

5. Download the transcripts of text chat from the google drive folder for this course into a folder. Create a script that can calculate your attendance as a percentage by looking for your roll number in each of the files.

The solution includes the script, total number of transcript files considered, lecture numbers missed and attendance percentage. [2 Marks]

Hint: Use the features of grep and wc for the task.

Application: When you run a large FEM program, you receive a log output in which certain warnings will be listed about change of algorithm or convergence failure etc., Looking for their occurrence and analysis is a part of engineering simulation work.

Link to the GitHub repository for this question: [GitHub](#)

This bash automatically processes the contents of the zip file and gives the attendance report.

The bash script can also take one parameter – the roll number and give the report of that person.

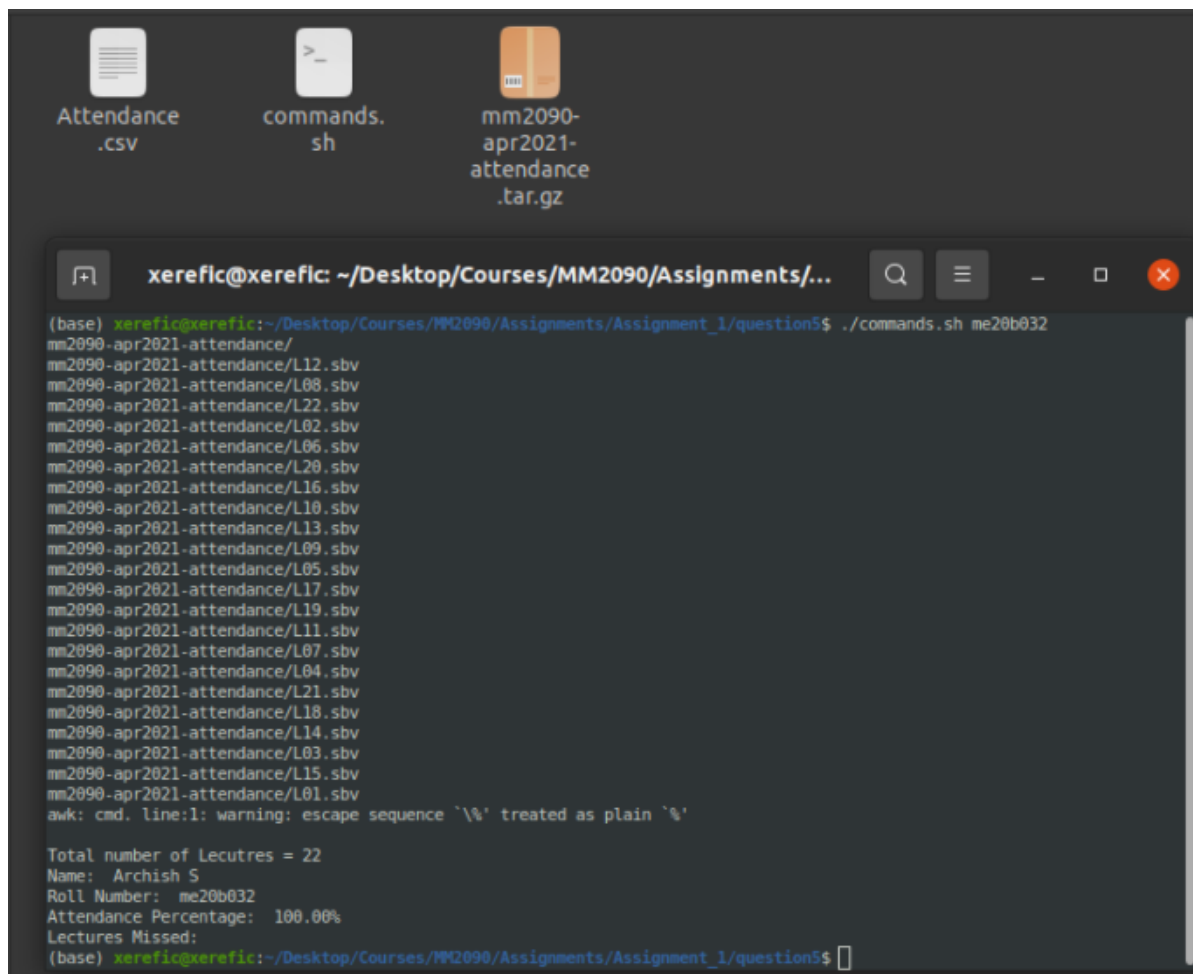
```

1. #!/bin/bash
2.
3. # Unzipping the chats
4. tar -xvzf mm2090-apr2021-attendance.tar.gz
5. mv mm2090-apr2021-attendance/ transcripts/
6.
7. # Grepping the name and roll number of students
8. mkdir registered
9. lectures=0;
10. for files in `ls transcripts/`;
11. do
12.     lectures=$(( $lectures+1 ));
13.     cat transcripts/$files | grep -oe '^(\.*\)[[:digit:]]{2}[:alpha:]*' >
        registered/$(basename $files .sbv).txt
14. done;
15. rm -r transcripts/
16.
17. # Pre-processing
18. for files in `ls registered/`;
19. do
20.     cat registered/$files | sed -e 's/(\.*\)[[:digit:]]{2}:/\2,\1/g' > registered/$(basename $files
        .txt).csv
21. done;
22. rm registered/*.txt
23.
24.
25. mkdir cache
26. for files in `ls registered/`;
27. do
28.     cat registered/$files | sed -e 's/ /_/g' > cache/$(basename $files .csv).csv
29. done;
30. rm -r registered/
31.
32. # Removing multiple instances for each lecture
33. mkdir attendance
34. for files in `ls cache/`;
35. do
36.     cat cache/$files | awk -F, '{name[$1]=$2;}END{for (id in name){printf("%s,%s\n", id, name[id]);}}' |
        sort -k1 -n > attendance/$(basename $files .csv).csv
37. done;
38. rm -r cache/
39.
40. # Finding the total registerants
41. tail -n+1 -q attendance/*.csv | awk -F, '{name[$1]=$2;}END{for (id in name){printf("%s,%s\n", id,
        name[id]);}}' | sort -k1 -n > registered.csv
42.
43. # Finding the percentage of attendance
44. tail -n+1 -q attendance/*.csv | awk -F, -v total=$lectures '{data[$1]=$0;name[$1]++;}END{for (id in
        name){printf("%s,%.2f\n", data[id], name[id]/total*100);}}' | sort -k1 -n > attendance.csv
45.
46. # Finding absentees in every lecture

