Best Practices in Data Visualization

Semester Project

Anouk Paulino
supervised by Johan Berdat

Spring semester 2023

Contents

Introduction

Data visualization is an important aspect of scientific research and analysis. Its effectiveness lies not only in its visual appeal but also in its ability to communicate complex information in a clear and concise manner.

As societies developed and data became more complex, we began to use increasingly sophisticated visualization techniques. In the 20th century, advancements in technology and computing power led to the creation of new tools serving this purpose, which even enable interactive and animated graphics. Today, data visualization has become an essential tool for conveying information at first sight, with applications in fields such as science, marketing, and journalism.

Throughout this project, we have investigated the best practices to adopt when making visualizations. We have also applied those pieces of advice to practical examples.

Best Practices

When designing a chart to convey information, we must therefore pay attention to different mechanisms, innate and acquired, in order for the result to be clear to our audience. We will see that ignoring those mechanisms can quickly lead to misinterpretation.

This part has been heavily inspired by Scott Berinato's book [1], which reviews how we perceive charts as human beings, what makes a good design, and what does not.

Human perception

We perceive our environment and, by extension, images, in a very specific way. There are some common grounds to each and everyone of us, but our culture, education, and personal experience also have their impact.

Innate

People's attention is often drawn to specific elements within a visualization, such as labels, titles, legends, and data points [2].

The visual hierarchy of a visualization can greatly affect how people interpret and understand the data. People tend to look at larger and more prominent elements first, such as titles and legends (useful to get the context of the graph), before moving on to smaller details like data points and labels [3]. Moreover, clean and simple visualisations are pleasing to the eye and make the reader feel confident about the conveyed message and their ability to understand it. In contrast, cluttered and overly complex visualizations can be overwhelming and difficult to interpret. It can confuse the viewer and obscure important trends and patterns.

There are different levels of information processing that happen when we look at a chart. The "blurry level" or "fast thinking" is an almost subconscious level that allows us to quickly pick out patterns. This level is important for identifying general trends. It occurs in dozens of milliseconds after exposure and can be described as reactive, automatic, instinctive and emotional. In comparison, the "high road" or "slow thinking" level of processing is a more deliberate parsing of information. This level allows us to identify specific data points and to make precise comparisons. It is more analytical, effortful and logical.

A great example of how this can impact visualizations are the Gestalt principles[4, 5]. Here are 6 of them (the community does not seem to agree on a specific number [6, 7]) and all are completely subconscious:

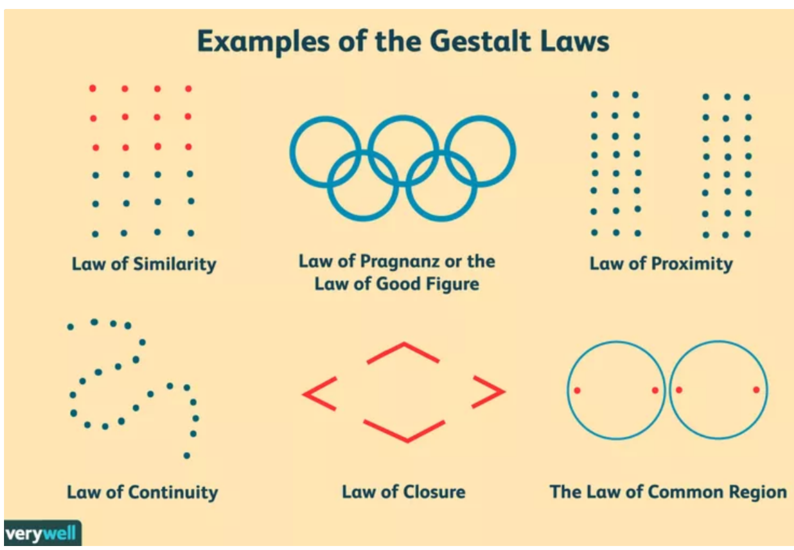


Figure 1: Illustration of the Gestalt Principles

- Similarity: if some elements have some common attribute, they are considered as part of a same group
- Pragnanz / Good Figure: unless the image is ambiguous, our brain recognizes the foreground first

- Proximity: if some elements are close to each other, they are considered as part of a same group, as opposed to points further away
- Continuity: if some elements form a continuous or are linked together, it is assumed there is a continuity between their values or attributes
- Closure: if some shapes are incomplete, our brain automatically fills them in
- Common Region: if some elements are placed in a region separated from other elements, they are considered as part of a same group

Hereby lies **saliency**. In the context of data visualization, saliency refers to the visual elements that draw the viewer's attention and stand out from the rest of the image. It can be achieved through various visual design attributes such as color, size, contrast, and shape. By emphasizing certain parts of the data, it can help guide the viewer's attention and focus on the most important information.

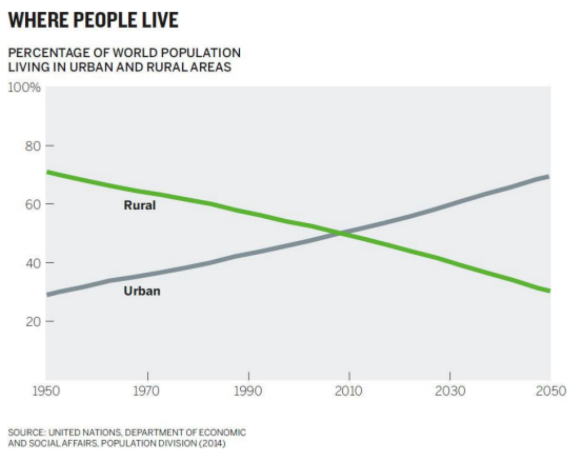


Figure 2: Example of obvious saliency
When first seeing this graph, our brain immediately notices the meeting point of the two lines.

Saliency helps make the information more understandable and memorable. By highlighting the most important ideas, the reader is better able to grasp the goal of the chart and remember it. In contrast, if all the visual elements are of equal weight and importance, the viewer may struggle to understand the main points.

However, it is important to ensure that the salient elements do not overwhelm the viewer, or distract from the main message of the visualization. Effective data visualization strikes a balance between saliency and coherence.

