



Creativity + Innovation & Technology

O-CITY PROJECT



Co-funded by the
Erasmus+ Programme
of the European Union

Project funded by: Erasmus+ / Key Action
2 -Cooperation for innovation and the exchange
of good practices, Knowledge Alliances.

Topic 2 - Lesson 2. Visual representation of data and information

This lesson has been prepared with the aim of facilitating students in choosing how to properly represent data and information in a graphical way, when creating an infographic.

Most common visualization techniques

When dealing with the representation of data and information there are several visualization techniques that can be used. They are an additional and more accurate way to graphically represent them, beyond the use of layout, colour, and typography, or other visual elements, that is usually made to convey meaning or for aesthetic purposes.

Visualization techniques organize data in a structured way with the aim of showing patterns and relationships that allow people to obtain information. This is particularly useful when we need to understand large quantities of data and the complex connections between them.

Systems to visualize information and data include different kinds of graphs, charts, diagrams, and maps.

Knowing the characteristics of each of them is important, not only to decide how to visualize data after collecting them, but also to learn how to read them, avoiding misleading.

Following, some most common visualization techniques are introduced.

For a complete description and more typologies of techniques, visit Dataviz project (<https://datavizproject.com/>) and Dataviz catalogue (<https://datavizcatalogue.com/>).

Bar chart

- It is useful to show discrete, numerical comparisons across categories, such as years or countries. It answers the question of "how many?" in each category.
- It can use either vertical or horizontal bars.
- The lengths of the bars are proportional to the values they represent.
- One axis of the chart represents the specific categories being compared and the second one a discrete value.
- You can order the bar in chronological order or not (depending on the category type).
- You can differentiate the colour of bars, but avoid creating confusion or misunderstanding.
- Be aware that labelling can be problematic when there are a large number of bars. By the way, horizontal bars better display labels.
- A similar type of chart is a stacked bar chart that represents multiple datasets on top of each other to show the subdivision of the larger category into smaller categories and their relations to the total amount.

I am interested in visiting a(n) ...

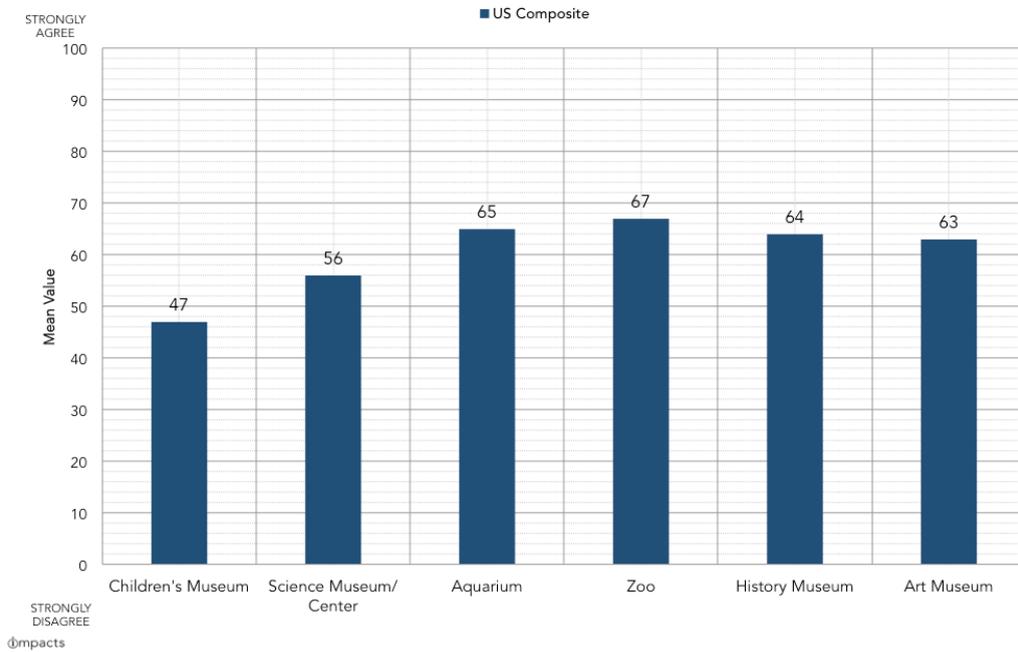


Fig. 1 - Example of bar chart (vertical). Source:

<https://www.colleendilen.com/2019/07/17/unwelcoming-uninteresting-and-unengaging-heres-whats-up-with-perceptions-of-childrens-museums-data/>

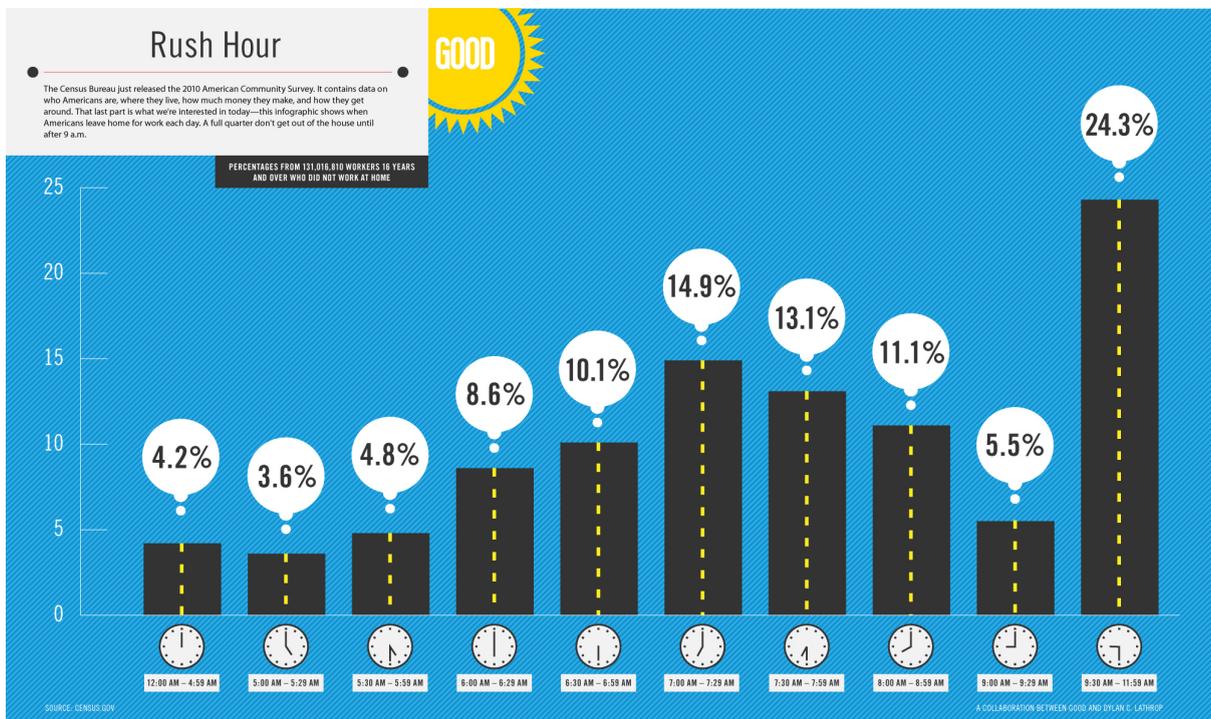


Fig. 2 - Example of bar chart (vertical). Source:

<https://www.good.is/infographics/infographic-out-the-door-late-americans-commute>

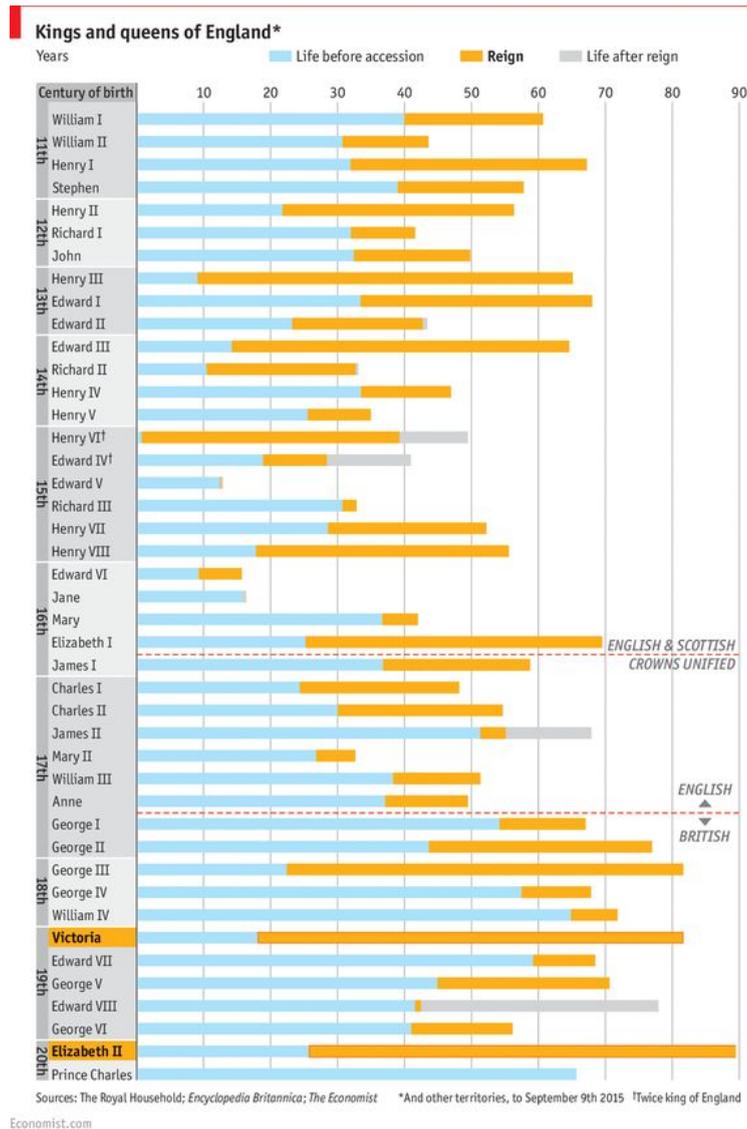


Fig. 3 - Example of stacked bar chart. Source: <https://ar.pinterest.com/pin/186477240799577645>; <https://www.economist.com/graphic-detail/2015/09/09/queen-elizabeth-ii-takes-the-crown>

Line graph

- It is useful to show quantitative values (typically on y-axis) over a continuous interval or time period (typically on x-axis), in order to better observe trends or variation over time.
- It represents information on a Cartesian coordinate grid as a series of data points called “markers” connected by straight lines. Markers can be put in evidence or not to highlight specific moments.
- It also allows us to represent negative values.
- When correlated with other data series individual lines can be compared to one another.
- Be aware that comparing more than 3-4 lines on the same chart can make reading difficult.

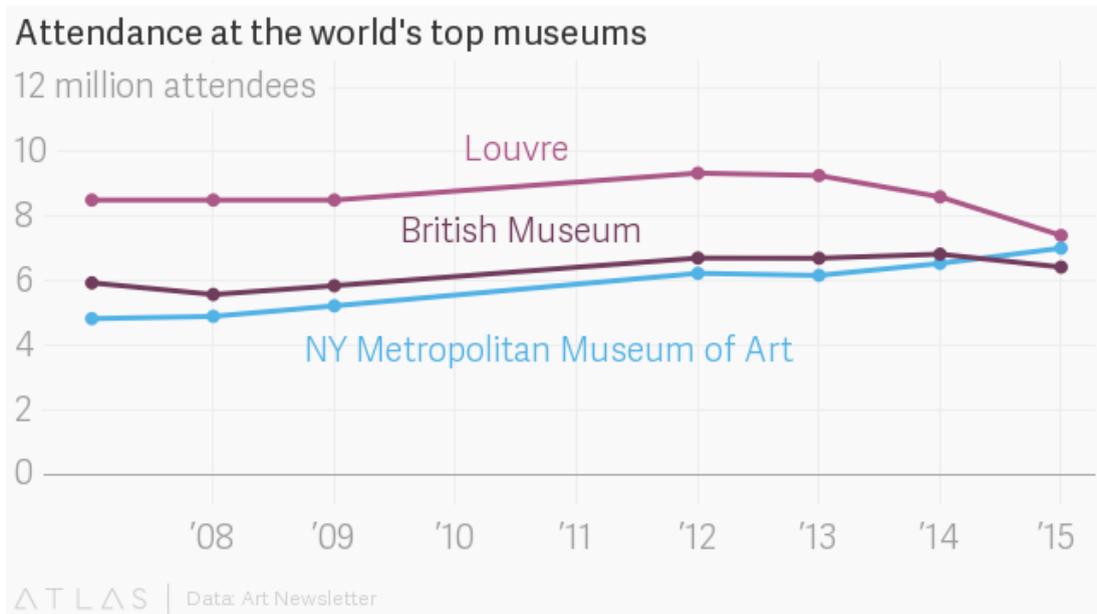


Fig. 4 - Example of line graph. Source:

<https://qz.com/1015412/new-yorks-metropolitan-museum-of-art-may-soon-surpass-the-louvre-as-the-worlds-most-popular-museum/>

