**Solution**

**Part 1: Corpus Analysis**

Sol - 1. Unigram analysis:

 (a) Mention the total unique unigrams present in the corpus.

 (b) Plot the distribution of the unigram frequencies.

 (c) How many (most frequent) uni-grams are required to cover the 90% of the complete corpus.

Q2. Bigram analysis:

 (a) Mention the total unique bigrams present in the corpus.

 (b) Plot the distribution of the bigram frequencies.