

CHALLENGES FACED WHILE WRITING THIS SCRIPT

Working on `split.py` required using many tools I had not worked with before, so I faced many challenges in the process. To organize them, I have split them into three categories:

1. Challenges faced coding:

- When I wrote my `write_tags()` function, I forgot to add a blank line between sentences. When I added `print("\n")` outside my loop to write individual tokens, I overcorrected and wound up with 2 blank lines, so I switched to `print()`.
- My text editor showed 2 lines at the end of each file, so I wrapped that `print()` statement in a condition to not execute after the last line. This solved it, but the total count for the `train.tag`, `dev.tag` and `test.tag` was 3 lines short now. I wrote a program to create a composite file with the content of all 3 and compare it to the original corpus, and I realized that the 2 lines at the end of each file were just in the editor and not the file. I removed the condition and the total counts matched.
- When I added `-v` as an optional argument to toggle between printing statistics with `print()` or as log messages, I had trouble getting it to take `-v` without an argument. I tried using `nargs='?', const=True` and `default=False` and it worked, but the argument looked funny in the Help text. More reading got me to `action="store_const"` instead.

2. Issues that stem from using an Online IDE:

- When I first tried running the script with just the `read_tags()` function, it was throwing a syntax error. I fixed this by changing my `#!` Statement to `python3`.
- Using `shasum -a256` gave an error. Google revealed that I had to use `sha256sum` instead.
- Installing `black` also required that I specify my python version to work.

3. A challenge not conquered (yet):

I made two very similar commits, so I decided to remove the one before last, but it was not very simple since I had already pushed them. I tried different things along the lines of branching and `cherry-picking` commits to link, but I had merge conflicts and I wasn't entirely sure how to get everything to work, so I decided to leave it as it was, having created even more unnecessary commits in the process. I will likely recreate and solve this issue on something other than this homework, but in the meantime, I have now learned not to commit compulsively.