

All You Need to Know About Data Visualization

23 MIN · YOUTUBE · [HTTPS://WWW.YOUTUBE.COM/WATCH?V=JHV9_QPWSY0](https://www.youtube.com/watch?v=JHV9_QPWSY0)

https://www.youtube.com/watch?v=jhv9_qPwSY0

SUMMARY

This video provides a comprehensive guide to mastering data visualization, emphasizing the importance of transforming raw data into actionable insights for stakeholders. Lorenzo, a data analytics lead, shares his expertise and practical tips for creating effective visualizations that communicate information clearly and engage audiences.

- Data visualization is crucial for making complex data understandable and actionable for stakeholders.*
- The goal is to transform data into knowledge that leads to informed decision-making and action.*
- Tailor visualizations to the audience's needs and understanding, avoiding one-size-fits-all solutions.*
- Utilize pre-attentive attributes (like color and shape) to enhance immediate comprehension of visual data.*
- Choose appropriate chart types based on data types (nominal, ordinal, discrete, continuous) and the visualization's purpose (comparison, relationship, distribution).*
- Apply design principles such as contrast, hierarchy, and white space to improve clarity and focus.*
- Avoid common pitfalls like 3D charts and pie charts, which can obscure data and mislead viewers.*
- Ensure visualizations start from zero on axes and maintain accurate proportions to prevent misinterpretation.*

My goal in this video is to give you the most comprehensive and practical guide that you can follow end to end to create visualizations that look like this and this that you see in the screen right now. Data visualization is a fascinating and super interesting topic, but so many people get it wrong. And so the promise that I have for you, if you follow this video until the end, you're going to be able to master this topic for good because these are also the principles and guides that I follow pretty much on a daily basis at work. And if you're new to the channel, I am Lorenzo. I'm a data analytics lead with over 7 years of experience in company like Deoid and AWS and I'm also the founder of the analytics and automation academy which is definitely not one of those video course on analytics that you see online but is a 101 journey that we do together until you land the job in data analytics or data science. So if you're interested in data analytics and you want to land the job the most efficient and fast way possible then definitely apply to my analytics and automation academy from the link that you will see in video description. And if you're interested in the topic of data visualization and you want to know all the concepts that you need to apply in your projects and also at work, then let's get straight into today's video. Okay, so as always, I created a bit of a mind map that I'm going to use for this video for our data visualization guide. And by the way, you are loving this type of format with the mind maps. So if you're interested to get in this mind map, send a comment in this video and I will try to export all of these into a PDF and send it to you. So the first thing that I want to cover in this mind map is actually starting from the the core basics. So why do we even talk about data visualization? I don't want to waste time with the definitions. I think we are all clear with what we mean by data visualization. But I think we should cover why we are visualizing data. So data visualization packages information into accessible messages that stakeholders can better process, retain and use to make well-informed decisions. Now to explain this concept in practical terms, I really like this kind of pyramid here that is basically telling us that data which is at the very bottom of the pyramid. This is information that is understandable by

computers. So that is our starting point. Now the goal of data visualization is actually to go up into this ladder and transform data into information. Why? Because information then becomes understandable by humans.

And so this is why we want to create tables and graphs to transform that data into information easily understood by humans. Now the thing that a lot of people get wrong is that a lot of people will actually stop at this level here. So you have these kind of massive reports or dashboards which show a lot of information and they basically stop here. Now if you want to go into the top 1% of data analyst or data visualization experts, what you want to do is actually bring information into knowledge. So that means that if you have stakeholders at work, if you have clients, you want to give them knowledge. Why? Because the ultimate goal for these people, for your clients, for your stakeholders, as soon as they get knowledge is to actually transform that knowledge into action.

And this to me is the part that is super super important whenever we talk about data visualization. Data visualization is only a tool. It's not the end point. It's not something that sits there on its own. is something that really becomes useful if you are able to create data visualization and discover insights that will translate into actions for the stakeholders for the whole business overall. I think this is self-explanatory. If you create a beautiful data visualization that sits on your laptop or maybe in a Tableau server and nobody really accessing that's pretty useless. Whereas if you're able to create a data visualization, maybe explain it to stakeholders, these stakeholders understand your message and then they start you know talking with other people maybe a higher level in the organization to actually take actions based on your insights. That is the ultimate goal that we want as a data analyst and data visualization experts.

