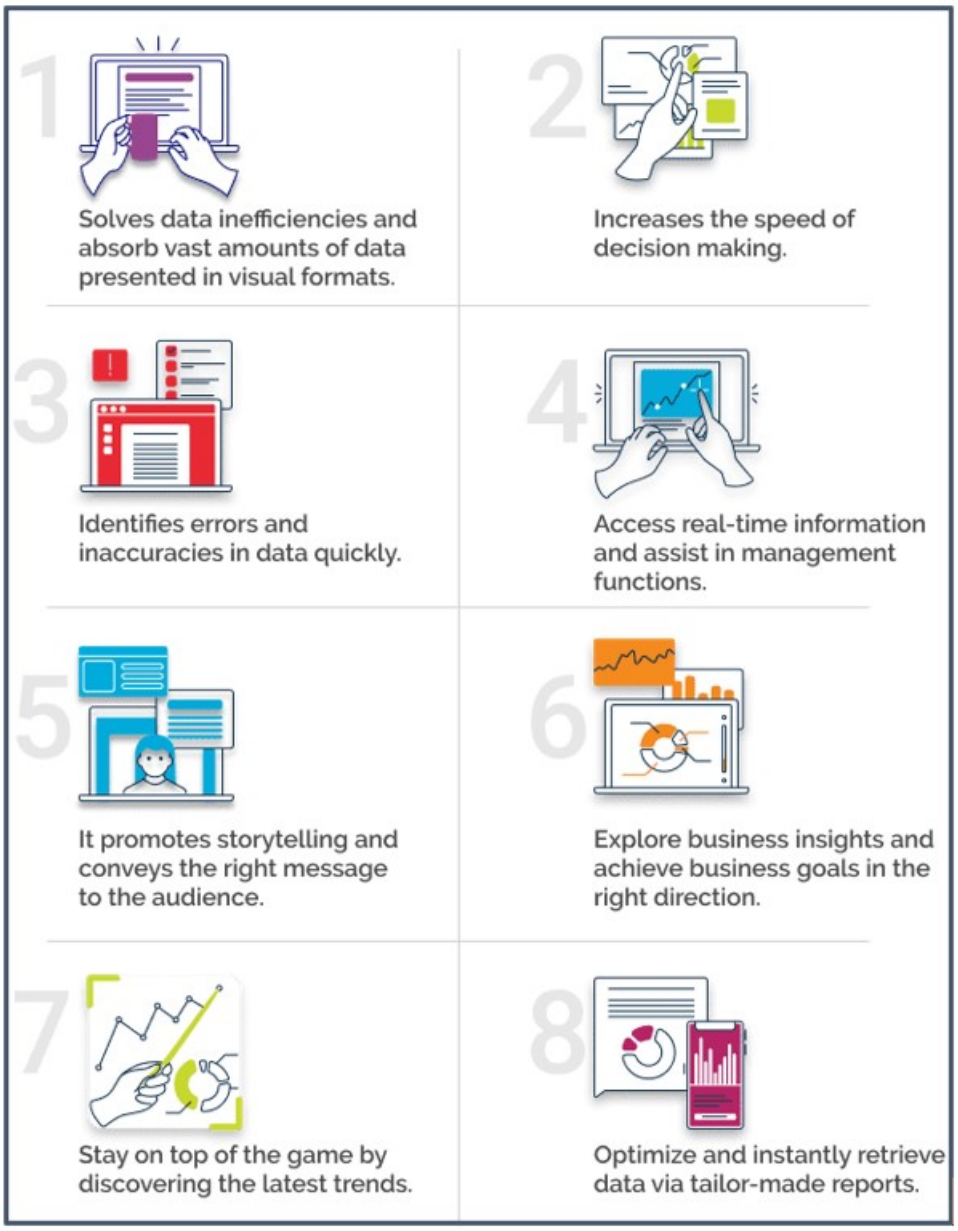
Data Visualisation Techniques

2.1 Data visualisation tools and techniques

Introduction

Last week, you delved into the various concepts of data visualisation. This week, you will explore the various data visualisation tools and techniques that can help businesses make sense of the humongous amount of data and statistics that they have at their disposal.

Human beings can grasp visual information better than data and numbers. These visualisations help decision-makers make better connections between the existing data and future trends. Business managers receive timely information without manually mining the data and then analysing it; this aids better decision-making. Further, it also eliminates the chance of errors and assumptions in the analysis.



Source: The benefit of data visualisation tools (Splashbi, 2020)