Consequently, we naturally seek meaning and try to make connections in everything we see.

This can lead to us over-analysing and jumping to conclusions based on previous experiences. This is called the confirmation bias.

Several other biases may be encountered and should be kept in mind when designing a graph [8]. Here are some of the most relevant:

- Availability bias: the tendency to rely on information that is easily available in memory, rather than seeking out additional information or considering less salient but equally important factors.
- Anchoring bias: the tendency to rely too heavily on the first piece of information encountered when making a decision, and to be insufficiently influenced by subsequent information. This is also heavily linked to saliency.
- Framing bias: the tendency to be influenced by the way information is presented, such as the order of presentation or the wording of a question. For instance, removing numerical information from a graph can also significantly influence decision-making and cause biases in the interpretation of the data.

Acquired

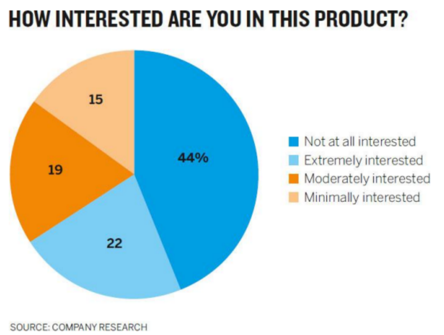
As cultural and educated individuals, we also rely on conventions and metaphors. Here are some examples of what can universally be assumed:

- like colors mean like items;
- color saturation hints value, lighter values meaning less dense or emptier;
- categories are arranged and plotted from one extreme to another;
- time is plotted from left to right;
- maps are displayed with the north "up".

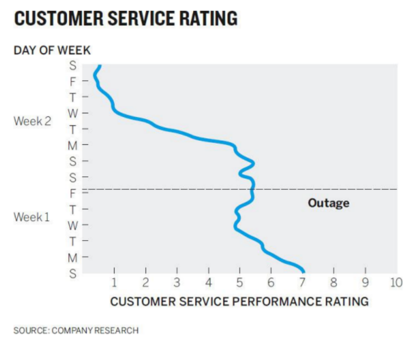
Some of those examples are tightly linked to the Gestalt Principles mentioned in the previous section, .

In addition, some other acquired beliefs may be influenced by our environment. Culture, education, personal experience can greatly affect our interpretation in several cases.

A well-known cultural bias is about the meaning of colors. For instance, in western countries white represents innocence and fertility, while in eastern countries it is symbolizes death and supernatural creatures. Colors also have a great impact on behavior and reaction to a graph as the color palette can impact the overall feeling of the visualization [9].



(a) different hues of the same color do not describe like patterns



(b) time is displayed from top to bottom

Figure 3: Examples of confusing graphs

Examples of poor cognition consideration

Taking the previous points into account, it is impossible to obtain an objective visualization, because our brains are wired to interpret what they see based on previously acquired knowledge. It means that a same chart can have as many interpretations as there are readers. Also the designer also being human, their work is necessarily subjective. The goal here is the become aware of those biases and try to reduce them as much as possible.

In the end, there is a blurred line between visual persuasion and visual dishonesty (c.f. this chart [10], where abortion and cancer rate are compared while not on the same scale). Other good examples are websites that use what are called dark patterns [11]. Those websites purposely may, for example, hide some type of information for the customer to give up before finding it. A well-known example is Amazon's website, which requires many confusing steps before allowing the user to delete their account.

Some practices are known to be misleading and must therefore be avoided:

- <u>Truncated y-axis:</u> the viewer will assume the axes are complete and will not necessarily pay attention to this change of representation.
- <u>Double y-axis:</u> this practice forces the viewer to compare two things that can not necessarily be compared (i.e. different orders of magnitude or even different units).
- <u>Geographical representation:</u> when the data is not relevant in terms of geography and is it represented as a map, it can lead to a real deception of the viewer.

(a) truncated y-axis *(b) full y-axis*

Figure 4: Both graph displaying the same data may lead to different interpretations
Notice how the left chart emphasizes patterns that are not necessarily visible on the right-hand side. If the viewer does not notice the truncated axis, they might misinterpret the shape of the line as a more dramatic event than it actually is.

Figure 5: Example of double y-axis misleading, displaying the price of gold and silver on the same graph
Both axes do not display comparable orders of magnitude. When reading the graph to quickly one could assume that there is a correlation, even though the data does not permit to establish any link.

Figure 6: Map of the votes for the scottish referendum for independence.
In reality, while less than 5% of land mass voted "yes", 38% of voters actually voted "yes".

We notice that there are many practices to avoid. To sum up, here is a list of honesty challenging (and interesting) questions to ask oneself to make sure the produced visualization is not unintentionally biased:

- Does my chart make it easier to see the idea, or is it actively changing it?
- If it is changing the idea, does the new idea contradict or fight with the one depicted in the less persuasive chart?
- Does eliminating information hide something that would rightfully challenge the idea I am showing?
- Would I feel duped if someone else presented me with a chart like this?

Having thought those questions through, we can assume that we are off to creating an honest and useful data visualization.

How to actually do it

We have reviewed how we perceive things as human beings, depending on our culture and education. Knowing the audience of a chart is therefore compulsory to make sure the visualization will achieve its goals.

We will know focus on how to implement this knowledge.

The importance of brainstorming

One can separate charts into four categories.

