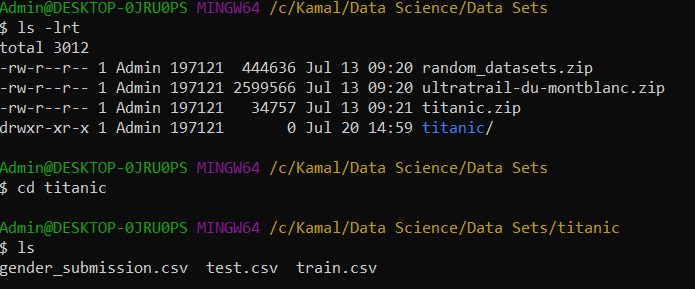
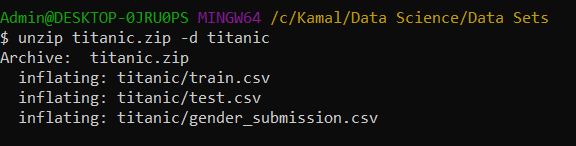
Basic

Use External Tools and Gnu Core Tools to enhance your shell skills. Compose them using pipes and filters.

1.

Unzip (using the terminal) our titanic.zip file to titanic

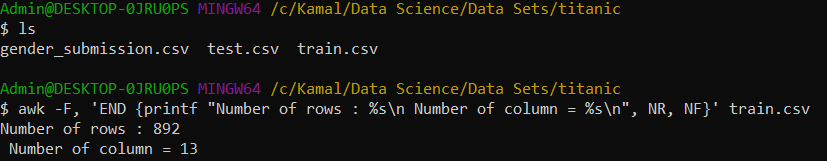
​

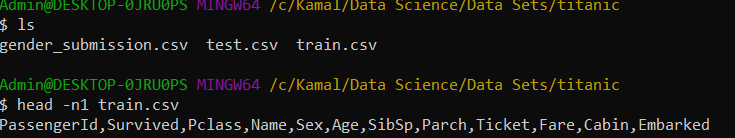
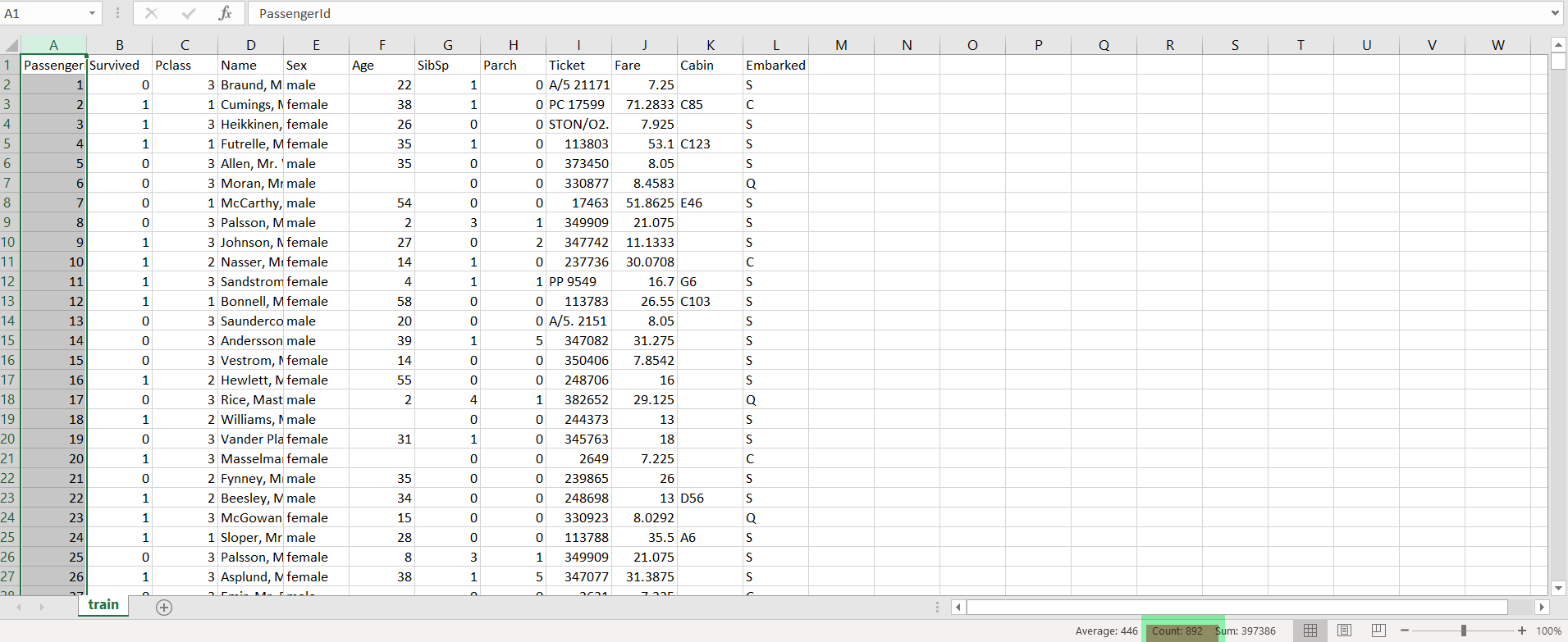


