Bobby Chisholm, rwc1007, 965700793
Bobby worked on the calcRunner, the U-L similarity, creating the graphs in the report, we both worked on answering the questions, and the interaction of the similarity with the searcher and both worked on tracking the total term vectors
Tony Barboza, ajb1047
Tony worked on the produceInput.sh, U JM implementation, UDS implementation, we both worked on answering the questions, both worked on tracking the total term vectors

Program 4 Analysis

Method	Precision@R	Ndcg 20	Map
Default(bm25)	0.5989 **SE=** 0.0011556994	0.7716 **SE=** 0.001039002	0.6034 **SE=** 0.0012251515
U-L	0.5187 **SE=** 0.0013678658	0.6625 **SE=** 0.001475285	0.5085 **SE=** 0.0014641082
U-JM	0.4992 **SE=** 0.0013397383	0.6280 **SE=** 0.0014432444	0.4773 **SE=** 0.0014214818
U-DS	0.5107 **SE=** 0.0013272078	0.6498 **SE=** 0.001433148	0.4907 **SE=** 0.0014577596

Key: **SE- Standard Error**

*We did not complete a successful TF-IDF implementation that is why we did not include it in the tables.

*The error bars are barely visible because the standard errors were so small.
Standard Error = N/sqrt(standard deviation)

MAP

NDCG 20

Precision@R
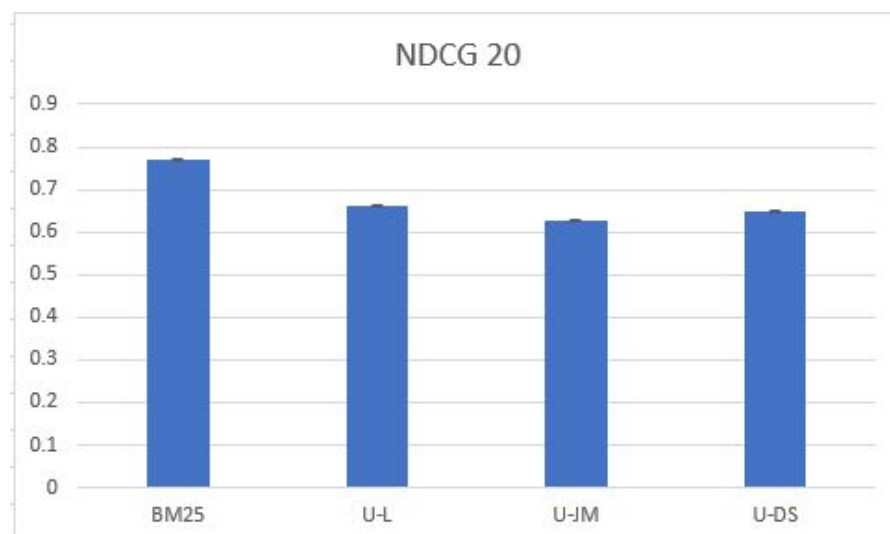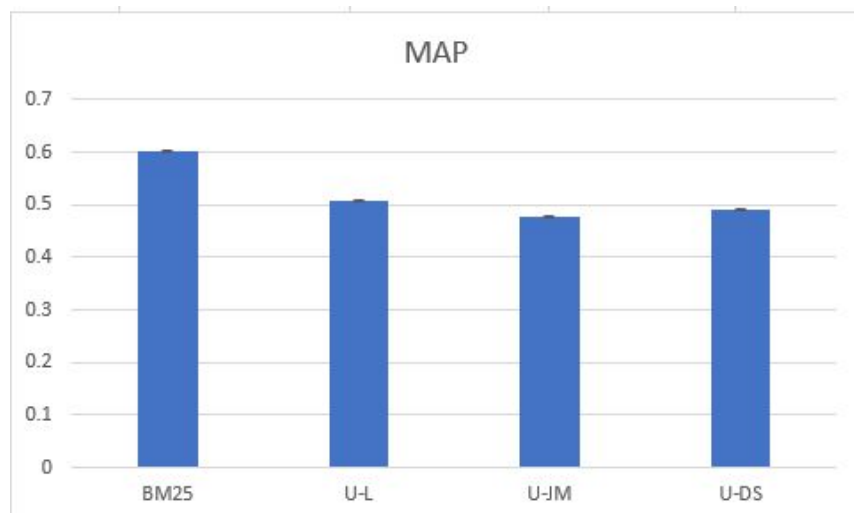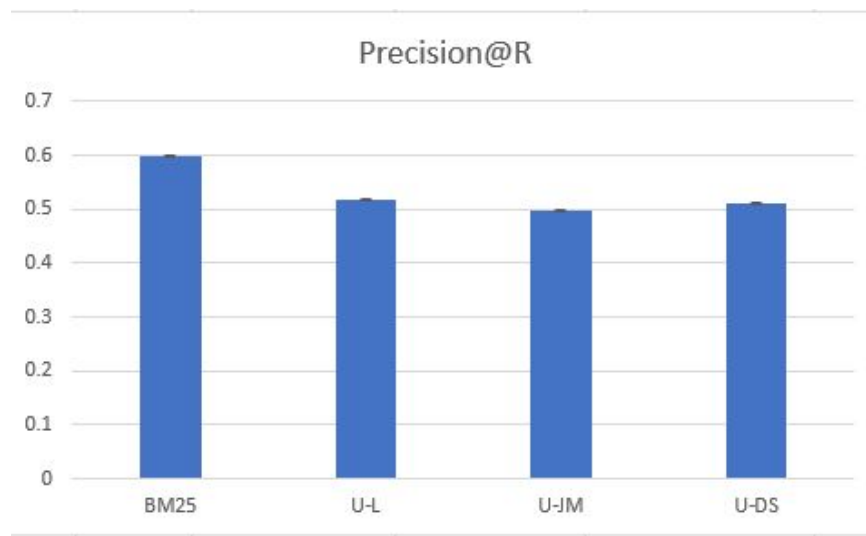
Questions