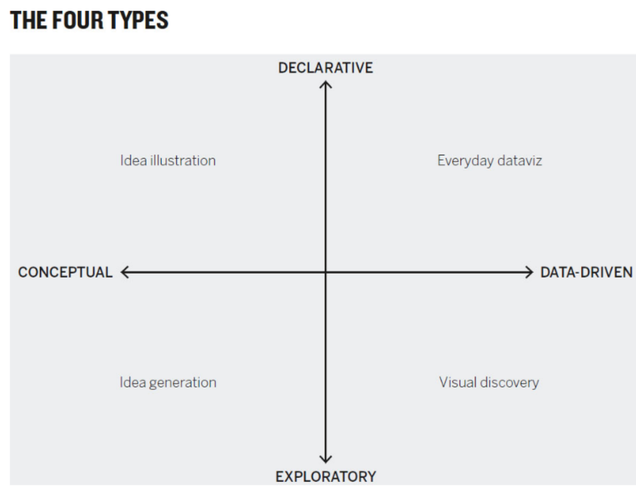


Figure 7: The four types of visualizations, according to Scott Berinato

Each of these categories represents a purpose for the chart:

- Everyday dataviz: express a given idea. The designer knows how the data is going to talk when making the graph. It is the most common type of graph.
- Visual discovery: confirm hypotheses or feelings about the data. One does not necessarily know what they are looking for in the first place when doing visual discovery.
  - *Visual confirmation:* declarative.
  - *Visual exploration:* use raw data to see patterns and trends emerge.
- Idea illustration: conceptually visualize ideas. It is the dataless version of everyday dataviz.
- Idea generation: sketch ideas, most often during in brainstorming sessions. It does not necessarily end up as a proper graph. It is used for reflection and is the dataless version of visual discovery.

Since our work is based on data, we will focus on the first two.

Let us actually get started with making the chart. As described above, we have to take into account the future graph's public in order to lessen the possible biases. To make a chart as useful and functional as possible, we have to go through the four following steps. Each step come with an estimated percentage of the total time spent on designing a chart.

Preparation (5%)

As a starting point, let us put the data aside. We will focus on the goal of the chart. A basic data visualization requires:

- a purpose (a story to tell);
- knowing its audience;
- knowing the setting in which it will be used in;
- knowing which of the four types it is;
- finding the balance between context and design.

Having those elements in mind helps determine which way to go.

Talk and Listen (15%)

We shall now talk the project over with a colleague, either external to the project if you want to set grounds for further thoughts, or internal to refine the ideas. This step is primordial.

- What am I working on?
- What am I trying to show (declarative)/prove (confirmatory)/learn (exploratory)?
- Why?

If you find yourself unable to answer, then the concept has not been properly thought through and requires some refining.

Sketch (55%)

Match keywords to approaches [2]:

- Line charts compare values overtime
- Pie charts cannot display time evolution, they rather show parts of a whole
- Bar charts compare quantitative data from different categories
- Scatter plots display relationships between different aspects of the data

However do not limit yourself to the above. Trying out new ways of representing the data can lead to a better understand the conveyed message.

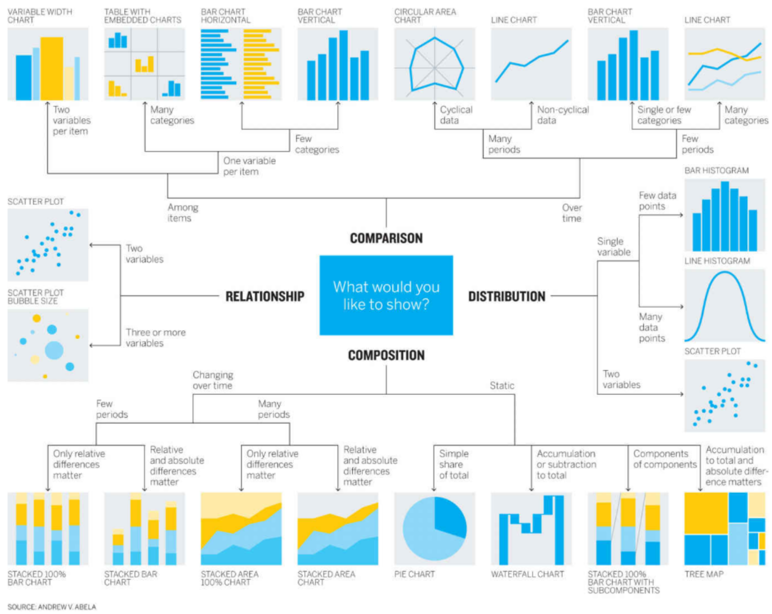


Figure 8: Andrew Abela's chart of which type of graph to use

Most importantly, we have to think of your audience and make the chart inclusive. For instance, colorblindness is rather common. Several tools exist to work around it (e.g. on the Adobe Suite tools, or using alternative color palettes). Also, since color is so important, take some time to thoroughly think through the overall look of the chart:

- High contrast makes a chart easy to read
- Pleasing colors make it less painful to look at
- Complementary colors help with making links between different elements on the image

We also shall make wise use of texts and icons to label our graph. Such additions may help the reader understand the chart, but may also confuse them if poorly arranged. We mention this in greater detail below.

Prototype (25%)

Progressively incorporate the real data into the graph. At this step, we can start using some online tools to produce our graphs.

An important thing to note is that the above steps are overlapping and the overall process is iterative. Trying to get a chart done in one shot will most likely give out poor results. Going back and forth between the steps is important to get feedback and think outside the box.

The execution

Our brain consumes lots of energy for visual recognition [2]. To increase clarity and therefore be more straight-forward, here are the main concerns to have in mind:

- Hone the main idea: make wise use of colors, contrast, highlights, etc. Do not overuse those elements. For clarity, a title and subtitle should differ in no more than two attributes. [12]

- Make the design easy to identify: simple designs are more comprehensible than overcrowded ones [13]. Deleting unnecessary elements can be rather hard when having spent a lot of time working on a chart. The following guide may help make it simpler.

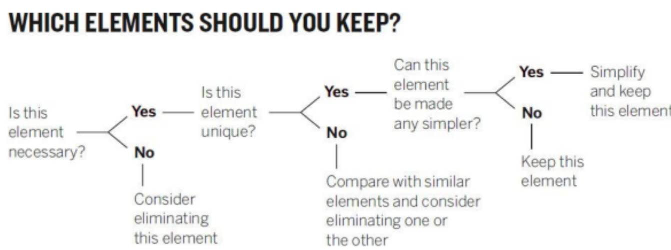


Figure 9: Guide to keeping only necessary elements

  This step is all the more important when one thinks everything in their visual is mandatory. It is normal to be reluctant to removing any part of a graph, but it should not prevent the chart from being clear.

- Minimize eye movements: some patterns are easier on the eyes. For instance, in website design, the F-pattern feels rather natural [2]. In charts however there is no such contract. The goal is to minimize eye movements to make the chart pleasing to the eye.

