ISR Assignment

1. Form a group having 5 students as members
2. Complete the vector space model given previously as an example to rank the documents for all queries (Query1 to Query 4)(**You may present this in class)**
3. Write a program (preferably Python) that implements vector space model using a document corpus(You can download from TREC conference website)
    - I can give you a downloaded text file if you like

The outputs of the program expected are:

- Number of index terms identified
    - After tokenization, stemming, normalization, stop words removal
- Document Length
- TF(Term Frequency in each document)
- IDF(Inverted Document Frequency)
- Weight(TF.IDF)
- Creating at least 3 Query Terms
- Display the probabilities of each query with respect of each document(using Cosine Similarity Measures)
- Display the ranks all documents with respect to each query

NB:

- **Due Date**
    - **Before 10 May 2025**
- **We will have a lab class starting this week**
- **Please remind me each group's lab session (Schedule) every time you have or at least for the first time)**