CS 446 Project 2 Analysis Questions

1. **A table of the "full" *ql*, *bm25*, and *dpr* runs:**

Query	QL	BM25		DPR	
135802	0.1164	0.1176	1.0%	0.0546	-53.1%
23849	0.0151	0.0186	23.2%	0.2391	1483.4%
336901	0.0634	0.0634	0.0%	0.1708	169.4%
768208	0.2353	0.2433	3.4%	0.0623	-73.5%
914916	0.3638	0.2860	-21.4%	0.4061	11.6%
1110678	0.4205	0.3262	-22.4%	0.0201	-95.2%
332593	0.2522	0.2310	-8.4%	0.2221	-11.9%
47210	0.1997	0.2021	1.2%	0.3692	84.9%
877809	0.1914	0.2516	31.5%	0.2097	9.6%
1106979	0.6340	0.5034	-20.6%	0.5401	-14.8%
583468	0.6706	0.7267	8.4%	0.7084	5.6%
640502	0.1199	0.0878	-26.8%	0.1880	56.8%
156498	0.0564	0.0701	24.3%	0.1348	139.0%
258062	0.0314	0.0317	1.0%	0.1640	422.3%
1071750	0.2587	0.2685	3.8%	0.2944	13.8%
1131069	0.0288	0.0856	197.2%	0.2143	644.1%
390360	0.3275	0.2741	-16.3%	0.2375	-27.5%
67316	0.0080	0.0151	88.8%	0.0853	966.3%
1056416	0.0000	0.0000	0.0%	0.0000	0.0%
1105792	0.3999	0.3840	-4.0%	0.1988	-50.3%
1113256	0.4953	0.4969	0.3%	0.4651	-6.1%
405163	0.0787	0.0735	-6.6%	0.0013	-98.3%
169208	0.1319	0.1075	-18.5%	0.1028	-22.1%
730539	0.2034	0.1356	-33.3%	0.1628	-20.0%
938400	0.1660	0.1043	-37.2%	0.3848	131.8%
1136769	0.0000	0.0000	0.0%	0.0000	0.0%
1133579	0.6677	0.6666	-0.2%	0.7530	12.8%
1051399	0.0201	0.0113	-43.8%	0.1348	570.6%
1136962	0.4689	0.4879	4.1%	0.4199	-10.4%
1136047	0.0666	0.0464	-30.3%	0.0623	-6.5%
1108651	0.0547	0.0250	-54.3%	0.2464	350.5%
1049519	0.0000	0.0000	0.0%	0.0000	0.0%
701453	0.5663	0.5531	-2.3%	0.3160	-44.2%
1037496	0.4259	0.3181	-25.3%	0.2945	-30.9%
1119543	0.0000	0.0000	0.0%	0.0000	0.0%
940548	0.0000	0.0000	0.0%	0.0000	0.0%
940547	0.0868	0.0892	2.8%	0.3152	263.1%
1121353	0.2557	0.2349	-8.1%	0.1002	-60.8%
911232	0.2038	0.1542	-24.3%	0.1592	-21.9%
1109707	0.1502	0.1750	16.5%	0.1376	-8.4%
1064670	0.2233	0.2312	3.5%	0.1521	-31.9%
1136043	0.0976	0.1569	60.8%	0.3695	278.6%
1043135	0.1128	0.1031	-8.6%	0.1281	13.6%
1122767	0.3460	0.3235	-6.5%	0.2052	-40.7%
174463	0.0041	0.0301	634.1%	0.1838	4382.9%
42255	0.1987	0.2625	32.1%	0.4411	122.0%
1132532	0.1666	0.1044	-37.3%	0.2442	46.6%
1030303	0.5014	0.5014	0.0%	0.1939	-61.3%
1127540	0.2693	0.2764	2.6%	0.1705	-36.7%
1103153	0.0000	0.0000	0.0%	0.0000	0.0%
555530	0.0084	0.0134	59.5%	0.2544	2928.6%
330975	0.1675	0.2357	40.7%	0.5441	224.8%
121171	0.6916	0.6956	0.6%	0.2195	-68.3%
118440	0.0041	0.0048	17.1%	0.0084	104.9%
141630	0.3959	0.4690	18.5%	0.4309	8.8%
1115210	0.0915	0.0887	-3.1%	0.0651	-28.9%
673670	0.0324	0.0416	28.4%	0.0009	-97.2%
1108729	0.0000	0.0000	0.0%	0.0000	0.0%
1116380	0.0396	0.0111	-72.0%	0.0587	48.2%
324585	0.0449	0.0369	-17.8%	0.3791	744.3%
997622	0.0748	0.0524	-29.9%	0.1313	75.5%
all	0.1952	0.1886	-3.4%	0.2091	7.1%

2. **Discussion about the table:**

   The table compares the effectiveness of three retrieval models on diverse queries: Query Likelihood (QL), BM25, and Dense Passage Retrieval (DPR). The query name, QL score, BM25 score, and DPR score are the four columns in the table.

   According to the findings, DPR consistently surpasses QL and BM25 on most queries. There are some queries, nevertheless, when QL or BM25 outperforms DPR. For instance, BM25 outperforms QL and DPR considerably on query 23849.

3. **MAP of a query with no retrieved documents:**

   Calculating the MAP for a query with no retrieved documents is a hard question because the standard formula for MAP involves dividing by the number of documents retrieved, which will be zero in this case.

   In order to calculate MAP, we need both the relevance judgments for each document and the rank of each document in the retrieved list. When there are no retrieved documents for a query, there are no ranks, and it's impossible to calculate the MAP.

   I feel like there are two options to handle this:

   a. MAP for a query with no retrieved documents should be undefined or not applicable. In this case, it would be better to report a separate metric such as recall, which measures the fraction of relevant documents that were retrieved among all the relevant documents in the collection.

b. Assign a score of 0 to the MAP in this case, as a way to indicate that the system did not retrieve any relevant documents. This approach assumes that the user is interested in both the presence and absence of relevant documents and that retrieving no documents is different from retrieving irrelevant documents.

I prefer option 1. This is mainly because the MAP can be 0 if no relevant documents are found. Then we won't be able to distinguish between cases when no relevant documents are retrieved and when no documents are retrieved. If we assign a NAN or NA value, we know that no documents are retrieved.

4. **Precision-Recall graph for query: "330975"**