1. Write a program in the language of your choice to compute the parameter estimates for the multinomial mixture model using the EM algorithm, beginning from arbitrary initial values. The formulas for these are in the class lecture slides.

   Data.zip contains four files:

   - doc_term_matrix.csv contains the document-term matrix for the corpus of US Presidential State of the Union address beginning in the 18th century through today.
   - index.csv indexes the rows of the matrix according to the year of the address and the President.
   - token_dict.csv provides the mapping between column indices (beginning from 0) and stems.
   - tfidf_ranking.csv provides the corpus-level tf-idf ranking of terms.

   Apply your code from above to the data contained in data.zip. Verify that after each iteration the observed data log-likelihood function (i.e. the last expression in slide 16 of lecture 2) increases. Note this property holds for any initial values for the parameters, so you can select whichever you like. A good rule-of-thumb is to set $\rho_k = 1/K$ and randomly draw $\boldsymbol{\beta}_k$, for example from a Dirichlet distribution.