The following assumes that you are in a Linux environment (you should be…)

Requirements: You need to have Maven installed (if sudo apt-get install maven, then follow the instructions at https://maven.apache.org/install.html).

Note: If you don't have Maven and want to be lazy, refer to the precompiled jar that we included (more on this later).

Step 1: Clone repository (git clone https://github.com/jramsdell/cs753_team2_assignment1.git)
Step 2: Compile jar using maven (if you want to skip this step, use the .jar in the bin/ directory). You can run the compile.sh script to compile the jar (no arguments). It will be located in target/.

Step 3: Run the indexer. Do jar -jar target/team2_1-1.0-SNAPSHOT-jar-with-dependencies.jar index CBOR

Where CBOR is the location of the paragraph cbor file that will be used to make a Lucene index out of. A new directory (paragraphs/) will be created.

If you can't compile, instead use jar -jar bin/backup_program.jar CBOR

Step 4: Run the queries. Do jar -jar target/team2_1-1.0-SNAPSHOT-jar-with-dependencies.jar search paragraphs/

This will point to the index directory and print out the query results (for all three queries).

Like before, you can use bin/backup_program.jar instead.

Step 4: Run the queries using the custom similarity function. Do jar -jar target/team2_1-1.0-SNAPSHOT-jar-with-dependencies.jar custom paragraphs/

This will point to the index directory and print out the query results (for all three queries). Ths uses the custom similarity function.