2.1.1 Data and design for the eyes

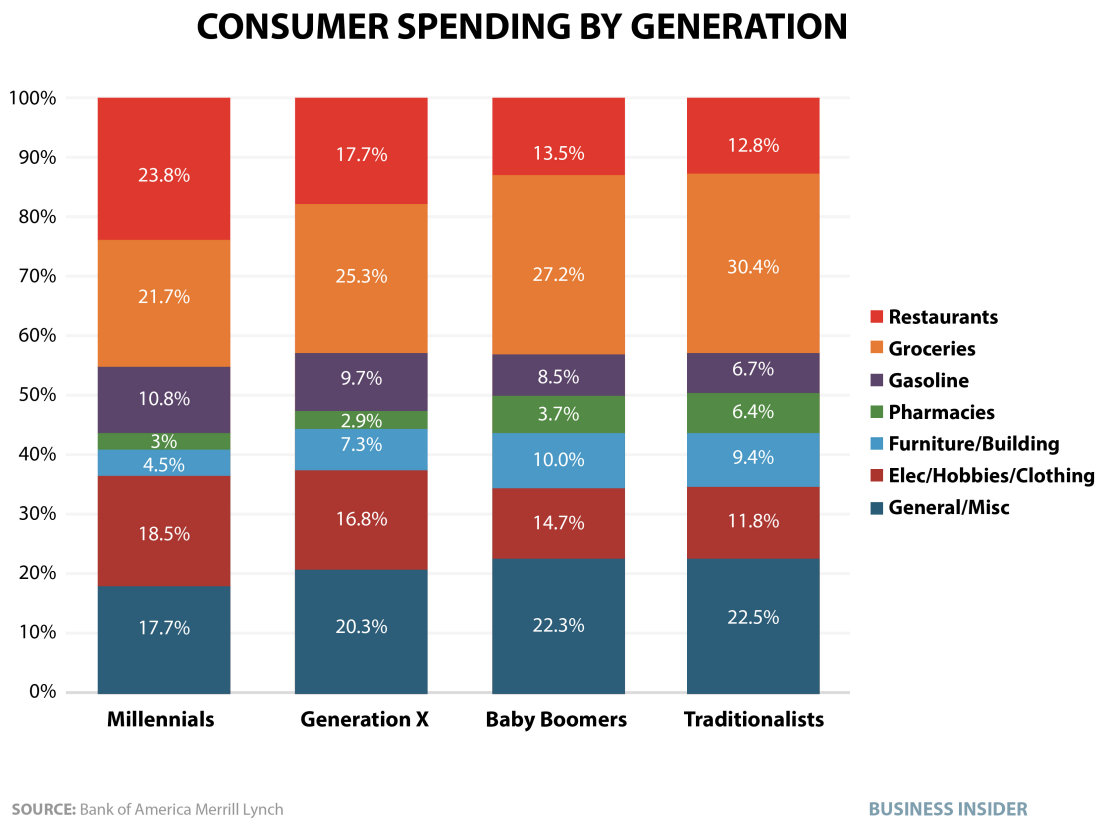
Activity: Data and design for the eyes

Time: 15 minutes.

Purpose: To analyse a data visualisation.

Task: Reflect on the visualisation below and analyse it based on the given questions.

Feedback: To help further the conversation, respond to the posts of at least 2 of your peers. You can ask them questions for clarification on their approach or provide them with feedback. Your facilitator will moderate this discussion.



Source: Consumer spending by generation (Business Insider, n.d.)

Questions:

1. Where are my eyes drawn?
2. What works in this visual?
3. Can you identify any clutter or confusing elements?
4. Is the take-away action clear?
5. Are pre-attentive attributes used to draw the user's attention?

2.1.2 Selecting the right visualisation and factors influencing them

Explore: Choosing the right visualisation

For most of us, it is much easier to read a colourful and organised chart than to scrutinise data in a spreadsheet or report. Making data highly visual - meaning clean, organised, colourful, and easy to read, makes it easier for our brains to process it.

Visualisations can vary depending on the category of data you are working on and when it is cleaned and sorted. You must understand what you want to show to simplify your selection and tell a compelling story with your data.

The following are a few categories to consider when approaching data:

a. Distributions

- If you want to compare multiple distributions of numeric data, you may consider using either a histogram or a boxplots.
- Histograms should be used when there is a similar sample size and only a few different comparisons to be made. Otherwise, the graph may appear cluttered and difficult to read. On the other hand, a box plot will work when your dataset is symmetrical or skewed.

b. Relationships:

- Relationships describe data connections (correlations) across different tables and spreadsheets. Hence for visualising the correlated data, you can use either a scatter plot, a bubble chart, or a heatmap. Relationships can also be described with Marimekko charts.

c. Compositions

- Data compositions, or part-to-whole relationships, can be communicated in many ways. Test your data to decide if it is dynamic or static and then limit the selection. Dynamic data could be visualised using stacked columns or area charts, and static data is best organised using column charts, area charts, and pie charts.

d. Data Comparisons

- Comparisons indicate differences across values (compare multiple items) which can be explored using tables or bar charts. Line and bar charts can also be considered if data must be presented over time.

Factors that influence data visualisation choices

There are a few factors that influence data visualisation choices.

1. **Audience**:

It is very important to adjust the data as per the audience's understanding. For example, investment/finance mobile app users who browse through their portfolios can work better with uncomplicated visualisations. On the other hand, if data insights are intended for researchers

and decision-makers who regularly work with data, you can use more sophisticated charts and graphs.

2. **Content**:
   The kind of data you are dealing with will determine the strategies for the visualisation. To show the relationship between two elements, scatter plots are often used. In turn, bar charts are good for comparative analysis.
3. **Context**:
   You can use different data visualisation approaches and read data depending on the context. To accentuate a certain figure, you can use one colour's shades on the chart, highlight the highest value with the brightest one, and differentiate elements; you can use contrasting colours.
4. **Dynamics**:
   There are various types of data, and each type has a different rate of change. Depending on the rate of change, you may consider dynamic depiction or static data visualisation techniques in data mining.
5. **Purpose**:
   To make a complex analysis, visualisations are compiled into dynamic and manageable dashboards fortified with different tools for visual

Refer to the following readings to further your understanding.

Boyer, S. (2022, June 9). Data visualization: how to choose a chart type. *Medium*. https://uxdesign.cc/data-visualization-how-to-choose-a-chart-type-7886dc0fe590

You are suggested to watch the following video from 1:18 - 13:50 to understand how to pick the right type of chart for your data.

https://youtu.be/aUk4npRmjL8 *Source: (Venngage - Visualize your ideas, 2019)*

The following two videos will teach you the 14 most popular chart types and help you understand when to use and avoid them. The first video is recommended to be watched from 00:00 - 11:52.
https://youtu.be/C07k0euBpr8 *Source: (365 Data Science, 2018)*
https://youtu.be/qGaIB-bRn-A *Source: (365 Data Science, 2018)*

2.1.3 Evaluate a superstore dataset

Activity: Gain insights by looking at an interactive dashboard

Time: 20 minutes

Purpose: To train Tableau users on Tableau tactics, data visualisation strategy, and design.

Task: Follow the instructions below:

Feedback: To help further the conversation, respond to the posts of at least 2 of your peers. You can ask them questions for clarification on their approach or provide them with feedback. Your facilitator will moderate this discussion.

Instructions

Step 1) Refer to this visualisation.

Step 2) To use the visualisation, choose a date range and a region to highlight.

