

Perl and Awk Scripts

Rwithik Manoj

April 24, 2019

Perl Scripts

Create a text file and answer the following queries :

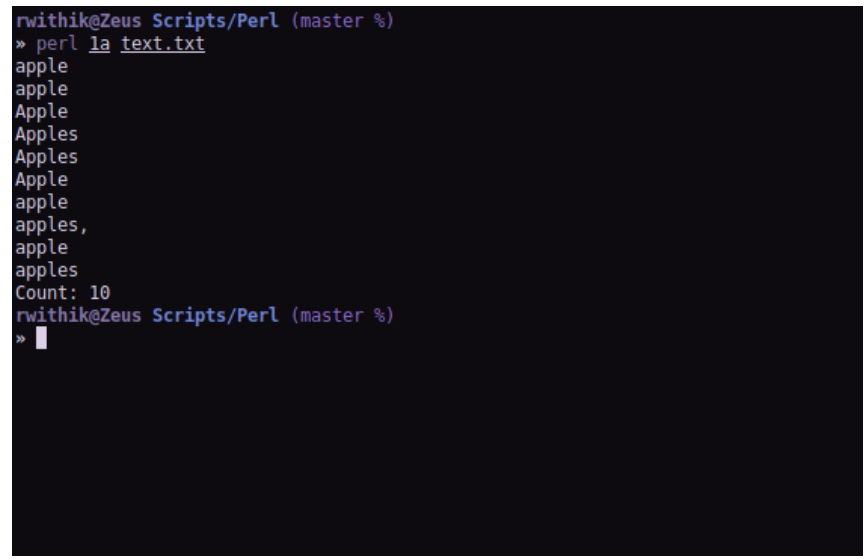
1. Search for the pattern 'apple' in the file and display the number of occurrences.

Script:

```
#!/bin/perl

my $count = 0;
while(my $line = <>){
    for (split ' ', $line) {
        if ($_ =~ /apple/i){
            printf("%s\n", $_);
            $count++;
        }
    }
}
printf("Count: %d\n", $count);
```

Output:



```
rwthik@Zeus Scripts/Perl (master %)
» perl la text.txt
apple
apple
Apple
Apples
Apples
Apple
apple
apples,
apple
apples
Count: 10
rwthik@Zeus Scripts/Perl (master %)
» █
```

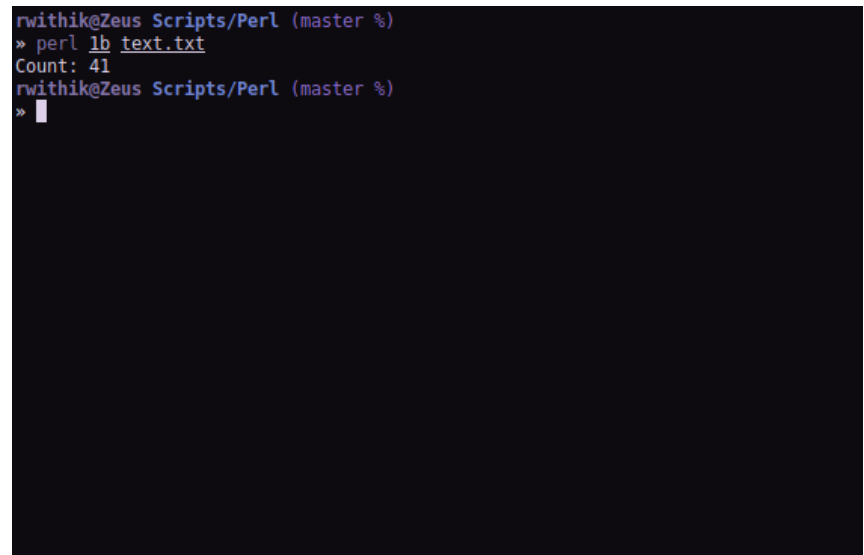
- Count the number of words that ends with 'e'.

Script:

```
#!/bin/perl

my $count = 0;
while(my $line = <>){
    for (split ' ', $line) {
        $count++ if $_ =~ /e$/i;
    }
}
printf("Count: %d\n", $count);
```

Output:



```
rwithik@Zeus Scripts/Perl (master %)
» perl lb text.txt
Count: 41
rwithik@Zeus Scripts/Perl (master %)
»
```

- Count the number of words that starts with 'ap'.

