SED & AWK QUIZ NOTE

A file called "grades.txt" contains the following 5 lines.

adam 10 14 09 18 12

brenda 11 12 19 10 08

charles 05 09 12 15 15

dora 12 13 14 15 16

evan 20 19 18 17 16

Show exactly what is displayed by each of the following awk commands.

[ 10 marks ]

1. awk 'END { print NR }' grades.txt

5

\* Basically asking what is number of row at the end : which is 5 row

2. awk 'NR == 3' grades.txt

charles 05 09 12 15 15

\*Asking What is row # 3 equal to ? or what is value of row #3

3. awk '$NF < 10' grades.txt

brenda 11 12 19 10 08

\*Asking What is NF(Last Column or field) value that is less than 10?

Then look of value of 12,08,15,16,16 and get the value of row that has 08, which is Brenda 11 12 19 10 08

4. awk '{ t += $2 } END{ print t }' grades.txt

58

\*$2 = is the value of field 2 (column 2) ,value of t is increment by the value of filed 2 (keep add) = 10 11 05 12 20 until END of field) then print the t value of field 2

5. awk 'NR == 2 { t = $2; $2 = $3; $3 = t; print }' grades.txt

brenda 12 11 19 10 08

\*NR == 2 is check the values of row 2 : Brenda **11** 12 09 10 08

T = is nothing in beginning :> put the value of field 2, row 2 ( 11 ) into T

$3 (row 2, field 3 value == 12) :> put the value $3(12) into $2(11), then value of #2 becomes 12 not 11. Then put value of t(11) into value of $3(12), then value of $3 becomes 11. Then PRINT all

6. awk '{ if(length($1) > 6) print $1 }' grades.txt

Charles

\*length is number of character in this case, $1 = is the value of column 1 : adam Brenda Charles dora evan. If there is character number greater than 6 characters then print that value in field 1.

Which is Charles ( 7 characters )

7. awk '{ t=0; for(i=2; i<7; i++) t=t+$i; print t }' grades.txt

63

60

56

70

90

\*t==0, for loop initialized by 2 and increment value of “I” by 1 up to 7th field from 2nd field, and then t= 0 + value of column 2,3,4,5,6, and add to 0 which will be 63 in first row. Then print total value of t

IF initialized by t=1, then +1 to all total value 64,61, 57, 71, 91

8. awk '$2 > m { m = $2; n = $1 } END{ print n,m }' grades.txt

evan 20

\*Look at m is holding (field 2 value) and n is holding (field 1 value), then look at **end of data** which is evan(1) 20(2) 19(3) 18(4) 17(5) 16(6), then print n value which is field 1 value(evan) and print m value which is field 2(20)

?? But what is $2 > m meaning?????

9. awk ' /^b/ { print $NF }' grades.txt

08

\*search the character “b” that starts in the beginning of data, which is Brenda row , then PRINT then value of last column brenda 11 12 19 10 **08**

10. awk '/a[^d n]/ { print $4 }' grades.txt ??????? \*\*\* Very difficult

09

12

\*First search pattern /a[^d n]/ : must start with a && **NOT** condition in the bracket ^d = means can not start with d after a.

White space : can not have white space after a.

“n” : can not have after a.

\*\*\* Only pattern we care is 2 characters in this scenario, a (must) and **one character** that does not correspond to pattern in bracket.

First look at adam: ad disqualify for condition but am does, so get the value of $4 column : 09

Second look at Brenda: a\_ disqualify condition, so move on to next row

Third look at Charles: a qualify, r qualify as it does not match any of characters condition in bracket, get value $4 column : 12

Fourth look at dora: a qualify but contains white space : disqualify the condition

Fifth look evan: a qualify but contains “n”: disqualify the condition.

adam 10 14 09 18 12

brenda 11 12 19 10 08

charles 05 09 12 15 15

dora 12 13 14 15 16

evan 20 19 18 17 16

A file named "elements.txt" begins with the following 10 lines. You do not

know how many lines are in the file but there are more than 10. The first

column is the atomic weight, the second is the name and the third is the

symbol of the element.

1.0079 Hydrogen H

4.0026 Helium He

6.941 Lithium Li

9.0122 Beryllium Be

10.811 Boron B

12.0107 Carbon C

14.0067 Nitrogen N

15.9994 Oxygen O

18.9984 Fluorine F

20.1797 Neon Ne

Write a complete sed command for each of the following. [ 10 marks ]

11. Delete the 3rd line.

sed '3d' elements.txt

12. Print only those lines whose element names begin with an uppercase 'A'.

sed '/[0-9] A/!d' elements.txt

13. Delete all lines after line 30.

sed '31,$d' elements.txt

14. Delete lines 40 through 50.

sed '40,50d' elements.txt

15. Only print line 12.

sed -n '12p' elements.txt OR sed '12!d' elements.txt

16. Delete all element names but not their symbols.

sed 's/ [A-Z].\* / /' elements.txt

17. Change Oxygen to uppercase.

sed 's/Oxygen/OXYGEN/' elements.txt

18. Change ALL uppercase B's to lowercase.

sed 's/B/b/g' elements.txt

19. Add a new line containing the string "33.123 Vulcan V" after Carbon.

sed '/Carbon/a\

33.123 Vulcan V' elements.txt

20. Change the atomic weight of Gold to 196.966. NOTE: its symbol is Au.

sed '/Gold/c\

185.966 Gold Au' elements.txt