Reviewing the cases of **EDWARD tufte’s “cholera epidemic in london”** and **“decision to launch the space shuttle challenger”**

# Overview

Data that can be presented visually through infographics and plots to help understand its significance is called as Data Visualization. Several patterns, trends and correlations that could go unnoticed in a table representation of data can be exposed and understood better through this process of data visualization.

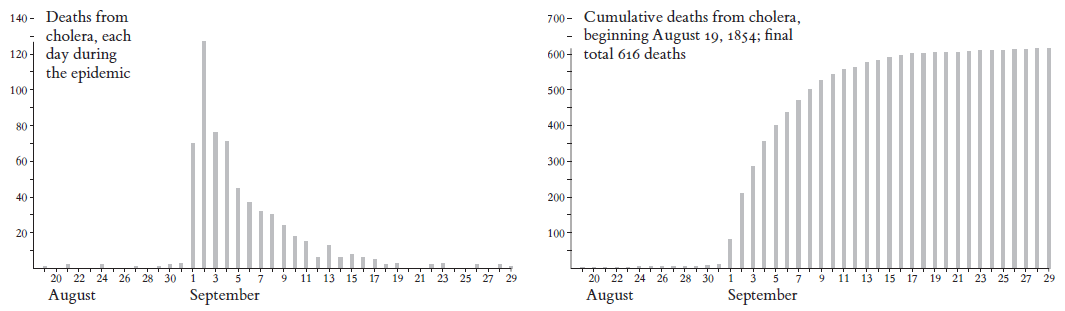
*“The greatest value of a picture is when it forces us to notice what we never expected to see”* *-- John W. Tukey*

In this article, we look at the case studies from Edward Tufte’s “Visual and Statistical Thinking: Displays of Evidence for Making Decisions” about the cholera epidemic that happened in London in 1854 and about the NASA’s decision to launch the space shuttle Challenger in 1986. We review the techniques and approaches that were handled back then and see what could have been done better.

# 1. Cholera Epidemic in london in 1854

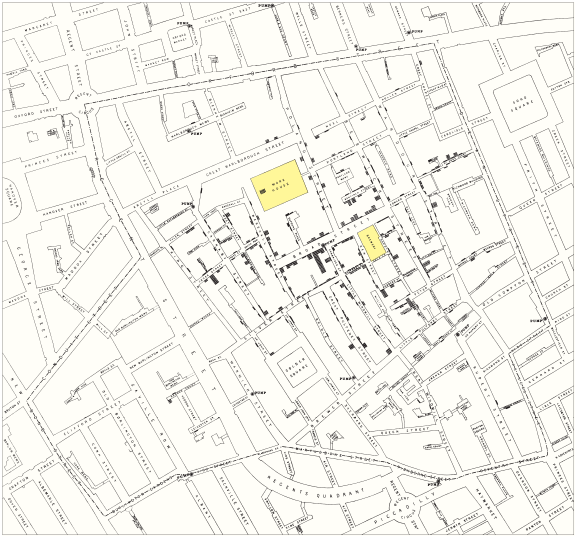
In 1854 in London, the fatality rate skyrocketed to 500 in numbers within a short span of just 10 days because of one of the most terrible outbreaks of cholera. John Snow who was a renowned physician, started investigating this outbreak and suspected that this could have been caused by the contaminated water from the community pump in Broad street. This lead to one of the greatest discoveries that the bacterium causing cholera, “Vibrio cholerae” is transmitted through water, which makes cholera a waterborne disease as opposed to the belief back in those days that it could be airborne.

The original data comprised of the list of victims’ names with their date of death, which was rendered useless by John to analyze the cause of outbreak to stop the spread of this epidemic disease.



In the above image, we can see the initial time-series analysis showing the count of casualties by date, which doesn’t provide any useful insights about the disease itself.

