1. Data.zip contains two files containing the output associated with a 45-topic version of LDA estimated on the state-of-the-union corpus from 1945 onwards.

   - topic_description.csv lists the top terms associated with each topic.
   - final_output_agg.csv contains the distribution of topics within each speech.

   Please answer the following questions:

   (a) Can you think of a way of ranking topics according to whether they are more "Democratic" or "Republican"? (The presidents in the sample are: Truman (D), Eisenhower (R), Kennedy (D), Johnson (D), Nixon (R), Ford (R), Carter (D), Reagan (R), Bush Senior (R), Clinton (D), Bush Jr (R), and Obama (D)).

   (b) Based on what you found in the question above, can you rank presidents on the basis of a Democrat-Republican scale? Who was the "most Democratic" president? Who was the "most Republican"?