



# ASTROINFORMATICS

---

## Project Practice 2

---

### Authors:

Jose Ruben Carcamo Portillo

[jose.carcamo.portillo@ua.cl](mailto:jose.carcamo.portillo@ua.cl)

**Professor: NINA HERNITSCHK**

**June 15, 2024**

# Contents

<b>1</b>	<b>Objective</b>	<b>2</b>
<b>2</b>	<b>Introduction</b>	<b>3</b>
<b>3</b>	<b>Task1</b>	<b>4</b>
3.1	Change delimiters . . . . .	4
3.2	Change the file extension . . . . .	5
3.3	Remove all columns that are not part of the light curve . . . . .	6
<b>4</b>	<b>Task 2</b>	<b>7</b>
<b>5</b>	<b>Task 3</b>	<b>8</b>
<b>6</b>	<b>Conclusion</b>	<b>8</b>

# 1 Objective

In this report our objective is put in practice what we learned in class, using different methods to achieve the different goals, using the latter csv file we will learn how; to change the delimiter from ',' to ' ', change the file extension, remove columns in our documents, also we are going to code the spectra for a star and return a julian date.

## 2 Introduction

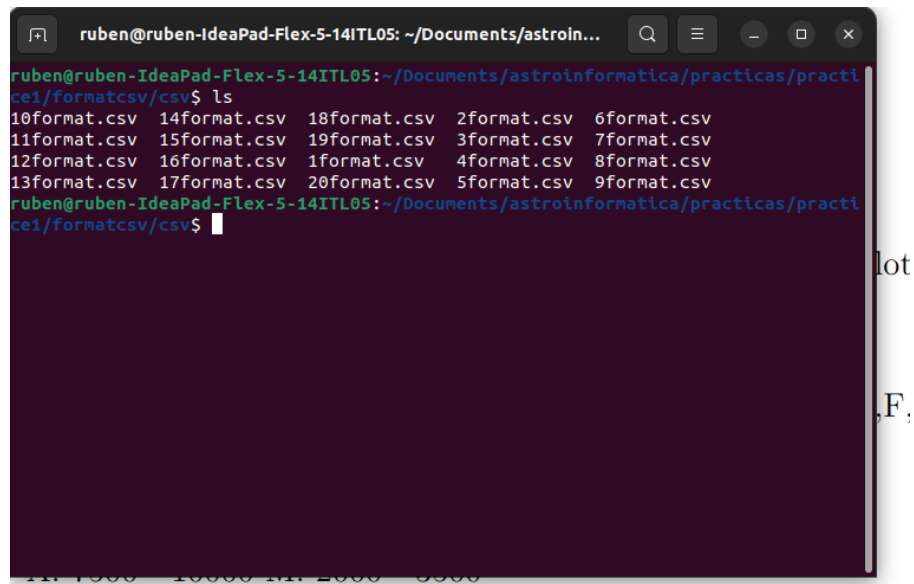
In this project practice 2 we are going to develop the first task in terminal because is more easily to management files and changes extension with basic command lines or scripts, in the first task we faced several challenges but with persistence and a deep search we could overcoming the challenges.

The second and third task was develop with a basic python script code.

## 3 Task1

### 3.1 Change delimiters

In the first task we was ask to change the delimiter from our latter csv file from "," to " " this can be easily done with the command sed

A terminal window with a dark background and light-colored text. The window title is 'ruben@ruben-IdeaPad-Flex-5-14ITL05: ~/Documents/astroin...'. The prompt is 'ruben@ruben-IdeaPad-Flex-5-14ITL05:~/Documents/astroinformatics/practicass/practice1/formatcsv/csv\$'. The command 'ls' has been entered, and the output is a list of 20 CSV files: 10format.csv, 11format.csv, 12format.csv, 13format.csv, 14format.csv, 15format.csv, 16format.csv, 17format.csv, 18format.csv, 19format.csv, 20format.csv, 2format.csv, 3format.csv, 4format.csv, 5format.csv, 6format.csv, 7format.csv, 8format.csv, 9format.csv, and 1format.csv. The prompt is now 'ruben@ruben-IdeaPad-Flex-5-14ITL05:~/Documents/astroinformatics/practicass/practice1/formatcsv/csv\$' with a cursor at the end.

```
ruben@ruben-IdeaPad-Flex-5-14ITL05:~/Documents/astroin...
ruben@ruben-IdeaPad-Flex-5-14ITL05:~/Documents/astroinformatics/practicass/practice1/formatcsv/csv$ ls
10format.csv 14format.csv 18format.csv 2format.csv 6format.csv
11format.csv 15format.csv 19format.csv 3format.csv 7format.csv
12format.csv 16format.csv 1format.csv 4format.csv 8format.csv
13format.csv 17format.csv 20format.csv 5format.csv 9format.csv
ruben@ruben-IdeaPad-Flex-5-14ITL05:~/Documents/astroinformatics/practicass/practice1/formatcsv/csv$
```

Figure 1: Csv files.

