Tracy Michaels

Tmichaels1 002430918

Big Data Programming Assigment 2

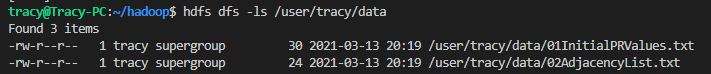
Creating the runnable jar file I used the command



Graphical user interface, text

Description automatically generated

After placing the 01InitialPRValues.txt and 02AdjacencyList.txt files into the hdfs using Hadoop commands hdfs dfs -put <local path> /user/tracy/data for both files



I was then able to run the jar file with the following command



Where the first parameter passed is the location of 01InitialPRValues.txt on the hdfs, the second parameter passed is the location of 02AdjacencyList.txt, the 3rd parameter passed is the local file location for the outputs and the last parameter is the number of iterations.

When the code if first run it takes those arguments in the first lines of code

Text

Description automatically generated

Then begins the main loop of calculating the page rank of the values and the output of those are put into separate folders in the output folder which are then fed to the next iteration

Text

Description automatically generated

Unfortunately the actual implementation of the pagerank algorithm ins this context I was unable to figure out using it in java like this. Even after pouring over the source code and other resources, asking for help, for weeks I just couldn’t figure out the actual implementation even though I understand the concepts behind the algorithm. Thank you for understanding.