You are viewing this thread in readonly mode.

Marks and channels for Categorical Data #15

Anonymous
7 months ago in General

339 VIEWS

I was wondering when we see some Data visualisation for Categorical data we always have some channel from the Ordered Attributes because even Data visualisation for Categorical data would include some numeric figures right?

lets say there are 2 categories for Laptops shown in a Diagram, which category is more favourite among people, of course this will include which category has more sales.

So in that situation should we criticise that they have used channels from Ordered Attribute even if they have data visualisation for categories?

or is it normal?

2 Answers

K

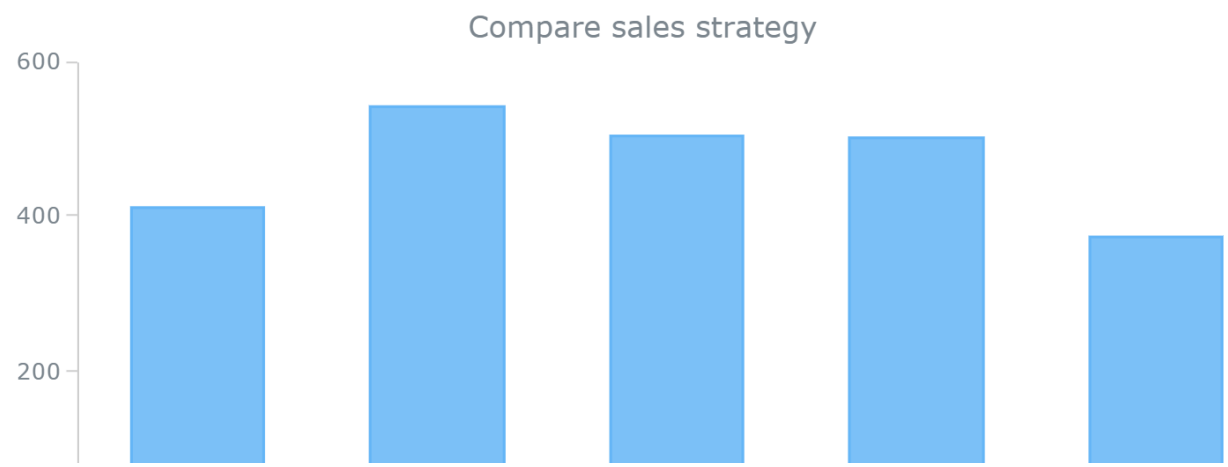
Kane Li STAFF
7 months ago

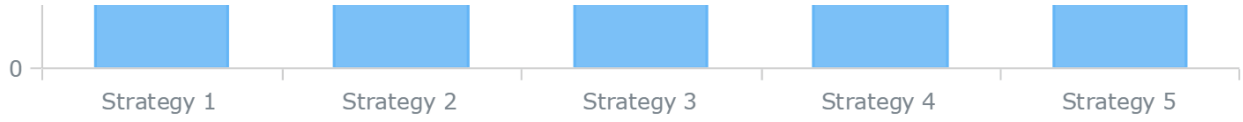
ENDORSED

Hi there,

There could be multiple attributes represented in a graph. For example, in the bar chart below, there are two attributes (variables):

- different strategies: categorical variable, represented as the x-axis (spatial region - identity channel)
- sales: numeric variable, represented as the y-axis (position on a common scale - magnitude channel)



