Map Reduce Report

**Instructions on how to run the program:**

Run the program as follows: ‘python3 filename’

**Problems and how I overcame them:**

The main problem I had was opening the files within the parallel section of the code. If I tried doing this, the code will be stuck with no error. To work around this, I opened the files before the parallel section and put them into lists and then distributed the lists between the processes.

**Time spent on the assignment:**

I spent about 3 and a half hours on this assignment with most of it trying to deal with distributing the files across the processes.

**Performance measurements:**

**1 thread:**

**Text

Description automatically generated**

**2 threads:**

**Text

Description automatically generated**

**4 threads:**

**Text

Description automatically generated**

**8 threads:**

**Text

Description automatically generated**

**Behavior analysis:**

This program increases roughly a second or so when doubling the number of threads. In a normal case this would likely be because of opening the files multiple times within the threads. However, in this case, I am dealing with files before the parallel section so the extra time may be from the distribution of the words between the threads as they become available.

**Output from the cpuInfo.sh program:**

model name : Intel(R) Core(TM) i9-9900K CPU @ 3.60GHz

4 36 216