A error who i found is put **sed '/,/ \ / g' format1.csv** because this will change the delimiter of the file just when you are reading but this not save the changes, the correct way that file can save the changes is with **sed -i -e \$'/s/ , \ / g' format1.csv** as we can see in [2 \[2\]](#).

Of course we check the file if they got a changes in the delimiter and the changes was save it.

```

ruben@ruben-IdeaPad-Flex-5-14ITL05: ~/Documents/astroin...
ce1/formatcsv/csv$ sed -i -e 's/,/\ /g' *.csv
ruben@ruben-IdeaPad-Flex-5-14ITL05:~/Documents/astroinformatica/practicas/practi
ce1/formatcsv/csv$ head -n 3 1format.csv
TIME TIMECORR CADENCENO SAP_FLUX SAP_FLUX_ERR SAP_BKG SAP_BKG_ERR PDCSAP_FLUX PD
CSAP_FLUX_ERR QUALITY PSF_CENTR1 PSF_CENTR1_ERR PSF_CENTR2 PSF_CENTR2_ERR MOM_CE
NTR1 MOM_CENTR1_ERR MOM_CENTR2 MOM_CENTR2_ERR POS_CORR1 POS_CORR2
3285.7989573595028 0.005495878 1482004 168
3285.8003462524134 0.0054958826 1482005 32
ruben@ruben-IdeaPad-Flex-5-14ITL05:~/Documents/astroinformatica/practicas/practi
ce1/formatcsv/csv$ head -n 3 2format.csv
TIME TIMECORR CADENCENO SAP_FLUX SAP_FLUX_ERR SAP_BKG SAP_BKG_ERR PDCSAP_FLUX PD
CSAP_FLUX_ERR QUALITY PSF_CENTR1 PSF_CENTR1_ERR PSF_CENTR2 PSF_CENTR2_ERR MOM_CE
NTR1 MOM_CENTR1_ERR MOM_CENTR2 MOM_CENTR2_ERR POS_CORR1 POS_CORR2
3285.7990169892932 0.0055555077 1482004 168
3285.8004058817382 0.005555512 1482005 32
ruben@ruben-IdeaPad-Flex-5-14ITL05:~/Documents/astroinformatica/practicas/practi
ce1/formatcsv/csv$ head -n 3 20format.csv
TIME TIMECORR CADENCENO SAP_FLUX SAP_FLUX_ERR SAP_BKG SAP_BKG_ERR PDCSAP_FLUX PD
CSAP_FLUX_ERR QUALITY PSF_CENTR1 PSF_CENTR1_ERR PSF_CENTR2 PSF_CENTR2_ERR MOM_CE
NTR1 MOM_CENTR1_ERR MOM_CENTR2 MOM_CENTR2_ERR POS_CORR1 POS_CORR2
3285.7990036215547 0.00554214 1482004 168
3285.800392516328 0.0055421465 1482005 32
ruben@ruben-IdeaPad-Flex-5-14ITL05:~/Documents/astroinformatica/practicas/practi
ce1/formatcsv/csv$

