CSCE 240 Project 1 Report

Charlie Grant

The assignment was to use any language to create a program to read a file and obtain a word count, character count, line count, and the amount of “parts”. I found the word, character, and line counts to be quite easy. I applied logic I would have used in Java and looked up how to translate these concepts into C++. I chose C++ simply because I wanted more practice with the language for the upcoming homeworks.

The hard part of the coding aspect is to separate by parts for the data because of the difference in formatting between the two documents. These formatting changes arise from the technicality that the Sony information is a 6-K document which is a 10-K report for companies whose headquarters are based outside of the US. I looked if there was a report of their information that could be transferred into a 10-K format but I found no evidence that they had any document in a similar format. Now theoretically I could just target the words they use to separate the parts but that requires a check of a string that must be carried out every advance of a single character. Instead I checked for if a line was blank. If it was, I stored it as a string to use for an if statement. Inside that if statement I checked the line after for the word part or any common synonyms, and if one of the words was present with a number after it I counted it as a new part.

Another difficulty came from my own stupidity. For at least an hour I struggled to get the expected result for the output of either document. This is because I had downloaded each document as code instead of a text document, meaning there were no lines that met the conditions set to increase the part counter. This was fixed by simply just downloading the files properly.