Script:

```
#!/bin/perl

my $count = 0;
while(my $line = <>){
    for (split ' ', $line) {
        $count++ if $_ =~ /^ap/i;
    }
}
printf("Count: %d\n", $count);
```

Output:

```
rwthik@Zeus Scripts/Perl (master %)  
» perl lc text.txt  
Count: 10  
rwthik@Zeus Scripts/Perl (master %)  
»
```

4. Search for words containing 'a' or 's'.

Script:

```
#!/bin/perl  
  
my $count = 0;  
while(my $line = <>){  
    for (split ' ', $line) {  
        if ($_ =~ /[a|s]/i){  
            printf("%s\n", $_);  
            $count++;  
        }  
    }  
}  
printf("Count: %d\n", $count);
```

Output:

```
a
organic
and
non-organic
means.
fruit's
was
sequenced
as
part
research
disease
and
selective
apple
apples
was
tonnes,
China
accounting
total.
Count: 115
rwithik@Zeus Scripts/Perl (master %)
> █
```

5. Search for words containing zero or more occurrence of 'e'.

Script:

```
#!/bin/perl

my $count = 0;
while(my $line = <>){
    for (split ' ', $line) {
        if ($_ =~ /e*/i){
            printf("%s\n", $_);
            $count++;
        }
    }
}
printf("Count: %d\n", $count);
```

Output:

```
in
apple
production.
Worldwide
production
of
apples
in
2017
was
83.1
million
tonnes,
with
China
accounting
for
49.8%
of
the
total.
Count: 204
rwithik@Zeus Scripts/Perl (master %)
>
```

6. Search for words containing one or more occurrence of 'e'.

Script:

```
#!/bin/perl

my $count = 0;
while(my $line = <>){
    for (split ' ', $line) {
        if ($_ =~ /e+/i){
            printf("%s\n", $_);
            $count++;
        }
    }
}
printf("Count: %d\n", $count);
```

Output:

```
prone
number
bacterial
pest
problems,
be
controlled
number
means.
the
genome
sequenced
research
disease
selective
breeding
apple
Worldwide
apples
tonnes,
the
Count: 92
rwithik@Zeus Scripts/Perl (master %)
> █
```

7. Search for words containing the letters 'l' and 'm', with any number of characters in between.

Script:

```
#!/bin/perl

my $count = 0;
while(my $line = <>){
    for (split ' ', $line) {
        if ($_ =~ /l.*m/i){
            printf("%s\n", $_);
            $count++;
        }
    }
}
printf("Count: %d\n", $count);
```

Output:

```
rwthik@Zeus Scripts/Perl (master %)  
» perl lg text.txt  
problems,  
Count: 1  
rwthik@Zeus Scripts/Perl (master %)  
» █
```


The text file used in the above screenshots:

```
rwithik@Zeus Scripts/Perl (master %)  
» cat text.txt  
An apple is a sweet, edible fruit produced by an apple tree (Malus pumila). Apple trees are cultivated worldwide and are the most widely grown species in the genus Malus. The tree originated in Central Asia, where its wild ancestor, Malus sieversii, is still found today. Apples have been grown for thousands of years in Asia and Europe and were brought to North America by European colonists. Apples have religious and mythological significance in many cultures, including Norse, Greek and European Christian traditions.  
  
Apple trees are large if grown from seed. Generally, apple cultivars are propagated by grafting onto rootstocks, which control the size of the resulting tree. There are more than 7,500 known cultivars of apples, resulting in a range of desired characteristics. Different cultivars are bred for various tastes and use, including cooking, eating raw and cider production. Trees and fruit are prone to a number of fungal, bacterial and pest problems, which can be controlled by a number of organic and non-organic means. In 2010, the fruit's genome was sequenced as part of research on disease control and selective breeding in apple production.  
  
Worldwide production of apples in 2017 was 83.1 million tonnes, with China accounting for 49.8% of the total.  
rwithik@Zeus Scripts/Perl (master %)  
»
```

Awk Scripts

1. Write a awk script that accepts date argument in the form of mm-dd-yy and displays it in the following format. The script should check the validity of the argument and in the case of error, display a suitable message.

Algorithm:

- (a) Start.
- (b) Check if the date is valid. Exit if its not.
- (c) If it is a valid date, find the corresponding month. And print the output in the format specified.
- (d) Stop.

Script:

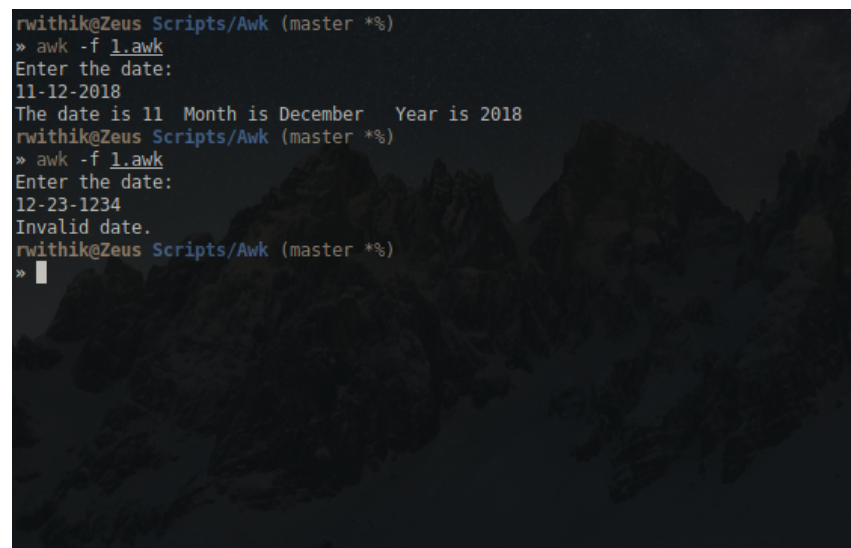
```
#!/usr/bin/awk  
BEGIN{  
FS="-"  
print "Enter the date:"  
getline < "/dev/tty"  
flag=1;  
if(((($3%4!=0) && ($2==2) && ($1>28)) || (($3%4==0) && ($2==2) && ($1>29)) || $2 > 12))  
flag=0;  
if(flag==0)  
print "Invalid date."  
else
```

```

{
    if($2==1)
        temp="January";
    else if($2==2)
        temp="February";
    else if($2==3)
        temp="March";
    else if($2==4)
        temp="April";
    else if($2==5)
        temp="May";
    else if($2==6)
        temp="June";
    else if($2==7)
        temp="July";
    else if($2==8)
        temp="August";
    else if($2==9)
        temp="September";
    else if($2==10)
        temp="October";
    else if($2==11)
        temp="November";
    else
        temp="December";
    print "The date is " $1 "\tMonth is " temp "\tYear is " $3
}
}

```

Output:



```

rwithik@Zeus Scripts/Awk (master *)
» awk -f 1.awk
Enter the date:
11-12-2018
The date is 11 Month is December Year is 2018
rwithik@Zeus Scripts/Awk (master *)
» awk -f 1.awk
Enter the date:
12-23-1234
Invalid date.
rwithik@Zeus Scripts/Awk (master *)
» █

```

2. Write an awk script to delete duplicated line from a text file. The order

of the original lines must remain unchanged.

Algorithm:

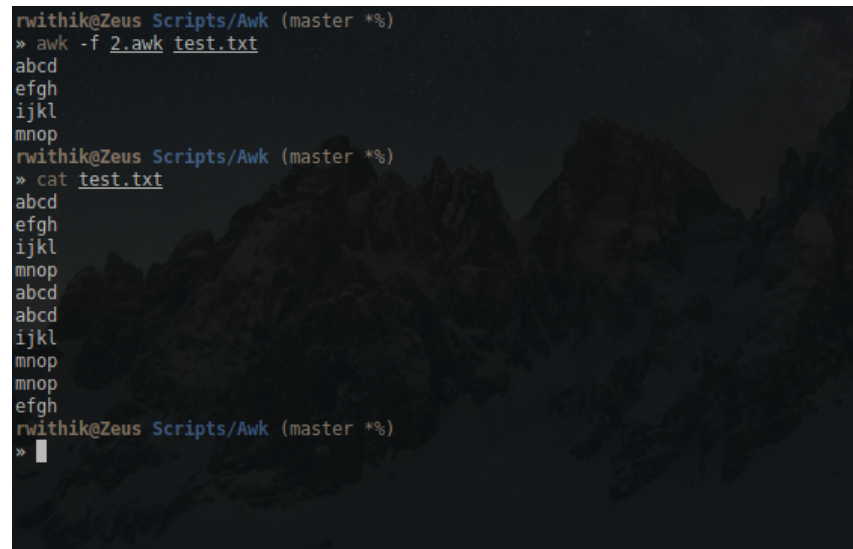
- (a) Start.
- (b) Print the line only if seen[\$0] is zero, ie the count of the current line is zero.
- (c) Stop.

Script:

```
#!/bin/awk

!seen[$0]++ {print $1}
```

Output:



```
rwithik@Zeus Scripts/awk (master *%)
» awk -f 2.awk test.txt
abcd
efgh
ijkl
mnop
rwithik@Zeus Scripts/awk (master *%)
» cat test.txt
abcd
efgh
ijkl
mnop
abcd
abcd
ijkl
mnop
mnop
efgh
rwithik@Zeus Scripts/awk (master *%)
»
```

3. Write an awk script to find out total number of books sold in each discipline as well as total book sold based on the given table.