2.

Provide the shape/dimensions of the file train.csv

​

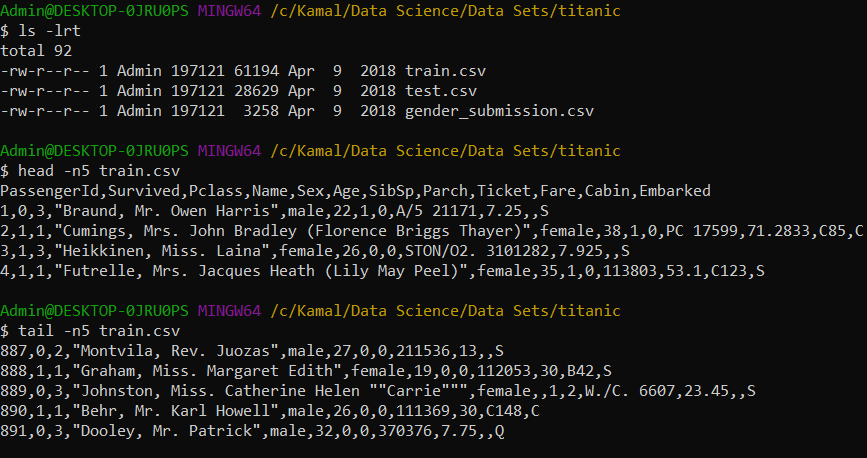




It showed columns returned as 13 but actually there are 12 columns. Though row count is correct.

3.

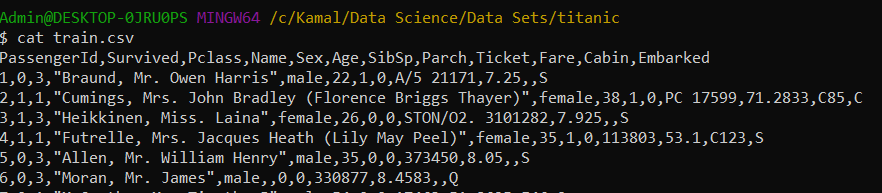
List the first 5 rows of the file. Now list the last 5.



