1 TF/IDF weighting and the Vector Space Model

Given is a document collection with two documents:

**Document 1:**
A good cook could cook as much cookies as a good cook who could cook cookies.

**Document 2:**
Best chocolate chip cookies recipe.

1. Index both documents using an inverted index. Apply case folding, use stop words list: \{a, as\}. Include TF and DF values at suitable positions in the index.

2. Compute the cosine similarity between these documents and the following queries. Apply the \text{ltc} weighting from the lecture.
   - $Q_1 =$ cookies
   - $Q_2 =$ how to cook chocolate cake

3. Explain the similarity scores! How would the result change with the \text{ln}c weighting?