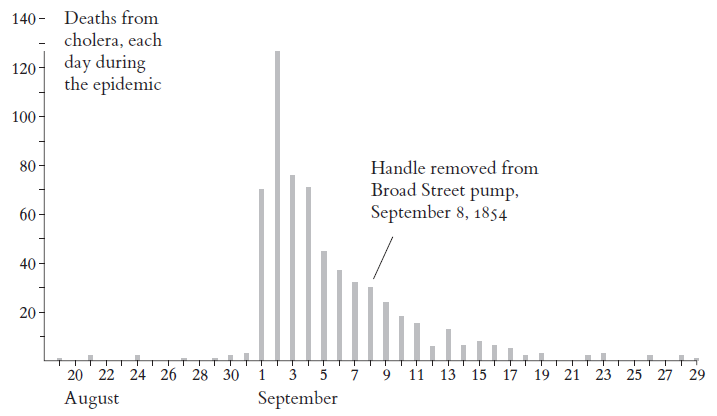
Later by recasting this data, he marked the places at which the death occurred on a map, which opened to a whole world of possibilities of the outbreak.



This map reveals a strong association between the count of deaths by cholera and the proximity to the pump in Broad Street, in a context of simultaneous comparison with other local water pumps and the surrounding neighborhoods without cholera.

John explained about this discovery to the concerned authorities of the community water supply. The management heeded to his analysis and had took measures to remove the pump-handle on the Broad Street well immediately. The epidemic soon ended.

However, it can’t be concluded that the epidemic ended because of the removal of the infected water pump. There isn’t enough evidence to state that.



In the above chart, we can see that the pump was only removed on September 8th, however, the fatality rate has been declining since September 3rd even before the pump was removed. Besides, the map that showcased the huge number of deaths around the well was highly populated as compared to the other areas surrounding it, yet again diminishing the suspicion because of the loss of evidence.

Though John had done a great job for his time by drawing a lot of interesting insights from the map he had plotted, we can see that his approach lacked enough evidence where there could have been possible alternatives for the outbreak. With the technological advancements we have today, it could have been analyzed and presented in a much better way.

As opposed to his conventional methods of plotting them on a map which could have a poor accuracy because of the human errors, we could have done it with the help of several advanced visualization software with a far better accuracy and thorough analysis that lacked evidence in his research. Though John had done both quantitative and qualitative analysis, he failed to prove a fundamental aspect of what he is comparing it with. More research should have been done on this by comparing his analysis with some of the other cholera outbreaks to see if his discovery and analysis stood true.

"It is such an irony that Snow's best thing about his work is also the most uncertain part of evidence that he has". - Tufte, E.R. (1997). While the data visualization techniques handled by John may have worked out for the best to end the epidemic in London, it still is not perfect because of the lack of concrete evidence. Yet we can’t deny the fact about some of his breaking discoveries which helped them in their crisis about the cholera’s transmission though.

# 2. The Decision to Launch the Space Shuttle Challenger in 1986

Seven astronauts died in the explosion of the space shuttle Challenger on January 28, 1986 due to the poor decision of the NASA officials to launch the space shuttle despite the temperature warnings by the rocket engineers. Although, in this case, the NASA officials aren’t the ones to be entirely blamed. It is also because of the poor report that was provided by the rocket engineers to stop the launch, which weakened their warning about not launching the space shuttle at such cold weather as it could affect the resiliency of the O-rings.

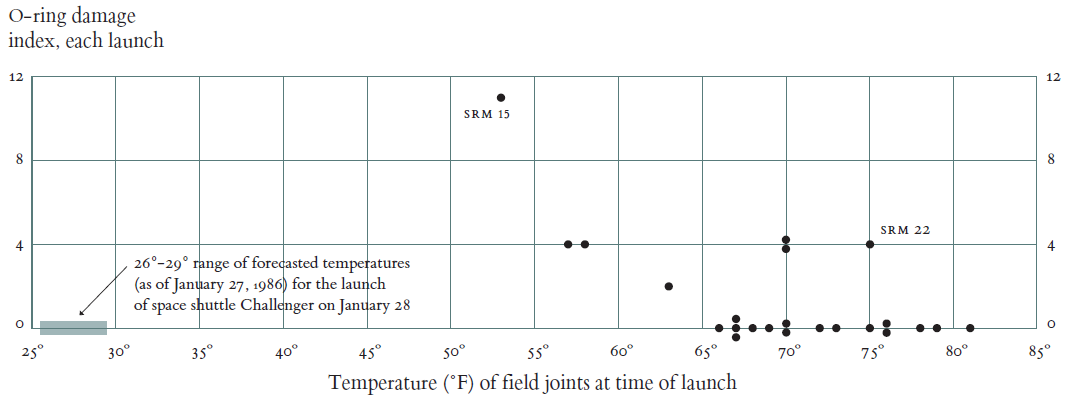
The cause of this accident was so evident that it was because of the failure of O-rings. Within a second after ignition, smoke emerged at the joint of the right booster which indicates that the O-rings have leaked and burned through and failed to seal. Later in few seconds, flames emerged from the leaky joint and engulfed the fuel tank, which resulted in the rupture of the tank and exploded, destroying the shuttle. The O-rings lost their resiliency due to the cold temperature as feared - Tufte, E. R. (1997).