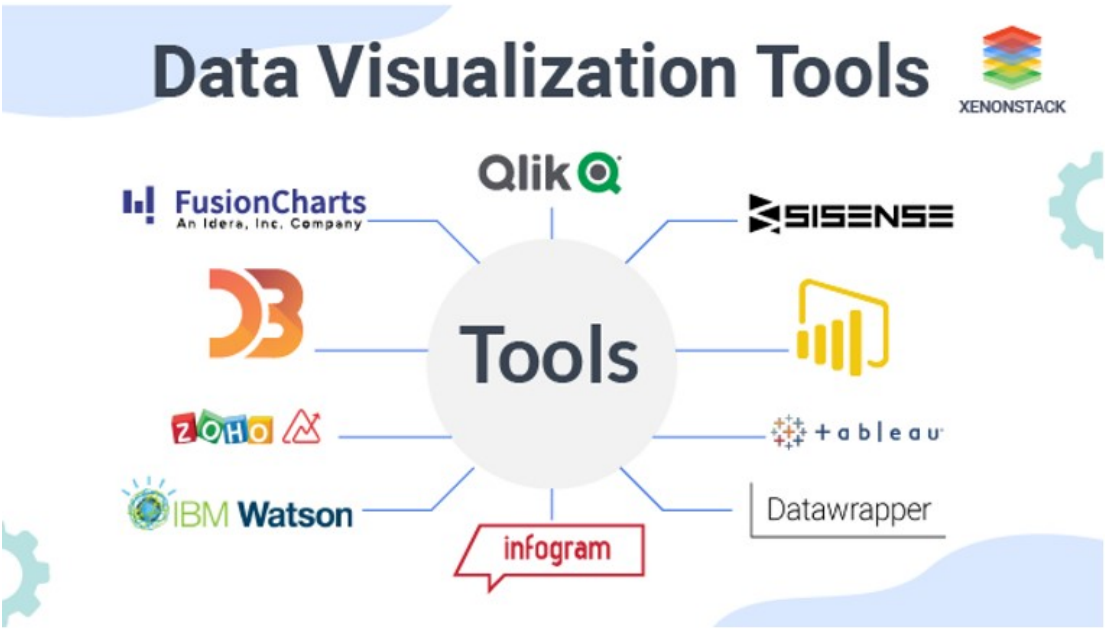
Step 3) You may click on a US state within a region to filter the descriptive dashboard once highlighted. You can also choose to look at the performance over equal periods or year over year and choose the trend lines' date granularity (i.e. year, quarter, month, week).

Step 4) You can use the prescriptive dashboard to gain insight into the performance shown on the descriptive dashboard. Try out the annotations tab to enter your insights. Share your findings in the discussion forum.

2.2 Data visualisation - Tools and techniques in depth

Introduction

Data visualisation is now being used in every field of knowledge. Data Scientists in various disciplines use computer techniques to model complex events and visualise phenomena that cannot be observed directly for gaining a qualitative understanding.



Activity: Creating a bar graph in tableau

Time: 30 minutes.

Purpose: To reproduce figures from a table visually.

Task: Follow the instructions present in the document below and share your visualisation in the discussion forum.

Feedback: To help further the conversation, respond to the posts of at least 2 of your peers. You can ask them questions for clarification on their approach or provide them with feedback. Your facilitator will moderate this discussion.

Document: Refer to Creating a bar graph in tableau.docx

2.2.1 Data visualisation techniques

Exploring data visualisation techniques

You can choose different data visualisation techniques and configure their features considering the factors listed above. The following are some of the most common types of data visualisation techniques.

1. **Charts**

   A chart is the easiest way to show the growth of one or several data sets. Bar and line charts show the relationship between elements over time, and pie charts show the

components or proportions between the elements of one whole.

2. **Plots**

Scatter and bubble plots are some of the most commonly used visualisations. Plots allow the allocation of two or more data sets over a 2D or even 3D space to show the relationship between these sets and the parameters on the plot. When it is big data, analysts often use more complex box plots to visualise the relationships between large volumes of data.

3. **Maps**

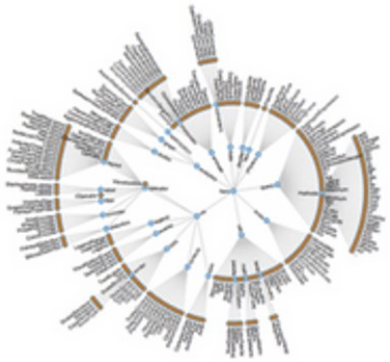
Maps are popular techniques used for data visualisation in several industries. The most popular map visualisations are heat maps, dot distribution maps, and cartograms.

4. Diagrams & Matrices

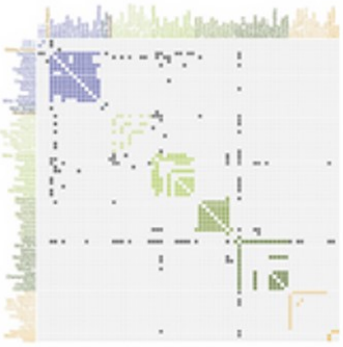
Diagrams are usually used to demonstrate complex data relationships and links and include various data types in one visual representation. They can be hierarchical, multidimensional, or tree-like.

Matrix is one of the advanced data visualisation techniques that help determine the correlation between multiple data sets. Correlation matrices enable quickly identifiable proof

of networks between variables by joining huge information and rapid response times.