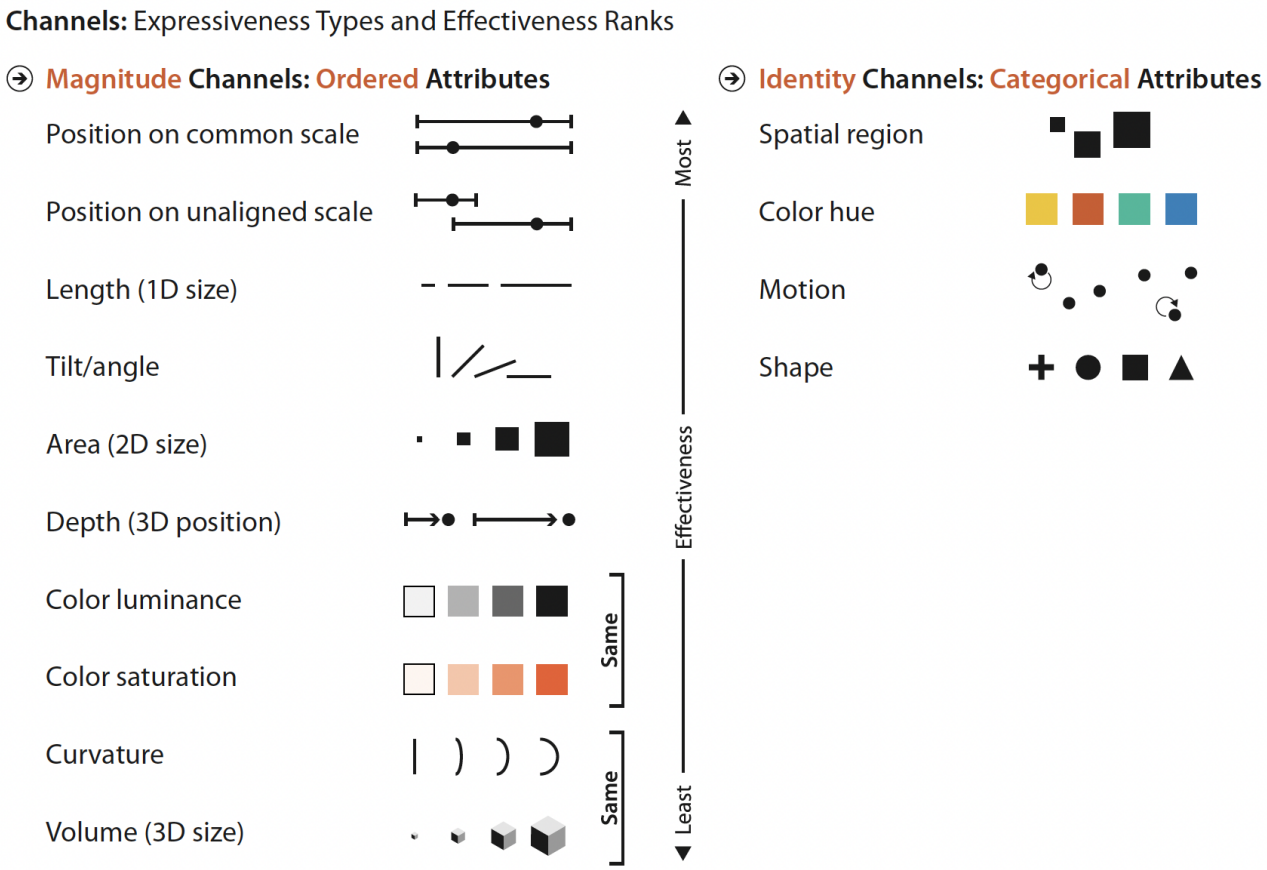


link: https://cloud.netlifyusercontent.com/assets/344dbf88-fdf9-42bb-adb4-46f01eedd629/fc51faeb-333e-4126-ab3d-4c0497eff395/4-single-series-bar-chart-large-opt.png

The difference between the magnitude channel and the identify channel is (Chapter 5.3.1 of the textbook):

- "The identity channels tell us information about what something is or where it is
- The magnitude channels tell us how much of something there is"

There is a channel that appears in both of the two ranked lists of channels: position. In the magnitude channels, we have "position on a common scale", or "position on an unaligned scale". In the identity channels, we have "spatial region". They are actually all related to the "spatial position" of the graph, which means - spatial position is the only channel that is both a magnitude channel and an identity channel, also it is the most effective one.



