Problem Set -7

Whenever we visualize, we are encoding data using visual cues, or “mapping” data onto variation in size, shape or color, and so on. Even the most important message would fall flat without appropriate visual cues to properly convey it. With the market’s oversaturation of data, plus a wealth of content clamoring for attention, you have your work cut out for you as a storyteller.

With the available default data sets within IBM Cognos , I have choosen Storm Events 2015 Data set. The Storm Events Database contains the records with significant weather phenomena having sufficient intensity to cause loss of life, injuries, significant property damage, and/or disruption to commerce.

Here I have developed three different kinds of visualizations.

1. **In this graph I have taken different kinds of events occured per each month.Based on event ID's it shows how many events occured and what type of events occured in each month.**

I have taken multiple types of graphs to see which graph gives the best visual representation. We should not overuse visual cues at the same time. We need to aware of which visual cue is best suited for our requirement.

Here first graph I have chosen is point graph, column graph and comparing these it looks clearly that point graph is the best option to understand the data. This will clearly represent the events month by month and each event type. We can say clearly “Winter Storm” and “Thunderstorm wind” events occurred every month and other events occurred on few months only.

In second graph, I used multiple visual cues in the same visualization, but it looks difficult to understand and read the story out of it. So never overuse visual cues even though they enhance the story.

Third graph I have taken is column graph. I didn’t used any visual cues in this. When I see, it didn’t represent the correct story. Months considered as numbers instead of months. So, if event occurred in December then column graph is showing as one month instead of December (12).

This visualization can be used by Data analysts to analyze the data in deep and they can give advice on which event can happen on which month and when government has to take precautionary measures to prevent the event to happen.

This graph is a point graph where I used visual cues Size and position. It represented correct visualization story and we can clearly say which event occurred in which month and how many events occurred per each month and that represented via size. Here each bubble shows how many events occurred. So, if bubble is huge then that means a greater number of events occurred.

A screenshot of a cell phone

Description automatically generated

Let’s look at below visualization. Apart from size I added color as well. Then I feel visualization looks clumsier and more difficult to read what that visualization is trying to say. So, do not overuse visualization cues.

A screenshot of a computer

Description automatically generated

Let’s look at column graph below and I didn’t use any visualization cues. When I see, it didn’t represent the correct story. Months considered as numbers instead of months. So, if event occurred in December then column graph is showing as one month instead of December. So we can say this graph is misleading the people by not showing correct information.

A screenshot of a computer

Description automatically generated

**2) Second one represents How many direct injuries, indirect injuries, direct deaths , indirect deaths happened in all four regions.**

For this I have taken line graph, so I can represent different injuries and deaths for each region at the same time so that we can look at different kinds of data at the same time.

This visualization can be used for internal management where management and government can visualize direct injusries , indirect injuries, direct death and indirect deaths for eahc region. This is like comparision graph where management can see which has high impact. We can clearly say direct injuries are more in number than any other incident. Management can use this data and concentrate on specific areas to work on.

Here I used normal bar graph to represent injuries and deaths without visual cues. It gives the information, but it is not giving the information as how deaths and injuries changes across different regions.

A screenshot of a social media post

Description automatically generated

Below visualization shows the timely manner on how data changes region wise with specific colors for direct injuries, indirect injuries, direct death and indirect death. So, we can clearly say how direct injuries changes across different regions and the same applies to other injuries and deaths.

So here Bar graph without visual cues and line graph with visual cues represent same story but line graph with visual cues gives more information.

A screenshot of a computer

Description automatically generated

**3) This one represents how many deaths and injuries occured in all four regions - Midwest , Northeast , South, west.**

Here I have taken Bar graph with two different colors which represents death and injuries. With this easily we can differentiate how many deaths and injuries occured.

This visualization can be displayed to lay public to know how many deaths and injuries occured due to this storm on region basis. This is very simple visualization where any public person can understand easily what is the storm damage to people based on region.

This bar chart is easy to understand and that is the strength of bar chart. We can get the easy differentiation block by block and its over all value for each block. Bar graphs are used to compare things between different groups or to track changes over time.

A screenshot of a cell phone

Description automatically generated

I tried to add visual cues and tried to tell the same story. So, I used point graph but if I add both deaths and injuries count in same graph then it looks clumsier and not a good option for this story. So, I used different point graphs for deaths and injuries separately.

Let’s look at below graphs. First graph shows injuries by region wise and second graph shows deaths per region wise with color bubbles. It is difficult to read from two graphs and the graphs individually also not giving proper information. So always visual cues are not an option when the story you are going to say is simple. We should not make it complex by adding cues unnecessarily.

A screenshot of a computer

Description automatically generated

A screenshot of a cell phone

Description automatically generated

So, for this story visual que is not a good option. Just use simple bar graph, which will show deaths and injuries clearly in a simple way.