Assignment 7: Burrows Wheeler Index
Computational Genomics

Objectives: Become familiar with using the Burrows Wheeler transform (BWT) to find matching substrings.

Tasks:

1. Accept the assignment at `https://classroom.github.com/a/XZ8DF6eW`
2. Clone the repository
3. In `src/bwt.py` implement
    a. `bwt`
        i. Takes a string and returns the BWT of that string
    b. `get_skip_list`
        i. Takes the BWT and returns the an array that gives the offset of each unique character in the first column of the BW matrix
    c. `get_count`
        i. Using the BWT, skip list, and wavelet tree(code provided), return the range of rows in the BW matrix that contain matching prefixes
    d. `get_sampled_sa`
        i. From the original string, return a suffix array that only contains subset of the elements based on the sampling rate (HINT, a sampling rate of 2 has ½ the elements)
    e. `get_offsets`
        i. Use the results from `get_count` and the sampled suffix array find the offsets of the matching substrings
4. Experiment with the size and speed of the BWT index and compare it to your suffix array.
5. Create figures that demonstrate the results of your experiments.
6. Update `README.md` and create a `doc/bwt.tex` (or similar) to include your new experiments.
7. Push your final code to GitHub.
8. Submit your final PDF to Canvas.