  The more complex the pattern, the more lost the viewer will be.

- Indicate what is going on: information that is not meant to be communicated can still impact interpretation. If the chart is ambiguous, another message may be conveyed instead of the original one.

As mentioned earlier, another important part about the design is colors. Colors allow us to distinguish elements. There is a huge bias in color interpretation based on culture. It is therefore necessary to pay great attention to the color palette. Also reduce the amount of different colors

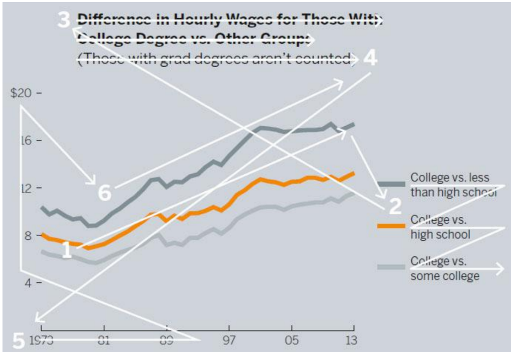


Figure 10: Example of typical eye movements when reading a chart (1 being the first step)

used when applicable, it makes the chart less overwhelming. Prefer using different hues (gray scaling) to inform about hierarchy.

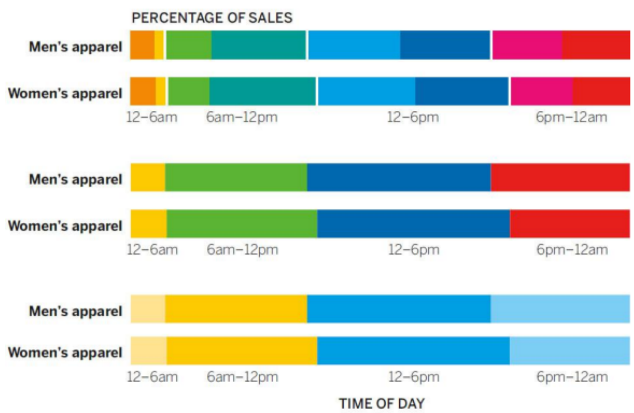


Figure 11: Comparison of several color palettes displaying sales with respect to gender and time of day. Notice how the first one's palette makes the chart look denser compared to the other ones. It contains too many different tones. The last one successfully distinguishes the morning and afternoon sales for each gender with different hues of the same color for early and late day periods, all the while remaining simple.

There exist several online tools to come up with a meaningful color palette (e.g. [14, 15]). Those tools suggest complementary colors to the ones already chosen.

As a(n) (almost) final step, it is always interesting to test one's design on a sample public, to get a completely new point of view. This step allows us to spot the eventual weaknesses of the chart.

Applications

We focused on four aspects that we wanted to investigate: a data set that showed various emissions related to food production, another set with food composition, one with geodata, a paper about the production of sugar, and finally a data set of the products sold at Coop supermarkets. We wanted to see what we could extract from those data sets or papers, having reviewed the previous section's theory.

Food

Food is a matter that concerns us all. We wanted to dive into this subject and try to answer some questions the average consumer could have.

Emissions

As a first approach, we decided to look into emissions related to food production. The idea was to get a better understanding of which agriculture sectors require the most resources (we were thinking of land, water, greenhouse gazes a.k.a GHG).

We used the Our World In Data data sets [16], which provided the following insights:

- land use (Swiss land use as a whole and per product);
- water use (per product, and there also exists a scarcity-weighted version, that balances the use of water with the amount of available water in the first place);
- carbon emissions (per product and stage of production);
- GHG emissions;
- eutrophication (i.e. how impacted is the groundwater table in said region, per product).

A few questions derived from those data sets:

- Which products consume the most resources (with respect to the ones cited above)?
- How to reduce food-related emissions of the average (Swiss or not) consumer?
- Which stages and/or products should we focus on to get the greatest impact on emission reduction with minimal modifications to the actual system?

We investigated the data in a visual exploration fashion. However, we did not manage to find proper answers to the above questions. The data set did not provide enough data for us to interpret.

Composition

Similarly, we tried to gain some insights into how different products were related to one another with respect to nutritional values. To compare such data, we tried out two reductions, PCA [17] (linear) and t-SNE [18] (non-linear). We wanted to see if we could get those reductions to put in light the product categories with only their nutritional values.

(a) PCA: no apparent group

(b) t-SNE: groups match category

Figure 12: Reductions mapped with category as color

In this example, the linear reduction has been less efficient at highlighting categories than the non-linear one. On the t-SNE mapping, we can clearly distinguish groups of the same color, showing that products of a same category have similar nutritional values, as we expected.

Other than that, those scatter plots are not really satisfactory in terms of the aforementioned good habits. There is no proper legend to them. Also, the colors are not one-time uses, i.e. some colors are used for several categories, which is clearly misleading as it goes against the Law of Similarity.

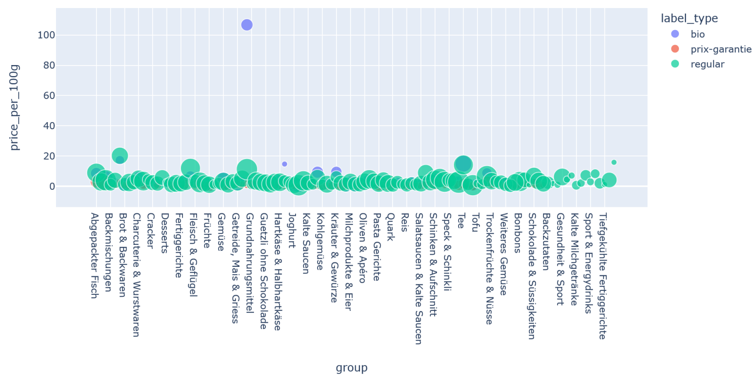
Coop products analysis

To play around a little more, we decided to get the data of the more than 10'000 products sold at Coop [19]. The data set provided some useful information about the brand of the product, as well as the weight, price, composition, etc.