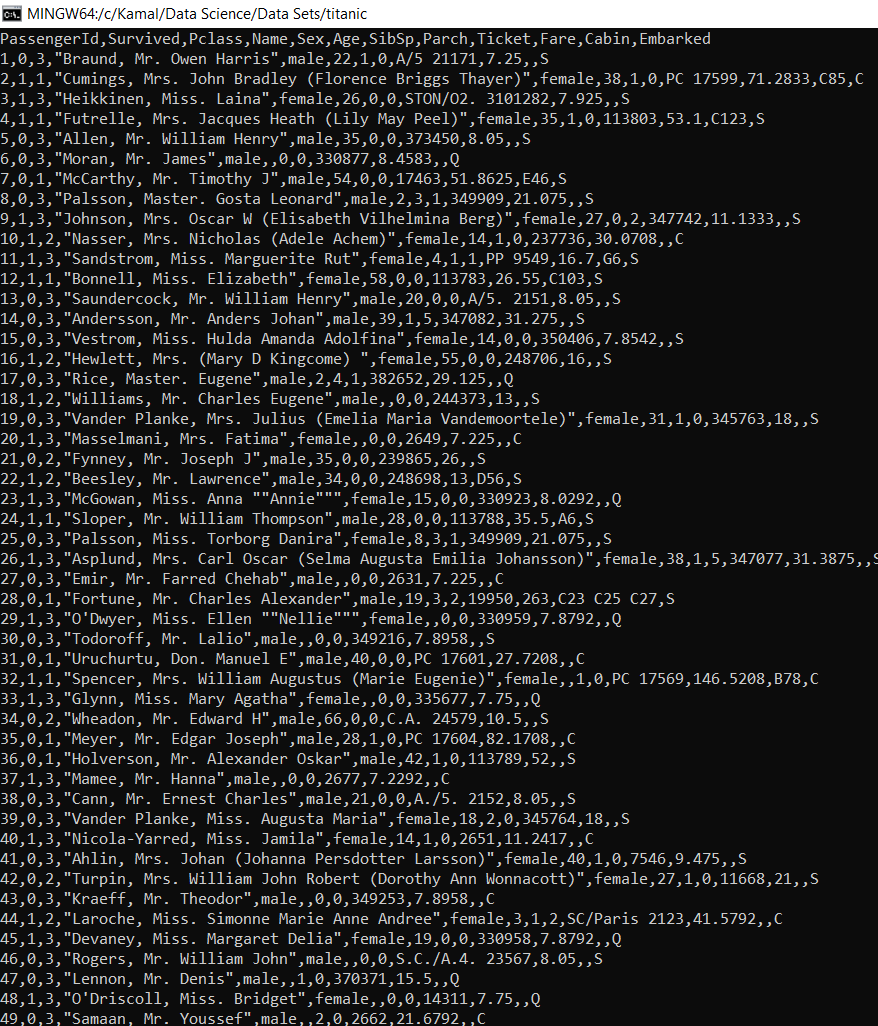
4.

Print this file in your screen using cat now use the less command.

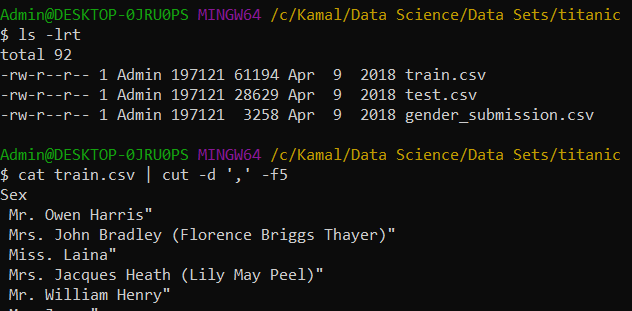
5.



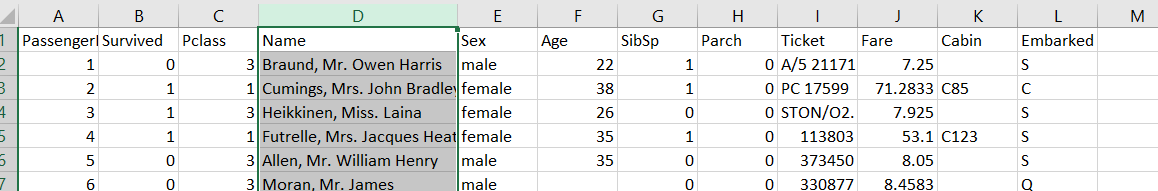
Less train.csv



Can you print only the names of all people in the file?