And so if this concept is clear, let's move on to the next part of this mind map. The other important thing that I wanted to cover very quickly because I think you're really familiar with this concept is that your information design should be adapted to each audience. So what does it mean? You have to understand who is your audience. What do they want to know and how will they communicate it? The best approach is not to create a one-size-fits-all solution, but really tailor what you are working on based on who is on the other side of the screen or who is on the other side of the presentation that you're delivering. And why is that? because research shows that visual information is processed more easily and stays with us longer than text. Now the next thing that I want to show you is actually something that I find super fascinating and is this concept called pre-attentive attributes. So these are visual properties that the brain the human brain processes instantly and unconsciously allowing for a rapid pattern recognition without conscious effort. So this is to me super interesting part of data visualization. There are these pre attentive attributes like the length, the width, the intensity, the hue, the orientation, the curvature, the shape and the position. And hopefully you see all the practical examples of each of these attributes that are super super simple for our brain to process. And therefore we should definitely use in our data visualizations cuz if we use these elements that automatically makes our data visualization way easier for the brain of our stakeholders to really understand and process. And so this is why using pre-attentive attributes is basically the building blocks of data visualization. Now if this is clear, let's move on to our mind map and so we covered kind of the foundation of why we're even talking about data visualization and some of the core elements. Now we want to start planning for our data visualization. We don't want to jump into you know creating charts and reports. We want to start planning for it of course beforehand and we made this concept clear already why we're doing this because different stakeholders have different data needs. Consider your stakeholders literacy, visual literacy and what data they need to make decisions. And if you are planning for data

visualizations and you want to use a framework that I think is pretty pretty interesting, this is the one that I think you can follow. And so basically this chart shows you four main ways to use visuals depending on what you're trying to do and who you are talking to. So we have two axis that I'm going to explain in just a second. But basically the idea here is that the first type of visualization and type of communication that you can use is the idea illustration. So this is where you already have an idea and just want to explain it clearly. And so basically you're not really doing exploration. You already know what you want to kind of declare with this visual. And actually it's not really datadriven is more conceptual meaning that you can use frameworks or flowcharts. So not something that is you know super data heavy let's say. So that is the first thing illustration. Then if we go to the top right you have static visualization. So you want to show facts or data in a clear finished way. So even here you have already the idea of what you want to represent and also you have all the data points and so this is actually where you can start creating basic chart graph design infographics or static maps. Now bottom left is what is called idea generation. So you are still thinking through ideas. So you don't have like a very very clear plan in mind. You are kind of a brainstorming and you actually use visual to do this type of brainstorming. So again you have framework development or graphic brainstorming and then bottom right is dynamic visualization. So you want to explore or interact with the data and you want to use a datadriven approach.

So in this way you want to create dashboards that are basically interactive visualization that users can use to dive deeper into your data and maybe potentially find interesting insights from that. And so the vertical axis here goes from exploring ideas which is the bottom part to explaining clearly which is the top part and the horizontal axis goes from conceptual left to databased which is on the right. So basically here you want to show what you already know.

Here you are still figuring things out. Here you have ideas and concepts and here you have numbers and data. And so if this is clear let's move on to the next part of our mind map on data visualization. And here I want to talk about design basics. So first of all a few design principles that we can follow pretty much every time that we create data visualization. So first one is called type face. So pick appropriate fonts for your audience. probably I will stick with two three fonts at most and use it pretty much everywhere in your data visualization. Then you definitely want to use contrast. So light, dark, big, small, thick and thin to highlight or emphasize. So here I also put some good example of contrast which is obviously based on color. So you immediately see that this one here bottom left and bottom right are pretty bad contrast and I would never use them. So we cover contrast and also color balance. So you have to know that the items to seek balance and will notice if your design is unbalanced and then obviously you have hierarchy. So you want to give your audience visual cues of what is the most important part of your message. Something also pretty important to cover on color is actually incorporating white space to give the eye a break. So this is a bad example.

This is actually a good one. And also you do want to avoid something that is called the skittletto effect which is basically using so many different colors all at once. So this is probably not a very good example of using colors the right way. So I think this is pretty self-explanatory. Now if I want to give you a very practical example of these few design principles applied in a real scenario, we can actually use a text as an example. So you have here the before and after. And here we are pretty much using a lot of these design principles. So we are using sizes, we are using types of fonts, we using colors, we are using contrast as well. And so you can immediately see the difference of presenting a lot of information in a not very well formatted way. And whereas we are actually

taking the same exact information but only changing you know the size of that main title and then a few paragraphs here.