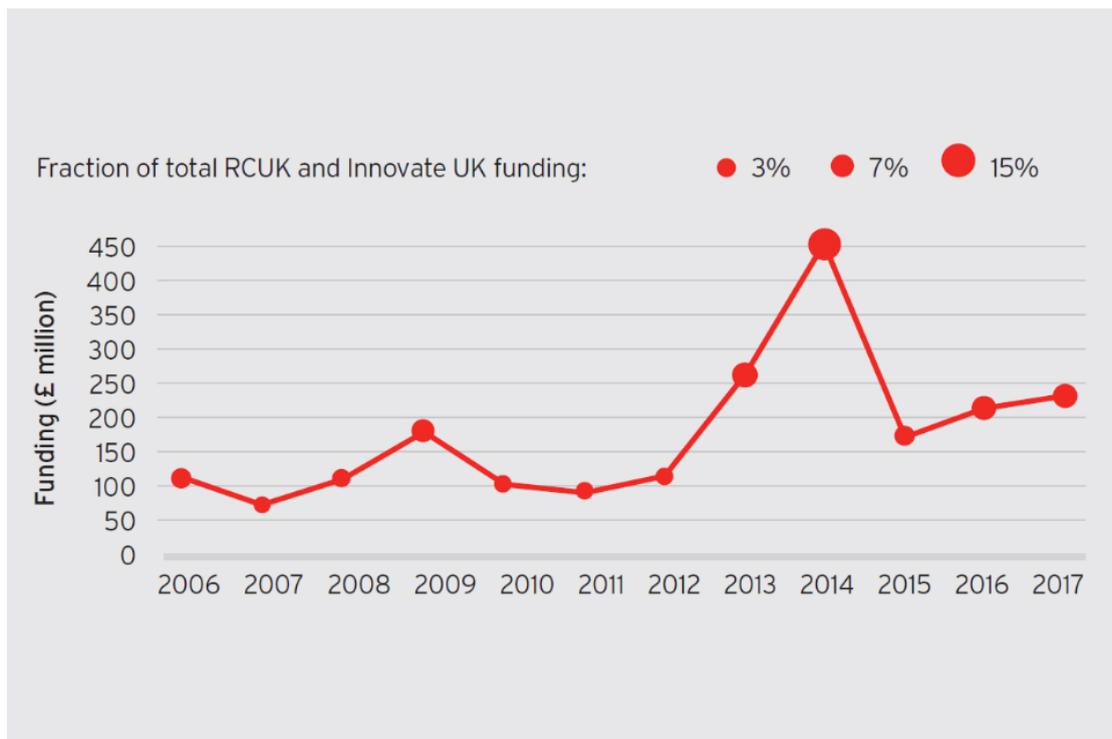


Fig. 5 - Example of line graph. Source:

<https://www.gov.uk/government/publications/creative-industries-sector-deal/creative-industries-sector-deal.html>

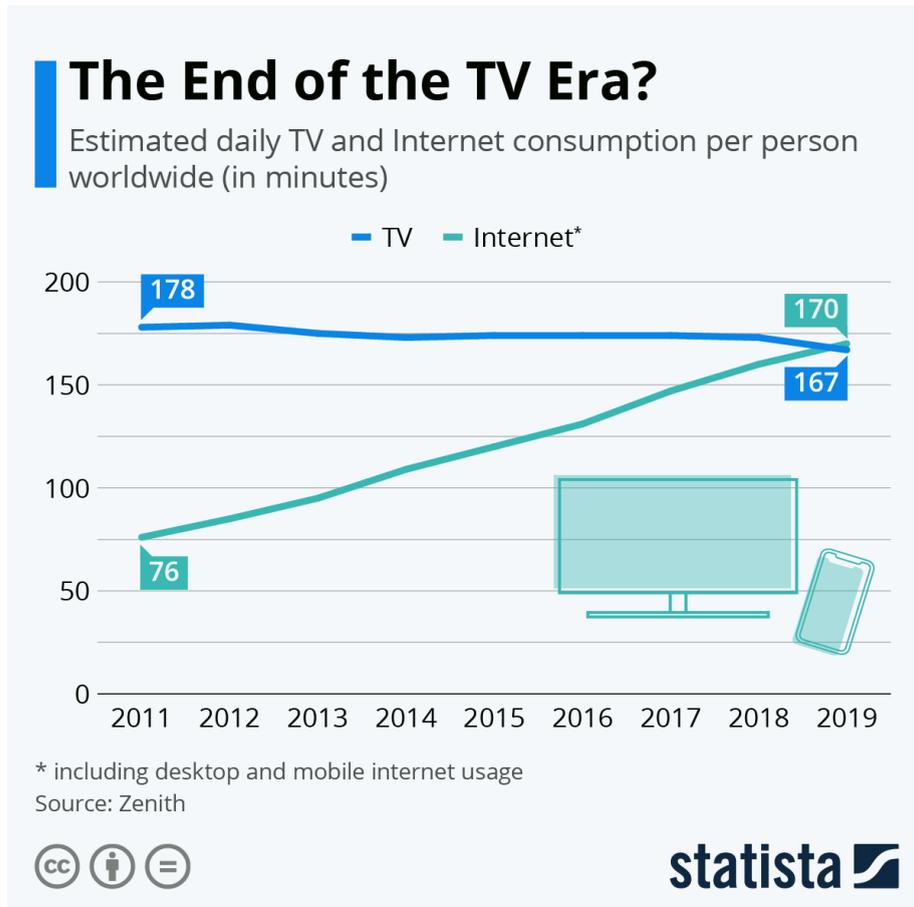


Fig. 6 - Example of line graph. Source: <https://www.statista.com/chart/9761/daily-tv-and-internet-consumption-worldwide/>

Pie chart

- It is useful to quickly show numerical proportions and percentages between categories by representing the proportional distribution of the data.
- The chart is a circle divided into proportional segments (sectors). The area of each sector is proportional to the quantity it represents of each category. The sum of all the sectors (then of all the data) results in the full circle (equal to 100%).
- Usually a legend is put aside to make more clear the categories represented in the chart.
- It is suited to show a few values, so avoid using it for large amounts of data.
- Be aware that it is not suited to make accurate comparisons between categories (especially without showing numerical values) or to compare different data series in the same chart.

Lastly, do you think it is important for your local town or city to have its own museum or art gallery?

(50 responses)

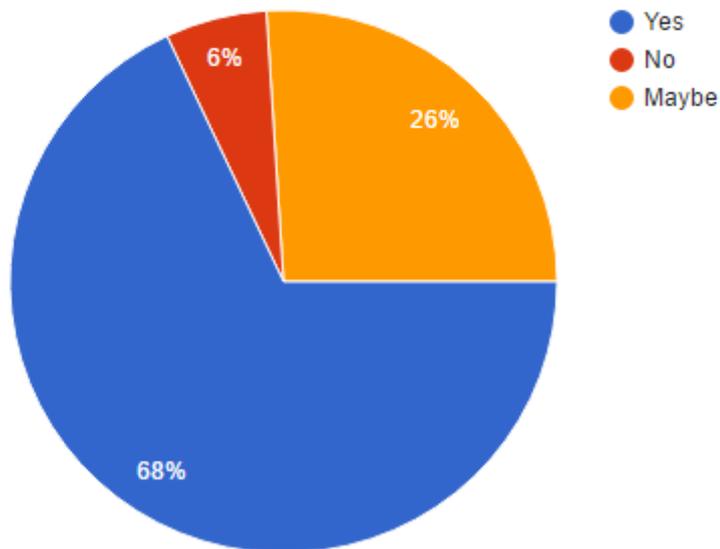
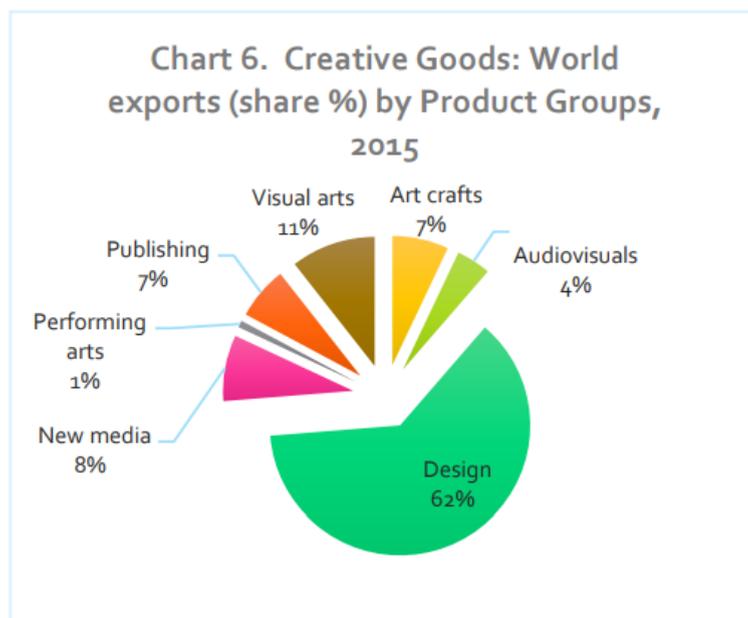


Fig. 7 - Example of pie chart. Source:

<https://medium.com/@miaeveliina/how-can-we-encourage-those-who-rarely-visit-museums-to-do-so-more-often-441c27cf4770>



Source: UNCTAD, based on official data reported to UN COMTRADE Database

Fig. 8 - Example of pie chart variation. Source:

<https://europeansting.com/2019/04/19/5-things-you-need-to-know-about-creativity/>;
https://unctad.org/system/files/official-document/ditcted2018d3_en.pdf

The Heritage Emergency Fund

Grants by heritage area

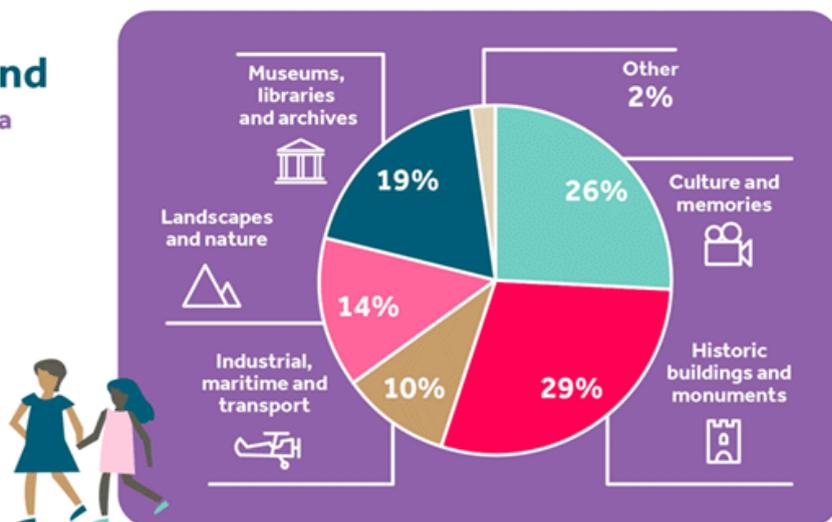


Fig. 9 - Example of pie chart. Source:

<https://www.heritagefund.org.uk/news/heritage-emergency-fund-supports-hundreds-organisations-across-uk>

Area chart

- It is useful to show the development of quantitative values over an interval or time period.
- It is like a line graph, but with the area below the line filled in with a certain colour or texture.
- A similar type of chart is a stacked area chart that represents multiple data series that start each point from the point left by the previous series in order to compare the variation of multiple data categories over time. It can be useful to use different colours to create contrast.