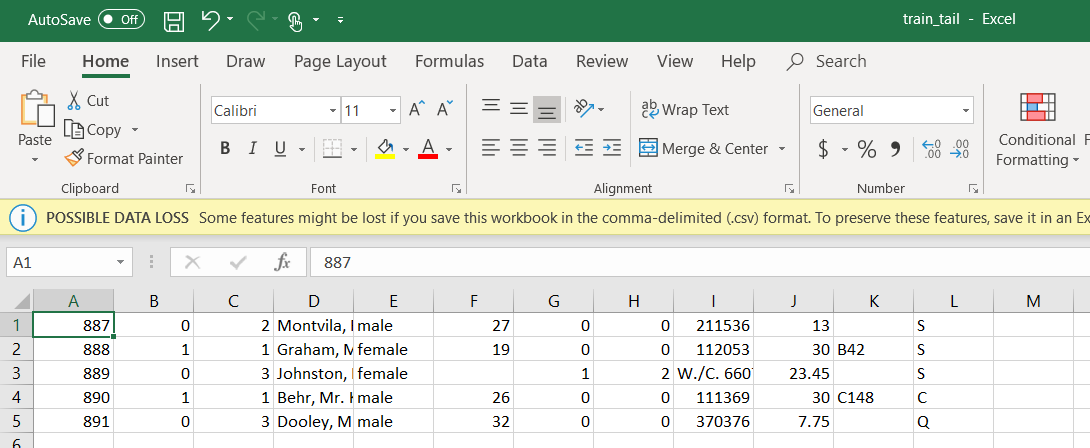
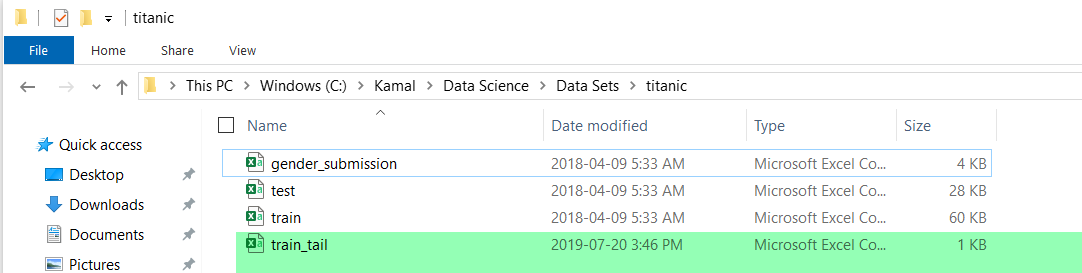
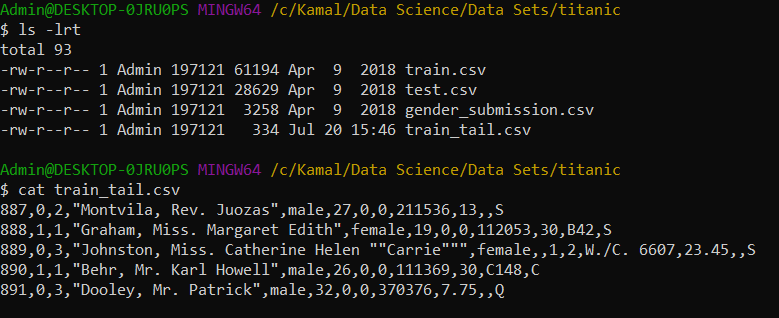


Not sure why it displayed the column header as ‘Sex’ instead of ‘Name’?

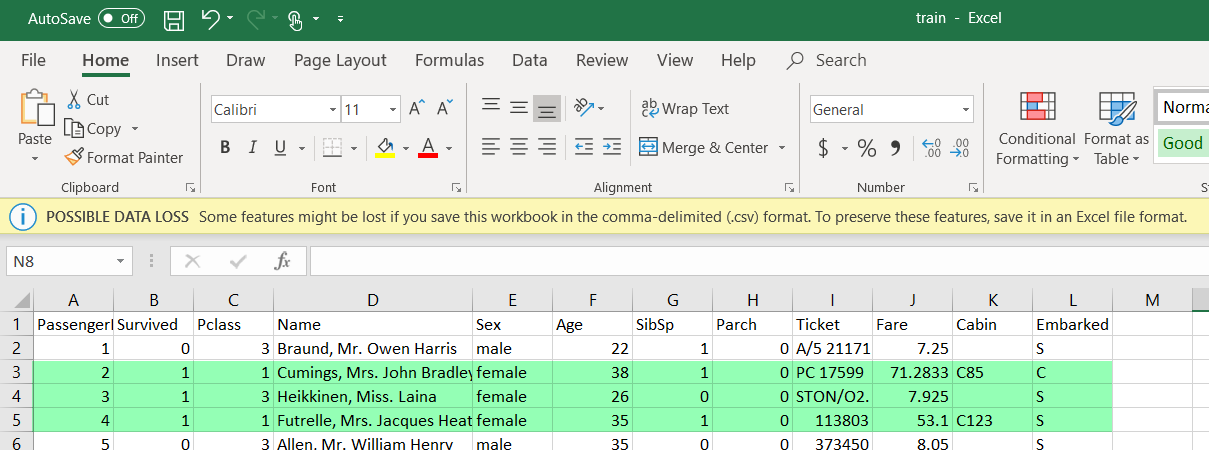
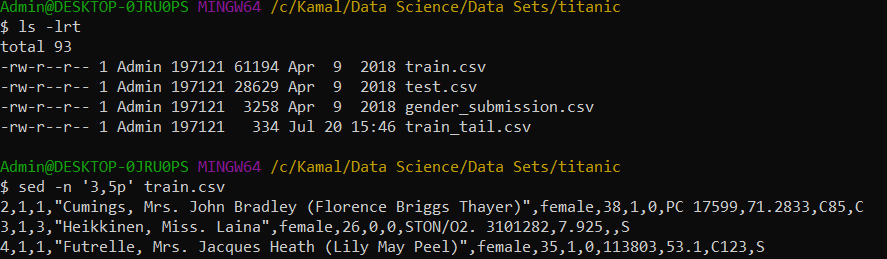