You can immediately see how easier it is to consume this type of information as opposed to what you see on the left. Now let's move on to the next part of our mind map. This is actually again something that I try to use pretty much all the time when creating a data visualization. Now a common issue that you might encounter when you plan your data visualization is actually choosing the right chart type. Why? because there are so many different charts that are available to you. And so the first thing that you want to do to choose the right chart type to use for your data visualization is actually to define your data types and then using and then using a decision tree that I'm going to show you in just a second. So first thing first, let's analyze all the different quantitative data types that you have at your disposition. So the first one is what is called the nominal and categorical data types. So this is where data can be sorted according to group or category. So here you can have the eye color, you can have the gender or type of cuisine. You're basically having different type of categories of a specific attribute. Then you have ordinal. So data of selected categories order along a numerical scale. Right here you have obviously depending on the location you have a primary school, a secondary school, maybe university, master degree, PhD and so on. And also star rating. So 1 2 3 4 five stars. Then you have discrete type of information.

So this is numerical data that has a number of possible values and units. So here you have for example number of children or number of items in shopping cart. And then you have continuous which is numerical data that is measured along a continuum and actually this can have infinite values depending on the precision that we use. So to understand the difference between these two you can immediately see that number of children here we're talking about a number of values. So you can think of zero 1 2 3 4 5 and and so on. Whereas weight depending on your precision you can have you know 12.1 kilos you can have 12.12 12.123.

So you can see that you can spread the weight into a infinite continuum. So that is kind of the clear way to distinguish between the two. And so why I'm talking about data types? Well, because the exact decision of the right data visualization to use can be based exactly on your data types, the type of data that you're dealing with. And so this is actually a guide that you can use to choose exactly the right visualization depending on what you're trying to visualize. And so as you can see here you have first of all four different elements that you can choose from. And so this is a composition probably a composition little typo there. So whenever you want to show the makeup of one or more variables usually in absolute numbers and normalized forms. So we are talking about this part at the bottom here. Then you have relationship. So it's a visualization method that shows relationships and connection between the data or show correlation between two or more variables. And we are talking about this one here. Then we have distribution.

probably I forgot how to write whenever I was writing this one. Another typo here. But basically these are visualization methods that display frequency. How data spread out over an interval or is grouped. And so this is this part here on the right. And then you have comparison. So visualization methods that help show the difference or similarities between values. And that's exactly our part here at the top. Now let's do a practical example here. So let's say that I want to see the correlation between height and weight.

So what am I trying to do here? I am trying to basically find the relationship between these two variables. And so I'm going to go here on the left hand side of the chart. How many variables do I have? I actually have two variables only. And so between these two I'm going to choose two variables. And then are these discrete or continuous values? actually height and weight are continuous because they can be put in a continuum and I can have infinite combinations and therefore the best practice for the type of information that I want to show is a scatter plot and if I want to give you a practical idea of what that looks like is exactly something like this. You have height and weight and basically the relationship between them. Okay. And so going back here you have a precise framework that you can use anytime to choose exactly what is the right visualization for your reporting or analysis. Now let's move on to the next part about data visualization which is okay we understand the right chart to use. We want to make sure that we communicate clearly with these charts and I want to show you exactly what is the rationale for these changes. So the first thing you want to make sure that you have a clean chart area that allows you to concentrate on the important information. So here what we've done is we kept 2005 and 2014 cuz that's important information that we are comparing. But actually we don't need all these tick axis to show all the type of percentages that we have in the chart because we are going to put the exact number in here in the chart itself and we only need four and we need the precise numbers. So a few reasons why it doesn't make any sense to put all of this information and also the lines here. Also we decide to get rid of the legend here. Why? because these legend information are actually again plotted in the visualization itself.