Tree



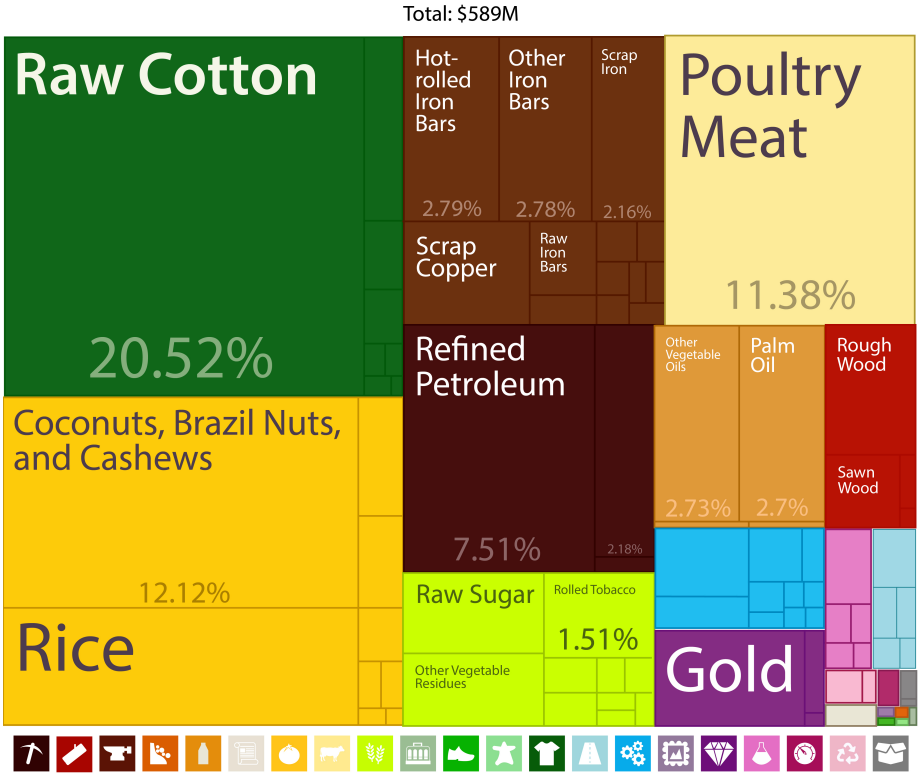
Matrix

5. **Word cloud and Network diagram for unstructured data**

A word cloud visual shows the frequency of a word inside a collection of text with its available size in the cloud. This technique is used on unstructured data to show high-or low-recurrence words.

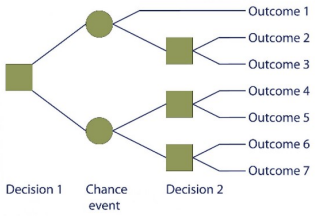
6. **Tree Maps**

This method indicates hierarchical data in a nested format.



Visualisation techniques for large data

When it comes to representing large data, you must take cognisance of a few critical aspects. You could create a Decision tree analysis [a graphic representation of various possible solutions available to solve a given problem, to determine the most effective sequences].



Ensure your information source is reliable and dates are relevant for the visualisation. You must also place the key data takeaways at the top of your data report to help the reader read it first and filter some of the data components to see how they affect the results.

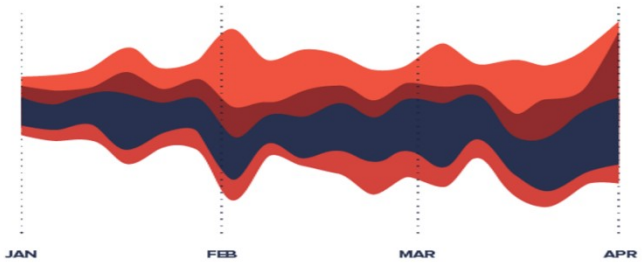
Although not necessary, you could share the process of collecting and storing data with your audience.

The following are two of the visualisation techniques for large data.

1. **Streamgraphs**

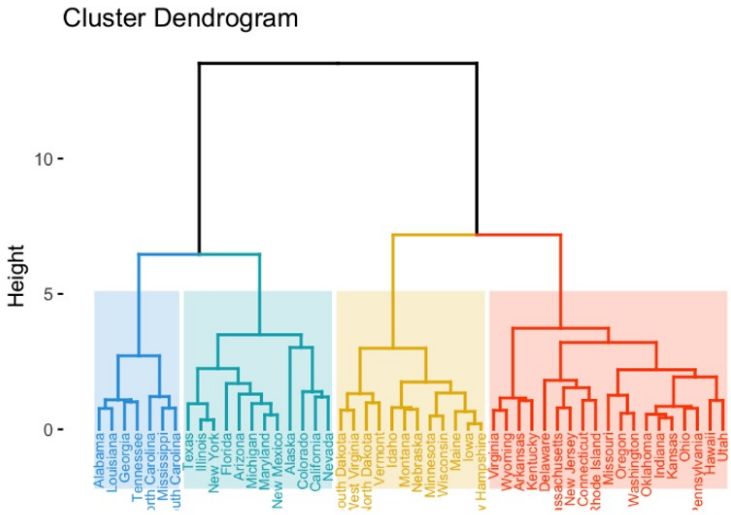
   A streamgraph is a variety of stacked area charts. Rather than plotting values against a customary y-axis, the streamgraph balances each 'stack' baseline to make it even around the

x-axis.



2. **Dendograms**

Dendrograms show the hierarchical connection between the objects. The major use is to figure out the best path to allocate objects to clusters.



You have explored various data visualisation techniques so far. Now, you will discover a few more data visualisation tools at our disposal.

1. **Tableau**

Tableau is one of the leaders in this field; it has a user-friendly interface and a well-equipped library of interactive visualisations and data representation techniques. Tableau stands out for its influential capabilities. The platform provides large integration options, including My SQL, Teradata, Hadoop and Amazon Web Services. This is a great tool for both data visualisations and specialised data analytics. The system can easily handle any data type, including streaming performance data, and allows combining visualisations into functional dashboards. Tableau invests in AI and augmented analytics and equips customers with advanced analytics and forecasting tools. (Source: Tableau.com)

Some of its best features are:

- customisable dashboards embedded with applications like Salesforce, SharePoint and Jive etc
- real-time interactive dashboards for filter on-demand and click to dive deeper into data
- plenty of data connections with live and in-memory data
- secure collaboration

- mobile augmented.

2. Power BI

Power BI is a suite of business analytics tools to analyse data and share insights. Power BI dashboards provide a 360-degree view for business users.

With Power BI, Microsoft is offering organisations the ability to have a single view of their most critical data and monitor the health of their businesses at every point in time.

Different Power BI versions like Desktop, Service-based (SaaS), and mobile Power BI apps are used for different platforms. It provides multiple software connectors and services for business intelligence.

Some of the Power BI features are listed below:

- build dashboards and reports and allow real-time updates
- offers secure and reliable connection to your data sources in the cloud or on-premises
- Power BI offers quick deployment, hybrid configuration, and a secure environment
- allows data exploration using natural language query
- offers features for dashboard visualisation regularly updated with the community.

(Source: Microsoft Power BI)

3. R Programming

The R programming language is the latest cutting-edge tool. It is an open-source programming language widely used as a statistical software and data analysis tool and is available across widely used platforms like Windows, Linux, and macOS. R generally comes with the Command-line interface.

R is a language and environment for statistical computing and graphics and provides a wide variety of statistical (linear and nonlinear modelling, classical statistical tests, time-series analysis, classification, clustering, etc.) and graphical techniques, and is highly extensible.

The S language is often the vehicle of choice for research in statistical methodology, and R provides an open-source route to participation in that activity.