6.

Print this file last 5 lines save the output to train\_tail.csv



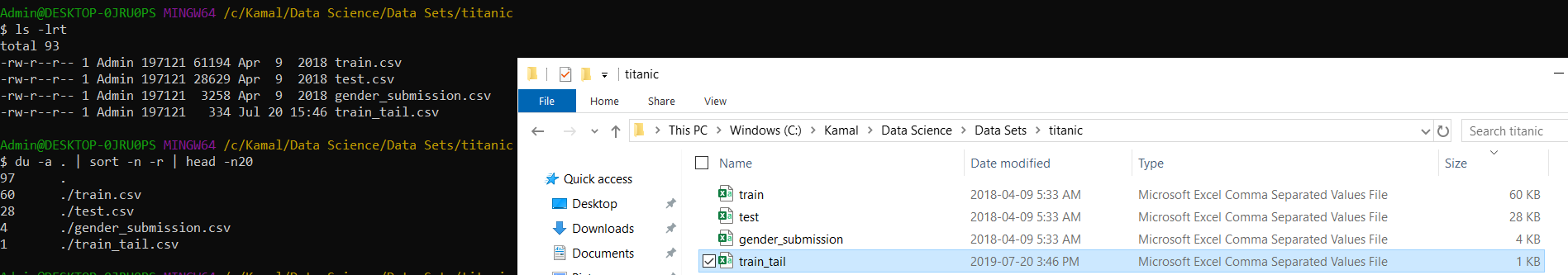
7.

Print only the lines 3 to 5 of the file? 

8.Can you explain the command and why would you use it?

