Jimmy Tran

CSC 4760

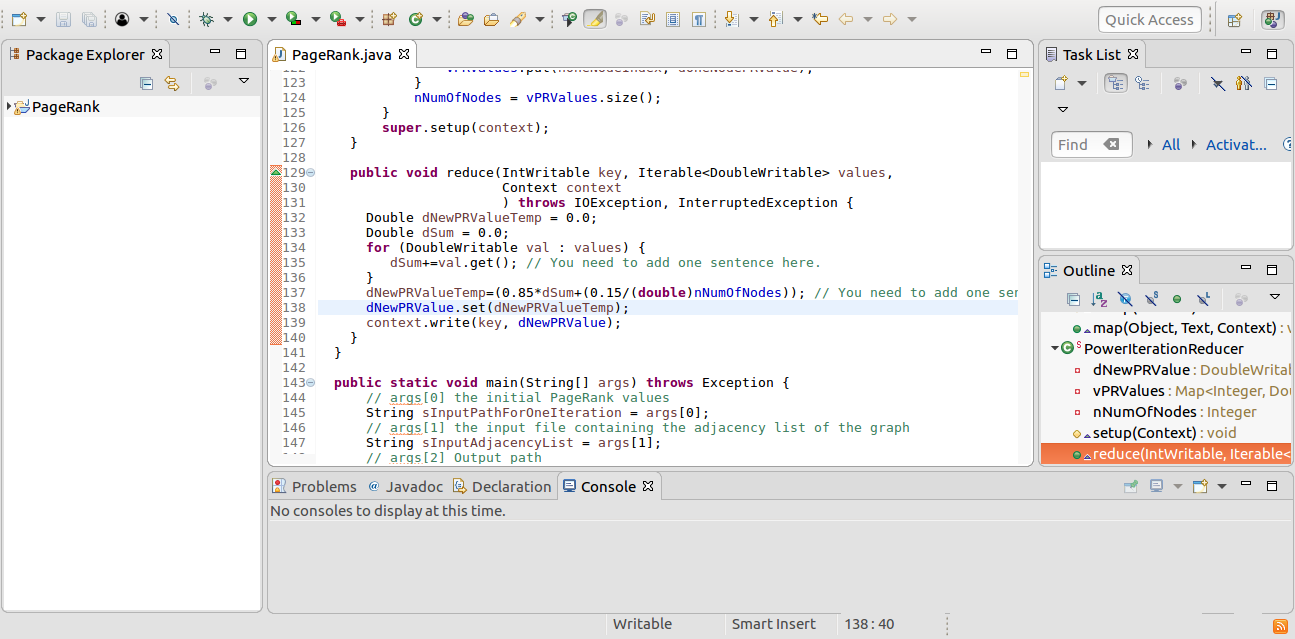
21 February 2018

Assignment 2 Report

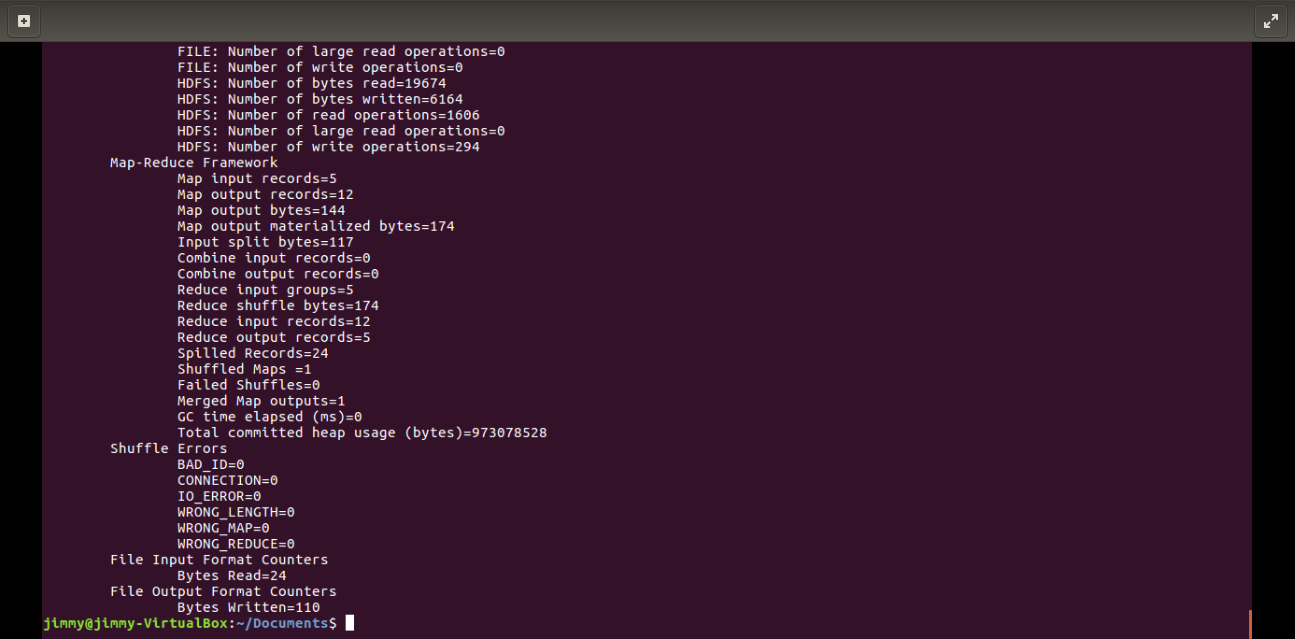
* 1. The mapper is defined with a key and value as input (which are the two text file inputs) and then a third parameter for the context in which it is used. It outputs intermediate results in the form of another key and value which are the nodes and respective PageRank value respectively. From the mapper, there may be multiple similar keys output.
  2. The reducer likewise is defined with a key, value, and context. The key and value are also the node and its Page Rank. Unlike the mapper though, the reducer uses a for-each loop to sum all the mapper outputs that share a key (node). Then it applies the Page Rank equation to the sum to compute the output value. As such, the reducer returns fewer <key, value> pairs than the mapper.
  3. No. There is no combiner in this implementation. A combiner could probably be implemented, but for such a small iteration number there probably would not be much of a performance improvement. Likewise, it would probably share its code with the reducer making it redundant.

2.1) Screenshots

Lines added to complete code



Execution and Output.





2.2) The exact Page Rank values are not given by this implementation, likely due to the low number of iterations (only 30). The errors are not very large though since adding outlinks doesn’t alter the page ranks much. This is shown by the error bounds for the power iteration method and its residuals: ||x(k)-π||1,∞­≤2ak, ||[x(k)]TG-[ x(k)]T||1,∞­≤2ak.