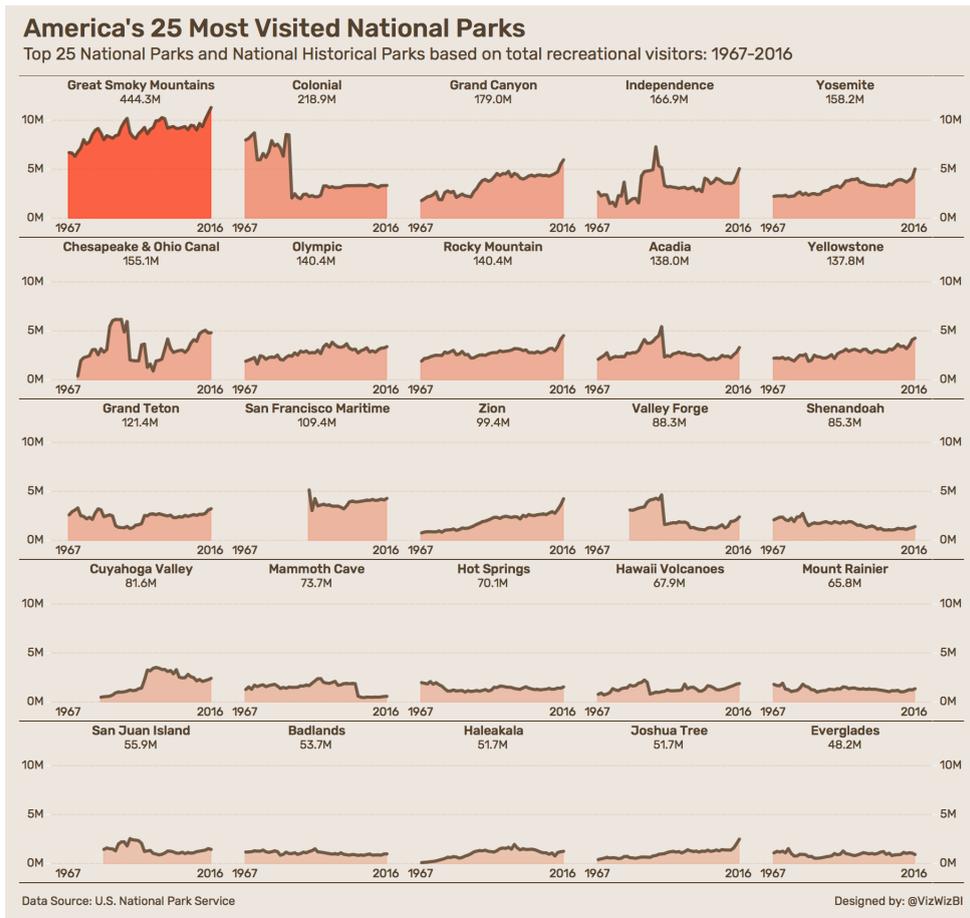


Fig. 10 - Example of area chart. Source: <https://www.vizwiz.com/2017/06/mm-23.html>

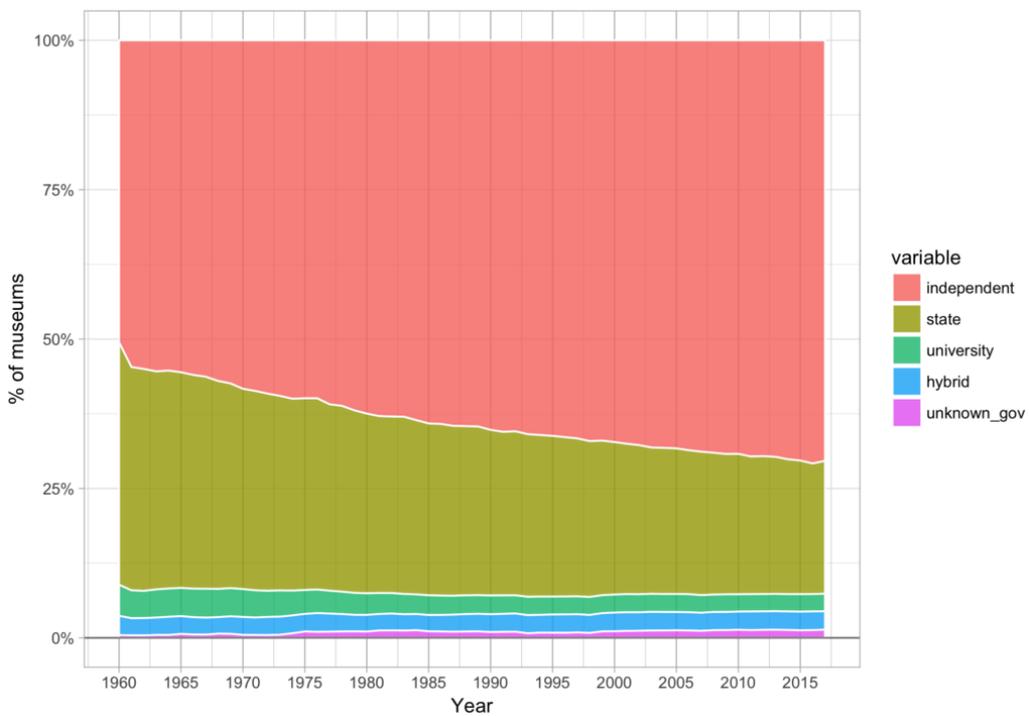


Fig. 11 - Example of stacked area chart. Source: <http://blogs.bbk.ac.uk/mapping-museums/2018/02/23/museum-closure-pre-findings/>

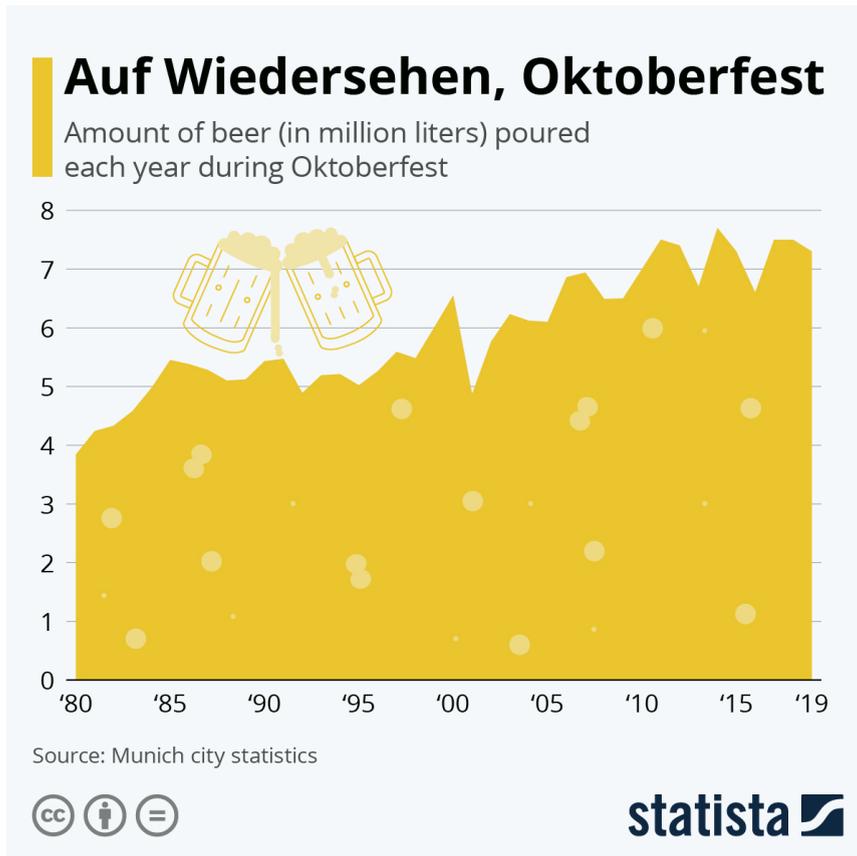


Fig. 12 - Example of area chart. Source: <https://www.statista.com/chart/21456/beer-poured-at-oktoberfest/>

Radar chart

- It is useful to compare multiple quantitative variables, allowing a quick view on variables that have similar values or any outliers amongst each variable. In detail, it is ideal for displaying performance, since it shows which variables are scoring high or low within a dataset.
- It displays multivariate data in the form of a two-dimensional chart of three or more quantitative variables represented on axes starting from the same centre. All axes are arranged radially.
- All the variables in a dataset, plotted along their individual axis, are connected together to form a polygon.
- It is suited for comparing values on a single straight axis (not for comparing values across each variable).
- Be aware that having multiple polygons or many axes (variables) in one chart makes it hard to read, confusing, and too cluttered, especially if the polygons are filled in.



Fig. 13 - Example of radar chart. Source: <http://www.thuynguyen.de/goodfood>

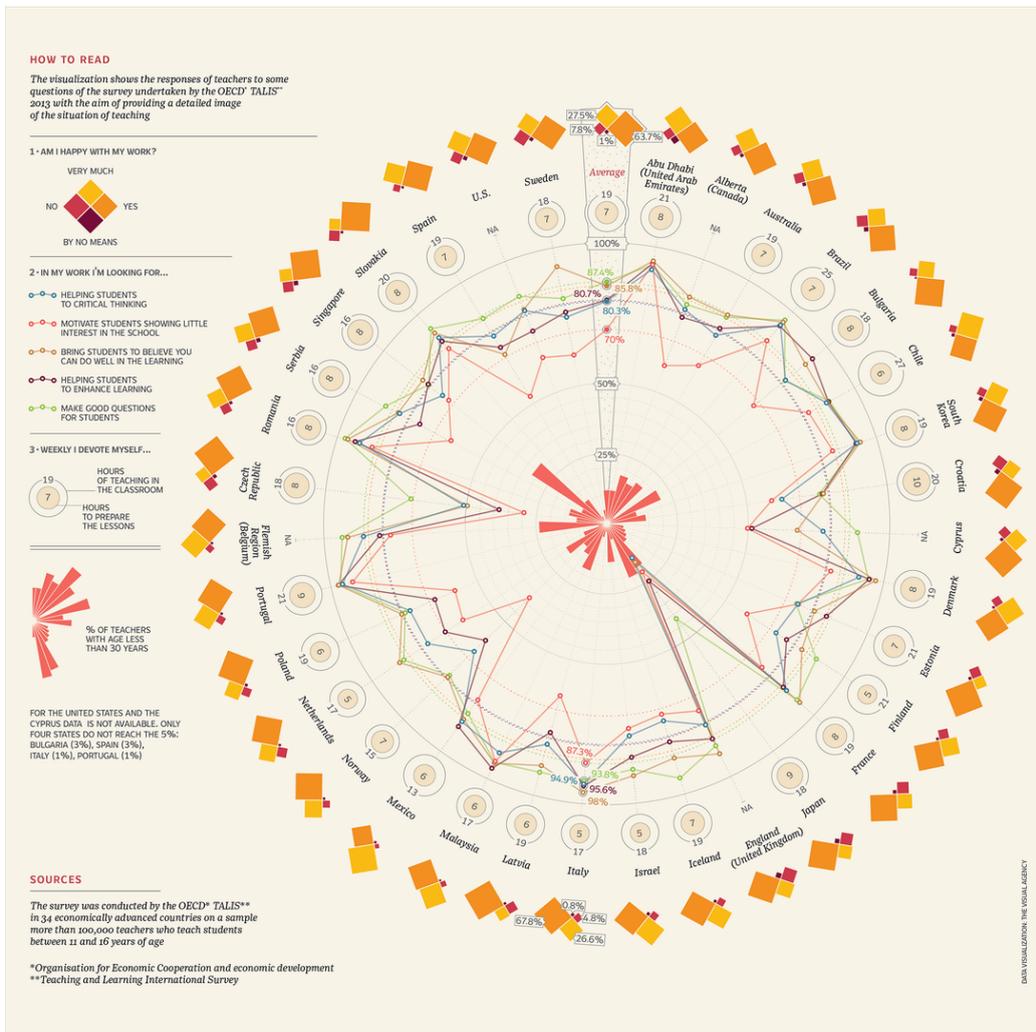


Fig. 14 - Example of radar chart. Source:

<https://www.informationisbeautifulawards.com/showcase/549-what-teachers-think>

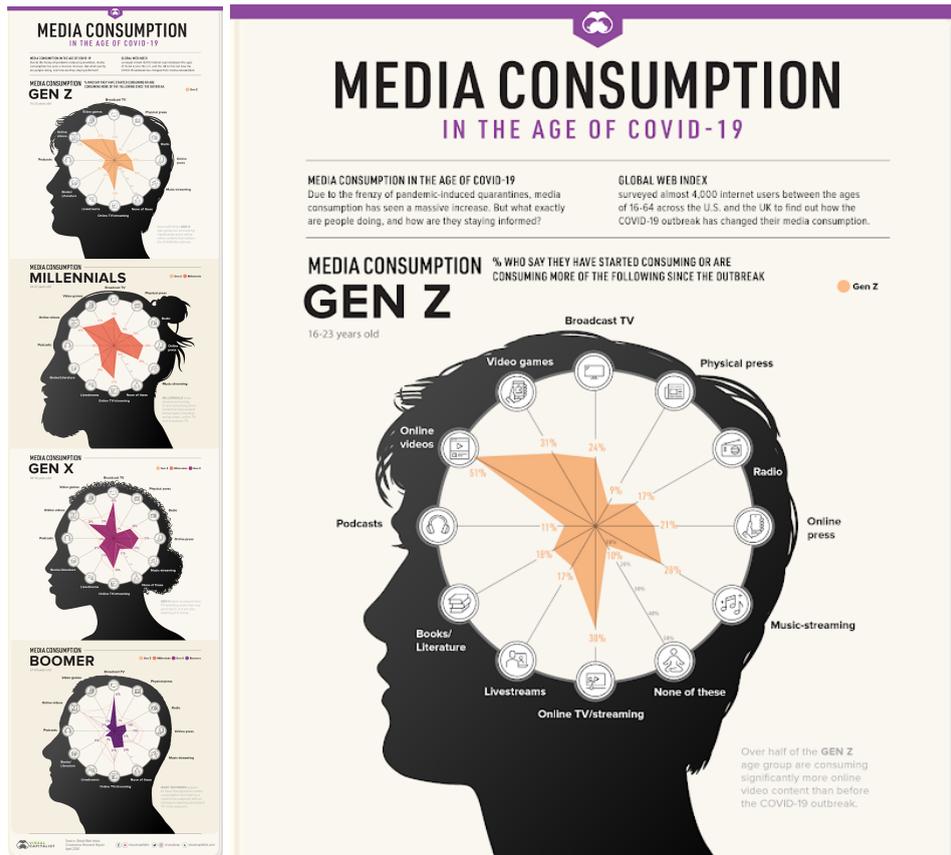


Fig. 15 - Example of radar chart (and detail). Source: <https://www.visualcapitalist.com/media-consumption-covid-19/>

Scatter plot

- It is useful to observe the distribution of data and find correlations between two variables.
- It is a type of mathematical diagram representing a dataset on a Cartesian coordinate grid as a collection of points.
- The position of each point is determined by the value of one variable on the horizontal axis (x-axis) and the value of the other variable on the vertical axis (y-axis).
- The use of colours and shapes can add a third variable to the graph (better if it is explained by a legend), or simply help visualizing the variation in the values.
- It best suits paired numerical data where one variable likely impacts the other one.
- Be aware that correlation is not causation.
- A similar type of graph is bubble chart that replaces data points with bubbles, displaying three dimensions of data (size in addition to the two variables determining position).
- Another chart with bubbles allowing to show proportion between elements, but without using a Cartesian coordinate system, is the clustered bubble chart, where nested circles are used to represent hierarchies and compare values.

