Data visualization essentials

Data visualization

A means to convey info using visual representation

Tells the story of the data

Allows the audience to gain an understanding of what the data is saying

Need to be useful for someone

It has to be explained logically & coherently

During data visualization we should never misrepresent or try to obscure the data. It has to be clear & concise.

Data visualization phases

Exploration phase – understand the data to find the story it’s telling

Explanation phase – tell the story

Know the audience

Who are you presenting to?

How well do they know the data?

What is their level of expertise?

What info do they require & do different audience members require different information?

Are they from different cultures or speak different languages?

Are there visual impairments? This affects the colors & sizes you can use.

After you answer the above questions:

Create data visualizations for the audience & ensure that they understand it & are engaged by it. Channel your narrative to them specifically.

Gestalt Principles of Perception

People digest data under the Gestalt theory which attempts to explain how we generate meaningful perceptions. Some principles are:

Principle of proximity – when individuals perceive an assortment of objects, they perceive objects close to each other as being part of the same group. In other words, if things are close to each other, people perceive them as being part of the same group.

Principle of similarity – if things look the same, people perceive them as the same type.

Enclosure & continuity – things with a boundary or continuation are seen as a group

Closure – individuals perceive structures as shapes, letters or pictures & even when these structures have gaps, they perceive them as being whole i.e, their brains fill in the gaps

Traits of good data visualization

Data presented is useful for the audience

Data presented is usable by the audience

They are pleasant to use

They are highly accessible

Visualization provides the right info to audience

Provide rapid response when users interact with it

Scalable – users can zoom/scale & easily explore your visualizations

Communicates & presents data in a comprehensive way

Enables analysis of data – supports decision making & answers the questions of users

Make large sets of data clear & coherent

Shows multiple levels – allows user to see the big picture while having the option to dig deeper

Don’t distort the data

A good data visualization makes the audience think about the data & not the way it was presented

Infographics – graphics are pre-generated, it’s more artistic, usually contains static data

Data visualization – the algorithm can crunch no:s in real-time to provide interactive visualizations, less artistic, can contain dynamic data

Choosing the type of visualization for the data

Graphs – Data must map 2 variables, at least 1 of them must be quantitative. Used to convey a trend or a pattern.

Timeline diagrams – for data over time

Template diagram – for showing a guide or a plan to be completed

Flowchart – for ordered set of instructions

Checklist – tasks to be crossed off as they are completed

Mind-map – represented the linking of ideas, with the subject (main topic) in the center which is connected to Ideas (sub-topics) which are then connected to sub-ideas. Can incorporate words, pictures, colors etc. & can present an overview of a central topic with large amounts of related information.

Interactive data visualization – users can interact to manipulate some aspect of the dataset or the way of presentation. It lets them gain insights from data at the level they require.

Importance of narrative

People care about stories

Data with a story attached to it is unforgettable

A good narrative captivates audiences

Images shown during the narrative should be aligned with the narrative so that the audience can relate to them easily

Adjust tone & delivery based on the intended audience

Image vs. text

We process all data from an image at the same time. We process text linearly. Images are more engaging & complements narratives, is easier for audience to relate to.

Reify data – take abstract concept represented by data & make it tangible using pictures, icons words etc.

Icons used during data visualization

Should be recognizable

Should convey meaning immediately but the meaning should be specific

Don’t overuse icons in your visuals

They should be of same size so that other images shouldn’t loose meaning

Lots of icons can appear facile – it can make your visualization look superficial

Visual aspects

Colors – end result must look aesthetic & pleasing

Contrast – use it to compare items

Position – use it to communicating quantitative, numerical & ordinal data

Shape – somewhat effective while communicating nominal data, avoid using quantitative data using it

Size – very effective in conveying quantitative data

Position & size – they are used together for quantitative data. Eg: Graphs, etc. Less relevant info is commonly put in the side & is relatively small in size

Colors

Color is powerful but subjective

Bold colors convey a message but could clash or be garish

Color theory is a good place to start

Shape

Shape is less effective than color for differentiating data. So they should be accompanied with color.

Shapes alone don’t make trends pop out.

Scale

Scale affects how we perceive relative values

The start value of an axis has a big impact on the relative differences

The same chart with different scales tells different stories.

If you use scale to distort user perception intentionally or un-intentionally, if found out, you will lose your credibility.

Scale exaggeration

Consider your goal & audiences

Suppose you want to convince an audience of a fact:

Skeptics are unlikely to believe & exaggerated scale

People already persuaded by you will loose their trust in you if they find out the scale is exaggerated

Be subtle when you use it

Sometimes it’s OK to exaggerate small differences

However, always choose a scale that gives an honest appraisal

It should still maintain a narrative

Legends

Legend/key explains data

Legend makes the data clearer

If the item labels are small, place it beside the chart. If they are long, place it below the chart

Complicated charts need detailed legends. In such cases legends could be broken out into a separate page

Sources

Adding sources gives credibility to your data visualizations

Adding good sources increases the trust users have in your data

Skeptics will cast doubt on un-sourced data