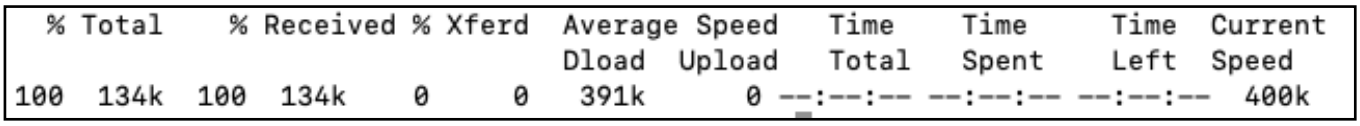
1. linux file manipulation:

1.

curl -O https://schlieplab.org/Static/Teaching/DIT852/private-healthcarepercent-gdp.csv

-it fetches the file located at the given URL (private-healthcarepercent-gdp.csv) and saves it in the current directory with the same filename.

Output:



2

sed -n '/AFG/,$p' private-healthcare-percent-gdp.csv | awk -F, '!/ABW/' > new_filtered_data.csv

this command does the following:

-It uses sed to extract lines from the input CSV file private-healthcare-percent-gdp.csv, starting from the line containing 'AFG' and continuing to the end of the file.

-It then uses awk to filter out lines that contain 'ABW' from the output of sed.

-Finally, it saves the filtered data into a new file named new_filtered_data.csv.

3

cut -f1 -d, new_filtered_data.csv | sort | uniq | wc -l

this Linux command does the following:

-It extracts the values from the first column of the CSV file (new_filtered_data.csv).

-It sorts these values alphabetically.

-It removes duplicate values, leaving only the unique values.

-It counts the number of unique values and prints the count to the terminal.

Output:

```
(base) Users-MacBook-Air:comp_assignment user$ cut -f1 -d, new_filtered_data.csv | sort | uniq | wc -l
     263
```

4

- cut -f1,15-25 -d, private-healthcare-percent-gdp.csv > newfile.csv
- tail -n +6 newfile.csv > finalfile.csv

the entire command sequence does the following:

-Extracts the first field (column) and fields 15 to 25 from the CSV file private-healthcare-percent-gdp.csv

-saves the result in a new file called newfile.csv.

-Takes the content of newfile.csv, starting from the 6th line, and saves this subset of data in a new file called finalfile.csv.

5

awk -F, 'NR>1 {print $1, $(NF-11)}' finalfile.csv | sort -t, -k2,2nr | head -n 10

this Linux command does the following:

It processes finalfile.csv, excluding the header row, and extracts the first field and the field located 11 positions from the end of each line.

It sorts the extracted data based on the second field in reverse numerical order.

It displays the top 10 lines of the sorted output, effectively showing the 10 largest values from the second field.

Output:

```
(base) Users-MacBook-Air:comp_assignment user$ awk -F, 'NR>1 {print $1, $(NF-11)}' finalfile.csv | sort -t, -k2,2nr | head -n 10
"AFG" "AFG"
"AGO" "AGO"
"ALB" "ALB"
"AND" "AND"
"ARB" "ARB"
"ARE" "ARE"
"ARG" "ARG"
"ARM" "ARM"
"ASM" "ASM"
"ATG" "ATG"
```

6

grep -E "Country Code|DNK|FIN|ISL|NOR|SWE" finalfile.csv > nordicfile.csv

this Linux command does the following:

-It searches the content of finalfile.csv for lines that contain either the text "Country Code" or any of the specified Nordic country codes (DNK, FIN, ISL, NOR, SWE) and saves the matching lines in a new file called nordicfile.csv.