1.) Which of these variants performs best?

BM25 performs the best across all evaluation metrics. Between the methods we implemented on this program U-L had the better overall scores than the other 2.

2.) Comparing to all variants including TF-IDF?

We did not create a successful implementation of TF-IDF last program but through research we found that BM25 performs the best out of all of the TF-IDF variants we were asked to implement.

3.) How does stemmer/tokenizer effect search results?

Grad student question.

4.) **Using the standard error is the difference significant?**

The standard error values throughout are all low in value, the difference is not significant. In addition, the standard error is really similar: standard error values are all within .00009 of each other.

5.) Evaluation measures focus on precision and recall. Do different evaluation measures agree on what method is best?

Different evaluation measures do not agree on what method is the best. Precision is a focus on the documents that you retrieved. It can be taken advantage of by retrieving no documents. Recall can be taken advantage of by retrieving all the documents in the corpus. They do not agree on what method is best because of this principle.

6.) For the query with ID Brush%20rabbit, display the contents of the top-ranked paragraph foreach method. Which of these paragraphs is most relevant?

BM25

The brush rabbit inhabits dense, brushy cover, most commonly in chaparral vegetation. It also occurs in oak and conifer habitats and it will live in brush or grassland, and form networks of runways through the vegetation. The brush rabbit does not dig its own burrow or den, but uses the burrow of other species, brush piles, or forms. In the San Francisco Bay Area, it was found that the brush rabbit concentrates its activities at the edge of brush and exhibits much less use of grassy areas. It uses the interior brush of the wilderness and it was also found that this may be a better environment for it than the chaparral one. Studies done on the brush rabbit in Oregon also showed that it rarely left the brushy areas it inhabits. Brush may be used more in the drier seasons while grasses are used in the wetter seasons in relation to growth of annual vegetation. Use of habitat also probably is related to the breeding season.

U-L

The brush rabbit inhabits dense, brushy cover, most commonly in chaparral vegetation. It also occurs in oak and conifer habitats and it will live in brush or grassland, and form networks of runways through the vegetation. The brush rabbit does not dig its own burrow or den, but uses the burrow of other species, brush piles, or forms. In the San Francisco Bay Area, it was found that the brush rabbit concentrates its activities at the edge of brush and exhibits much less use of grassy areas. It uses the interior brush of the wilderness and it was also found that this may be a better environment for it than the chaparral one. Studies done on the brush rabbit in Oregon also showed that it rarely left the brushy areas it inhabits. Brush may be used more in the drier seasons while grasses are used in the wetter seasons in relation to growth of annual vegetation. Use of habitat also probably is related to the breeding season.

U-DS

The brush rabbit inhabits dense, brushy cover, most commonly in chaparral vegetation. It also occurs in oak and conifer habitats and it will live in brush or grassland, and form networks of runways through the vegetation. The brush rabbit does not dig its own burrow or den, but uses the burrow of other species, brush piles, or forms. In the San Francisco Bay Area, it was found that the brush rabbit concentrates its activities at the edge of brush and exhibits much less use of grassy areas. It uses the interior brush of the wilderness and it was also found that this may be a better environment for it than the chaparral one. Studies done

on the brush rabbit in Oregon also showed that it rarely left the brushy areas it inhabits. Brush may be used more in the drier seasons while grasses are used in the wetter seasons in relation to growth of annual vegetation. Use of habitat also probably is related to the breeding season.

U-JM

The brush rabbit feeds mainly on grasses and forbs, especially green clover, though it will also take berries and browse from bushes.

Which is most relevant?

The document returned by BM25, U-L and U-DS is the most relevant because it gives a very descriptive document on the query and 3 of the retrieval methods retrieved it. U-JM retrieved a much shorter document that is less relevant, it was retrieved because the document is so short that the query terms are a large portion of the document.