```

```
47. mkdir absent
48. for files in `ls attendance/`
49. do
50.   diff attendance/$files registered.csv | egrep '....b...' | sed 's/> \(.*\)/\1/g' | awk -v
    lecture=$(basename $files .csv) '{printf("%s,%s\n", $0, lecture)}' > absent/$(basename $files .csv).csv;
51. done;
52. rm registered.csv
53. rm -r attendance/
54.
55. # Processing the absentee list
56. tail -n+1 -q absent/*.csv | awk -F, '{name[$1]=$2;absent[$1]=absent[$1]$3;}END{for (id in
    name){printf("%s,%s,%s\n", id, name[id], absent[id]);}}' | sort -k1 -n > absent.csv
57. rm -r absent/
58.
59. # Combining the data
60. awk -F, 'NR==FNR {absent[$1]=$3; next}{printf("%s,%s\n", $0, absent[$1]);}' absent.csv attendance.csv >
    cache.csv
61. rm absent.csv
62. rm attendance.csv
63.
64. cat cache.csv | sed 's/_/ /g' | sed 's/\L/ L/g' > final.csv
65. rm cache.csv
66.
67. # Adding header
68. awk -F, 'BEGIN{printf("Roll,Name,Percentage,Missed Lectures\n");}{print $0;}' < final.csv >
    Attendance.csv
69. rm final.csv
70.
71. echo
72.
73. echo "Total number of Lecutres = $lectures"
74.
75. if [ -z "$1" ]; then
76.   echo
77. else
78.   cat Attendance.csv | grep $1 | gawk -F, '{print "Name: ", $2; print "Roll Number: ", $1; print
    "Attendance Percentage: ", $3; print "Lectures Missed: ", $4;}'
79. fi
80.
```

TERMINAL:



The terminal window shows a file manager interface at the top with three icons: a document for 'Attendance.csv', a terminal icon for 'commands.sh', and a tar file icon for 'mm2090-apr2021-attendance.tar.gz'. Below the icons, the terminal window title is 'xerefic@xerefic: ~/Desktop/Courses/MM2090/Assignments/...'. The command prompt shows the user running './commands.sh me20b032'. The output lists 22 lecture files (L01.sbv to L22.sbv) and a warning from 'awk' about an escape sequence. The final output shows the total number of lectures (22), the name (Archish S), the roll number (me20b032), the attendance percentage (100.00%), and the number of lectures missed (0).

```
(base) xerefic@xerefic:~/Desktop/Courses/MM2090/Assignments/Assignment_1/question$ ./commands.sh me20b032
mm2090-apr2021-attendance/
mm2090-apr2021-attendance/L12.sbv
mm2090-apr2021-attendance/L08.sbv
mm2090-apr2021-attendance/L22.sbv
mm2090-apr2021-attendance/L02.sbv
mm2090-apr2021-attendance/L06.sbv
mm2090-apr2021-attendance/L20.sbv
mm2090-apr2021-attendance/L16.sbv
mm2090-apr2021-attendance/L10.sbv
mm2090-apr2021-attendance/L13.sbv
mm2090-apr2021-attendance/L09.sbv
mm2090-apr2021-attendance/L05.sbv
mm2090-apr2021-attendance/L17.sbv
mm2090-apr2021-attendance/L19.sbv
mm2090-apr2021-attendance/L11.sbv
mm2090-apr2021-attendance/L07.sbv
mm2090-apr2021-attendance/L04.sbv
mm2090-apr2021-attendance/L21.sbv
mm2090-apr2021-attendance/L18.sbv
mm2090-apr2021-attendance/L14.sbv
mm2090-apr2021-attendance/L03.sbv
mm2090-apr2021-attendance/L15.sbv
mm2090-apr2021-attendance/L01.sbv
awk: cmd. line:1: warning: escape sequence `\\%' treated as plain `%'

Total number of Lecutres = 22
Name: Archish S
Roll Number: me20b032
Attendance Percentage: 100.00%
Lectures Missed:
(base) xerefic@xerefic:~/Desktop/Courses/MM2090/Assignments/Assignment_1/question$
```

