

Homework Week 3

Anders Hjulmand

16/09/2021

Github: <https://github.com/ah140797/CDS>

- 1 Identify the names and format of the 3 biggest files. Can you come up with a command to generate a numerically ordered list of 3 biggest files? (hint: try using the wc command)**

I found two ways of identifying the names and format of the 3 biggest files. The first is `ls -s | sort -n -r | head -n 3`. The second is more efficient `ls -S | head -n 3`. The 3 largest files are 9240_Overview_S.RW2, 9247_Overview_SW-RW2 and 9237_Overview_W.RW2.

- 2 Some of the image files are empty, a sign of error in the data processing or corruption. Generate a list of empty photo filenames to facilitate their later replacement easier? (hint: look up find)**

I detected the empty image files using the find command and saved the filenames in a txt-file. `find -name *.JPG -size 0 > empty_images.txt` Using `cat empty_images.txt` i saw that the first two characters of each line were `"/`. I removed these characters with `cat empty_images.txt | cut -c 3- > empty_images.txt`.

- 3 Write a shell script called `emptyfiles.sh` that takes the name of a directory and generates a list of empty photo filenames (remember to pick up all extensions) . So that, for example, running `$ bash emptyfiles.sh Downloads/badphotos` produces a list of empty file filenames either in the terminal or in an external textfile. Can you eliminate the `./` notation at the start of the filename?

First I created a shell script file called "emptyfiles.sh" using `nano emptyfiles.sh`. This opens a text editor where i wrote the following script:

```
for file in $@
do
    find -name "*.JPG" -size 0 | cut -c 3- > list_empty_photos.txt
done
```

`for file in $@` means that the script searches through files in a folder that is specified when executing the script. This makes the script versatile. `find -name "*.JPG"` finds filetypes of JPG, `cut -c 3-` cuts out the first two characters `"/` of each filename and `> list_empty_photos.txt` saves it all in a txt-file. I then executed the script using `bash emptyfiles.sh ~/Desktop/Hw`. This gave me a list of the empty photo filenames in the folder `~/Desktop/Hw`.

- 4 Imagine you have a directory `goodphotos/` (same password as above) with original non-zero-length files sitting at the same level as the current directory. How would you write a loop to replace the zero-length files?

The following for loop iterates through the name of all JPG-files in the Hw-directory. The If-statement test if the filename matches a string in `list_empty_photos.txt`. If it matches, the corrupted photo in the folder `Hw/` is replaced by the original photo from the folder `goodphotos/`. If the file doesnt match, the name of the file is printed instead.

```
for file in ~/Desktop/Hw/*.JPG
do
    filename=$(basename $file)
    if grep -Fxql "$filename" ~/Desktop/goodphotos/list_empty_photos.txt;
    then cp -f ~/Desktop/goodphotos/"$filename" ~/Desktop/Hw ;
    else echo "$filename"; fi; done
```

To verify that the corrupted photos have indeed been replaced, i used the command `ls -s | sort -n -r`. From the output i can see that none of the JPG-files are empty. So it worked i think :)