

Angela Alaisa

CSCE 240, Section 001

Assignment 2 Report

For programming assignment two, we had to create a program that extracts data based on questions.

The first part of the program prompts the user to enter "District 74" which verifies that they want to access information for the representative of District 74, then asks for an information type. There are six that the user can choose from: Contact Information (I1), Personal Information (I2), Committee Assignments (I3), Sponsored Bills (I4), Voting Record (I5), and Service in Public Office (I6). I created multiple variables in order to store the information once it is extracted. This information type that the user inputted is then passed to a method that does the following:

- Reads District 74 website source code line by line
- Outputs information to console based on user input
- Writes the same information to an output file

In order to sort through the code, I created regular expressions each information type that looked for patterns that these types of information had in the source code for the website. One tool that I used which I found really helpful was regex101.com. I used this website to formulate and look at what regular expressions to use in order to match a certain set of words. If a line matched any of the regular expressions, it would then store that match into the corresponding variable. This process would repeat until the source code does not have any more lines left to read. If the user gives an input that is not one of the accepted information types, then the program will terminate and ask them to give a valid input.