Data Analytics Project Deliverable Pt. 2

Introduction

In the United States there is over 600 amusements parks. With just one of those amusement parks being one of the most visited parks in the country and world, having over 17.7 million visitors in just 2023. Having this highly visited park within an hour drive from campus I thought that it might be interesting to see how all these amusement parks perform in terms of safety. This will show whether there are some amusement parks or rides much more dangerous than others, and even if an amusement park itself is fairly dangerous or not.

Objectives

The main goal of this project is to see if there are any connections or anything that stands out in the data of amusement park accidents. This is to try and see if there is any proof of a certain ride being too dangerous or an underlying problem that may be happening at an amusement park that has not been caught. Ideally the project will show that all the rides and parks are safe with the rare possibility of an accident. However, this could also show evidence of an amusement park being much more dangerous than the other parks.

Methodology

I plan to first go through the data determining where everything is and what it all means. I will then go through all of the data cleaning it up and making sure that everything is the proper format and units. To do this I will be using R to help make sure the data is tidy and useable. I will also use excel for certain equations and statistics. This will be done along side with R based on what statistic I am looking at. Using these two softwares, I will create different data columns for a variety of probabilities. This would include the amount of amusement park accidents based on geographical areas, based on specific rides, and based on frequency of injuries.

For a rough timeline of the project:

November 10th: All my data cleaned and tidy with initial basic statistics. Along with a brief summary explaining some of those statistics.

November 17th: A few visualizations showing data in a clear and concise manner. A few charts for explaining and helping to show some analysis and a few charts for comparisons. This would be done along with any revisions that need to be made to any of the data or statistics done before.

November 24th: A final draft of all my results found with my conclusion based on any statistics or data I discovered during my research and time working.

December 3rd – 5th: Presentation to the class about my data and final report.

December 6th: Final zip file with proper file structure.

Expected Outcomes

The main output of the project will be a conducive report explaining any conclusions made in the project. Supporting this report will be charts and graphs showing detailed insights and comparisons across the data. Making it easier for others to view the data at a glance and make their own analyses. A short presentation will also be made and presented in a five to eight minute presentation. This will give a very brief overview of how I conducted my research and a brief conclusion of my final report. Hopefully there by showing in a clear manor if there are any true dangers to riding an amusement park ride.

Challenges and Limitations

The first logistics challenge for the project will be trying to clean the data and make it all tidy. This is partially due to the fact that most of the data is in completely different units. Another factor that could cause issues is differing different companies. The accident reports do not say exactly what company or ride an accident happened on, but instead is just the general information concerning the accident. This could make it very hard to distinguish some rides given that some amusements parks are built fairly close to each other.

Conclusion

Amusement parks are loved by almost everyone, having something to enjoy for all kinda of different people. This love and fondness for these parks and the thrill of some rides can be amazing and a little blinding. Making the need for so many safety inspection jobs, and why safety reports like the one that is proving data for this project, are incredibly important. These reports allow us to know that someone is checking and watching, but it is also important to know that everything is actually safe. Which is why I hope that this project shows that there is nothing dangerous about amusement park rides.

Sources:

https://www.statista.com/topics/10431/amusement-and-theme-parks-in-the-us/#topicOverview