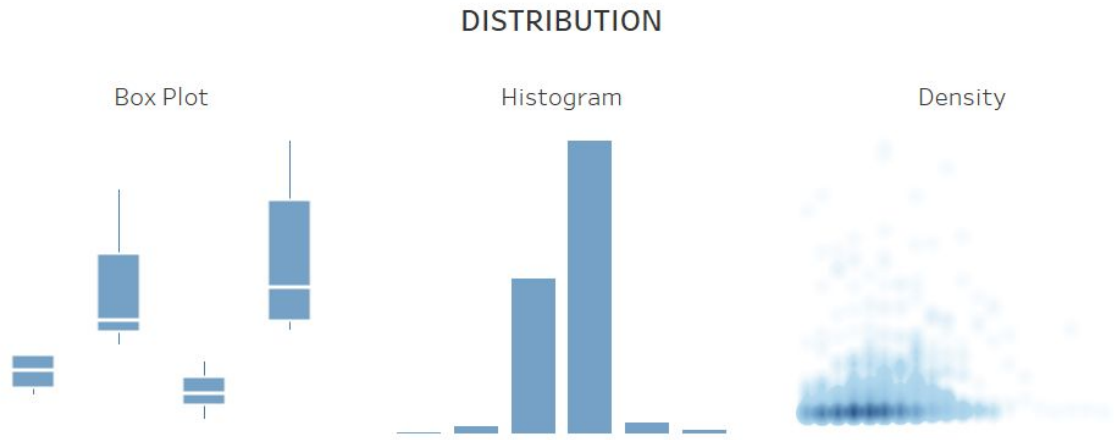
4. Part to Whole

In this chart, all of your partial subjects are depicted to make up one big subject. For example, if you want to dissect the entire population across various ethnicities, this chart will help you represent this.



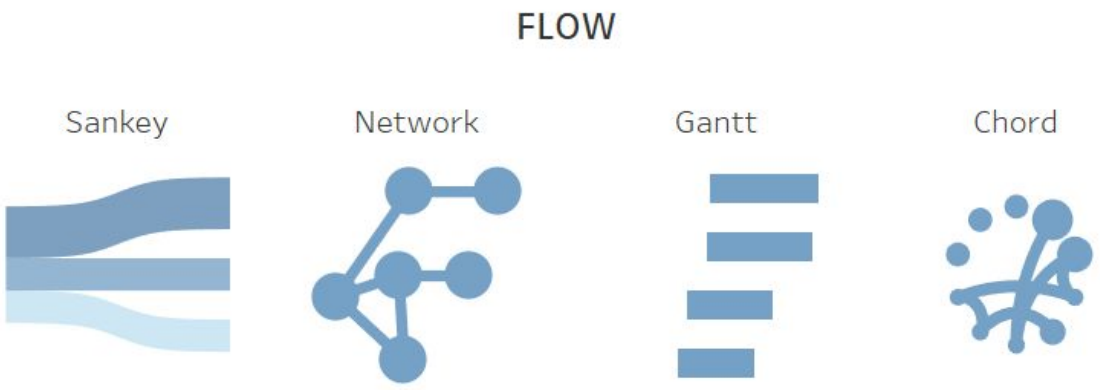
5. Distribution

The distribution chart reflects the occurrence of a value within a particular data set.