Ecological Footprint and Human Development

A low average Ecological Footprint and high score on the UN Human Development Index are the minimum conditions for global sustainable human development. By learning to “think inside the (blue) box,” we can strive toward a world where everyone lives well, within the means of one planet. At Global Footprint Network, we believe this is humanity’s shared goal.

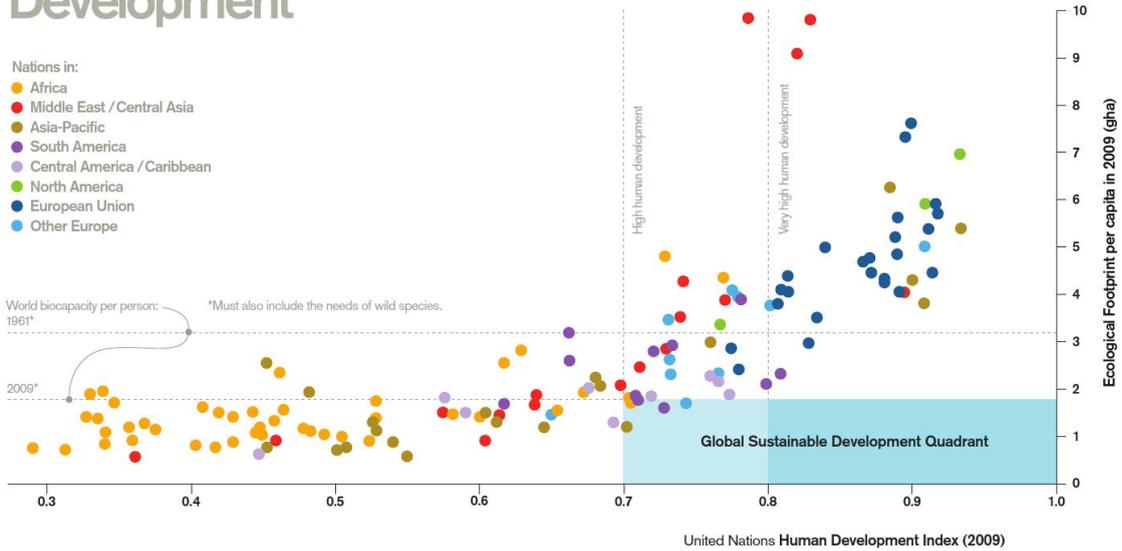


Fig. 16 - Example of scatter plot. Source:

https://issuu.com/globalfootprintnetwork/docs/2012_annual_report_new_110613a/22

Speech Length

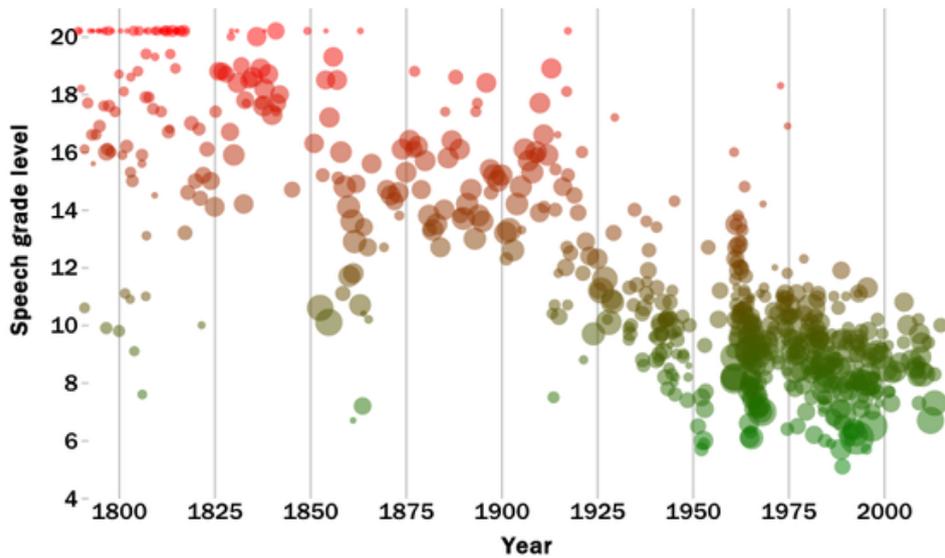


Fig. 17 - Example of bubble chart. Source:

<https://www.theatlantic.com/politics/archive/2014/10/have-presidential-speeches-gotten-less-sophisticated-over-time/381410/>

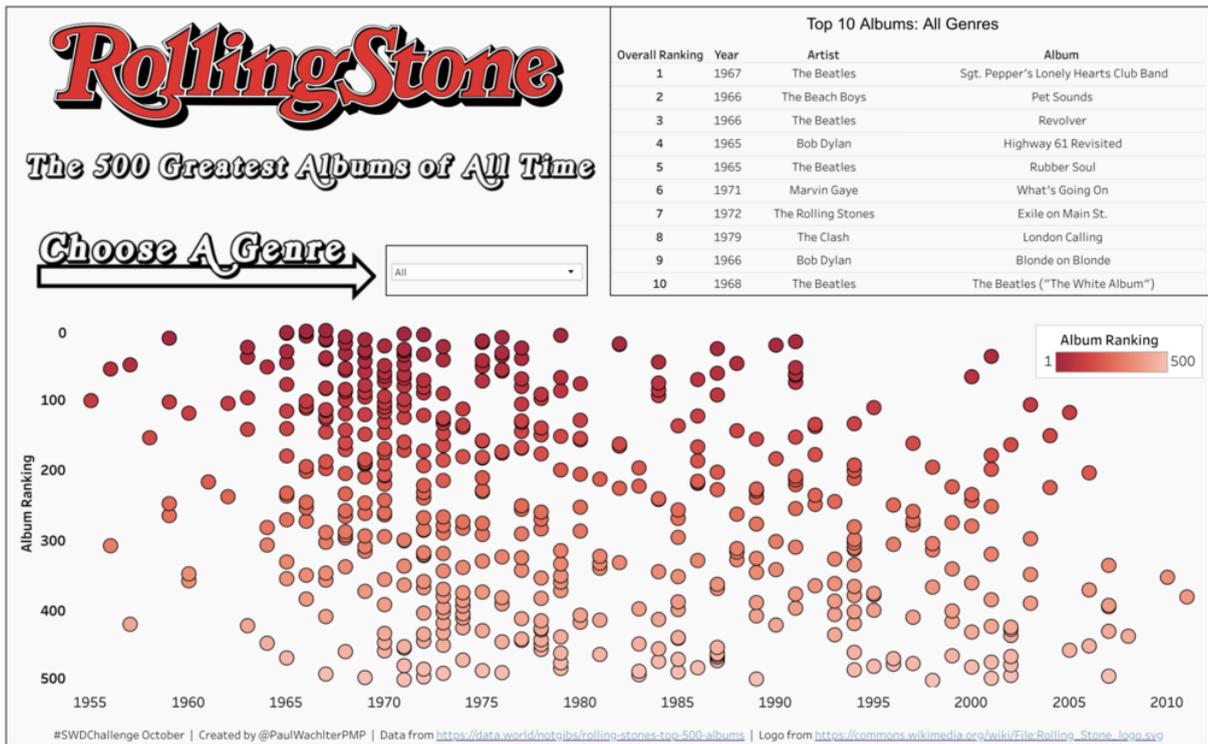


Fig. 18 - Example of scatter plot. Source:

<https://public.tableau.com/profile/paul.wachtler#!/vizhome/SWDCChallengeOctoberRollingStone500GreatestAlbums/500GreatestAlbums>

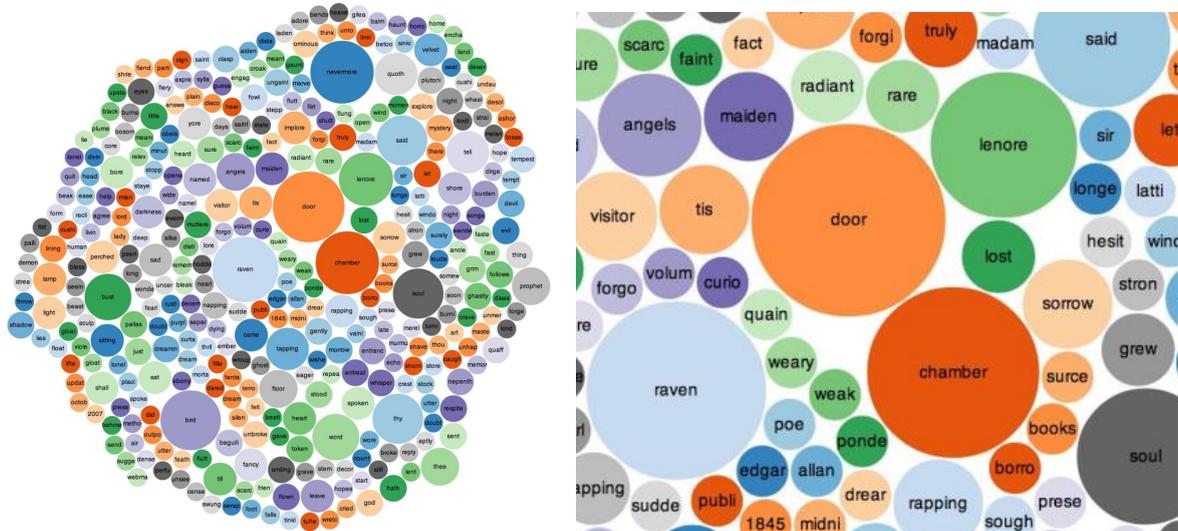


Fig. 19 - Example of clustered bubble chart (and detail). Source:

<https://datarep.tumblr.com/post/56879205177/nevermore-word-density-of-edgar-allan-poes-the>

Network diagram

- It is useful to represent the relationships and interconnections between specific elements that are part of a network in order to study the structure of a network.
- It connects nodes (typically little dots or circles, but icons can be used too) through links (typically lines) with other nodes or clusters to represent their connections and identify the type of relationships between them.

- An additional variable can be introduced by varying the node size or link stroke weight proportion.
- It can also use arrows instead of simple lines to show the direction of the relationship (one-way or two-way).
- Be aware that it is suited for a limited dataset, since it is hard to read when there are too many nodes.

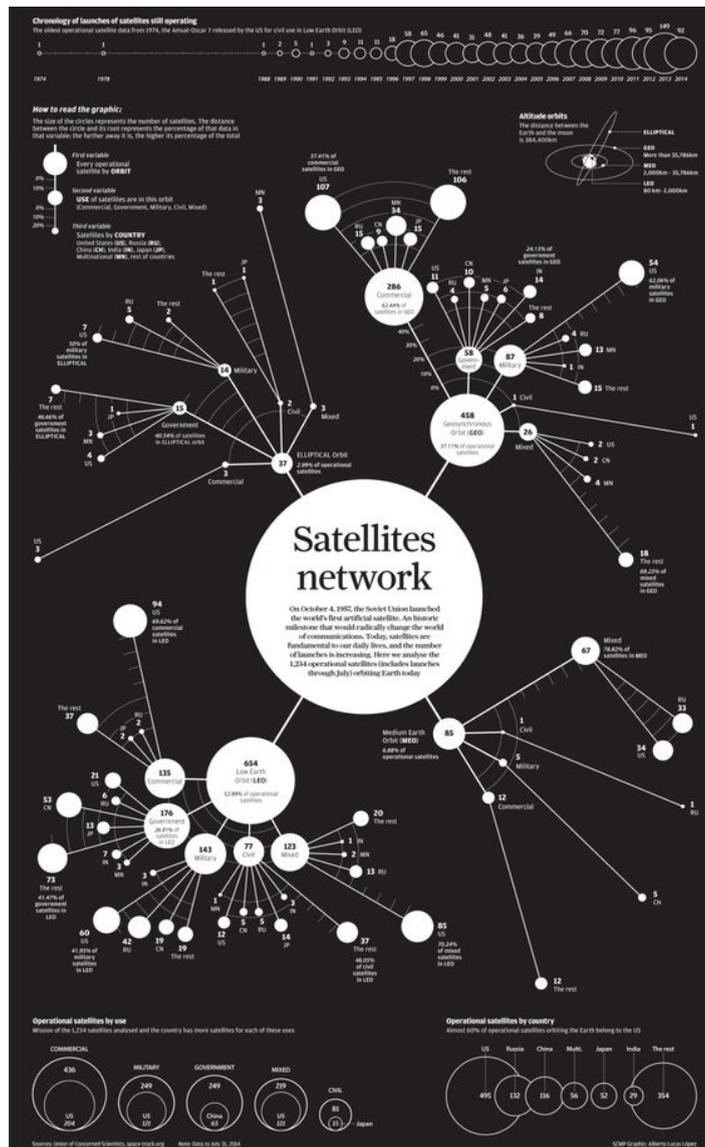


Fig. 20 - Example of network diagram. Source:

<https://www.scmp.com/infographics/article/1670384/infographic-satellites-network>

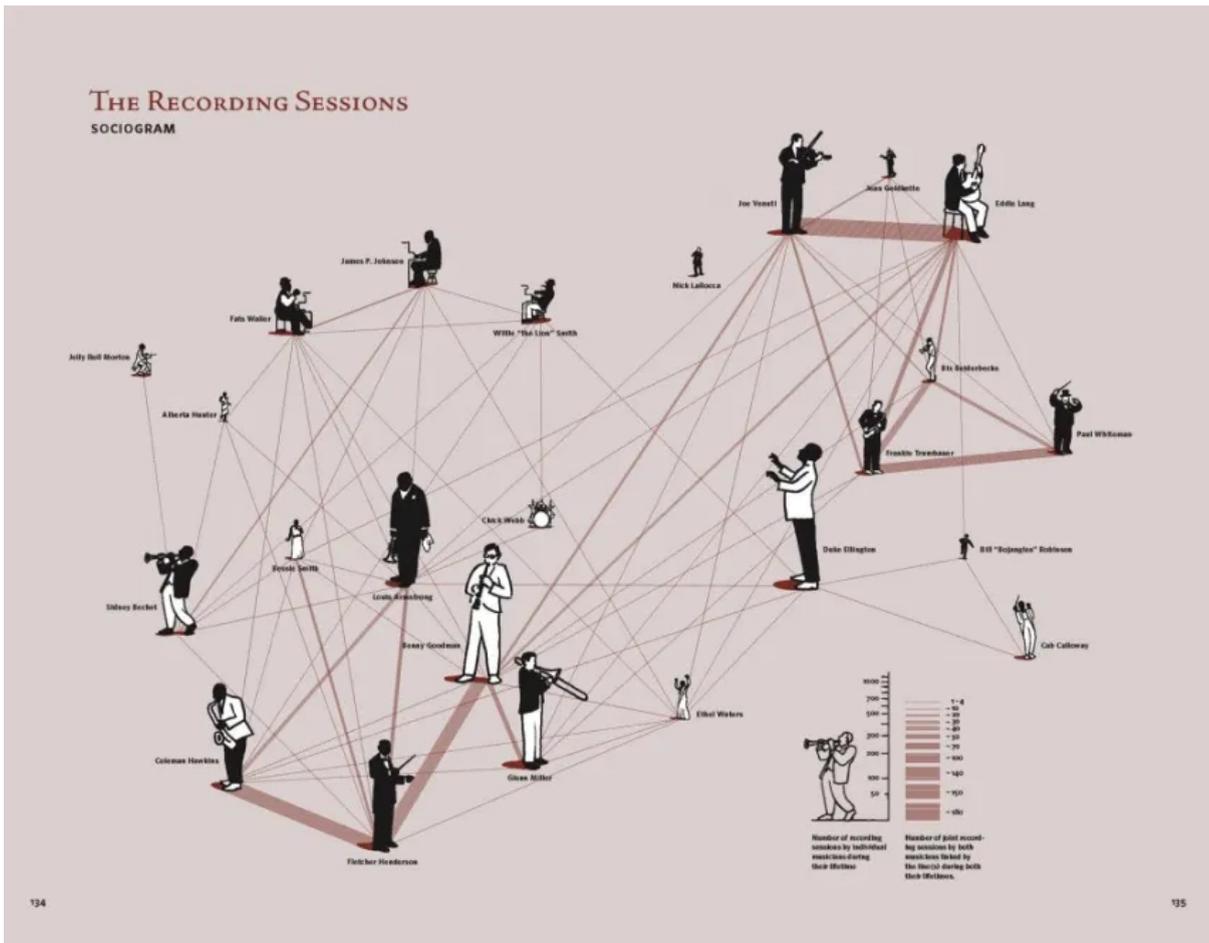


Fig. 21 - Example of network diagram (sociogram). Source: <https://www.nippoldt.de/en/illustration/info-graphics>

Tree diagram

- It is useful to represent the hierarchical relationships between different elements of an organization, a group (e.g. a group of languages), or a field of knowledge.
- It uses a tree-like structure.
- It usually start from a root node (i.e. the element that has no superior/parent) and pass through branches (i.e. line connections or relationships) to the other nodes (i.e. members that has both superior/parent and child nodes), till arrive to the leaf nodes or end-nodes (i.e. members that have no children or child nodes).
- It is suited to show family relation and descent, to classify elements (taxonomy) or species, for organizing the structure of computer nets, organizations, etc.
- Be aware that it is not suited to represent quantitative variables.

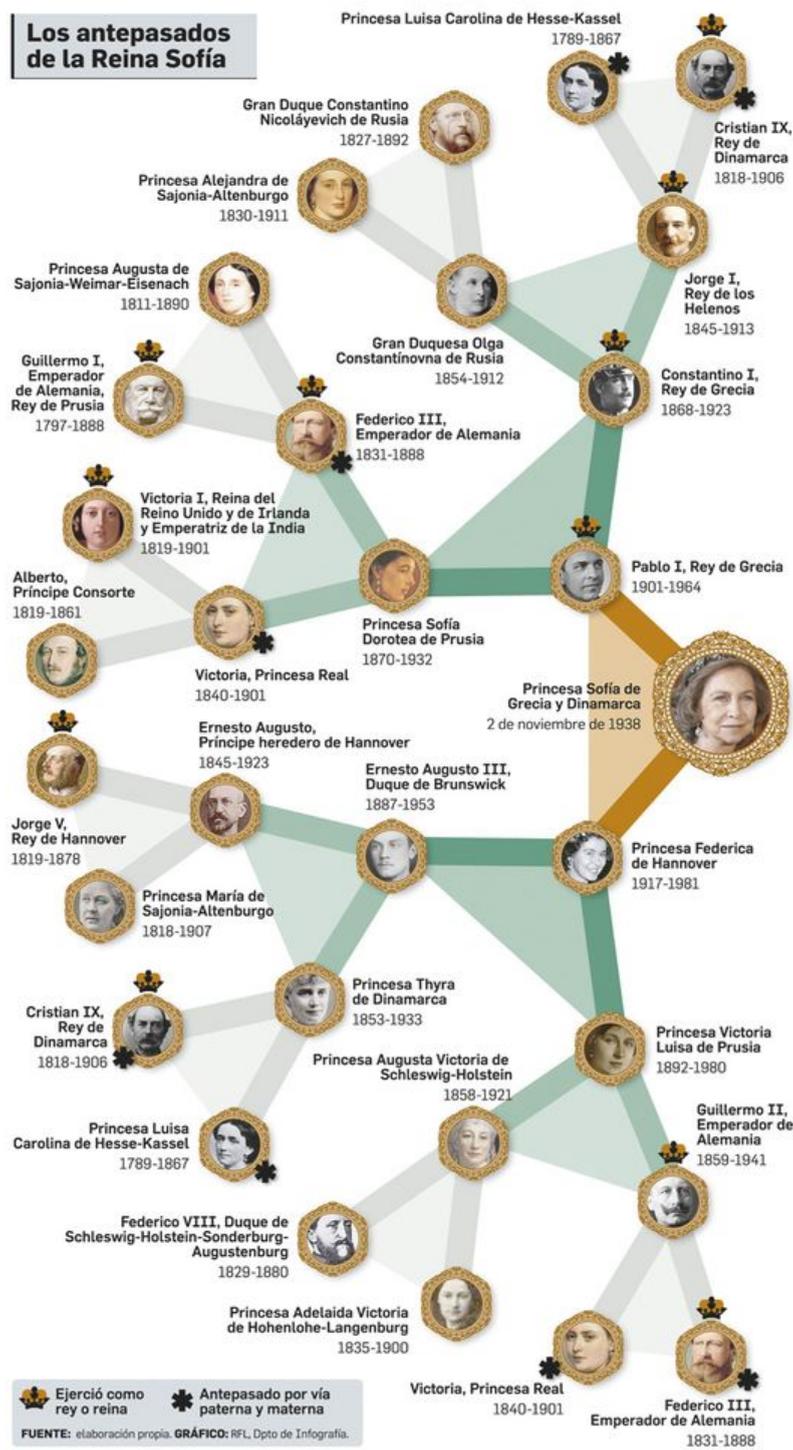


Fig. 23 - Example of tree diagram. Source: <https://www.pinterest.it/pin/609111918343841828/>; <https://www.raquelferia.com/>

Pictorial chart or Pictogram

- It is useful to give a more engaging and representational overall view of data or to help overcoming linguistic, cultural, or educational barriers, since icons can easily convey meaning.
- It uses icons, pictures, or pictograms, etc., to represent the data's subject or category.

- The number of repeated pictograms or icons (placed side-by-side in either columns or rows) or their relative size, indicate a quantitative value compared to that of another category. Each icon can represent one unit or any number of units.
- It suits a small set of discrete data, since in large datasets it could be hard to count the values (i.e. counting each icon).
- Be aware that displaying partial icons can add confusion and make the icon hard to understand.
- It can present variants or be combined with other visualization techniques, for example bar chart

Worker productivity: Lazy ants win

After watching five ant colonies for 2 weeks, researchers came to the following conclusions:

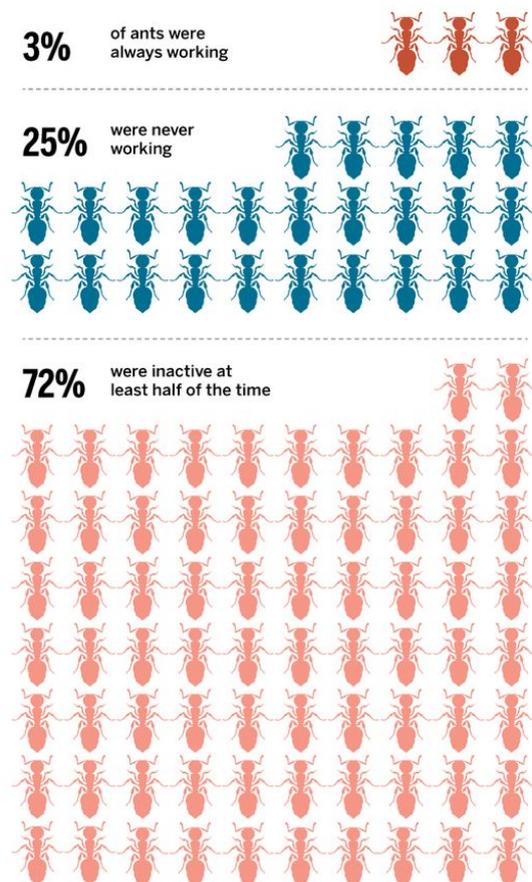


Fig. 24 - Example of pictorial unit chart. Source: <https://www.sciencemag.org/news/2015/10/most-worker-ants-are-slackers>

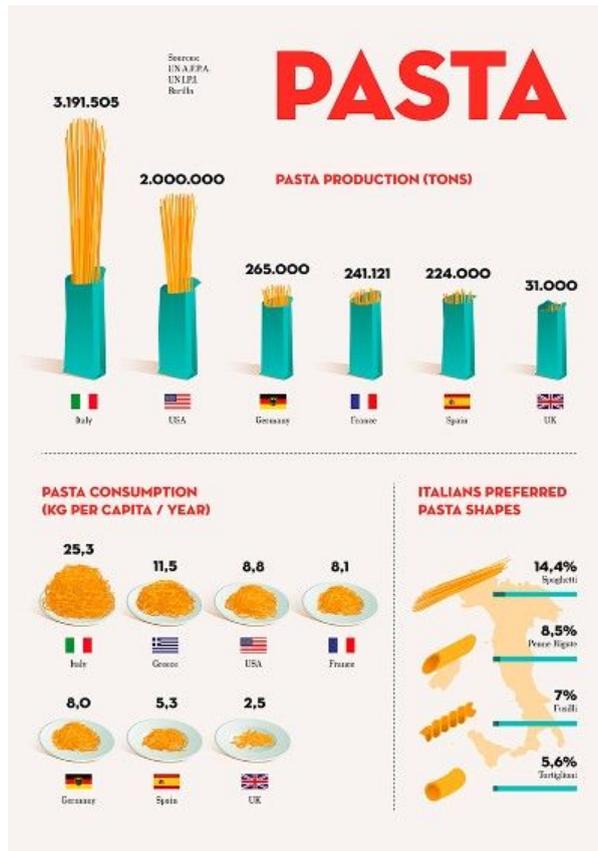


Fig. 25 - Example of pictorial bar chart. Source: <https://iwantdata.tumblr.com/post/127783551330/pasta-infographics>