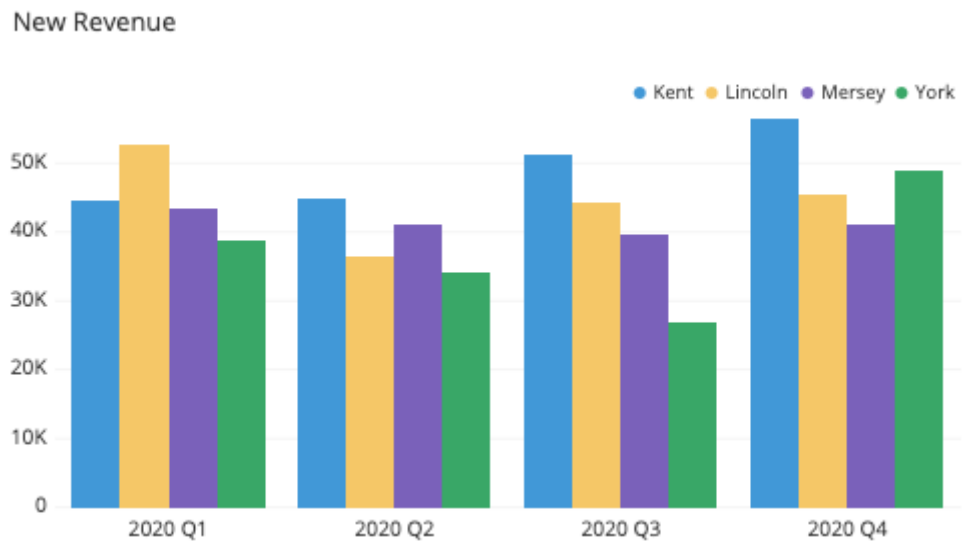
Getting back to the bar chart, it uses the x-axis and the y-axis to represent two attributes (a categorical attribute + a numeric attribute).

- y-axis, this is a magnitude channel, as we care about the order of the data, a sale of 400 is larger than a sale of 200; so you can't move around the data on the y-axis.

- x-axis, this is a identity channel, as we do not care the order of the data - you can move around different categories here.

Why it is called "spatial region" instead of "position"? First, it is the same thing - you can just say "position" for the categorical variable (which is totally fine). Second, this is my own understanding - if you think about the grouped bar chart below, on the x-axis, the data for 2020 Q1 are all placed in the same region. So the attribute (the quarter of the year) is represented as the spatial region. The colour hue is used to represent different towns.

If we talked about the x-axis position, it actually has two variables: the quarter of the year, and different towns. However, when we look at the grouped bar chart and want to compare the new revenue of the same town (e.g., Kent) at different times, we are actually looking at the colour.



link: https://chartio.com/assets/dfd59f/tutorials/charts/grouped-bar-charts/c1fde6017511bbef7ba9bb245a113c07f8ff32173a7c0d742a4e1eac1930a3c5/grouped-bar-example-1.png

So "spatial region" might be a more accurate word here than "position".

Hope this helps answer your questions.

Cheers,

Kane

S Sitian Huang 7mth

Hello Kane, I have another question about the bar charts above. Is it the line mark(says the textbook) or area mark? I am quite confused with these 2 concepts and could you give me some examples that are area mark except geographical maps or trees?

Thank you.

