

Data Visualization

Michael Tennant, Michael Fahr, Bryce Mendenhall, Pao Yang, Rafael Del Carmen
University of Arkansas, College of Engineering: Computer Science and Computer Engineering



Introduction

Large amounts of data are generated on an everyday basis and can be identified in different categories. At Sorcero, it could be about the corpus or the user interactions with the corpi. With such a large amount of data, it is hard to navigate and be able to find the information that is needed.

Purpose

The purpose behind this project was to help find insights on key performance indicators (KPIs). It is important because it will assist the user in processing and understanding the data shown. By being able to navigate through the data in an understandable way, it becomes easier to detect trends or patterns. It also communicates the data quickly and effectively to other people who may not be familiar with the information

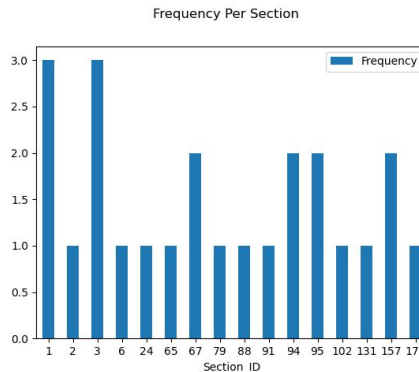
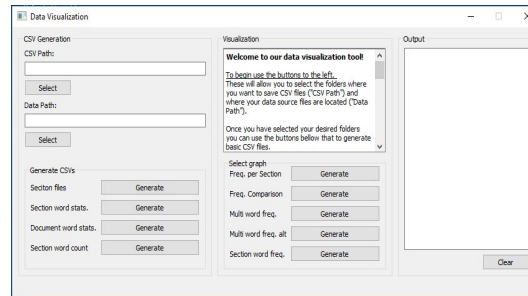
Methods

The approach is to design and implement a program that collects the data, develops relations in the data using a set of algorithms, and outputs a straightforward visual context based on the information provided.

Results

CSV Generation: The ability to generate a CSV based on certain characteristics desired.

Graph Generation: Different graphs could be created depending on the certain output that is desired.



Conclusion

The goal of the Data Visualization project was to develop an application that could effectively parse large volumes of data gathered from corpus, analyze the data using statistical analysis and natural language processing, create meaningful visualizations from the data, and present these visualizations to the user. Overall, we were successfully able to create this application, and could perform the data analysis and visualization on any provided data file.

Multi Word Frequency