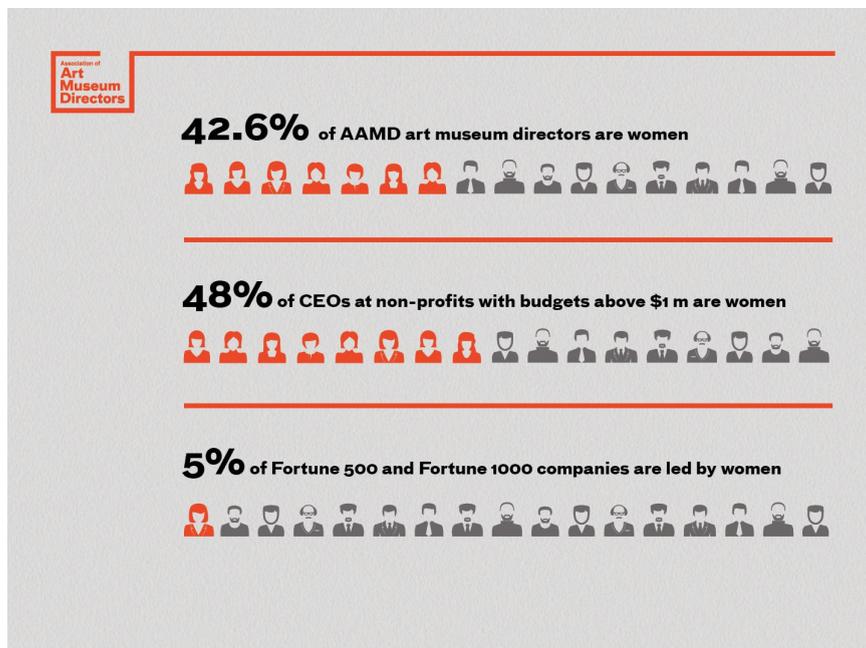


Fig. 26 - Example of pictorial fraction chart. Source: <https://www.pinterest.it/pin/141793088242335582/>; <https://www.stonesoupcreative.com/portfolio/association-of-art-museum-directors>

Predicting Best Picture: What other nominations do winners have?

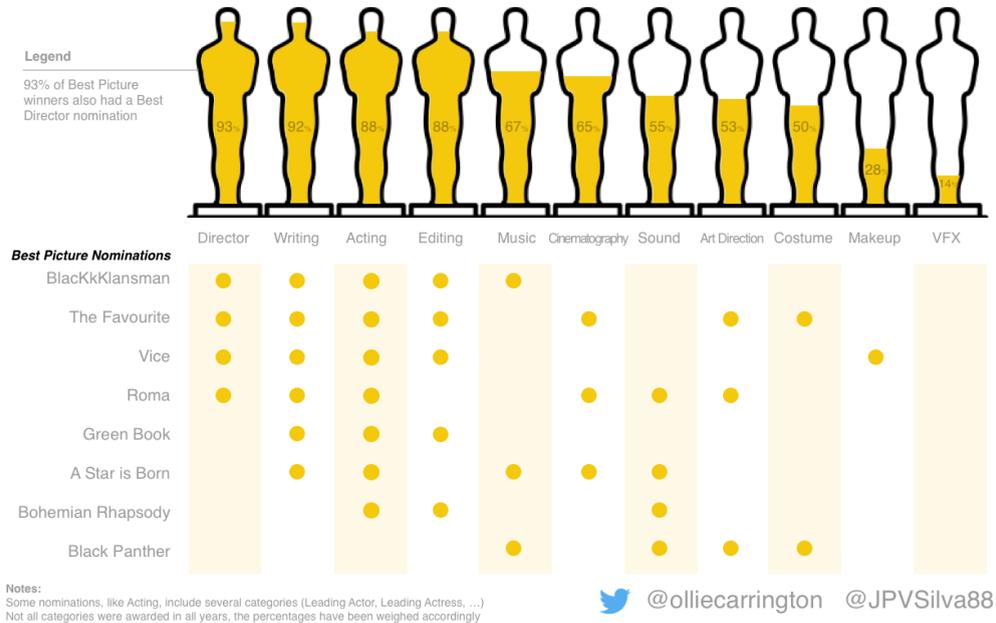


Fig. 27 - Example of pictorial fraction chart. Source:

<https://towardsdatascience.com/predicting-the-oscars-with-data-viz-e5bb0fd01766>

Word cloud or tag cloud

- It is useful to display a large amount of keyword metadata (tags) from a given body of text, for example from a website, and quickly perceive the most prominent terms or analyse speeches.
- It usually visualizes free form text as single words arranged in the shape of a cloud, but they can be arranged in any format.
- It visually represents text data using font size or colour to show importance. Colour can also be used to display another data variable.
- It is especially used for aesthetic reasons.
- Be aware that long words are emphasised over short words and that it is not suitable for an accurate analysis.



Fig. 28 - Example of word cloud. Source: <http://marcinignac.com/projects/mela-representing-museum-technologies/>

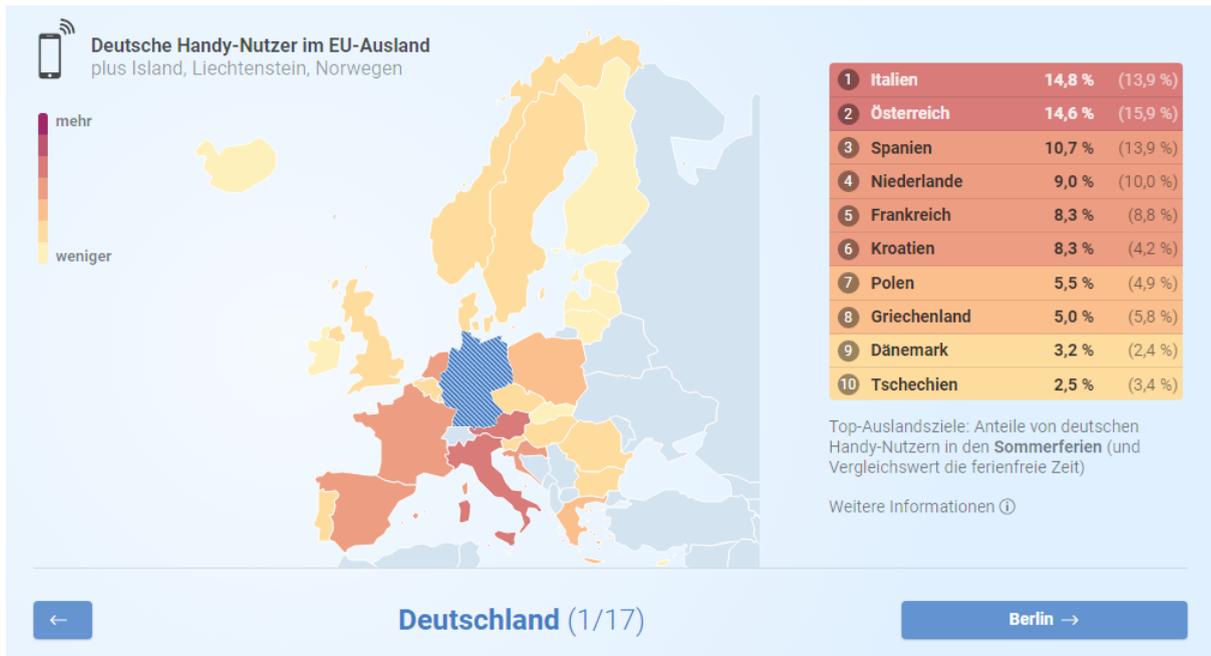


Fig. 30 - Example of choropleth map. (Title translation: Top Summer Travel Destinations for Germans). Source: <https://interaktiv.morgenpost.de/eu-roaming-ferien/>

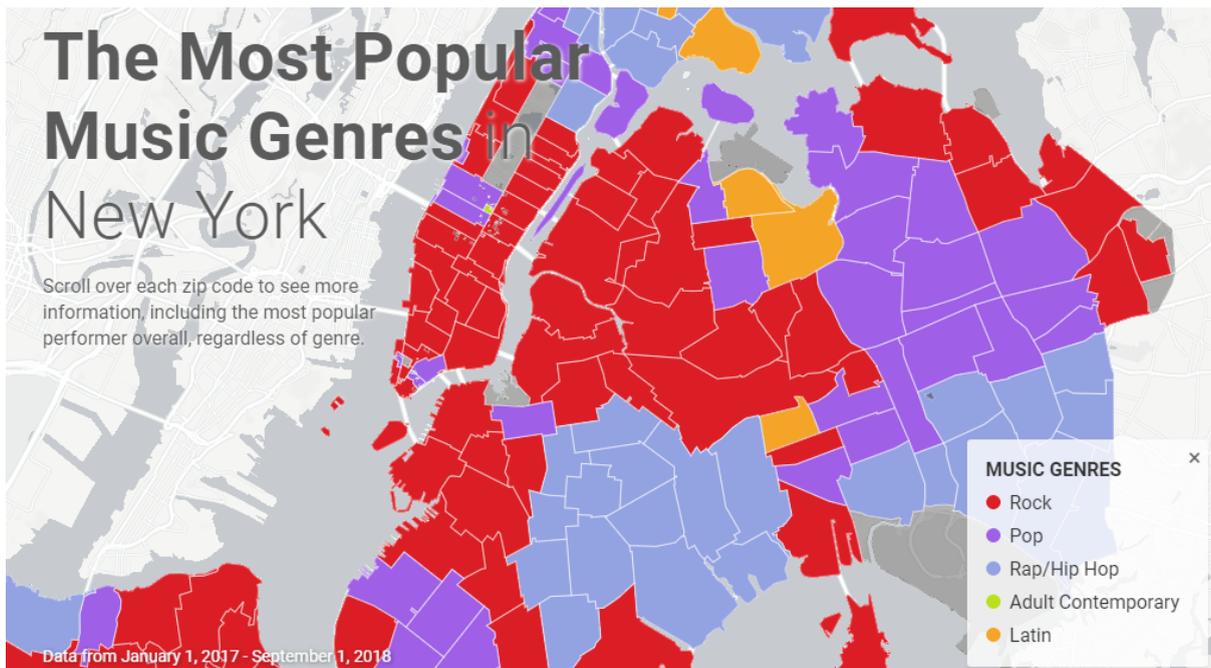


Fig. 31 - Example of choropleth map. Source: <https://www.vividseats.com/blog/most-popular-music-new-york/map>

- **Pin map** displays geospatial data on a map by pinning locations and giving them labels/descriptions. It can be combined with colours, icons, or other charts.
- It is different from dot map (or point map), which is instead a technique for representing the geographical distribution of data by plotting points of the same size on a geographical region.

Festivals : le bilan de la saison 2018

Les 25 festivals les plus fréquentés, en milliers de spectateurs

Organisateur

● Association ● Live Nation ● Entreprise ● Autres

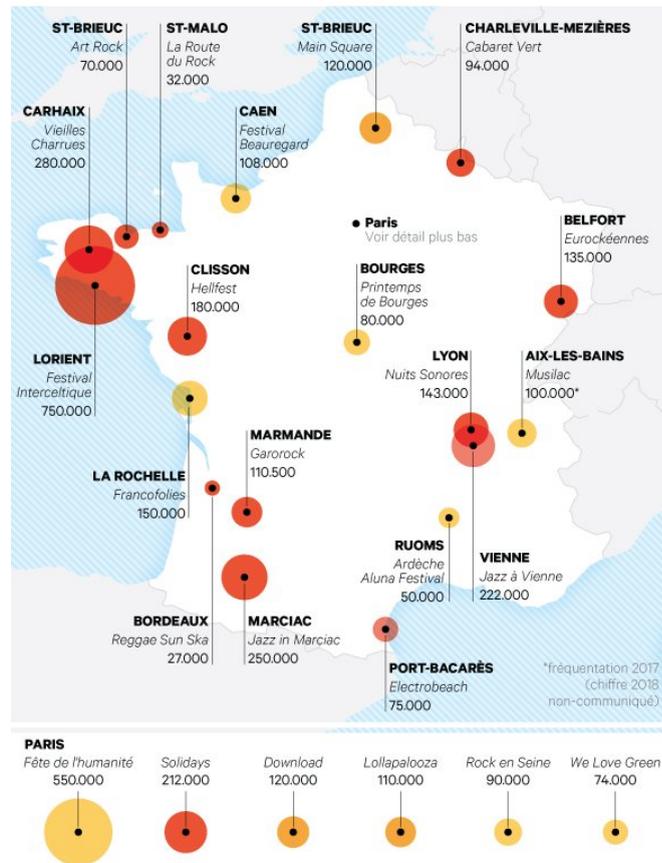


Fig. 32 - Example of pin map. Source: <https://twitter.com/EchosGraphiques/status/1034786035438374912/photo/1>; <https://start.lesechos.fr/societe/culture-tendances/se-reperer-dans-la-galaxie-des-festivals-en-une-infographie-117741>



Fig. 33 - Example of pin map (and detail). Source: <https://www.charmingpuglia.com/en/best-beaches-in-puglia.php>

- **Connection Map** displays networks combined with geographical data for visualising flows and any kind of connections between different locations. It can be combined with thickness of lines, colours or patterns, or other charts to add numeric values to the representation.
- It is different from flow map, since the connection map displays no quantitative values between the connections.

