Sheet0

Q1

Draw the inverted index that would be built for the following document collection. (See Figure 1.3 for an example.)

Doc 1   new home sales top forecasts
Doc 2   home sales rise in july
Doc 3   increase in home sales in july
Doc 4   july new home sales rise

Q2

Consider these documents:

Doc 1   breakthrough drug for schizophrenia
Doc 2   new schizophrenia drug
Doc 3   new approach for treatment of schizophrenia
Doc 4   new hopes for schizophrenia patients

a. Draw the term-document incidence matrix for this document collection.
 b. Draw the inverted index representation for this collection, as in Figure 1.3 (page 7).

Q3

For the document collection shown in question 2, what are the returned results for these queries:

a. schizophrenia AND drug
b. for AND NOT (drug OR approach)

Q4

Implement a simple Boolean information retrieval for indexed data question 2 and verify the results of question 3