•••

B Bernhard Jenny STAFF

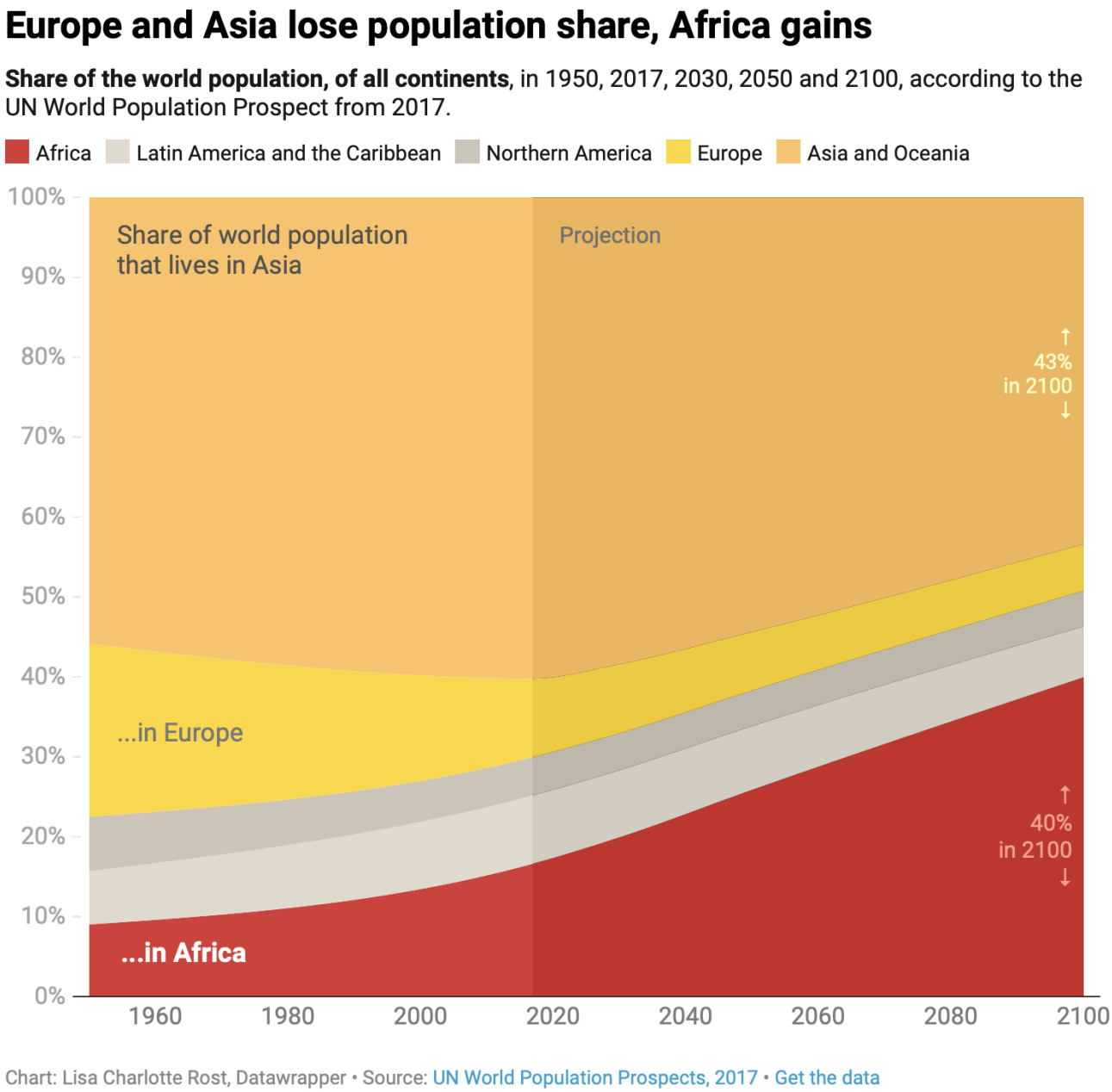
7 months ago

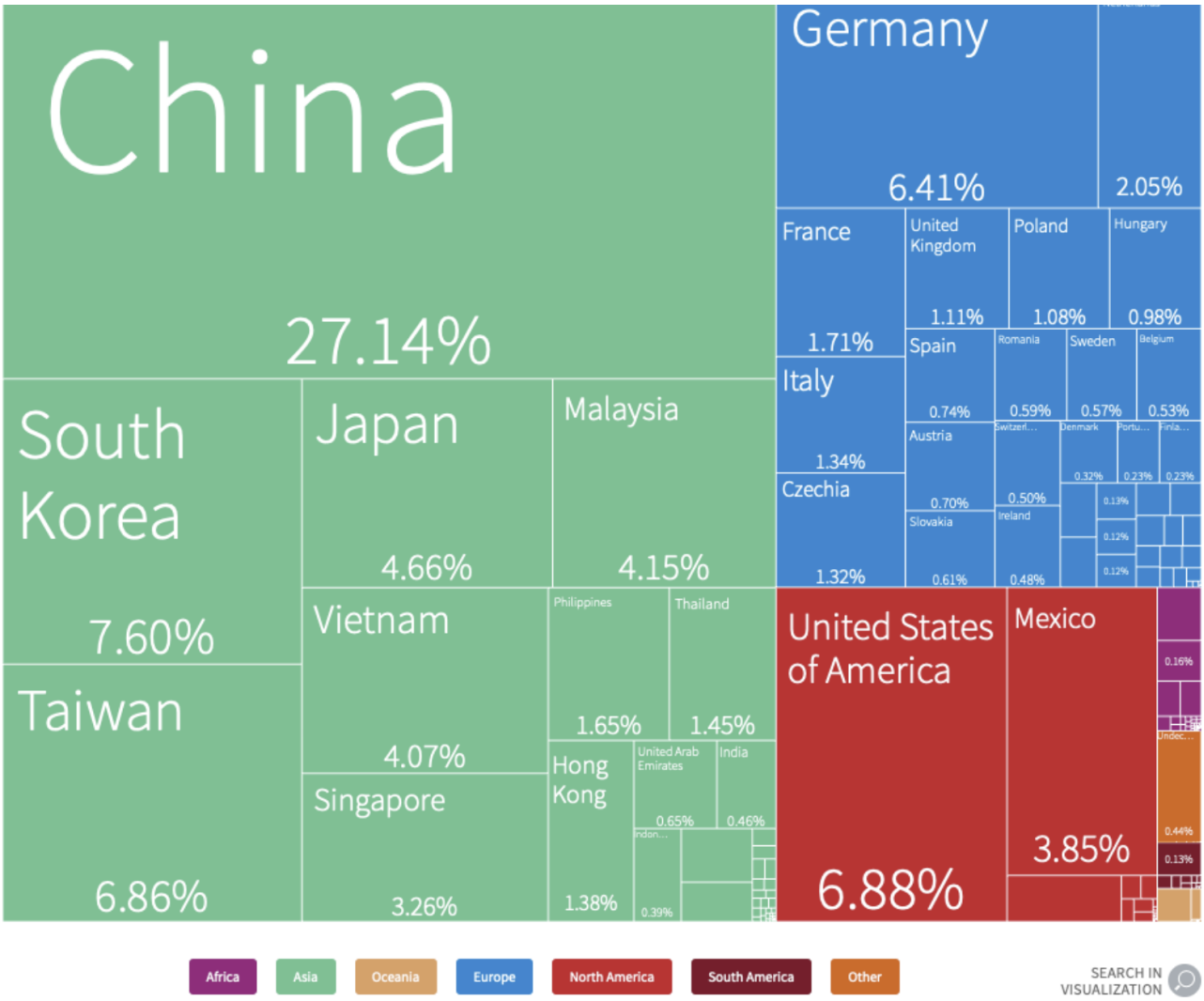
Hello Sitian,

The marks in bar charts are lines. The lines obviously need to have a width to become visible; you can imagine that each line is drawn with a single stroke of a wide pen.

Attached are an area chart (for table datasets) and a treemap (for tree datasets), which are two visualisation idioms that use area marks.

Cheers,
Bernie





S Sitian Huang 7mth

Thanks so much!

•••