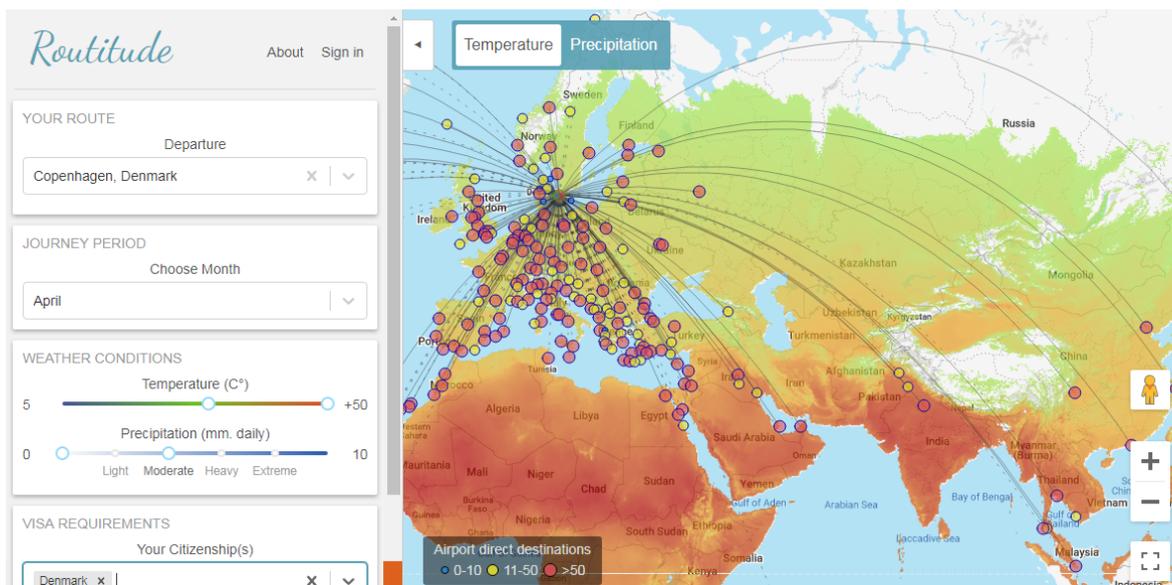


Fig. 34 - Example of connection map. Source: <https://www.anychart.com/blog/2019/08/09/top-data-visualizations-dataviz-weekly/>; <https://www.routitude.com/map>



Fig. 35 - Example of connection map. Source: <https://www.ricksteves.com/europe/spain/itinerary>

Timeline

- It is useful to display a list of events in chronological order to communicate time-related information for analysis, tell a story or view of history.
- It can work on a scale or simply displays a sequence of events either for analysis or to visually present a story or view of history.
- It can be combined with other visualization techniques, for example to show how quantitative data changes over time.

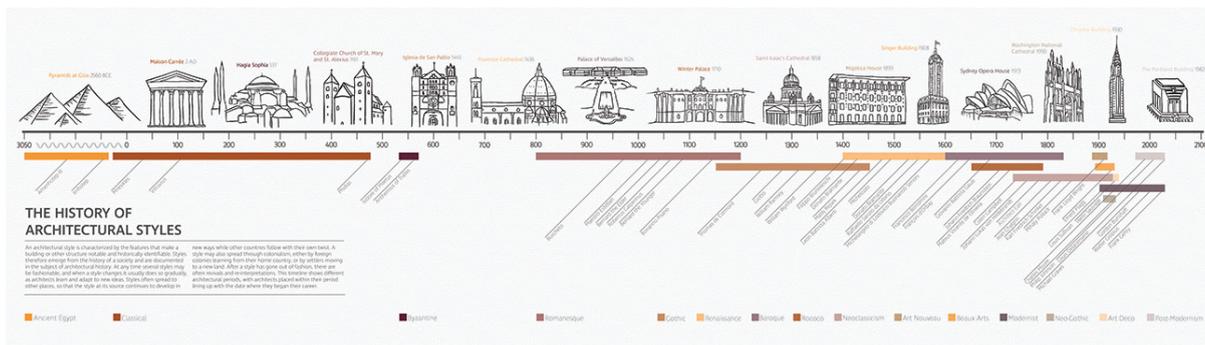


Fig. 36 - Example of timeline. Source: <https://www.behance.net/gallery/59219581/Architecture-History-Timeline>

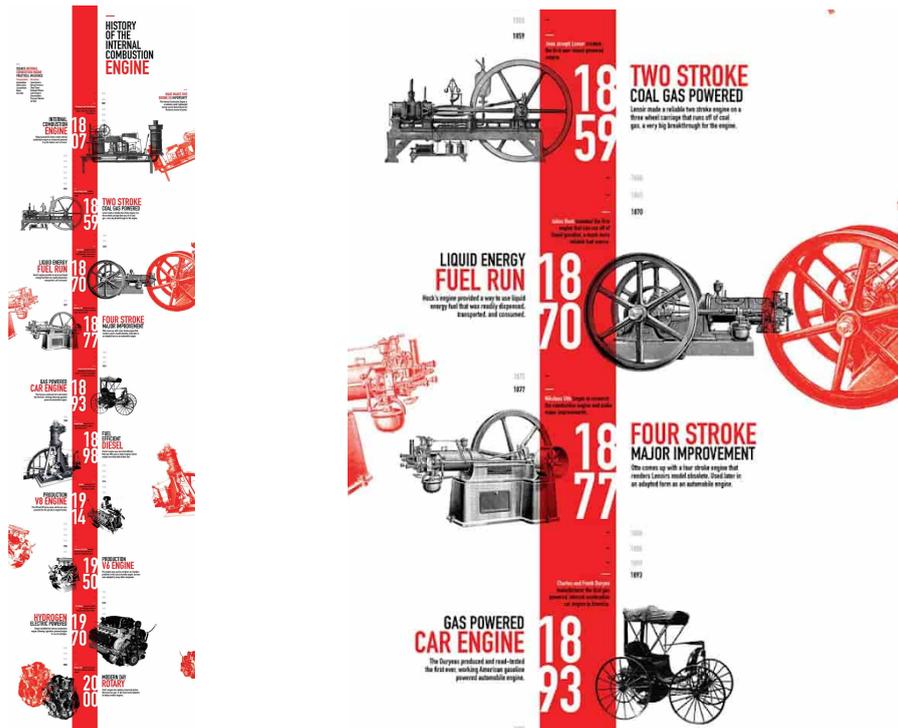


Fig. 37 - Example of timeline. Source:

<https://www.bluecompass.com/blog/best-infographic-designs-of-2015-that-are-sure-to-be-a-hit-in-2016-infographic>

Five decades of David Bowie's creative output

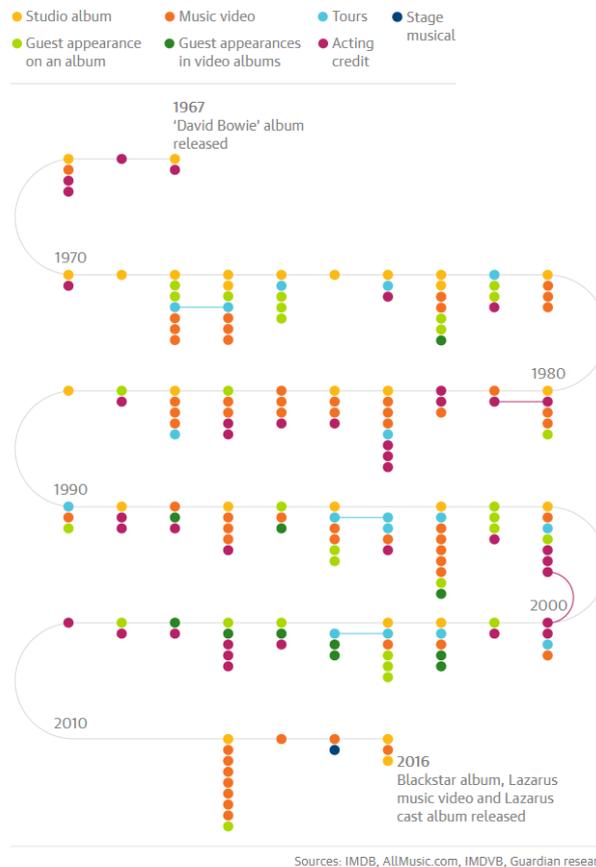


Fig. 38 - Example of timeline. Source:

<https://www.informationisbeautifulawards.com/showcase/1809-david-bowie-s-50-year-creative-output-at-a-glance>

<https://www.theguardian.com/music/datablog/2016/oct/22/david-bowies-50-year-creative-output-at-a-glance>

There are many other graphical techniques that you can use to encode numbers in pictures, including some very suitable for infographics, such as typography with icons (Figures 39 and 40), especially useful to help understanding the text, and illustration diagrams (Figures 41 and 42).

Be aware that in infographics the use of typography, especially for numbers, is better suited for representing standalone information (a number is frequently used to emphasize a single data) and it should be well visible and readable.

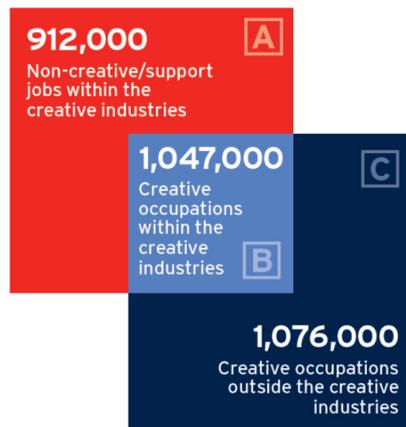


Fig. 39 - Example of text and numbers. Source:

<https://www.gov.uk/government/publications/creative-industries-sector-deal/creative-industries-sector-deal.html>

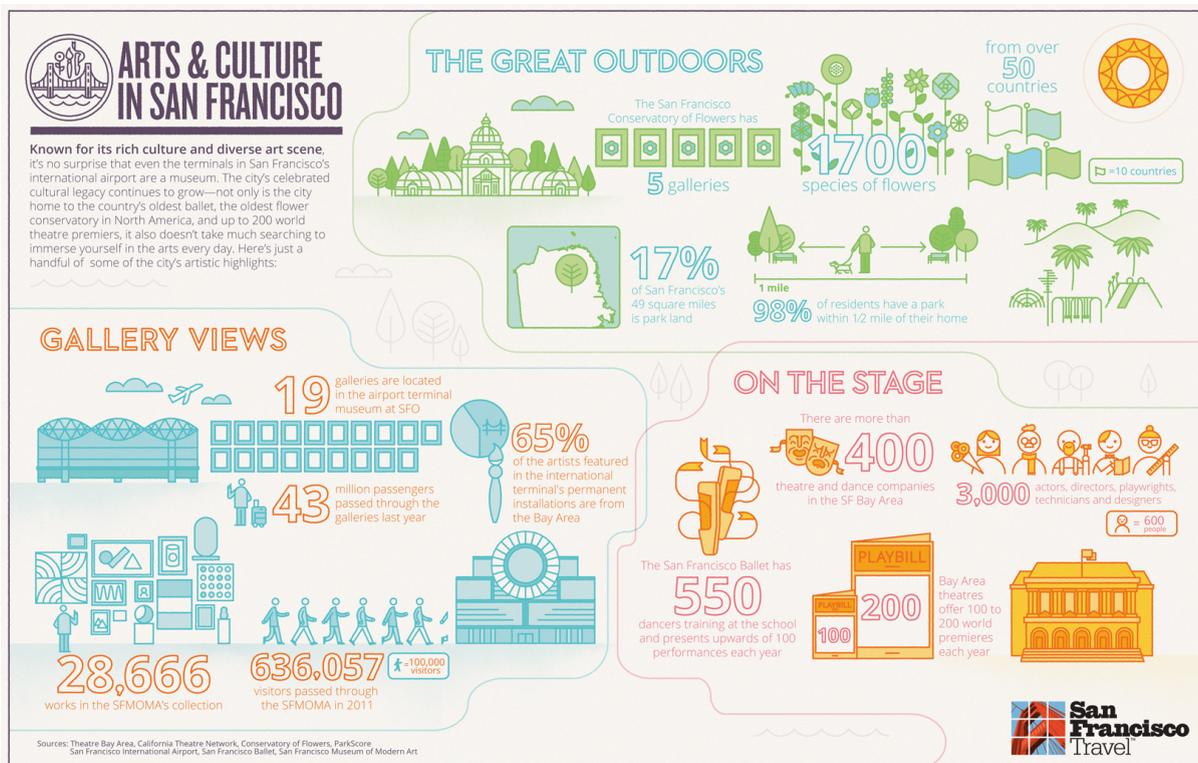


Fig. 40 - Example of text, numbers, and icons. Source:

<https://www.good.is/infographics/infographic-arts-and-culture-in-san-francisco1>

Illustration diagrams can be used to explain concepts or describe something. They are usually accompanied by notes, labels, or a legend.