Algorithm:

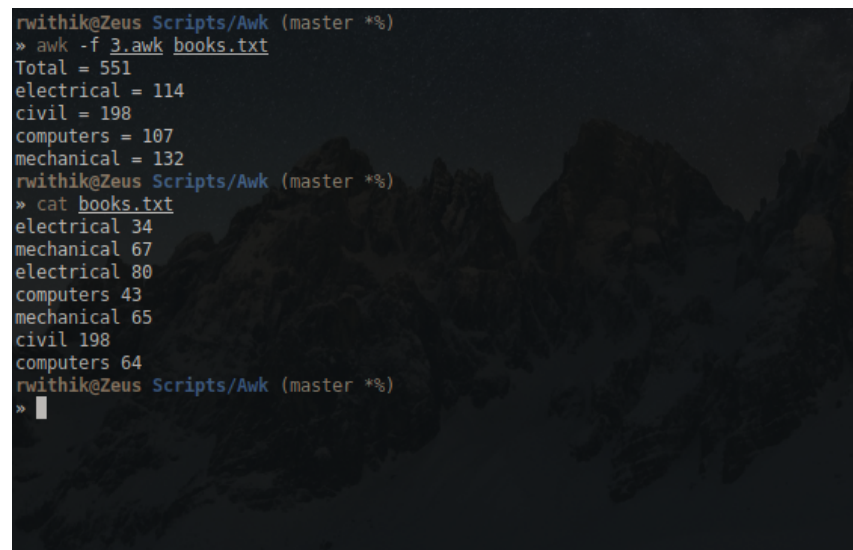
- (a) Start.
- (b) Initialise sum to 0.
- (c) Add the second field to the sum.
- (d) Add the Second field to the corresponding array element.
- (e) Print the values.
- (f) Stop.

Script:

```
#!/bin/awk

BEGIN{
    sum = 0
}
{
    sum = sum + $2;
    subject[$1] = subject[$1] + $2;
}
END{
    printf "Total = " sum "\n"
    for(word in subject) printf word " = " subject[word] "\n"
}
```

Output:



```
rwthik@Zeus Scripts/awk (master *)
» awk -f 3.awk books.txt
Total = 551
electrical = 114
civil = 198
computers = 107
mechanical = 132
rwthik@Zeus Scripts/awk (master *)
» cat books.txt
electrical 34
mechanical 67
electrical 80
computers 43
mechanical 65
civil 198
computers 64
rwthik@Zeus Scripts/awk (master *)
»
```

4. Write an awk script to compute gross salary of an employee accordingly to rule given below: If basic salary is less than 10000, then DA = 45% of the basic and HRA = 15% of basic. If basic salary is greater than 10000, then DA = 50% of the basic and HRA = 20% of basic.

Algorithm:

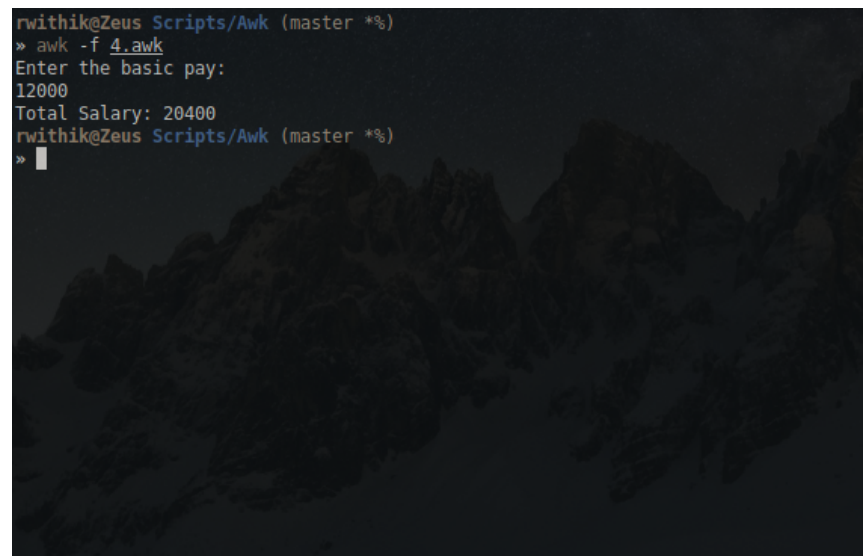
- (a) Start.
- (b) Check if the basic pay is less than 10000. If it is, calculate the total salary as $BP + 0.45BP + 0.15BP$.
- (c) Otherwise, the total salary is $BP + 0.5BP + 0.2BP$.
- (d) Stop.

Script:

```
#!/bin/awk

BEGIN{
    print "Enter the basic pay:"
    getline < "/dev/tty"
    totalSal = $1
    if ($1 < 10000)
        totalSal = 0.45 * $1 + 0.15 * $1 + $1
    else
        totalSal = 0.50 * $1 + 0.20 * $1 + $1
    print "Total Salary: " totalSal
}
```

Output:



```
rwithik@Zeus Scripts/Awk (master *)
» awk -f 4.awk
Enter the basic pay:
12000
Total Salary: 20400
rwithik@Zeus Scripts/Awk (master *)
» █
```