One of R's strengths is the ease with which well-designed publication-quality plots can be produced, including mathematical symbols and formulae where needed.

The following are a few statistical features of R:

- The most common basic statistics terms are the mean, mode, and median. These are all known as 'Measures of Central Tendency.'
- R is rich with facilities for creating and developing interesting static graphics and contains functionality for many plot types, including graphic maps, mosaic plots, biplots, etc.
- Probability distributions play a vital role in statistics. Using R, we can easily handle various types of probability distribution such as Binomial Distribution, Normal Distribution, Chi-squared Distribution, and many more.
- It provides a large, coherent and integrated collection of tools for data analysis.

(Source: Geekforgeeks, 2021)

Features of data visualisation tools

Data visualisation software increases the efficiency and productivity of business functions by providing valuable data insights for better investment and operational decisions. For example, the marketing department of any business needs to analyse the demographics of their possible customers and their preferences before launching a new product or service. The analysis then helps to define the product's specifications, design, and most competitive price point.
To choose the best data visualisation tools, consider the following features:

1. Capacity to embed data

    Data visualisation software should have the capacity to embed/extract the visualised data into your existing networks, interfaces, tools, and systems for the operational teams to detect important insights and patterns for improved decision-making. This also helps teams judge the efficiency of the investment and take corrective action if required.

2. Integration with AI & ML

    Visualisation software is beginning to integrate artificial intelligence and machine learning tools to make it more feature-rich. This integration helps in easier analysis, integration, and delivery of insights to the ultimate users of this data.

3. Collaborative

    One of the most important features of data visualisation tools is that they should be collaborative. The teams using a data set can share it in real-time to collude with other teams and gain collective insights.

4. Performance, Analytics & Insights

    The data analysis prompts, and collaboration should be seamless and high performing. The tools must be in-built with analytics and insights to present well-visualised data. This will help the working teams to predict the market trends and advise the users on their future course of action.

5. Simplify the Data Streaming

    Visualisation tools should assimilate and analyse data from several data sources as businesses today also collect data from sources like social media platforms, IOT (internet of things), mobile applications, etc.

Data visualisation tools help businesses keep up with competition and sync their decisions with the consumers' needs and preferences. They can quickly analyse shifts in trends, patterns, and consumer preferences using predictive analytics that comes bundled with the software being used.

The tools and techniques help businesses customise the data they have at their disposal. They can clean, sort and update data which will better guide their decision-making.

Visual tools convert data into understandable and usable formats compared to static and stagnant data reports and make them interactive. These visualisations can be edited, formatted, and converted into the most usable formats.

The following videos are recommended to further your understanding.

- Watch this 11-minute 29-second video to explore eight visually explained plots. https://youtu.be/uZIMNUPA26E *Source: (Kimberly Fessel, 2021)*

- Watch this 7-minute 8-second video to understand why you should make use of each of the top five visualisation tools. https://youtu.be/g_mLR3hI1Kg *Source: (DataCatchup, 2021)*

2.2.2 Data visualisation techniques for dashboard

Activity: Video case discussion

Time: 25 minutes

Purpose: To understand the utility of the data visualisation techniques and their correct application.

Task: Watch the videos below and answer the given questions.

Feedback: To help further the conversation, respond to the posts of at least 2 of your peers. You can ask them questions for clarification on their approach or provide them with feedback. Your facilitator will moderate this discussion.

Watch the following videos and reflect on the arguments and ideas presented in them.

- Watch this 3-minute 41-second video on data visualisations for dashboards. https://www.youtube.com/watch?v=6arCpof4XIM&ab_channel=NNgroup *Source: (NNgroup, 2020)*
- Watch this 4-minute 17-second video on the art of visualising data. https://www.youtube.com/watch?v=DzMYt_kGUmM&ab_channel=nullQueries *Source: (nullQueries, 2021)*

Questions:

1. Identify the pre-attentive attributes that the speaker mentions while designing a dashboard.

2. What kind of challenges do the techniques face when thinking of design?

2.3 Visual perception

Introduction

Visual perception refers to the brain's ability to process what the eyes see. The visual perception of colours, patterns, and structures has been of interest to graphical user interfaces (GUIs) because these are perceived exclusively through vision. Therefore, the knowledge of visual perception enables us to create more real and impactful user interfaces.

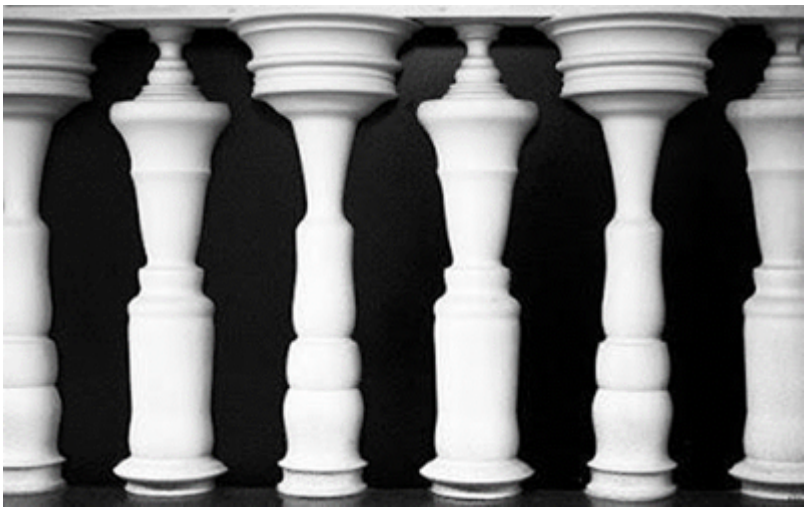
Activity: Observe and tell

Time: 10 minutes

Purpose: To reflect on the ways our brain perceives an illustration.

Task: Observe the illustration and share what you infer from it in the discussion forum.

Feedback: To help further the conversation, respond to the posts of at least 2 of your peers. You can ask them questions for clarification on their approach or provide them with feedback. Your facilitator will moderate this discussion.



Source: Optical illusion: Pillars or figures (Genius Puzzles, n.d.)

2.3.1 A memory process

Visual perception - A memory process

Visual perception, the act of seeing a visual or an image, is handled by the visual cortex is extremely fast and efficient.

Cognition is the act of thinking, processing information, making comparisons and examining relationships. This is handled by the cerebral cortex located at the front of the brain. The cerebral cortex is much slower and less efficient.