And so we are using a combination of text and also colors to distinguish between the two. And lastly, the other thing that we added in the after version is this text here. So exactly the explanation of what we are trying to say here because in the before version we have two lines that are going down. So it's not super clear what's the information here. And so here the core message that we want to highlight is that for both married and unmarried women these unmet need for family planning is declining and basically reaching the pretty much the same level in 2014. So this is clear. Let's also check this part here. So this is also some practical examples on how to use color to add impact and accessibility.

and to make things super clear. I included a before and after implementing these best practices. So here we talk about the skittle effect which basically means using different colors all at once in a visualization. And so the problem here is that I'm not sure with my eye where I should focus on. Whereas the after where the best practices of colors are implemented, my eye knows exactly where to focus on because this different color here means that I'm going to highlight on this category because this is the value of interest. Also, this is a bit of a subtle one, but you have to consider that about 8% of men are color blind. So that means one out of 12 men are color blind and so ideally you want to use colors that are colorblind friendly like this version here rather than mixing together primary colors like what's done in here. Then the other aspect that I want to show you is also a few additional best practices that we should be careful about. So in here I put that you should definitely avoid 3D charts and using a flat design instead. I can agree that sometimes having 3D charts looks a bit more advanced. But actually the point that I want to make here is that this is not really adding anything extra to our chart and can actually make it more tricky to understand cuz if you see here this bar that is in 3D, I'm not entirely sure probably what's exactly the value that it corresponds to cuz there is this extra complex element in this chart that makes it more complicated to read.

And so in this case because the 3D version doesn't really add any value to it. It's way more straightforward to u

visualize and understand from the human brain to have in this flat version. Also the other thing that is super important to understand is that pie charts should be avoided at all cost. I still see a lot of people using pie charts in their visualization and this is definitely not best practice. Why this is the case? Well, because if I remove the numbers from this pie chart here, it is actually very difficult for our brain to compute the degree of for example this slice of the pie. And same for the orange, red and blue, it's actually very very difficult to understand the right comparison between these values. Whereas to make this information way more clear for our brain to digest then a simple bar chart is definitely the best practice and best visualization to use because I can immediately see the massive discrepancies between the different variables here. Also another thing that sometimes is misunderstood. You want your axis to all the time being full and so starting at zero. You can see this bar chart here and focus on the different values and then actually realize that our chart is not starting from zero by starting from 25 in this case. And so again that's an extra level of complexity that first of all can be completely ignored by our eye if we don't actually check the axis and also can definitely create some misunderstanding. Whereas what we want in the after version here is definitely start from zero. Also let me show you some very practical examples of stuff to avoid at all cost or best practices to follow. And so in this case you have an example of step one 2 3 and four that a person wanted to visualize in the chart. Now the massive problem here is that if you check the percentages used here they don't add up to 100. And so this is of course very misleading. And so if you want to use this type of visualization and we are splitting a shape into different sections then obviously we want the whole thing to be represented to 100%. Now the other thing that you see in here that is actually something that you see a lot done by magazines or newspapers online is actually representing the right proportions of your data. So you can see here the 25 and 30% this is the difference and then pretty much the same difference here is actually representing 75% difference and then obviously have these two NAS that completely don't make sense and so again this is super misleading because our I immediately goes into the size of the bar rather than checking the actual number. So definitely something to be fixed here. And lastly, the other massive pitfall that a lot of people are actually implementing in the visualization is something like this.

So I'm not sure if you're having the same effect as me, but I have no idea on where to start reading information in this type of visualization. I have u years here. So 2016 and 2017, but I also have years here in this circle. So 2018, 19, 20, and 21. Also, I have the same symbol here and here. Again, I'm not even sure what that represents. That is actually shown as violet and blue. We have so many text as well. So, not really clear what's exactly the focus. And the last thing, we also have this X and tick symbol, but then the tick symbol is also represented with a dollar amount over there and other variation here and here. So I guess the core point here is that this is a type of chart that is super super overwhelming and so immediately loses efficiency because if there is no person next to me explaining exactly what's going on, then I will be completely lost. And there you have it. This is a full end to end guidelines on data visualization. If you found at least one useful information in this video, make sure to like and subscribe to my channel so that I can help you even further in the next videos. And if you want to learn data analytics and everything that you need for a real workplace and having me helping you until you land a job, then definitely make sure to check out my analytics and automation academy at the link that you see in video description. We cover together Excel, SQL, Tableau, Python, AI agents and also domain knowledge. And as I said, I work together with you until you land a real job. Also, I will leave here in the screen some of the other most popular videos that I made lately. So in case you're interested on those topics, make sure to check them out. And as always, enjoy the rest of your day.