Dataviz exam simulation - Practice 11

Analysis

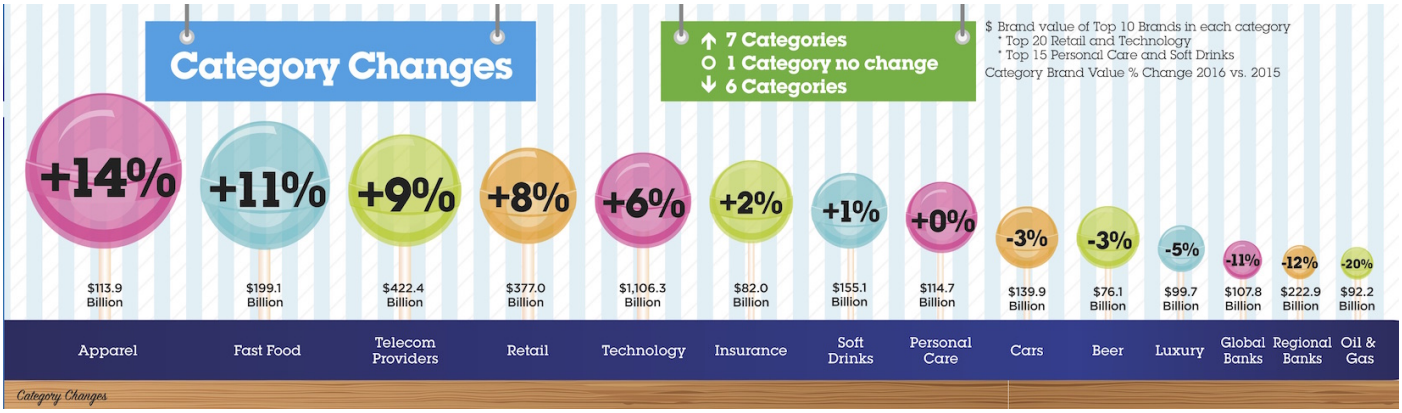

Figure 1: BrandZ

Analyze the above graph according to the following criteria.

Question

Which one of the following questions represents the purpose of this visualization?

- *What is the change in the value of some categories of brands from 2015 to 2016?*
- What is the absolute value of some categories of brands in 2015?
- What is the trend of the value of some categories of brands over several years?
- What is the category of business associated with the highest revenues?
- What is the most representative color associated with some categories of brands?

Data

Is the data quality appropriate? Select true answers only.

- The values associated with each category are too similar to be accurate.
  - Incorrect because the visualization shows different percentage changes and absolute values for different categories, which suggests variation in the data. Without additional information on the expected variability of these values, we cannot deem them as inaccurately similar.
- *The data accuracy cannot be verified because the visualization does not cite sources or methodology.*
  - Correct because without a cited source or collection method, there is no way to confirm whether the values are accurate.
- *The data is not complete because it reports only the top brands in each category rather than the full category total.*
  - Correct because the visualization specifies that it is showing brand values for top brands in each category, not the total value of the entire category.
- The data is complete because all possible categories of brands have been reported.
  - Incorrect because the visualization does not claim to report all possible categories of brands. It is evident that only selected categories are shown, and the title suggests that there may be changes in other categories not reported here.
- The data is consistent as similar categories of brands are considered.
  - Incorrect because consistency in data refers to the use of data collection and reporting methods that are uniform across different datasets or categories. Since the visualization only includes the top brands from each category, it may not be consistent if the number of brands considered for each category is different, or if the method for determining these top brands varies across categories.
- *The data is not consistent because different categories aggregate different numbers of top brands (10/15/20).*
  - Correct because the varying top-N thresholds mean the categories are not summarized on the same basis.
- The data used in this visualization has been collected before 2015.
  - Incorrect because the visualization notes the brand value percentage change from 2015 to 2016, which implies that the data was collected at least in 2016 or later to be able to report the change over the previous year.
- The visualization clearly explains what the sources of the data are.
  - Incorrect because there is no explicit mention of data sources in the visualization. Without a clear indication of where the data comes from, we cannot say that the sources are clearly explained.
- *Understandability is not adequate because the explanatory note about top brands is easy to miss and does not clearly define what the values represent.*
  - Correct because the small, dense annotation can obscure the fact that the numbers are sums of top brands rather than full category totals.

- Precision is not appropriate; percentages should have at least two decimal digits.
  - Incorrect because the level of precision for percentages in data visualization should match the needs of the task and the audience's understanding. There is no universal rule that percentages must have two decimal digits; one decimal or even whole numbers can be appropriate depending on the context and the data's variability. The precision used in this visualization appears to be in whole numbers, which can be perfectly acceptable for an overview chart like this.

Visual

Proportionality Are the values encoded in a uniformly proportional way?

The percentage change values are not reported with a uniform proportion in the size of the bubbles. There is a zero value that theoretically should correspond to a zero dimension. Negative values are represented as smaller, without a clear proportionality or coding rule. The position of the bubble centers relative to the base of the graph could potentially represent the value; however, this is not explicit.

Utility All the elements in the graph convey useful information?