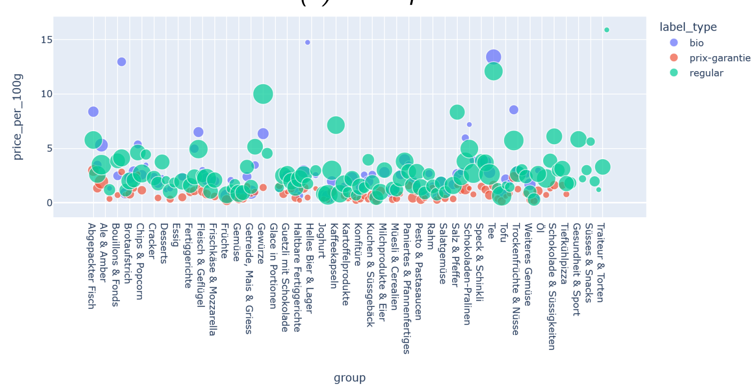
We asked ourselves a couple of questions:

- Are the products which are labeled "Prix Garanti" (supposedly first price) really more affordable than their regular counterparts?
- Are the products which are labeled "Organic" more expensive than their regular counterparts?

We ended up with some interesting graphs.



(a) mean price



(b) median price

Figure 13: Scatter plots displaying the price per 100g of product, per category and label (organic, Prix Garanti, none)

With such charts, we see how the number of products impacts both results.

On a visual level, they provide with both qualitative and quantitative aspects. The "bubbles", being of different colors, allow the viewer to distinguish each group of labels, per category. Additionally, the Jupyter Notebook versions have a hover option which displays some complementary information about a specific "bubble".

These graphs still lack of a scale for the size of the "bubbles". There is no numerical values associated with this metric. This can mislead the reader, since they may assume some wrong information.

Geodata

An interesting point that we wanted to look into was geodata. It is much graphical and allows for a wide variety of visual interpretation.

We have chosen to use a basic set of population density in Switzerland (with the highest precision possible), among the many provided by Kontur [20]. In this set, the data is organised in hexagons. It appears to be the most optimal way of representing a surface. This way of organising geodata has been imagined by Uber to analyze traffic [21]. We decided to turn this data set into a *qualitative* map. Therefore, we have tried several ways of displaying it.

Kepler.gl

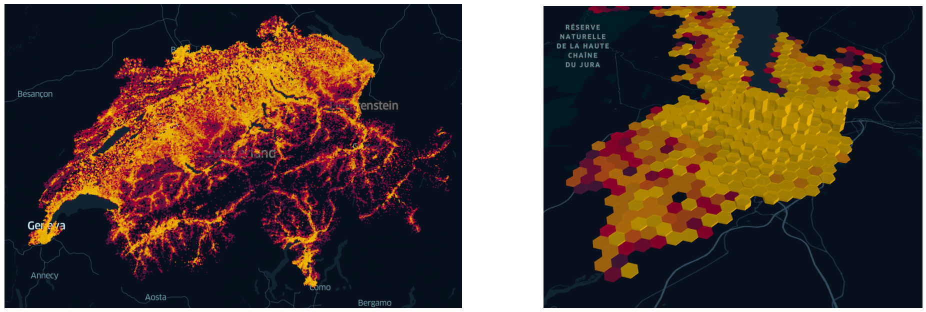
Starting at the source, we investigated the Kepler.gl tool [22] also developped by Uber to make use of the hexmaps.

The online interface allows the user to upload a data set and tweak some parameters to alter the look of the resulting map. The tool is flexible and allow the data to be of several kind, including hex points.

As a result, Kepler.gl is practical if one wants to produce a quick visualization. Importing the data is really straight-forward, and changing some parameters is equally simple. However, the tool seems a bit limited when using hex data sets. We could only change the height of the spikes and the color palette. For instance, it did not permit us to modify the look of the spikes. However, it appears that the Python extension of Kepler.gl is more customizable [23], but we did not have the time to properly look into it.

Rayshader and Rayrender (RStudio)

Many map visualizations appear to be made using R packages, Rayshader [24] and Rayrender [25], so we decided to dig into those practices.



(a) default settings when inserting the dat

(b) focus on Geneva, 3D visualization with height proportional to population density

Figure 14: Visuals of Kontur data using the Kepler.gl tool

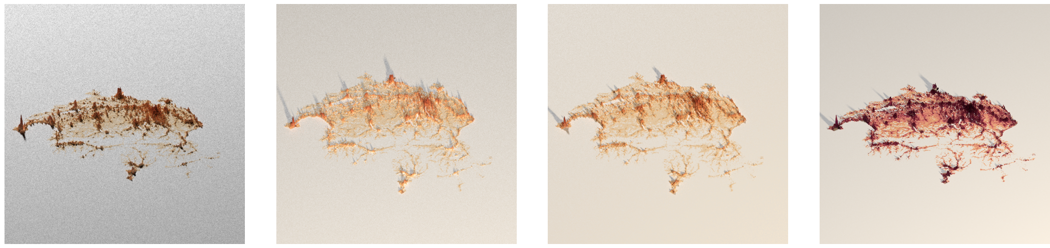
Rayshader is a tool that allows us to control the light sources within the 3D render. Rayrender is used by Rayshader to define the way an object interacts with light. Combined together these packages allow us to create complex 3D visualizations. We decided to go for RStudio as an IDE, since it is widely used and provides a useful visualization interface.

Having encountered a tutorial that rendered a map of Florida in RStudio (which also used a Kontur data set [26]), we decided to implement the same process to our Swiss data. The author of the tutorial also wrote an article about this method [27], allowing us to go back and forth between both media, which prove to be complementary. Since we had to adapt to the Swiss map, we looked for a hexmap displaying the borders of Switzerland [28]. This map would serve as a ground map for the data to be displayed.

We tried out different camera angles and lights. Changing the material of the spikes (i.e. change the way it interacts with light using Rayrender) also impacts the visual and its clarity. As a matter of example, we have tried to render glass-like spikes, but it would make the chart all the more confusing because it became difficult to distinguish the topography of the map.

R also provides a package called Magick [29] that is mostly used to annotate images. We decided to use it to complete the freshly created map with some of the city names to make the chart more readable.

