Data Visualization Critique Paper

## 

A) What you think the story the visualization is trying to tell:

This visualization displays the top four defects, with an additional “other” category for the remaining defects. Each category is represented by the number of defects and their percentage of the total defects. I came across this graph when it was presented in my Operations and Supply Management class. It was taken from a study conducted by a packaging company. The purpose of the study was to identify areas in the packaging process that require improvement. To do this, the company collected customer complaints over a two-week period and analyzed the data, grouping similar complaints together, to determine the most common issues that customers came across after receiving their product. This visualization is trying to communicate the findings of the study and be used as a tool to help guide the company in their attempt to improve their packaging process.

B) How effectively you think it conveys that story:

In my opinion, the visualization inadequately conveys the story it is trying to get across. For example, the title of the graph, "Current State- Defects", is vague and fails to accurately convey the focus and scope of the data presented in the graph. After analyzing the graph and comparing it to my knowledge of the study, there seems to be many aspects that might cause some confusion for the reader and may lead them to draw different conclusions than the graph intended.

C) Whether it’s being misleading in some way:

Some aspects of the visualization are misleading. The y-axis is labeled as “Number of Defects,” which can be misleading because the graph provides no other context, which could cause the reader to make an incorrect conclusion. The y-axis actually means to show the number of customer complaints about the type of defect. Since the company got this data from their customers who reported defects, this means that the data is not representative of all the defects in the company's packaging. For example, it makes it seem like there are only 20 “smeared paint” defects, but this is actually showing that 20 out of the 48 complaints were about the smeared paint, but there could be many more unreported. Similarly, the use of percentages on the x-axis is misleading because this is not calculated from the actual number of total defects the company had. The use of different colors for each bar can be misleading as well because it could cause the reader to believe that there is something relevant to the story they are missing. The different colors draw importance to each different type of defect, but there is no reason to do so. The title "Current State – Defect" can be confusing as well because it doesn't accurately communicate what the graph is trying to convey. It implies that the graph provides a thorough overview of the current state of the company’s defect rate, when in reality it is only showing the top four categories of defects.

D) How well it holds up the visualization principles we’ve learned about:

This visualization violates many of the visualization principles we have learned in class. First off, there is a lot of clutter present. The percentages under the different types of defects do not need to be there, and they are adding clutter to the x-axis. There is also no true focus anywhere in the graph. The different colored bars in the graph do not highlight any key takeaways from the data, and the reader is left guessing at what is important in the visualization. However, it does use the principle of proximity to order the information in descending order, which makes comparisons easier. The “other” category also does decrease clutter since it is there for comparison, and there are only 2 complaints related to it. The graph also labels the data directly, which makes the information easier to follow.

E) How it could be improved:

Starting off with the title, it can be revised to better explain what the graph is trying to depict. Something could also be added to the title that indicates this graph is related to a certain company's packaging process, since the audience currently has no context. The use of multiple colors for the bars and the percentages can be eliminated to decrease the clutter. Instead, a single color could be used for all the bars except the "Smeared Print" category, since this is what we want the audience to focus on. I noticed all the data points are multiples of two and only go up to twenty, which leaves a lot of white space at the top of the graph. To improve this, the y-axis could be increased by two instead of five to align with the data points, and the range could be decreased from thirty. The x-axis label should also be changed to a simpler title like "Typed of Defects" since I recommended to take out the percentages. The bars in the graph should be spaced out to improve readability and avoid overcrowding.

F) Why your improved version improves it:

My version of the visualization improves it because it helps the reader to better understand the message the visualization is trying to get across. I have eliminated the misleading and vague aspects of the original graph, so the reader doesn’t need to be guessing at what they are supposed to learn from the graph. I have also decreased the clutter of the original graph, so it is easier for the reader to follow. Overall, my changes better follow the visualization principles and properly communicate the intended message.