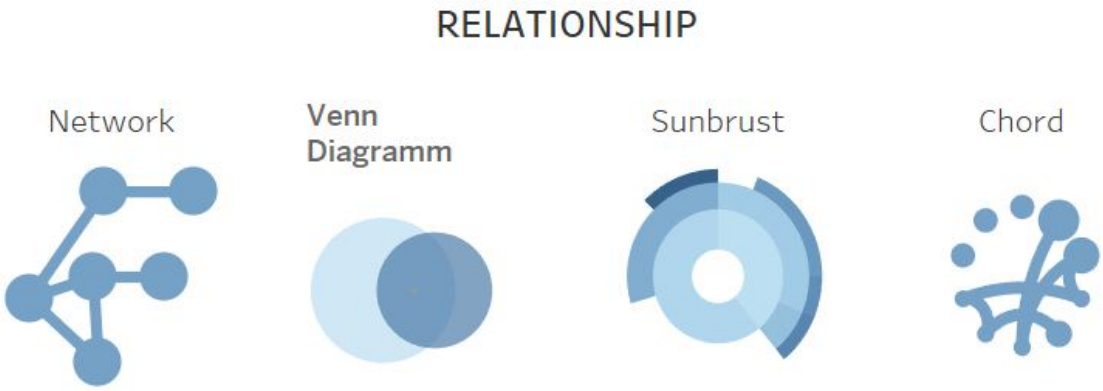
6. Flow

A flow chart shows the flow of data from one state to another.



7. Relationship

A relationship chart shows relationships between different subjects that make up a network.



Let's review the practical implementation of most common chart types.

- Tables

  Some may not include tables in "visuals," although tables are quite a common element of dashboards and complex reports. They are good for the representation of detailed data, and, with the use of additional features like colors, conditional formatting, and sizing, they can be a worthwhile element of the dashboard.

- Bar Charts

  These charts present categorical data using rectangular bars with heights or lengths proportional to the values that they represent. Horizontal bar charts can be called “column” charts in some sources or tools. Histograms are a subtype of a bar chart which show the distribution of the data inside a category, rather than the values of different categories.

- Area and Line Charts

  Most commonly, these charts show the dynamic of data change over time by connecting each data point with lines, which helps to easily identify a trend. Area charts fill the area below the line with color. This can be useful in the case of stacked area charts.

- Pie Charts

  These are for the depiction of categorical data in a circular visual, divided into slices to illustrate the proportion of each category. This type of chart should be used cautiously for data with a lot of categories, as it may be hard to distinguish the size of each category.

  Donut charts are a subtype of pie charts, with a blank central area which is sometimes filled with data (like titles or information related to the whole donut chart).

- Map Charts

  These are the best way to represent geo data visually. The most common types of map visuals are:

  - filled map: area (country, region, etc.) filled with conditional color
  - maps with marks: in most cases, circular elements or pies placed on map elements and conditional colors/sizes help to read the data more easily.

- TreeMap

  These display hierarchical (tree-structured) data as a set of nested rectangles. Each branch of the tree is given a rectangle, which is then tiled with smaller rectangles representing sub-branches. This hierarchical view of your data makes it easy to spot patterns.

- Other Charts

  There are many other chart types that depend on the tool you use for visualization. They include:

  - scatterplots
  - Gantt bars
  - timeline charts
  - waterfall charts
  - Sankey diagrams
  - sunburst diagrams

Style

Data visualization uses shapes and styles to make data comprehensible for the user. Various attributes can be used in order to develop a context for the data.

- Graphical Elements
  This includes, but is not limited to, shapes and sizes of the subjects in the chart. Images can be included in this category.
- Typography
  This includes the type of word representation, such as titles and labels.
- Iconography
  This improves a chart's overall usability by differentiating categories, defining states, and providing UI controls.
- Axes and Labels
  Proper axes and labels give good representation of the data that we want to show. Sometimes it is really important to show the axes; sometimes we can hide them, so, in practice, it depends on a use case.
- Annotations and Legends
  This describes a chart's information, with annotations pointing to a specific part to provide more emphasis.

Behavior

There are various interaction tools that help to refine the data in a graphical data set. This helps the user put emphasis on a particular part of the data set to extract a particular data value. These interaction patterns include but are not limited to the following:

- Progressive Disclosure provides the ability to disclose a part of data on demand from the user.
- Manipulation enables the user to operate UI elements to limit the number of actions needed on the screen.
- Altering Perspective enables the user to change the perspective of the data set by changing its design, through altering data types and value.

As described, data visualization is quite complex and includes many elements that need to be considered, as well as a good understanding of domain and business needs. Nonetheless, we should strive to create simple and standardized but informative visuals that are more flexible, easy to maintain, and perform well (time to load and render).