du -a . | sort -n -r | head -n 20

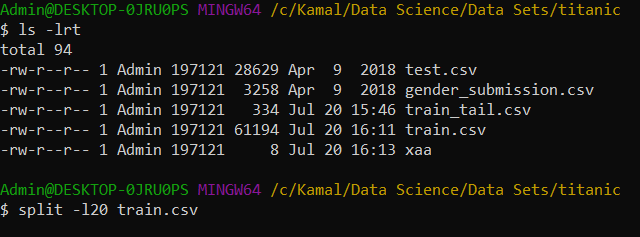
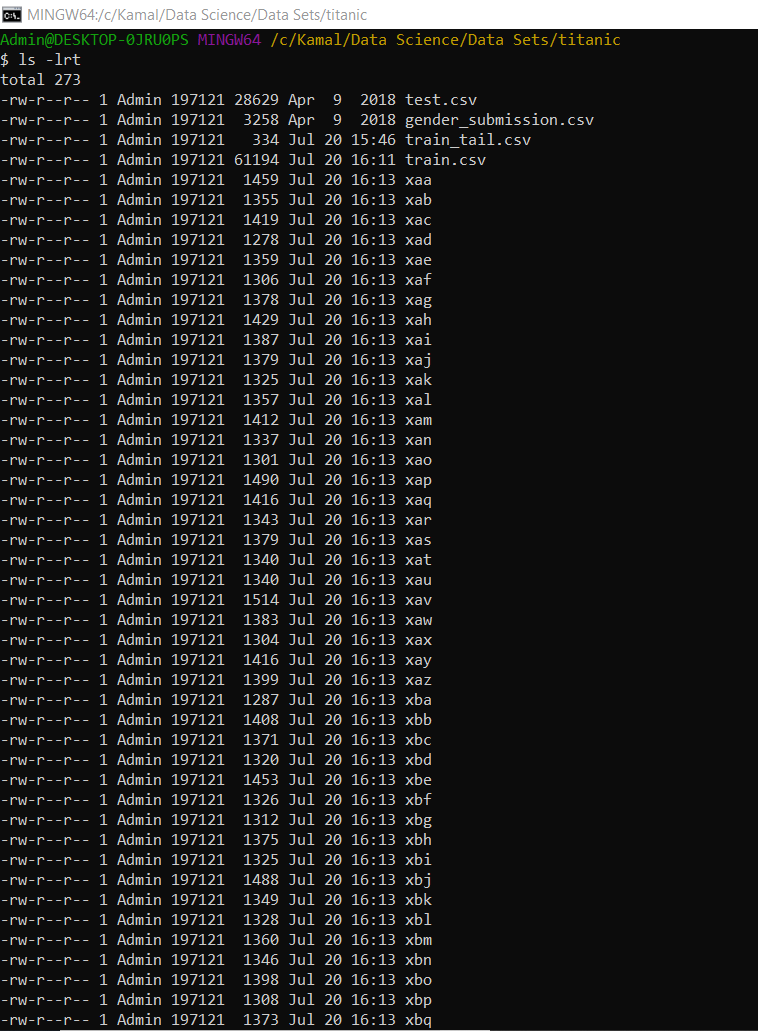
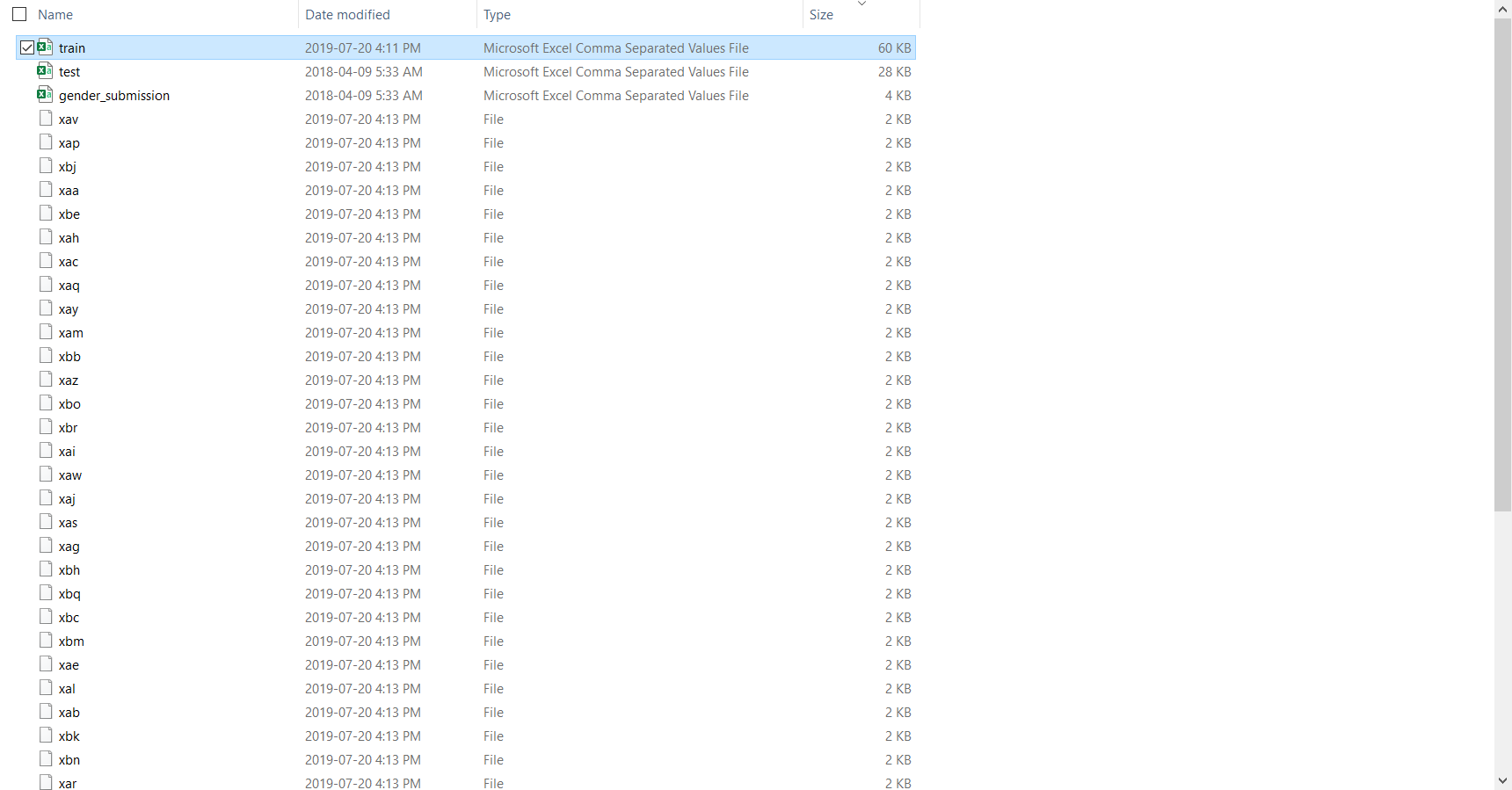
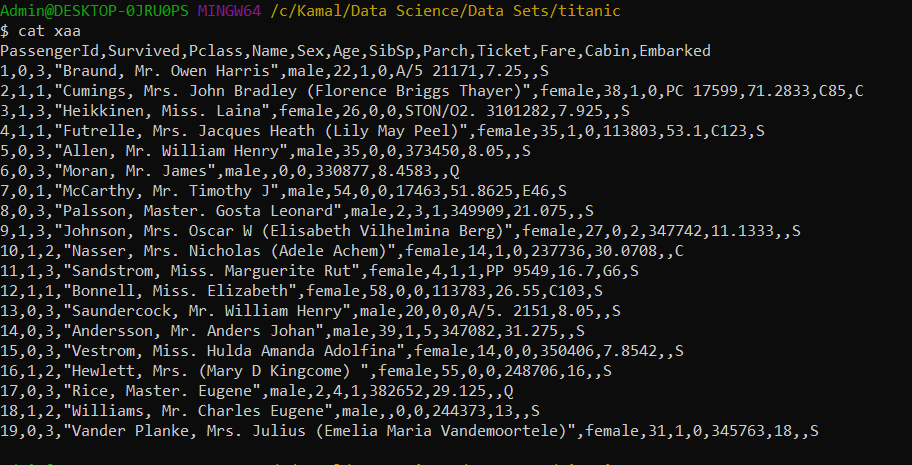
This command is used to provide top 20 largest files in a given directory size wise. We will use this to find out the disk space accumulated by the different files sorted by size. For example as shown below in the snippet.



​

9.

Split the train.csv file in multiple files with 20 lines each.

Write loops to iterate over lists

1.

Download the

​

ultratrail-du-montblanc.zip

​

file from Slack and unzip it to

/Users/<myusername>/ultratrail

2.

Write a loop that prints the name, dimension and first 2 lines for each of the

​

.csv

files.

3.

Write a loop that copies each of the

​

.csv

​

files with the prefix

​

bkp-

​

to a folder

/Users/<myusername>/ultratrail/backups

​

.

Reach

Create scripts to automate basic processes

1.

Write a script that suggests the data formats: csv, xlsx, pdf, doc and txt. It should

allow the user to pick their desired extension then create a file named

selected.<extension selected>

​

. Use the

​

read

​

command to read the user input!

2.

Write a script that keeps only the first N number of lines of all files in

'/Users//files\_to\_clean/\*.csv'. N should be an argument passed before starting

the script! If other people depend on this being done daily, how can we automate

it's daily execution at 8:00AM?