Since the elements (text, segment, etc.) are positioned programmatically using this package, we spent a lot of time determining the correct coordinates for each of city name. It was an iterative process, i.e. we had to render the annotated image, make a guess of how many pixels were missing in each direction, change the coordinates accordingly, and repeat until we were satisfied with the result. However, We believe that we might have not used the tool to its full potential. Magick might allow for automatically determining the local coordinates of the different elements with some parameters (say camera angle, size of the image, actual



(a) default light sources *(b) two light sources* *(c) two light sources but with a different camera angle* *(d) glass-like material for the spikes*

Figure 15: Rayshader visuals with different light sources/reflection and camera angles

coordinates of Switzerland borders and cities, etc.). Nevertheless, using it as we did, we think it would have been more efficient, had we used a graphic design software to do annotate the image (e.g. Inkscape [30]).

In any case, the resulting visualization fulfills its purpose. Cities are displayed in a clear manner. However, this map is purely qualitative, as we had decided in the first place. It is impossible to deduce any of the data from this visual (no scale is present to help the viewer do so). Such maps could be combined with more practical ones generated with matplotlib [31], for instance, to get a better grasp of the data.

As a side note, if one only wanted to get a quick understanding of the data, such a map is not the most straight-forward way to achieve this goal. As mentioned above, a simple matplotlib generated map would have been more useful in that case.

Sugar

Finally, we were interested in the matter of sugar, sugar processing, sugar production, etc. Sugar production looked like a widely used process that involved a lot of waste. We wanted to see to what extent this belief was true.

We found a paper that described how sugar production wastes could be reused in several ways [32]. We completed the knowledge from the paper with a short documentary about sugar production [33].

We chose to turn this data into a poster, which would clearly conveyed the idea we wanted to give out: sugar processing has, in fact, a very poor yield. Following the aforementioned iterative brainstorming process, we spent a lot of time trying to figure out how to organize the poster, which information to add, and how to make it look appealing. Regarding the information

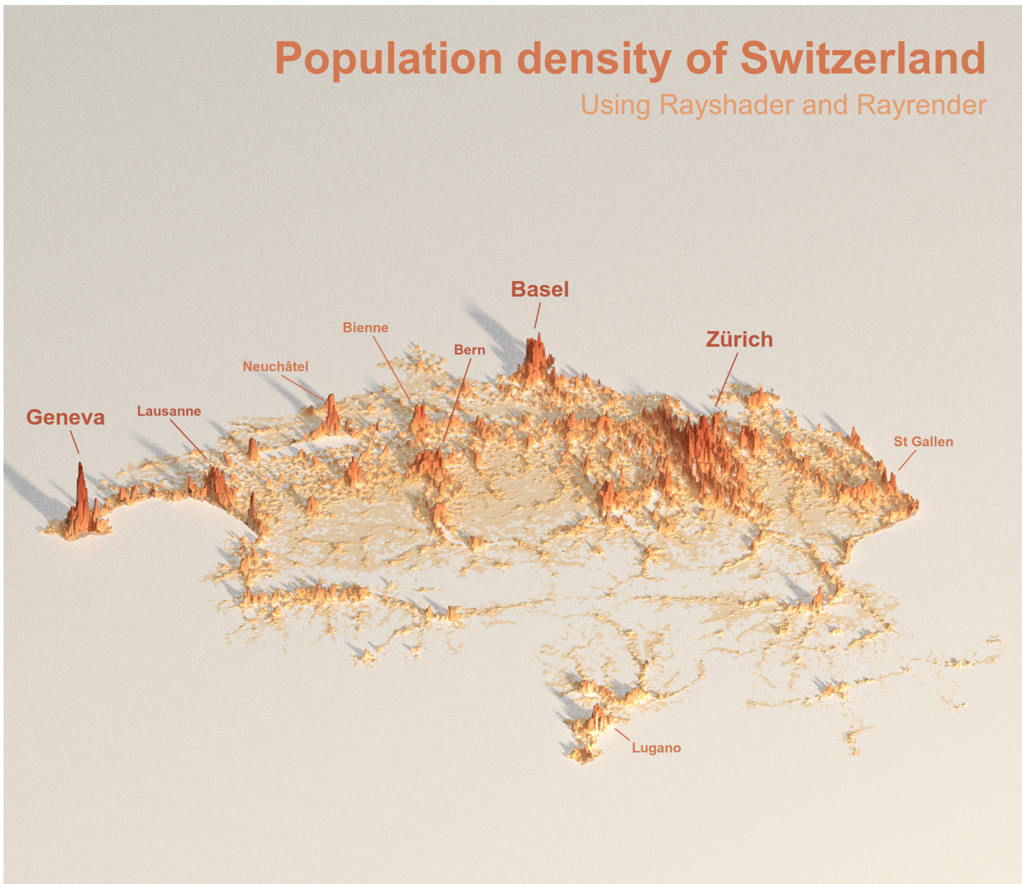


Figure 16: Map of population density in Switzerland, annotated.

encountered in the paper, we decided to go with a Sankey plot, which is of representing the flow of the different products.

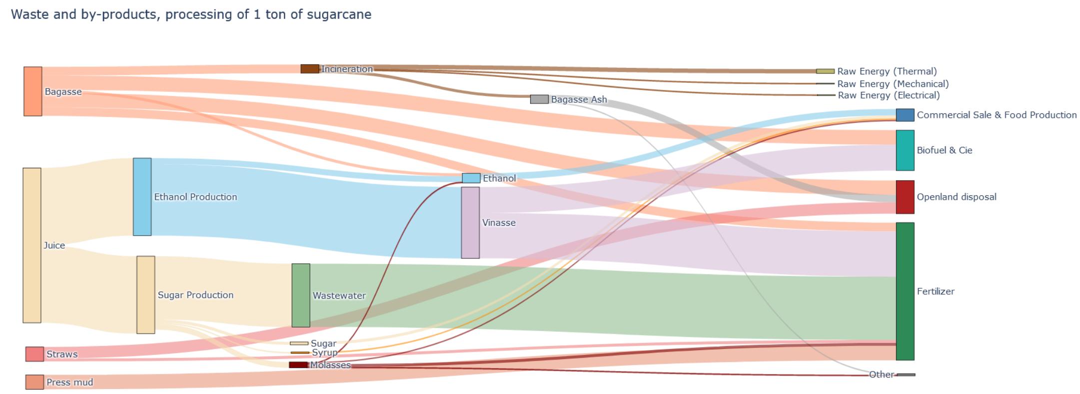


Figure 17: Sankey plot with the data found in [32]

The diagram is functional: it properly shows the proportion of the different products in each process. However, it is hardly readable and no clear point stands out. As a result, even though the proposals for waste recycling were interesting, we decided to focus on the actual generated wastes nowadays. After a few more iterations of the brainstorming process, we ended up with a much clearer plot, simplified and with selected harmonious colors for the user to be attracted to it.

