day36 article
Due Thursday 11/5/20 by midnight

For today's program, open this repl. Put your name and the date in the comment at the top.

This program is a modified version of the pi frequency calculator from Monday that counted numbers. This program counts how many of each letter shows up. When you run today's version, this is the output:

```
[5, 0, 1, 0, 2, 0, 0, 1, 0, 0, 0, 4, 0, 2, 0, 0, 0, 1, 3, 1, 0, 0, 2, 0, 2, 0]
```

This represents 5 letter As, no letter Bs, 1 C, etc.

You will modify this program to improve it: You will make the output print nicely like we did with the pi digits program on Monday then you will change the text being analyzed to be an article from the Tiger Times, our very own Analy student-run newspaper.

1. First, modify the program so that it considers capital and lower case letters the same using a .lower() method before the for loop. Do this by adding this line:

```
text=text.lower()
```

before the for loop. Run the modified code and you should get the next line here (the A count is now 6, meaning it included the capital A from Analy where before it did not.)

```
[6, 0, 1, 0, 2, 0, 0, 1, 0, 0, 0, 4, 0, 2, 0, 0, 0, 1, 3, 1, 0, 0, 2, 0, 2, 0]
```

2. Next, remove the line "print(tally)" from the program. Then add a for loop to the end of the program that prints the results nicely in 3 columns like we did with the pi digits program. Make the output look like this:

```
a      6   24.0%
b      0    0.0%
c      1    4.0%
( ... I cut out the lines up through the letter z to save space here )
```

Do this by adding the following lines at the end of your program:

```
for x in range(26):
  letter = letters[x]
  count = tally[x]
  perc = count/sum(tally)*100
  perc = round(perc,2)
  print(letter, str(count).rjust(5), str(perc).rjust(6)+"%")
```

These lines go through the letters a through z, pull the count out of the tally list variable, then calculate the percentage that letter represents, then prints it all on a single line spaced out nicely.

3. Add this line to the end:

```
print(sum(tally), "letters processed.")
```

Run your program to make sure it runs well.

(continued on next page)

4. When you have it working, find an article from our school newspaper, The Tiger Times, and run your report on that. Click here to view the newspaper. Select the article on the web site, view the page with the article on it. Select the entire article from the headline down to the last word of the text, then copy the text. Come back to repl, and remove the text "Analy always answers the call!" from between the triple quotes, and then paste the article between the two sets of triple quotes in your program.

In other words, take this line:

```
text="""Analy always answers the call!"""
```

And replace the part between the two sets of triple quotes with the text you copied from the Tiger Times web site.

5. Run the program. It should now present the frequency of every letter from a through z as each showed up in the article. If it doesn't work, take a moment to fix things before trying to move forward.

6. Once it is working, take a quick look at the results. What letter or letters show up the most? Does this surprise you? You don't have to answer anywhere, I'm just asking you to think about it.

7. Go to this web page on Wikipedia https://en.wikipedia.org/wiki/Letter_frequency. This page discusses the relative frequency of each letter of the English language in normal text. Look at the chart on the top right.

Compare the chart with your results. Are they identical? Probably not, but they should be relatively close (for example, it is likely that "e" is the most common letter in your article as well as on the Wikipedia page. Do one or more letters in your article occur dramatically more or less than Wikipedia predicts? Hold this thought, I'm asking you to leave me a private comment about it after you turn in the repl link.

Turn in your program on the Google Classroom.

After you turn it in, make a private comment on this assignment and tell me how your results compare with the Wikipedia results. Are there any interesting differences? I'm asking for a few lines of commentary from you here as a private comment. This is part of the assignment.

If you are having trouble with this program, here is a link to a video where I work it all out.