There are several purely decorative elements that distract from the quantitative message: the striped background, the stylized reflections on the bubbles (lollipops), the sticks that support the bubbles, the wood-effect strip at the base of the graphic. The sign with information on how many rise/fall or are stable could be deduced from the graph and is essentially redundant. The variable font size with the percentages is of little use in transmitting relative values.

Clarity Are the data in the graph clearly identifiable and understandable (properly described)?

The legend on the values shows the two measures in inverse order: at the top it mentions the absolute value in $, which is at the bottom, while at the bottom it shows the % of variation that is at the top. Negative percentages are reported with a very small font size. In the hypothesis that the position of the bubbles represents the variation values, a possible vertical axis and a line corresponding to zero are missing. Furthermore, the label with the absolute value, shown on the vertical section, may suggest that the height represents this value instead (but it is not).

Design

Design the visualization based on the following data structure.

Field	Dim./Measure
Category	Dimension
Percentage	Measure
Value	Measure

Design schema

Schema	Details
Rows	Category
Columns	SUM(Percentage), SUM(Value)
Type	Bar
Color	Stepped color with 2 steps for SUM(Percentage), one color for SUM(Value)
Size	Default value
Label	No label

Sketch of the resulting graph

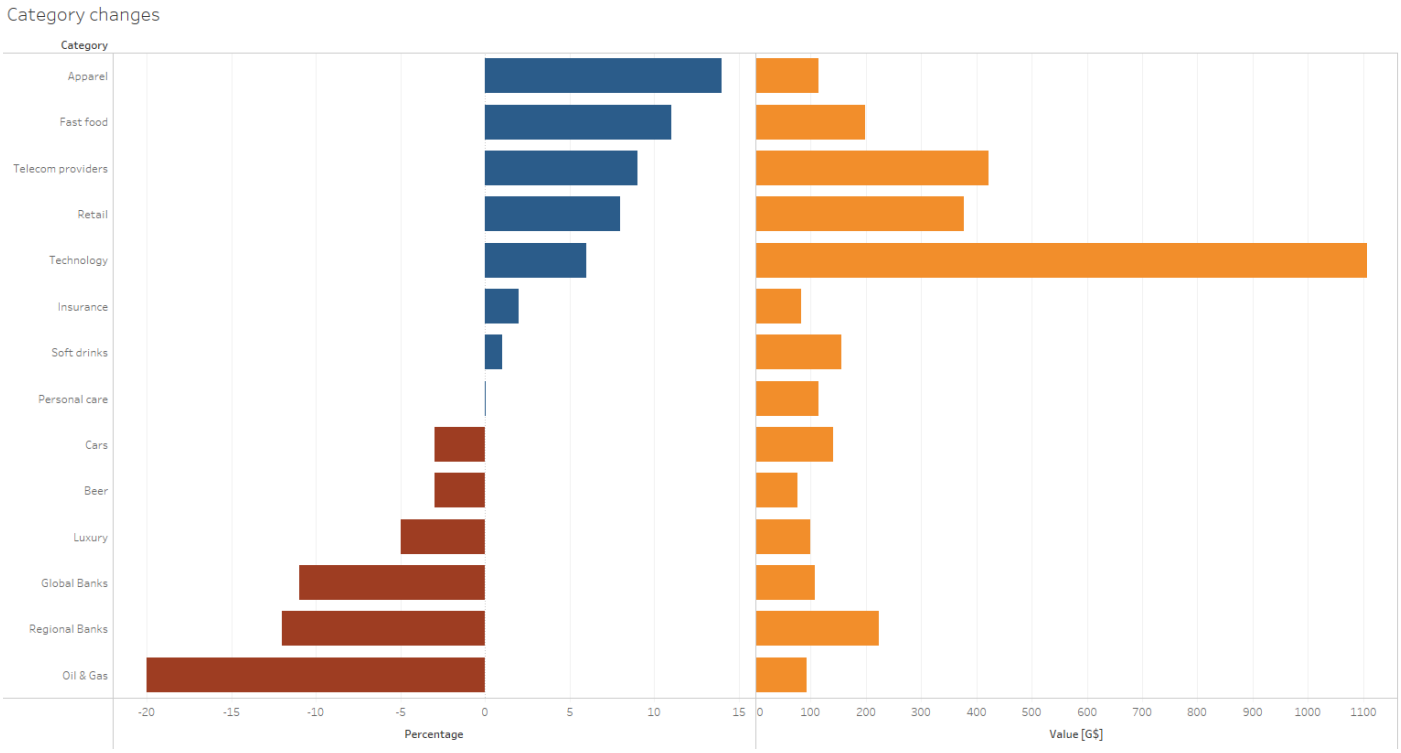

Figure 2: Redesign

The proposed redesign is two aligned horizontal bar charts by category, with percentage change on the left and absolute value in G$ on the right, sharing the same category axis. Positive percentage changes extend to the right in blue while negative changes extend to the left in brown, making direction clear. Absolute values are encoded as orange bars on a separate scale, allowing a straightforward comparison of category size without mixing units. This layout removes decorative clutter, restores proportional bar encodings, and separates measures into clear, labeled axes to address the ambiguity, proportionality, and clarity issues noted in the analysis.

Theory

If a variable represents heights of people and a data point is "0.002 km", we are observing an issue of:

- *Precision*
- Accuracy
- Understandability
- Consistency
- Completeness