Information is clearly delimited by the frames, titles stand out. The graph on the left is much simpler than the original Sankey plot and, therefore, makes the viewer feel like they can understand it at first sight.

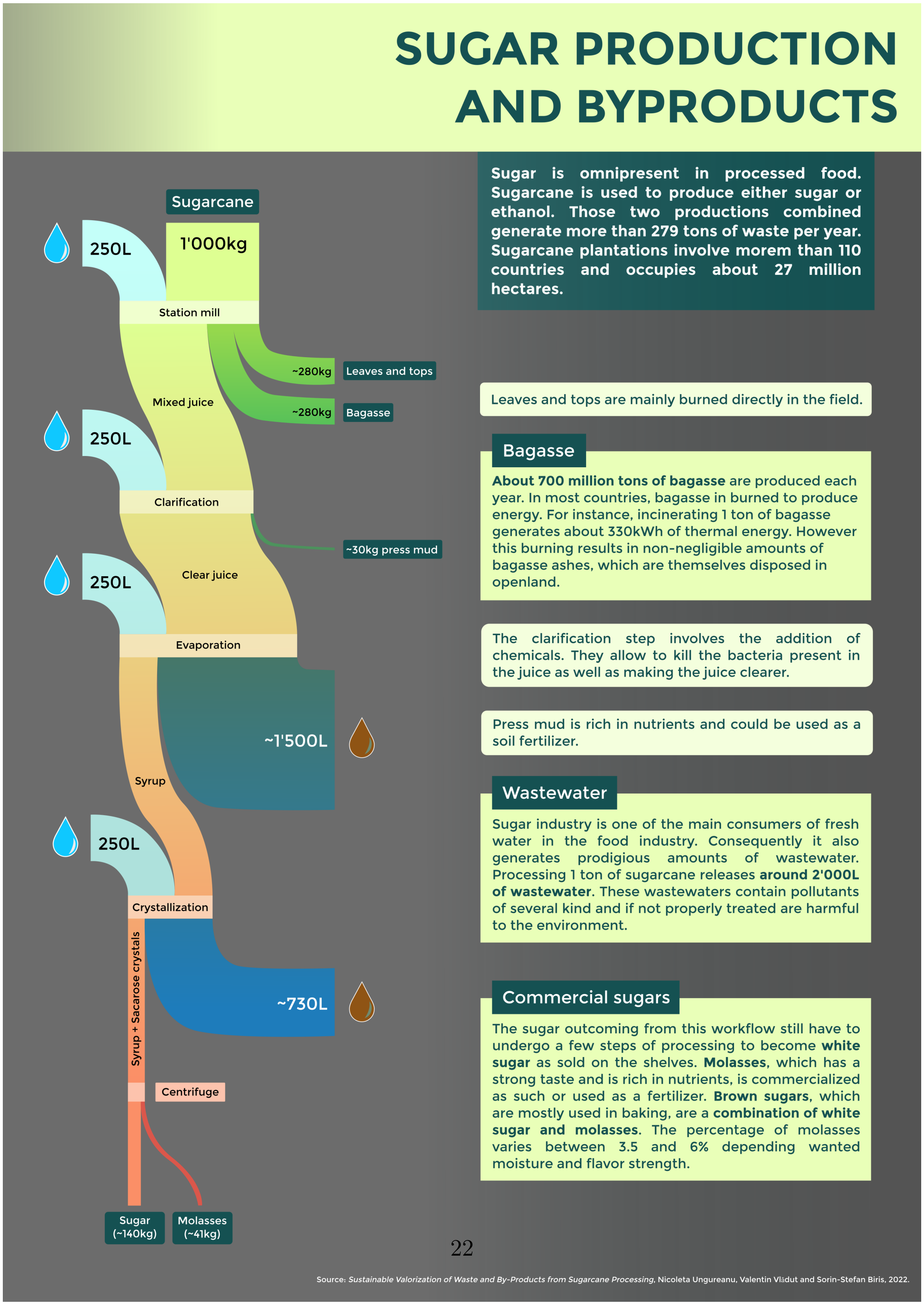


Figure 18: Sugar Production and Byproducts, realized with Inkscape

Conclusion

Throughout this project, we have gathered a good intuition of how the human brain reacts to given images and how we all are affected by our environment when it comes to analysing a graph. We have therefore adapted our way of thinking charts to meet those prerequisites and make the most out the data, so as to best convey our message.

The best practices discussed in this report emphasize the importance of simplicity, clarity, and context in visual representations of data.

The case studies reveal how poor design choices can hinder understanding and distort the intended message. These challenges highlight the need for adherence to those practices and careful consideration of the target audience and objectives.

In the age of information overload, data visualization has become an indispensable tool for transforming complex data into actionable insights. With those best practices in mind, organizations and individuals can effectively communicate their data and drive informed decision-making.

Further Reading

Here are some additional media that we did not find relevant enough to mention. They either did not fit scope of this project or were simply redundant with other sources. However we found them very enlightening on a personal level.

[34] A forum where people debate on visualizations practices.

[35] Yet another article with data viz tips.

[36] An article about color palettes in geovisualizations.

[37] A psychological approach to UI.

[38] How to avoid lies in data vizualisation.

[39] Recommendations of the best data visualization books in 2022.

[40] An article about bullet graphs.

[41] An AI to analyse how efficient a design is with respect to UX.

[42] A list of recommended tools which can be used to make visualizations.

