ALY 6080 Module 2 Assignment

Student's name: Mohammad Hossein Movahedi

Assignment title: Module 2 Assignment — XN Project: Pros & Cons

Course number and title: ALY6080 71445 Integrated Experiential Learn SEC 12 Fall 2022 CPS [TOR-1-OL]

Term: 202315_1 Fall 2022 CPS Quarter Full Term

Instructor's name: Matthew Goodwin, Ph.D.

Oct 2, 2022,

Introduction

In this assignment, I will go through the following:
What are the strengths and weaknesses of the data set the sponsor provided as it relates to Business Performance Impact? We want you to consider the good and bad of data use.

How can data be used to tell the story positively or negatively?
How can statistics be skewed?
What are the Causes of Bad Data?
How Can You Manage Bad Data Efficiently?

How can data be used to tell the story positively or negatively?

With infographics and data visualization, journalists have effectively conveyed stories. The spectator is better able to understand the data through storytelling. If so, then perhaps the purpose of this visualization should be to enable exploratory data analysis (EDA) rather than to provide information. What is the audience's knowledge of the subject? The content of a visualization ought to be determined by the facts, not by your preferences.

The amount of the effect in the graphic divided by the size of the effect in the data is what John Tufte called the "Lie Factor." Alternate representations of the data in the same visualization, such as multiple clustering, confidence intervals in place of lines, shifting timeframes, and variable scale, can add balance. Unless you are certain that you are providing your audience with the greatest picture of what the data "says," don't be picky about the data you include or exclude. Avoid falling into the "that looks cool" trap since it might not be the greatest way to use compelling storytelling to explain the data (Harvard Business Review, 2013).

How can statistics be skewed?

Data skewness is one of the essential characteristics, especially when determining the data distribution. There are numerous sorts of skews and distributions. A longer or fatter tail on the left side of the distribution is referred to as negative skew. Positively skewed data will have a mean that is higher than the median. The situation is completely reversed in a negatively skewed distribution. Zero skews happen when a data graph is symmetrical (Investopedia, 2022).
 I used to work with some highly skewed datasets, and the transformation and reasoning for those transformations I had to use gave me a severe headache at the time.

A dataset with a natural source of data is usually very little skewed; however financial data and other computer and human-influenced data are usually, in my experience, more skewed.
After checking the dataset that our sponsor Donfass provided during the initial EDA, no case of extreme skewness has been detected, which is very promising.

What are the Causes of Bad Data?

Bad data is most often caused by human error. Employee data input mistakes reduce the quality of the data. Users adhere to varying data collection procedures in various locations. Data quality is at risk while moving to new systems, especially if the data is dynamic (Caous, 2020).

How Can You Manage Bad Data Efficiently?

Human error cannot be eliminated, but you may start by establishing explicit data quality policies. Investing in a complete Data Management solution is one of the greatest methods of managing problematic data. Access sophisticated technologies that can manage master data effectively and give your data stewards their use (Caous, 2020).

References

Investopedia. (2022). *Skewness: Positively and Negatively Skewed Defined with Formula*. [online] Available at:
https://www.investopedia.com/terms/s/skewness.asp#:~:text=Skewness%20refers%20to%20a%20distortion,is%20said%20to%20be%20skewed. [Accessed 30 Sep. 2022].

Harvard Business Review. (2013). *How to Tell a Story with Data*. [online] Available at:

https://hbr.org/2013/04/how-to-tell-a-story-with-data [Accessed 3 Oct. 2022].

Caous, J. (2020). *How Does Bad Data Affect Your Business?* [online] To-increase.com. Available at:

https://www.to-increase.com/master-data-management/blog/how-does-bad-data-affect-your-business [Accessed 3 Oct. 2022].