Data visualisation shifts the balance between perception and cognition to use our brain's capabilities to its advantage. This means more use of visual perception and lesser use of cognition.

How do we achieve this? Before we answer this question, we need to understand how our memory works.

Type of memory	Iconic memory	Working memory	Long term
Duration	About a second	About 1 minute	A second to lifetime
How processing happens	Pre-attentive processing even before we pay attention	Can hold and process between 5 - 9 chunks of information	Information is stored by repeated application or through rehearsal

Refer to the following readings to further your understanding.

- D'Agord, J. (2019). *Improve Your Dashboard By Understanding Visual Perception*. Zencos. https://www.zencos.com/blog/data-visualization-improve-perception/
- Few, S. (n.d.). *35. Data Visualization for Human Perception*. Interaction Design Foundation. https://www.interaction-design.org/literature/book/the-encyclopedia-of-human-computer-interaction-2nd-ed/data-visualization-for-human-perception

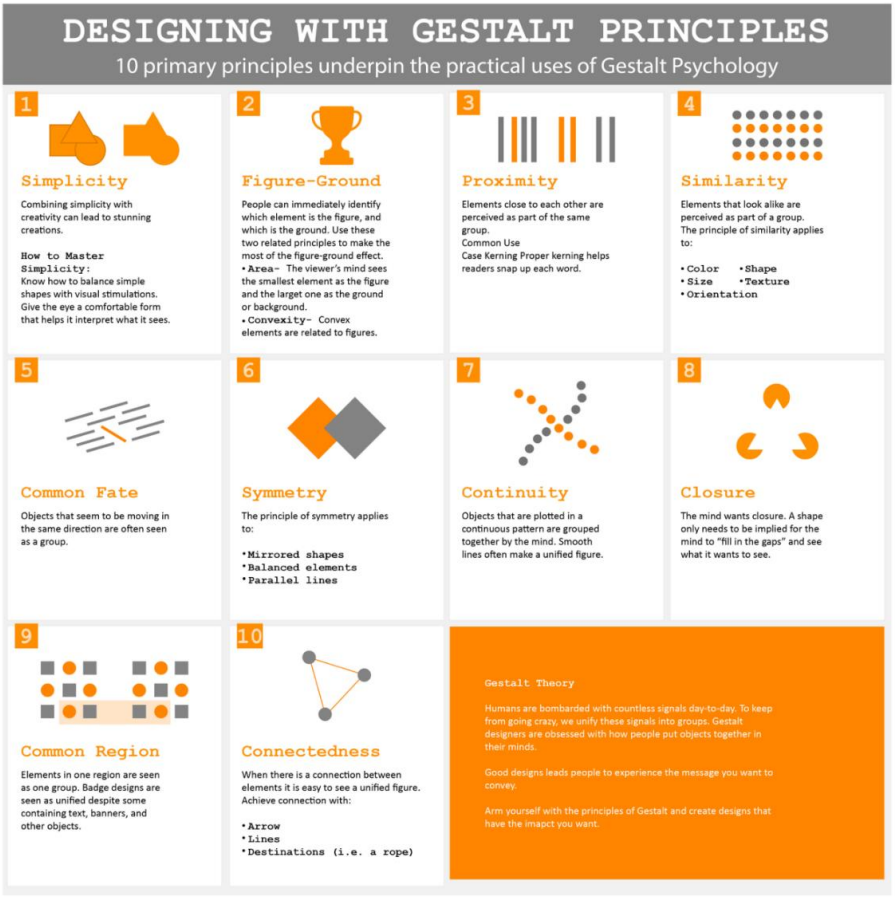
Watch this 6-minute 20-second video to understand the principles of visual perception. https://www.youtube.com/watch?v=ZIkcuS5iw4w&ab_channel=ArtwithKunstler *Source: (Art with Kunstler, 2020)*

Gestalt principles of visual perception and how to apply them to visualisation

Data is simply an assemblage of many individual elements, and visualisation groups these elements in a meaningful way to highlight patterns and variances. Principles by Gestalt are a good set of guidelines to assemble different elements into groups.

The *Gestalt psychologists* (Wolfgang Köhler, Max Wertheimer and Kurt Koffka) found that humans have a habit of grouping elements together, identifying patterns and finding order in chaos.

Remarkably, they found that humans don't do this casually – there are several principles that our brains use. The design field began to use the gestalt psychology principles to create designs that are more than the sum of their parts. Refer to the image below to grasp some of the primary principles.



Source: Designing With Gestalt Principles (Design Contest, 2017)

Refer to the following reading to further your understanding of using Gestalt principles of design.

- Craig, W. (n.d.). *How to Make Data Visualization Better with Gestalt Laws*. WebFX. https://www.webfx.com/blog/web-design/data-visualization-gestalt-laws/

Watch this 6-minute 50-second video to learn how to analyse and apply essential design principles to your Tableau visualisations. https://youtu.be/kgVVgfT49vU *Source: (Rachelle Galin, 2020)*

2.3.3 Application of Gestalt Principles

Activity: Gestalt principles in action

Time: 15 minutes

Purpose: To understand the Gestalt principles and their application in the practical world

Task: Observe an interactive visual from Tableau Public Website and answer the given question. Visit the URL.

Feedback: To help further the conversation, respond to the posts of at least 2 of your peers. You can ask them questions for clarification on their approach or provide them with feedback. Your facilitator will moderate this discussion.

Instruction

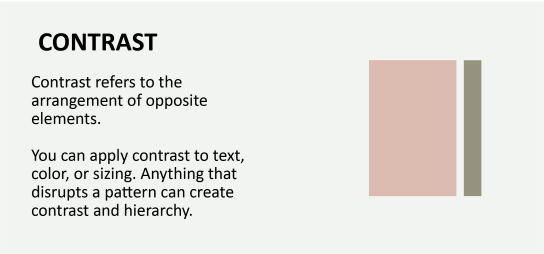
Step 1) Analyse this simple visualisation from Tableau public.

Step 2) Identify the elements it uses from the ones listed below and explain how.