[43] A wide panel of visualizations.

[44] Some templates for infographics.

Figure 19: Color meaning by culture

References

[1] Scott Berinato. *Good charts : the HBR guide to making smarter, more persuasive data visualizations*. 2016. ISBN: 978-1-63369-070-7.

[2] Various Toptal authors. *The Psychology of Design, A Collection of our Top Articles on Behavioral Design*.

[3] Yolanda Gomez Jose Vila. *Extracting business information from graphs: An eye tracking experiment*. 2015. URL: https://www.sciencedirect.com/science/article/pii/S0148296315004713.

[4] James Enns Christopher Healey. *Attention and Visual Memory in Visualization and Computer Graphics*. 2012. URL: https://ieeexplore.ieee.org/abstract/document/5963660.

[5] Kendra Cherry. *What Are the Gestalt Principles?* March 2023. URL: https://www.verywellmind.com/gestalt-laws-of-perceptual-organization-2795835.

[6] Maria de la Riva. *What Are The 5 Gestalt Principles?* May 2023. URL: https://careerfoundry.com/en/blog/ui-design/what-are-gestalt-principles/.

[7] User Testing. *7 Gestalt principles of visual perception*. URL: https://www.usertesting.com/resources/topics/gestalt-principles.

[8] Michel Fabre Jean-Pierre Caverni Jean-Francois Gonzalez. *Cognitive Bias*. Ed. by North Holland. 2004.

[9] Cameron Chapman. *Cause and Effect: Exploring Color Psychology*. URL: https://www.toptal.com/designers/ux/color-psychology.

[10] Timothy B. Lee. *Whatever you think of Planned Parenthood, this is a terrible and dishonest chart*. September 2015. URL: https://www.vox.com/2015/9/29/9417845/planned-parenthood-terrible-chart.

[11] Nerdwriter1. *How Dark Patterns Trick You Online*. March 2018. URL: https://www.youtube.com/watch?v=kxkrdLI6e6M.

[12] Tony Hammond. *9 best practices and tips to follow for effective data visualization*. January 2023. URL: https://www.thoughtspot.com/data-trends/data-visualization/best-practices-and-tips-for-effective-data-visualization.

[13] Tommy Walker. *Why Simple Website Design is The Best: The Scientific Reasons*. May 2022. URL: https://cxl.com/blog/why-simple-websites-are-scientifically-better/.

[14] *paletton.com*. URL: https://paletton.com/.

[15] *Coolors*. URL: https://coolors.co/.

[16] *Our World In Data*. URL: https://ourworldindata.org/.

[17] *Analyse en composantes principales*. URL: https://fr.wikipedia.org/wiki/Analyse_en_composantes_principales.

[18] *Algorithme t-SNE*. URL: https://fr.wikipedia.org/wiki/Algorithme_t-SNE.

[19] *Coop online supermarket*. URL: https://www.coop.ch/en/.

[20] *Geospatial Data and real-time risk management for business, government and humanitarian organizations*. URL: https://www.kontur.io/portfolio/population-dataset/.

[21] Uber Engineering. *[Uber Open Summit 2018] Tiling the Earth with Hexagons*. November 2018. URL: https://www.youtube.com/watch?v=_-265mfMzl4.

[22] Uber Engineering. *Kepler Gl*. URL: https://kepler.gl/.

[23] Uber Engineering. *kepler.gl for Jupyter User Guide*. URL: https://docs.kepler.gl/docs/keplergl-jupyter#install.

[24] *Rayshader Github repository*. URL: https://github.com/tylermorganwall/rayshader.

[25] *Rayrender Github repository*. URL: https://github.com/tylermorganwall/rayrender.

[26] Spencer Schien. *Rayshader Tutorial—Florida Population Density*. December 2022. URL: https://www.youtube.com/watch?v=zgFXVhmKNbU.

[27] Spencer Schien. *Creating High-Quality 3D Visuals with Rayshader*. January 2023. URL: https://spencerschien.info/post/data_viz_how_to/high_quality_rayshader_visuals/.

[28] Cartography Vectors. *Switzerland Detailed Boundary*. URL: https://cartographyvectors.com/map/1043-switzerland-detailed-boundary.

[29] *Magick Github repository*. URL: https://github.com/ropensci/magick.

[30] *Inkscape*. URL: https://inkscape.org/.

[31] *Matplotlib*. URL: https://matplotlib.org/.

[32] Valentin Vlădut Nicoleta Ungureanu and Sorin-Stefan Biris. *Sustainable Valorization of Waste and By-Products from Sugarcane Processing*. 2022.

[33] Discovery UK. *How Sugar Is Made — How It's Made*. July 2022. URL: https://www.youtube.com/watch?v=wvuiavj0pIo.

[34] URL: https://chandoo.org/forum/forums/discuss-data-visualizations-and-charting/.

[35] *Data Visualization Tips For More Effective And Engaging Design*. URL: https://www.tableau.com/learn/articles/data-visualization-tips.

[36] Peter Kovesi. *Perceptually Uniform Colour Maps*. URL: https://colorcet.com/.

[37] Jakob Nielsen. *Mental Models*. October 2010. URL: https://www.nngroup.com/articles/mental-models/.

[38] Jakob Nielsen. *Mental Models*. October 2010. URL: https://pubs.aip.org/aip/cip/article-abstract/10/3/268/136498/How-Not-to-Lie-with-Visualization?redirectedFrom=fulltext.

[39] Evelina Judeikyte. *2022 in books*. December 2022. URL: https://www.theplot.media/p/2022-books.

[40] *Bullet Graphs*. URL: https://www.tableau.com/data-insights/reference-library/visual-analytics/charts/bullet-graphs.

[41] *Data driven design*. URL: https://www.eyequant.com/.

[42] Simplylearn. *23 Best Data Visualization Tools for 2023*. URL: https://www.simplilearn.com/data-visualization-tools-article.

[43] Anna Glivinska. *The 30 Best Data Visualizations of 2023*. May 2023. URL: https://visme.co/blog/best-data-visualizations/.

[44] *Infographic Templates*. URL: https://www.visme.co/templates/infographics/.