```

Figure 2: A different way to change the delimiters

### 3.2 Change the file extension

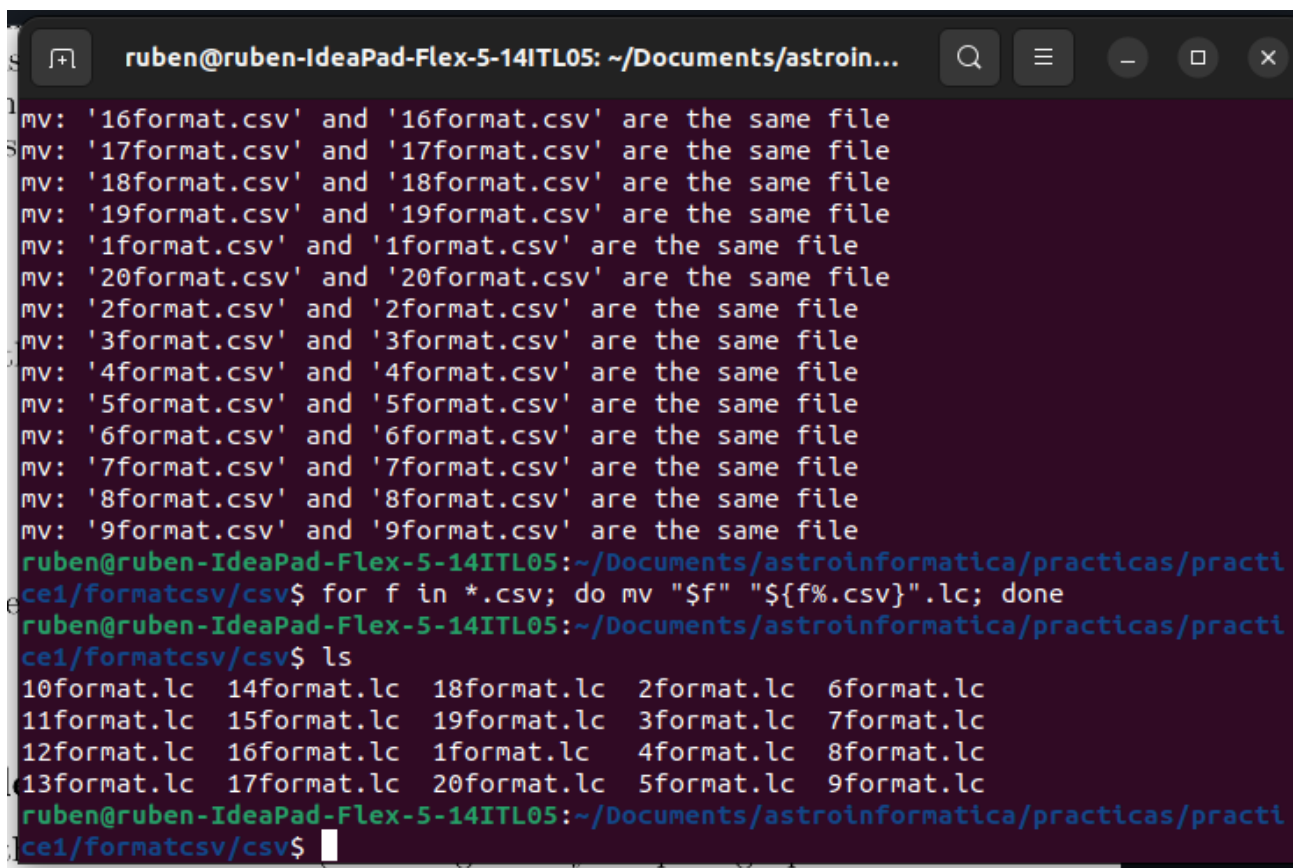
We have to be careful when going to change the extension, because is easily to get confused in add extension instead of changing it like 3 so to avoid this error we need to add `"${f%.csv} ".lc` [1] in order to change the file extension like 4

```

ruben@ruben-IdeaPad-Flex-5-14ITL05: ~/Documents/astroin...
ruben@ruben-IdeaPad-Flex-5-14ITL05:~/Documents/astroinformatica/practicas/practi
ce1/formatcsv/csv$ cd ..
ruben@ruben-IdeaPad-Flex-5-14ITL05:~/Documents/astroinformatica/practicas/practi
ce1/formatcsv$ ls
10format.csv 16format.csv 3format.csv ch.bash split_file.txtaa
11format.csv 17format.csv 4format.csv csv split_file.txtab
1.1format.csv 18format.csv 5format.csv csv1 split_file.txtac
12format.csv 19format.csv 6format.csv file1 split_file.txtad
13format.csv 1format.csv 7format.csv file2.txt splitscript.sh
14format.csv 20format.csv 8format.csv output1_files.txt
15format.csv 2format.csv 9format.csv output1script.sh
ruben@ruben-IdeaPad-Flex-5-14ITL05:~/Documents/astroinformatica/practicas/practi
ce1/formatcsv$ for f in csv/*.csv; do mv "$f" "${f%.csv} ".lc; done
ruben@ruben-IdeaPad-Flex-5-14ITL05:~/Documents/astroinformatica/practicas/practi
ce1/formatcsv$ cd csv
ruben@ruben-IdeaPad-Flex-5-14ITL05:~/Documents/astroinformatica/practicas/practi
ce1/formatcsv/csv$ ls
10format.csv.lc 15format.csv.lc 1format.csv.lc 5format.csv.lc
11format.csv.lc 16format.csv.lc 20format.csv.lc 6format.csv.lc
12format.csv.lc 17format.csv.lc 2format.csv.lc 7format.csv.lc
13format.csv.lc 18format.csv.lc 3format.csv.lc 8format.csv.lc
14format.csv.lc 19format.csv.lc 4format.csv.lc 9format.csv.lc
ruben@ruben-IdeaPad-Flex-5-14ITL05:~/Documents/astroinformatica/practicas/practi
ce1/formatcsv/csv$