Despite their serious debate on this matter before the launch, the NASA officials pointed out the serious weaknesses in the charts that the engineers had provided which recommended against the launch of the shuttle. The engineers reconsidered their suggestion and changed their mind to launch the shuttle, as the team felt that the evidence they provided was inconclusive that cold temperatures had a link with the O-ring problem and were convinced to launch the shuttle. A better data visualization and report analysis by the rocket engineers showing the criticality in the link between the O-rings and the temperature could have prevented this mishap. The 13 charts that were prepared by the engineers and Thiokol to show the temperature concerns failed to convince the NASA officials from launching as it didn’t provide any evidence of correlation between the temperature and resiliency of the O-rings.

The report had its own weaknesses like, the first title-chart did not provide the names and details of the people who had done the research to prepare the material. This affects the credibility of the source, and the author’s reputation could have played a major role, as it could provoke some doubts about the evidence. The second chart details about the history of the erosion of O-rings. Though it displayed the varying damages to the O-rings, none showcased how catastrophic it could be with regards to the effect of temperature on these O-rings, thereby failing to prove the link between the temperature and the resiliency of the O-rings. It is to be also noted that the rockets had different names. The third chart details about the interaction between the primary and secondary O-rings as a back-up during the erosion. Although it denotes the severity of the degradation of the O-ring seal due to rotation, it stands irrelevant to support the main cause that is the temperature. Two more charts talked about the “Blow-by history”. Although it showed a detailed damage report, it was irrelevant to address the main concern.

However, in the case of SRM 15, a substantial O-ring damage was reported due to the cold weather conditions. This, however, was not significant enough to be considered with just one case of evidence. Besides, there was also another case to weaken this further, where the SRM 22 that was launched on a warm day also had a blow-by history.

This argument could have still been won if the blow-by histories were carefully examined, as they profoundly differed. A clear indication of cold launch danger could have been seen, as these charts define the database for decision. Even the temperature chart had several missing values, thereby diminishing the evidence. The recommended temperature for launch, ≥ 53º F stood insignificant as well, because the 53º launch itself had a considerable erosion. These 13 charts failed to abort the launch, as they couldn’t showcase the correlation between O-ring distress and temperature. They didn’t have sufficient data to quantify the cold.



In the above scatterplot prepared by Tufte, we can see the serious risk of launch at 26º -29º. Every launch below 66º resulted in damaged O-rings, while on warmer days, only few resulted in erosion - Tufte, E. R. (1997).

There were major discrepancies between the tasks in hand and failed to show the current risk. The data and chart visualizations failed the case, as the chart-makers should have done a better data representation to make their points clear with enough evidence to support.

# Conclusion

In both the cases, data visualization played a major role. They could aid in the understanding of data in a right or wrong manner, which in turn could have severe consequences like the case studies we reviewed above. It could be as helpful as it was in the cholera epidemic case, if done right or it could be as catastrophic as it was in the explosion of space shuttle case, if done wrong. Hence, it plays a crucial role.

# REFERENCES

1. Tufte, E. R. (1997). Visual and statistical thinking: displays of evidence for making decisions. Cheshire, Conn, Graphics Press.
2. Tukey, J., (1970). Exploratory Data Analysis. 2nd ed. Boston, Massachusetts, United States: Addison Wesley Publishing Company