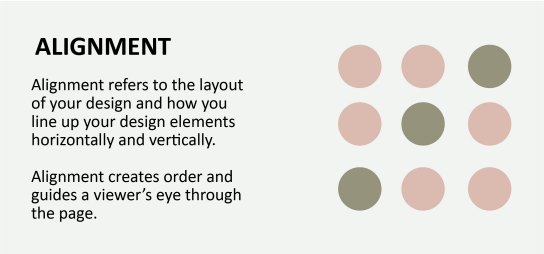
Gestalt Principles: Figure & Ground, Symmetry, Similarity

2.4 Alignment, white space and contrast

Introduction



Design is not merely a decoration and aims to make communication as easy and clear for the viewer as possible. You must learn to see and manipulate space to lend greater association, clarity, and interest to your visualisation.



The alignment principle connects elements visually (through invisible lines) on a visualisation. Using the principles of contrast, one can create strong dynamic variances among elements that are different. On the other hand, grids are very useful for achieving good alignment. This gives the graph/image a clean, well-organised look.

Source: Learn the Basics of Visual Hierarchy in Graphic Design (Jordan Prindle Designs, 2018)

Activity: Explore a dashboard for the use design principles

Explore this Tableau's Sample KPI dashboard which shares insights for sales and shipping data. Take a few minutes to reflect and share your opinions on the use of empty space, alignment and contrast. This will help you understand the concept better for the forthcoming activities.

Tableau's Sample KPI dashboard URL : https://www.tableau.com/dashboard-examples

Principles of design

What are the principles of design?

The principles of design are the rules one must follow to create an effective and attractive composition. The fundamental principles of design are Alignment, Contrast, Repetition, Proximity and White Space.

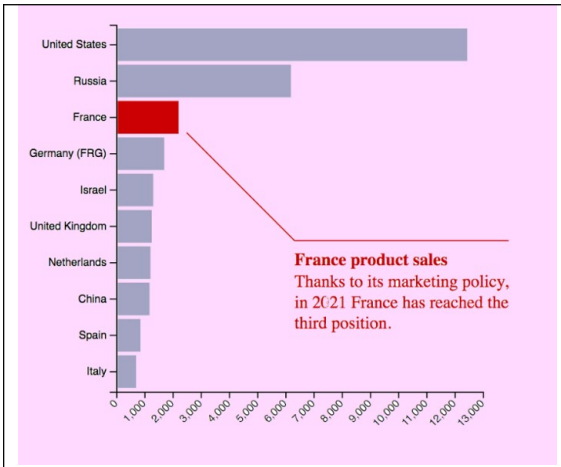
As for Data Visualisation, there are three primary principles at work: Alignment, White space and Contrast.

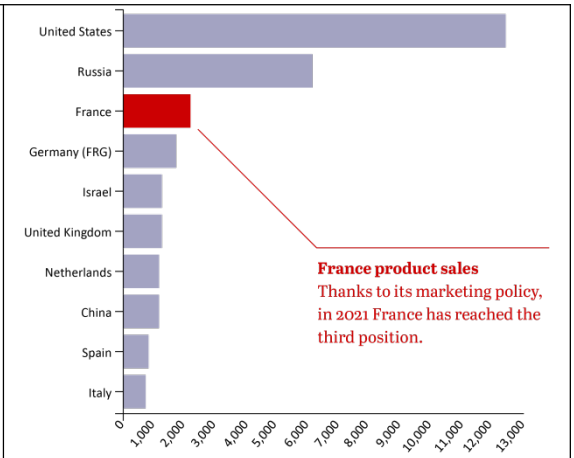
a. **White space**:

White space is the space among the elements of a visual composition. Good use of white spaces will increase readability and focus the reader's attention.

Two types of white spaces exist within a visualisation:

Macro white space	Micro white space
Macro white spaces - Refer to all the spaces surrounding the main content, such as the space behind a figure. In the example below, the macro white spaces are marked in pink	Micro white spaces - Refers to all the spaces between digits in text, between an axis and text, or between the bars in a bar chart. The following figure shows some micro white spaces in pink





b. **Alignment**:

The Principle of alignment states that nothing should be placed on the page randomly and that every item should have a visual connection with something else on the page. The basic purpose of alignment is to unify and organise the page. When items are aligned, the result is a robust, consistent unit.

c. **Contrast**:

Contrast is an important design principle because it captures attention towards the most important design elements and adds emphasis. Contrast happens when two essential elements oppose each other, like black and white, thick and thin, modern and traditional. High contrast can help guide the observer's eyes to the most vital parts of visualisation.

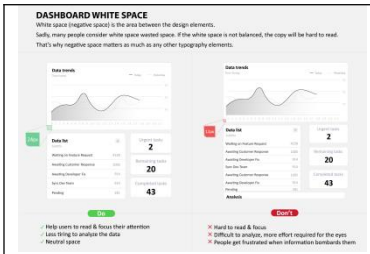
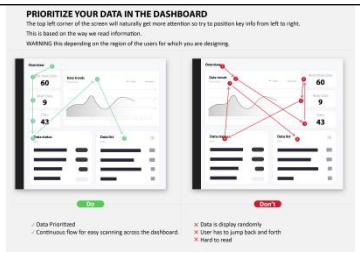
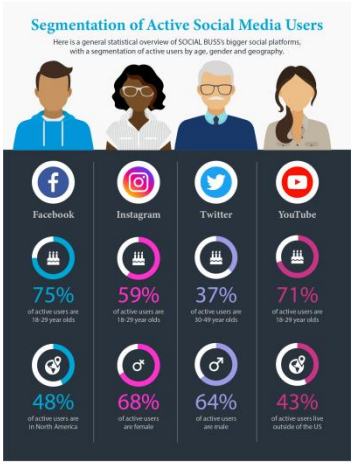
The following two videos are recommended for your further understanding.

- Watch this 10-minute 33-second video to learn more about Contrast, Repetition, Alignment, Proximity (C.R.A.P.) https://www.youtube.com/watch?v=AEt4wAllxAc&ab_channel=MsuTism
  Source: (Msu Tism, 2010)

- Watch this video of 3 minutes on the use of white space in infographic design. https://www.youtube.com/watch?v=npt3ipuJJBU&ab_channel=Easelly%3ADesignTips%26Tutorials *Source: (Easelly: Infographic Design Tips & Tutorials, 2019)*

Application examples of alignment, white space and contrast

Example 1	Example 2	Example 3

 *Source: Dashboard White Space (Do-Dont-UX, n.d.)*	 *Source: Prioritise your data in the dashboard (Do-Dont-UX, n.d.)*	Observe the following visualisation and reflect on the use of design principles of white space, alignment and contrast.  *Source: Segmentation of Active Social Media Users (Intellspot, 2020)*

2.4.1 Do's and Don'ts of a good visualisation

Activity: Identify the Do's and Don'ts of a good visualisation

Time: 20 minutes.