```

Figure 3: Common error in adding extension



```

ruben@ruben-IdeaPad-Flex-5-14ITL05: ~/Documents/astroin...
mv: '16format.csv' and '16format.csv' are the same file
mv: '17format.csv' and '17format.csv' are the same file
mv: '18format.csv' and '18format.csv' are the same file
mv: '19format.csv' and '19format.csv' are the same file
mv: '1format.csv' and '1format.csv' are the same file
mv: '20format.csv' and '20format.csv' are the same file
mv: '2format.csv' and '2format.csv' are the same file
mv: '3format.csv' and '3format.csv' are the same file
mv: '4format.csv' and '4format.csv' are the same file
mv: '5format.csv' and '5format.csv' are the same file
mv: '6format.csv' and '6format.csv' are the same file
mv: '7format.csv' and '7format.csv' are the same file
mv: '8format.csv' and '8format.csv' are the same file
mv: '9format.csv' and '9format.csv' are the same file
ruben@ruben-IdeaPad-Flex-5-14ITL05:~/Documents/astroinformatica/practicas/practi
ce1/formatcsv/csv$ for f in *.csv; do mv "$f" "${f%.csv}".lc; done
ruben@ruben-IdeaPad-Flex-5-14ITL05:~/Documents/astroinformatica/practicas/practi
ce1/formatcsv/csv$ ls
10format.lc  14format.lc  18format.lc  2format.lc  6format.lc
11format.lc  15format.lc  19format.lc  3format.lc  7format.lc
12format.lc  16format.lc  1format.lc   4format.lc  8format.lc
13format.lc  17format.lc  20format.lc  5format.lc  9format.lc
ruben@ruben-IdeaPad-Flex-5-14ITL05:~/Documents/astroinformatica/practicas/practi
ce1/formatcsv/csv$

```

Figure 4: Changing extension

### 3.3 Remove all columns that are not part of the light curve

To remove the columns that do not need we can use the command `cut`, but we need to do it with all 20 files so do it one by one is a harder task so for that reason we will use a for loop we write ***for file in \*.lc; do cut -d\ -f1,4,5,8,9 "\$file" > new\_"\$file"; done*** so we need to remember the \*.lc is refer that we want all files with extension .lc and the -f1,4,5,etc are the columns we want to keep it and will save it in a new file call new format, we check with command `head` to verify if the columns were remove like 5

```

ruben@ruben-IdeaPad-Flex-5-14ITL05: ~/Documents/astroinformatica/practicas/practice1/formatc
sv/csv2$ for file in *.lc; do cut -d\ -f1,4,5,8,9 "$file" > new_"$file"; done
ruben@ruben-IdeaPad-Flex-5-14ITL05: ~/Documents/astroinformatica/practicas/practice1/formatc
sv/csv2$ ls
10format.lc  18format.lc  6format.lc      new_14format.lc  new_2format.lc
11format.lc  19format.lc  7format.lc      new_15format.lc  new_3format.lc
12format.lc  1format.lc   8format.lc      new_16format.lc  new_4format.lc
13format.lc  20format.lc  9format.lc      new_17format.lc  new_5format.lc
14format.lc  2format.lc  new_10format.lc new_18format.lc  new_6format.lc
15format.lc  3format.lc  new_11format.lc new_19format.lc  new_7format.lc
16format.lc  4format.lc  new_12format.lc new_1format.lc   new_8format.lc
17format.lc  5format.lc  new_13format.lc new_20format.lc  new_9format.lc
ruben@ruben-IdeaPad-Flex-5-14ITL05: ~/Documents/astroinformatica/practicas/practice1/formatc
sv/csv2$ head -n 5 new_1format.lc
TIME  SAP_FLUX  SAP_FLUX_ERR  PDCSAP_FLUX  PDCSAP_FLUX_ERR
3285.7989573595028
3285.8003462524134
3285.801735145324
3285.8031240382356
ruben@ruben-IdeaPad-Flex-5-14ITL05: ~/Documents/astroinformatica/practicas/practice1/formatc
sv/csv2$