OUTPUT:

Total number of Lectures = 22
Name: Archish S
Roll Number: me20b032
Attendance Percentage: 100%
Lectures Missed:

OUTPUT:

Roll	Name	Percentage	Missed Lectures
me20b001	Aahan Bhargava	77.27%	L15 L16 L18 L21 L22
me20b004	Abhaumika Bijudith	22.73%	L01 L02 L05 L07 L08 L09 L10 L11 L12 L13 L14 L16 L17 L18 L19 L21 L22
me20b005	Abheshek Paramanand Kamble	86.36%	L06 L17 L19
me20b008	Abhishek Yadav	81.82%	L04 L06 L17 L22
me20b014	Aditya Kishore Dhoke	77.27%	L02 L04 L13 L17 L19
me20b017	Akhil Koshy Rajesh	81.82%	L10 L15 L19 L21
me20b020	Akshat Rakesh Garhwal	86.36%	L08 L09 L21
me20b021	Alpha P Jose	72.73%	L01 L03 L04 L10 L14 L19
me20b022	Amar Muhammed	9.09%	L01 L03 L04 L05 L06 L07 L08 L09 L10 L11 L12 L13 L14 L15 L16 L17 L18 L19 L21 L22
me20b024	Ankit Kumar	86.36%	L02 L03 L04
me20b027	Anshid K	31.82%	L06 L07 L08 L09 L11 L13 L14 L15 L16 L17 L18 L19 L20 L21 L22
me20b028	Anushka Asit Vadhavkar	100.00%	
me20b029	Anushka S	86.36%	L01 L15 L22
me20b032	Archish S	100.00%	
me20b036	Arun Palaniappan	95.45%	L03
me20b050	Cecil Jacob Thomas	95.45%	L04
me20b053	Chinmayee Tushar Kolhe	86.36%	L05 L08 L13
me20b055	Chris Joy Beck	40.91%	L01 L06 L08 L10 L12 L13 L15 L16 L17 L18 L19 L20 L21
me20b088	Jay Harish Shah	100.00%	
me20b112	Monisha C	90.91%	L03 L16
me20b132	Prabhat Bedida	86.36%	L04 L16 L22
me20b150	Rithwin K Ashraf	54.55%	L03 L05 L07 L10 L14 L15 L16 L18 L19 L21
me20b162	Shrid Suresh	27.27%	L04 L05 L06 L07 L08 L09 L10 L11 L12 L13 L16 L18 L19 L20 L21 L22
me20b163	Shriya Shukla	100.00%	
me20b166	Siddhagavali Shital Bhiku	90.91%	L05 L17
me20b178	Sukeerth Ramkumar	100.00%	
me20b183	Swapnil Pares Mehta	100.00%	
mm20b005	Albin George	50.00%	L04 L06 L09 L11 L12 L13 L15 L17 L18 L21 L22
mm20b009	Bankar Niranjan Janardan	4.55%	L02 L03 L04 L05 L06 L07 L08 L09 L10 L11 L12 L13 L14 L15 L16 L17 L18 L19 L20 L21 L22
mm20b011	Bhagat Singh S	27.27%	L04 L05 L08 L09 L10 L11 L12 L13 L14 L15 L16 L17 L19 L20 L21 L22
mm20b017	Divya Jyothi D	86.36%	L01 L08 L16
mm20b019	Gatkal Siddhesh Sarjerao	95.45%	L07
mm20b020	Gokul C	81.82%	L04 L16 L18 L22
mm20b042	Nayanatara Deepak	81.82%	L01 L04 L13 L17
mm20b043	Nedunchezhiyan K	59.09%	L01 L02 L05 L07 L09 L10 L15 L19 L20
mm20b049	Prithviraj Pratap Bhosle	77.27%	L06 L09 L10 L13 L22
mm20b057	Shreya Rajesh	81.82%	L01 L02 L04 L17
mm20b059	Sumanth Manjunath Hegde	77.27%	L06 L08 L10 L16 L21