

5:W35: Managing Files on Steroids with Shell

DESCRIPTION

Your supervisor has shared a [folder of photos on Sciencedata.dk](#) with you (password is 2020CDS, folder is 500Mb and contains 189 images) and needs your help with a couple diagnostics:

Identify 3 biggest files

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: consider using `wc` to gauge image size)

I use the following commands (see screenshot) that are also explicitly described in the table below. The names and formats of the 3 biggest files are:

- 1) 9240_Overview_S.RW2 *
- 2) 9247_Overview_SW.RW2
- 3) 9237_Overview_W.RW2.

All 3 files have the format .RW2 which is a raw camera image taken by a Parasonic LUMIX series camera.

```
photos — zsh — 174x49
Last login: Fri Sep  2 14:28:13 on ttys000
(base) emmaolsen@Emmas-MacBook-Pro ~ % cd /Users/emmaolsen/Desktop/cognitive_science/5th_semester/cultural_datascience/au650627_olsen_emma/hw_4/photos
(base) emmaolsen@Emmas-MacBook-Pro photos % pwd
/Users/emmaolsen/Desktop/cognitive_science/5th_semester/cultural_datascience/au650627_olsen_emma/hw_4/photos
(base) emmaolsen@Emmas-MacBook-Pro photos % ls -laSh
total 1021160
-rw-rw-r--@  1 emmaolsen  staff   14M Jan 18  2017 9240_Overview_S.RW2
-rw-rw-r--@  1 emmaolsen  staff   14M Jan 18  2017 9247_Overview_SW.RW2
-rw-rw-r--@  1 emmaolsen  staff   14M Jan 18  2017 9237_Overview_W.RW2
-rw-rw-r--@  1 emmaolsen  staff   14M Jan 18  2017 9242_Overview_S.RW2
-rw-rw-r--@  1 emmaolsen  staff   14M Jan 18  2017 9241_Overview_S.RW2
-rw-rw-r--@  1 emmaolsen  staff   14M Jan 18  2017 9236_Overview_W.RW2
-rw-rw-r--@  1 emmaolsen  staff   14M Jan 18  2017 9245_Overview_S.RW2
-rw-rw-r--@  1 emmaolsen  staff   14M Jan 18  2017 9246_Overview_S.RW2
-rw-rw-r--@  1 emmaolsen  staff   14M Jan 18  2017 9243_Detail_Trench.RW2
-rw-rw-r--@  1 emmaolsen  staff   14M Jan 18  2017 9243_Overview_N.RW2
-rw-rw-r--@  1 emmaolsen  staff   14M Jan 18  2017 9244_Overview_S.RW2
-rw-rw-r--@  1 emmaolsen  staff   14M Jan 18  2017 9238_Overview_S.RW2
-rw-rw-r--@  1 emmaolsen  staff   14M Jan 18  2017 9236_Detail_Trench.RW2
-rw-rw-r--@  1 emmaolsen  staff   14M Jan 18  2017 9239_Overview_S.RW2
-rw-rw-r--@  1 emmaolsen  staff   14M Jan 18  2017 9236_Detail_Stratigraphy.RW2
[... 186 files ...]
(base) emmaolsen@Emmas-MacBook-Pro photos % wc -c 9240_Overview_S.RW2
14761472 9240_Overview_S.RW2
(base) emmaolsen@Emmas-MacBook-Pro photos % wc -c 9247_Overview_SW.RW2
14733312 9247_Overview_SW.RW2
(base) emmaolsen@Emmas-MacBook-Pro photos % wc -c 9237_Overview_W.RW2
14713856 9237_Overview_W.RW2
(base) emmaolsen@Emmas-MacBook-Pro photos %
```

Step order	Command	What for?
1	<code>cd</code>	Set directory to <i>photos</i> folder
2	<code>pwd</code>	Check that I am in the correct directory
3	<code>ls -laSh</code>	the flag <code>-l</code> means long listing, and <code>-a</code> means that I want to include all files, i.e., also the hidden files. The <code>s</code> indicates that I want to list all files and sort them by size in descending order. The <code>h</code> indicates that I want to output the file sizes in a human-readable format
4	<code>wc -c</code> <code>^name_of_biggest_file^</code>	To get the count of bytes present in a file

2. Some of the image files are empty, a sign of corruption. Can you **find** the empty photo files (0 kb size) , count them, and generate a list of their filenames to make their later replacement easier?

Pipe all empty files to a text (one file name on different line)
`wc` to get lines in the file name

Finding empty photo files and generate list of their filenames: `find . -name '*..*' -size 0 > empty.txt` adding list of empty files to .txt file by piping.

Counting empty photo files: `wc -l empty.txt` to check number of lines (i.e. count number of empty photo files) in the .txt file.

<https://swcarpentry.github.io/shell-novice/04-pipefilter/index.html>

I get 74 from the above command, but the file empty.txt is also in the photos folder, so that means that there are 73 empty image files.

3) **Optional/Advanced:** 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?

```
if [ ! -s /tmp/foo ] ; then
    rm /tmp/foo
fi
```