Purpose: To understand the basic do's and don'ts when creating a good visualisation.

Task: Identify a list of features that can turn a plot into poor visualisation. Refer to the previous videos/journals in the sections above or think carefully about the plots you have recently encountered in scientific publications, text books or lab reports.

Feedback: To help further the conversation, respond to the posts of at least 2 of your peers. You can ask them questions for clarification on their approach or provide them with feedback. Your facilitator will moderate this discussion.

2.4.2 Activity: Analyse dataset and select the most suitable data visualisation by using Gestalt Principles

Analyse dataset and select the most suitable data visualisation by using Gestalt Principles

Time: 40 minutes

Purpose: To analyse the given dataset and select the most suitable data visualisation by using Gestalt Principles

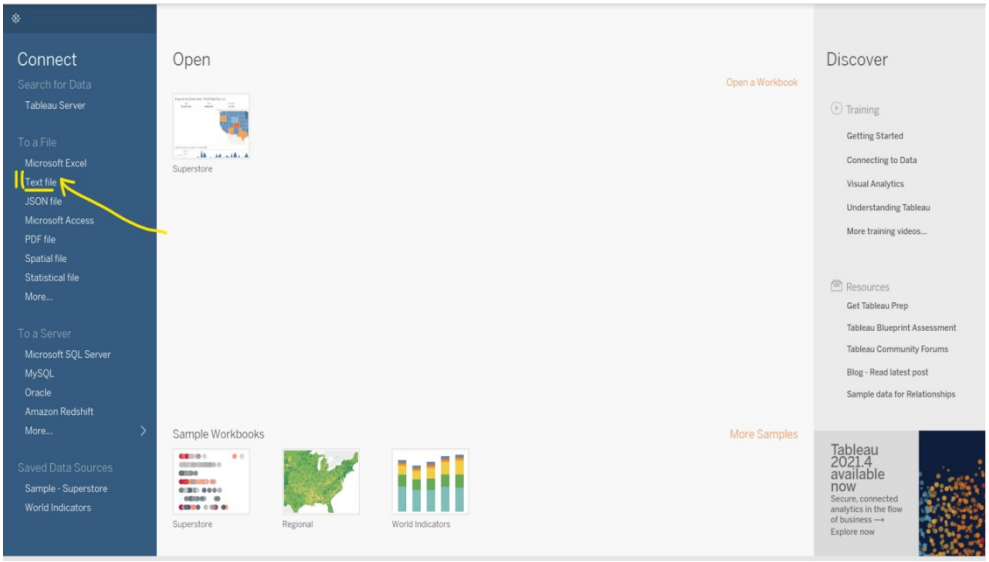
Task: In this assignment, you will design a visualisation (preferably a Tableau Dashboard) for a small data set of Walmart stores and provide the rationale for your design choices using the design principles from Gestalt's Theory. (Example area to explore: the effect on weather sales, temperature, fuel consumption, the holiday season, and other factors.)

1. Download the dataset (https://www.kaggle.com/c/walmart-recruiting-store-sales-forecasting/data). (Click on Download All)
2. Design at least three different visualisation types in the dashboard. (e.g. bar chart, scatterplot, map, etc.)
3. Each visualisation should illustrate different relationships (e.g. Don't make a pie chart and a bar chart from the same data).

Feedback: To help further the conversation, respond to the posts of at least 2 of your peers. You can ask them questions for clarification on their approach or provide them with feedback. Your facilitator will moderate this discussion.

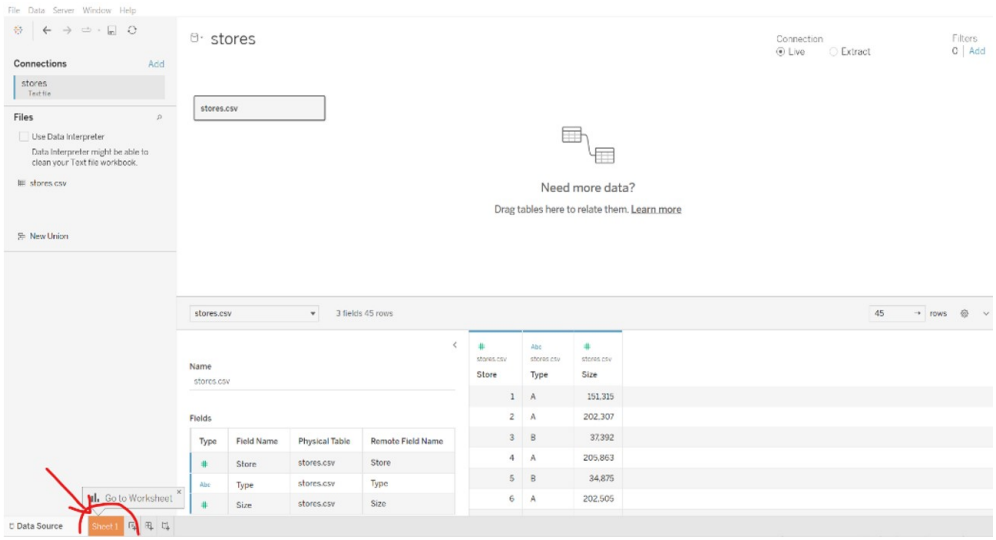
Instruction

1. Download the dataset by clicking on the link in the above section.
2. Open Tableau Desktop. Use Tableau Online if you have not yet installed Tableau.
3. To upload a CSV file to Tableau, on the left menu bar, click the "Text File" link under "Connect", navigate to the location of downloaded files, select the CSV file you wish to work with, and click "Open" to import.



Source: Tableau Desktop (Tableau, n.d.)

4. Once you've uploaded your CSV, the screen should display your data. In the bottom left corner of the screen, click on the tab that says, "Sheet 1".



Source: Tableau Desktop (Tableau, n.d.)

5. You're now ready to start building your visualisation.