CECS 451
Assignment 8
Total: 60 Points

General Instruction

- Submit your work in the Dropbox folder via BeachBoard. (Not email or in class)

- Submit the separate files as they are. (no zip file)

1. Evaluate the performance of `Google Web Speech API`.

    (a) Read `How Speech Recognition Works.txt` which includes 24 sentences, and record your speech as separate `WAV(PCM)` files using the nomenclature 'Group ID-Native Language-gender-Sent#.wav'. For instance, `01-English-female-Sent01.wav`. I recommend you to use `Audacity` to record and edit your speech.

    (b) (10 points) Upload all audio files to 'BeachBoard - Discussions - Lab - How Speech Recognition Works' by clicking 'Start a New Thread'. Write your **Group ID** at the subject line and attach the audio files.

    (c) (5 points) Complete the `read_original` method that imports a text file into a list of strings, `self.original`, in the order of the sentence number.

    (d) (20 points) Refer this *site*, and complete the `conv_audio` method that converts audio files into a list of strings, `self.recognized`. The method should convert all audio files (`.wav`) in the folder, `inDir`, in the order of `Sent#`.

    (e) (10 points) Complete the `comp_string` method that compares two lists of strings, `self.original` and `self.recognized`, and calculates the similarities of two strings by using *Levenshtein Distance*. (You need to convert the strings into the lists of words. For instance, 'I love AI' to ['I', 'love', 'AI']. Please refer this *site*.) This method should store the separate similarity scores in `self.similarity` in the order of `Sent#`.

    (f) Visit 'BeachBoard - Discussions - Lab - How Speech Recognition Works' again, and select two other threads than yours which include;

        - the same native language but distinct gender of yours
        - distinct native language but same gender of yours

    (g) (15 points) Write a report which includes comparison results using *box-and-whisker plots* by native languages and genders as shown in Figure 1 and Figure 2. You may use `Seaborn`, `Pandas`, or `Matplotlib` to draw the plots.

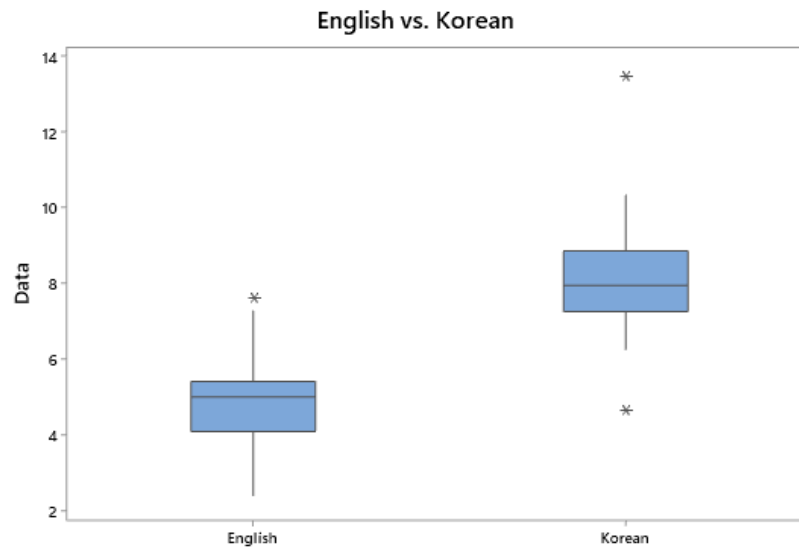   (h) Submit `speech.py` and your report.
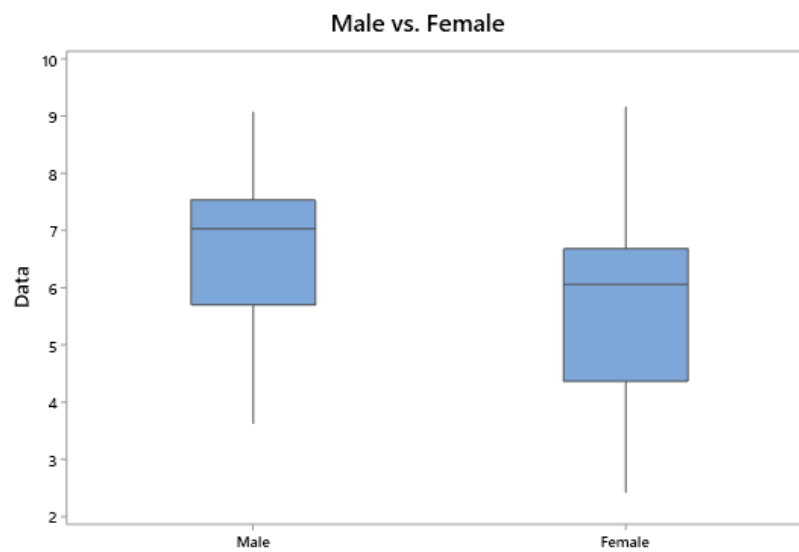
Figure 1: An example of the comparison by native languages

Figure 2: An example of the comparison by genders