```

Figure 5: Removing columns

## 4 Task 2

To solve this task we will support of python in this task ask about spectra classification so in order to do this we need to use the condition if else and elif of python to distinguish the wavelength, so we name a variable 'temp' that is the temperature that user will enter, the variable 'esp' was defined to put the corresponding spectra like O,B,A, etc, so after that we use if conditional to achieve the task.

```

Task #2

Spectra of the stars are classified according to the letters O,B,F,G,K and M. These correspond to the following temperature ranges (in degrees K):

*O: 30000 - 60000 K: 5000 - 6000
*B: 10000 - 30000 K: 3500 - 5000
*A: 7500 - 10000 M: 2000 - 3500
*F: 6000 - 7500

Write a program which takes the temperature as a command line argument and prints out the spectral class. Print a suitable message if the temperature is out of range.

[1]: temp=int(input("Enter the Temperature in K"))
    esp=""
    if temp >20000 and temp<=35000:
        esp='M'
        print(f"The spectral line is {esp}")
    elif temp >35000 and temp<=50000:
        esp='K'
        print(f"The spectral line is {esp}")
    elif temp >50000 and temp<=60000:
        esp='G'
        print(f"The spectral line is {esp}")
    elif temp >60000 and temp<=75000:
        esp='F'
        print(f"The spectral line is {esp}")
    elif temp >75000 and temp<=100000:
        esp='A'
        print(f"The spectral line is {esp}")
    elif temp >100000 and temp<=300000:
        esp='B'
        print(f"The spectral line is {esp}")
    else:
        print(f"The spectral line is {esp}")
        print("Temperature out of range")
Enter the Temperature in K 7000
The spectral line is F

```

Figure 6: Conditions if in spectra wavelength



## 5 Task 3

For the task three we use dictionary cause we need to relate month with number to make a calculation of the julian date, to achieve this we create 3 variables corresponding to the date(day, month, year) that is the user will enter, beside that we are going to create a dictionary to relate the number 13,14,etc to the corresponding month, after that wil put the julian date formula and will print the julian date corresponding to the latter data.



```

Task #3

Given the year, month and day of the month, the Julian day is calculated as follows: Julian = (36525*year)/100 + (306001*(month+1))/10000 + day + 1720981 where month is 13 for Jan, 14 for Feb, 3 for Mar, 4 for Apr etc. For Jan and Feb, the year is reduced by 1.

Write a script which asks for the day, month and year and calculates the Julian day. All variables must be of integer type. What is the Julian day for 7 Jun 2008?.

[45]: day =int(input("enter the day"))
      month =int(input("enter the month"))
      year =int(input("enter the year"))
      month_list={'january':13,'february':14,'marh':4,'april':5,'may':6,'june':7,'july':8,'august':9,'september':10,'october':11,'november':12,'december':13}
      julian_month=month_list[month]
      julian_date = (36525*year/100)+(306001*(julian_month+1)/10000) +day+1720981
      print(f'the julian date is {julian_date}')

enter the day 1
enter the month january
enter the year 300
the julian date is 1830985.4014

[ ]:

```

Figure 7: Using dictionary in python

## 6 Conclusion

In this practice we learned how to delimiter, change extension and delete columns this will help us to sort astronomy data in larger survey, in python we use different methods to solve the task like conditional and dictionary.

## References

- [1] <https://www.unix.com/unix-for-dummies-questions-and-answers/79091-changing-extension->
- [2] <https://www.youtube.com/watch?v=iyXCzAYNmSs>.

## List of Figures

1	Csv files. . . . .	4
2	A different way to change the delimiters . . . . .	5
3	Common error in adding extension . . . . .	5
4	Changing extension . . . . .	6
5	Removing columns . . . . .	7
6	Conditions if in spectra wavelength . . . . .	7
7	Using dictionary in python . . . . .	8