Fig. 41 - Example of illustration diagram. Source: <https://www.iberdrola.com/culture/sustainable-museum>

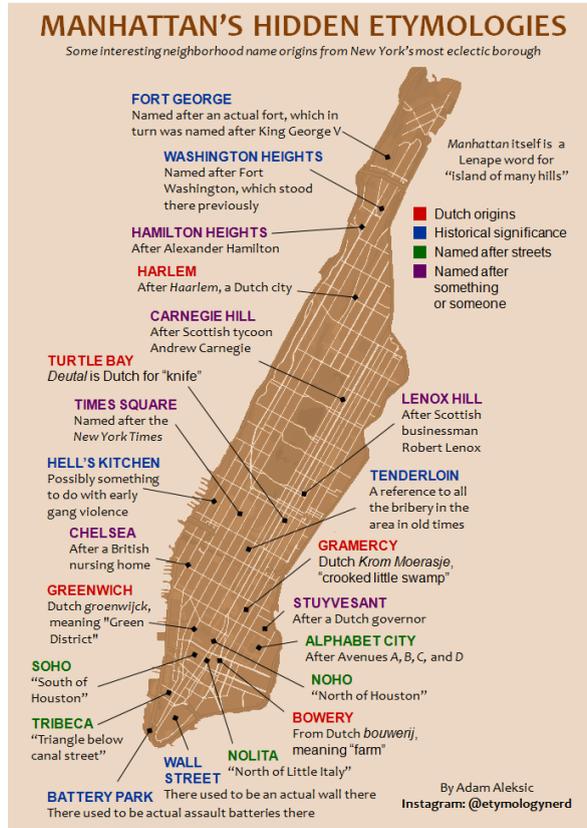


Fig. 42 - Example of illustration diagram. Source: <https://vividmaps.com/manhattans-hidden-etymologies/>

Even the use of metaphors is widespread for representing information and data in a visual way, since they translate complex and abstract models in familiar and easy-to-understand symbols that are pleasant to see, too.

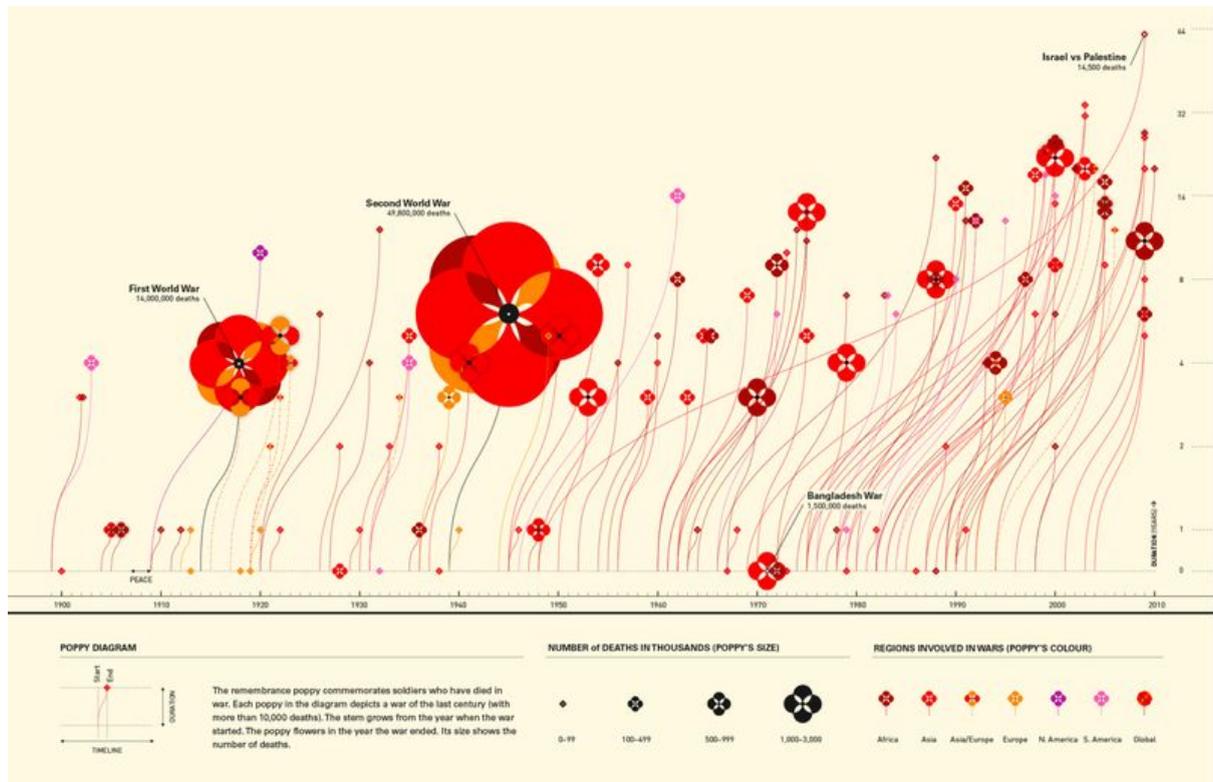


Fig. 43 - Example of the use of a metaphor. Source: <http://www.poppyfield.org/>; <https://centralillustration.com/illustration/valentina-defilippo#portfolio-10>

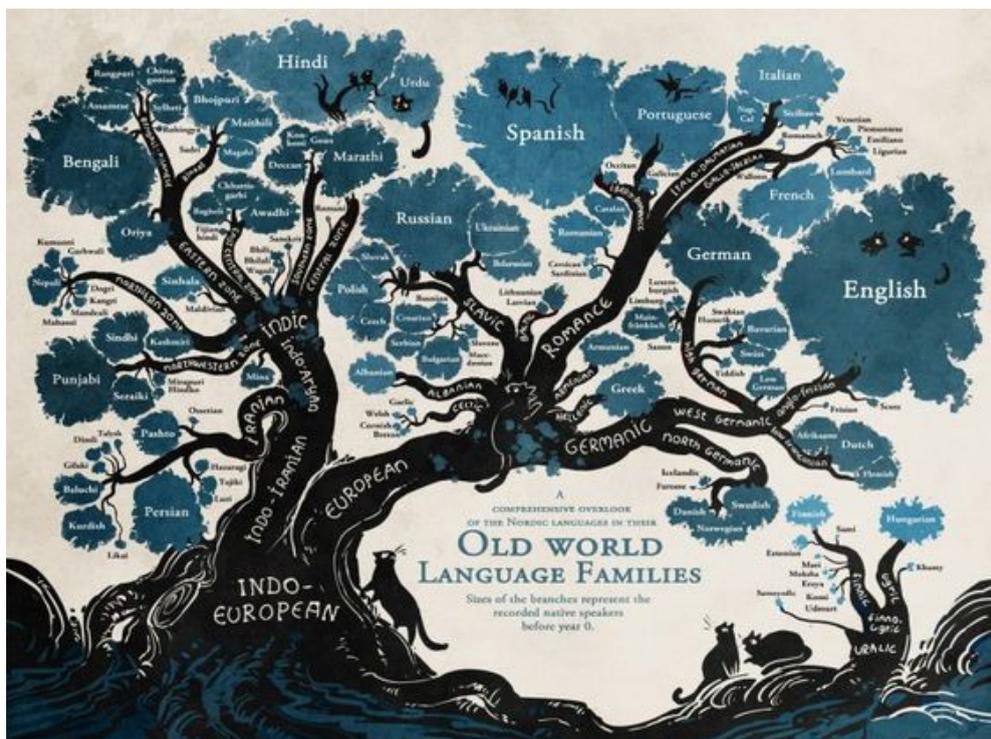


Fig. 44 - Example of the use of a metaphor. Source: <https://audreydriscoll.com/2015/08/15/a-language-tree/>

A few examples of errors to avoid in visualizing data and information

Following a few examples of infographics and data visualizations where the visualization techniques are applied in a misleading and incorrect way.

In Figure 45, for example, there are too many categories represented in the pie chart and the same colours are repeated for different slices without any relation. Moreover, no real value is given to understand the numerical proportion of the data. Indeed, the chart can not be read or even seen in its littlest parts.

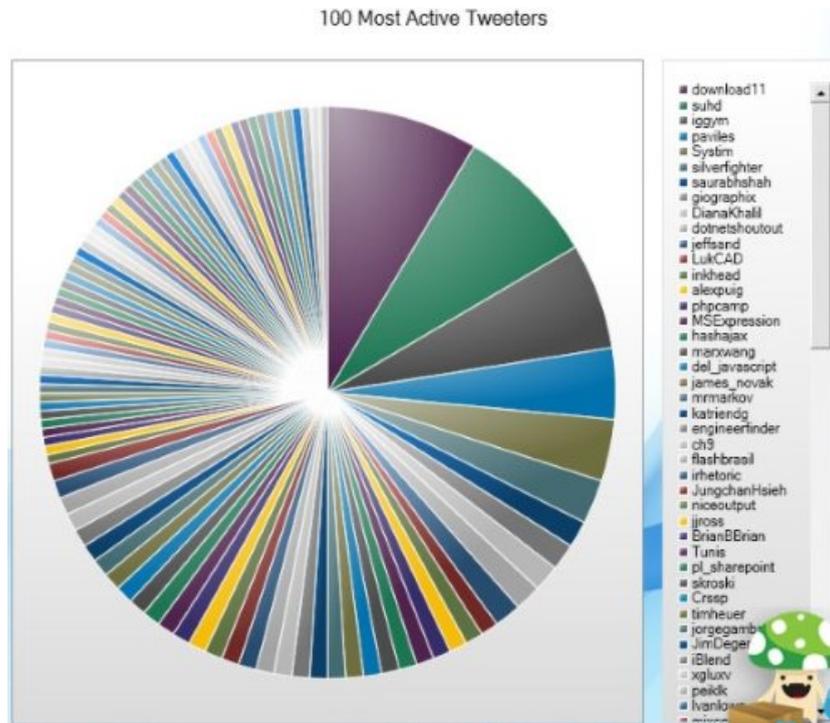


Fig. 45 - Example of a bad use of the pie chart. Source: <http://livingqlikview.com/the-9-worst-data-visualizations-ever-created/>

In Figure 46, it is hard to match icons and percentages at a glance. Moreover, the two half coloured icons create some confusion and lack of clarity on how they should be read.

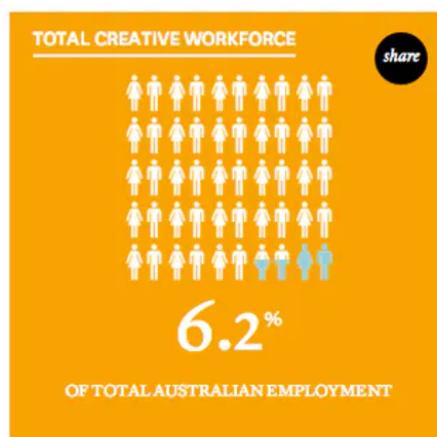


Fig. 46 - Example of a bad use of the pictorial fraction chart. Source: <https://theconversation.com/we-can-rebalance-australias-economy-with-creative-industries-23458>

In Figure 47, another pictorial fraction charts with a lot of mistakes both in data visualization and visual design. One above all: the sum of the percentages shown is greater than the whole (243% instead of 100%), as the graph requires! Moreover, there is no relation between the shape used (a walking man) and the topic of the survey.

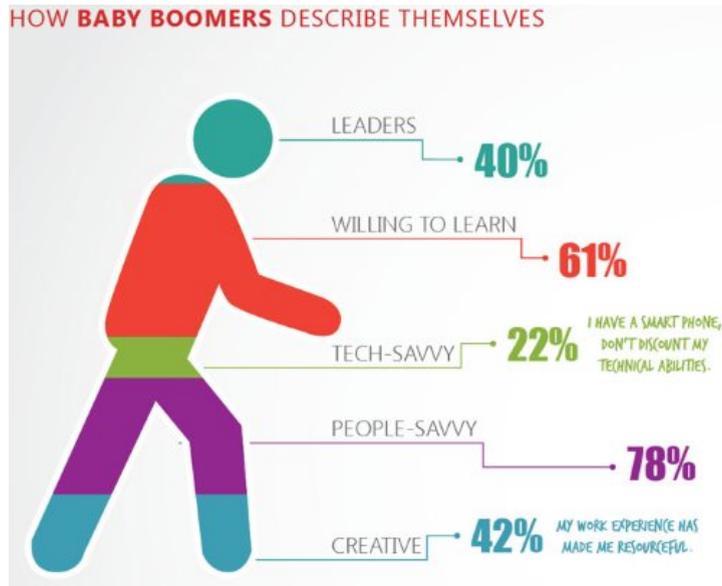


Fig.47 - Example of a bad use of the pictorial fraction chart. Source: <http://livingqlikview.com/the-9-worst-data-visualizations-ever-created/>

In Figure 48, there are no values for Y-axis, then we can not know the real numbers of market share increase/decrease.



Fig.48 - Example of a bad use of the line graph. Source: <https://badvisualisations.tumblr.com/>

In the end, being aware of different ways to graphically represent data and information, allows to choose which type best suits a certain need, taking into account that graphs, and statistics in general, can be misleading if incorrectly used.

Conclusions

The knowledge learned in this lesson will allow students to learn how to encode data and information in pictures, in order to correctly visualize and present a topic when creating an infographic.