IT Workshop II

Regex Practice Questions

The following questions on `grep`, `sed` and `awk` filters are meant for practice only. Attempt each yourself at your own pace. The sooner the better.

1 grep/egrep

1. Read the `info/man` page of `grep/egrep/fgrep/rgrep` commands. Understand the differences. Try out `grep`'s various flags/options. (`apropos grep` to see more variants).

2. For the following questions, use the `/usr/share/dict/words` dictionary file.

   (a) Find all words that have the string "India" or "Africa " in them.

   (b) Find all words that begin with a vowel. How many are there?

   (c) Now, find the words that begin with a vowel and end with a vowel (not necessarily the same vowel).

   (d) Find all words that have two or more consecutive 'a' in them.

   (e) Find all words that have any punctuation marks (such as apostrophe) in them.

   (f) Find words that contain an upper case letter:

   i) at the beginning ii) at the end iii) anywhere

   (g) Find all words that are twenty characters or more in length.

   (h) Find all five-character or ten-character words that do not contain any upper case letter.

   (i) Find all words that begin and end with the same letter.

2 `sed`

1. Redo the above questions under (2) using `sed` only.
2. You are given the following address book details (in CSV format) in the order: `given name, surname, address, city, DoB` and `e-mail`.

```
Alice, George, 83 First St, Howard, 12/05/1980, awatson@example.com
Bob, Charlie, 40 West Ave, Anycity, 02/11/1995, bcharlie@example.com
Steve, Park, 1981 Songdo, Smartcity, 24/1/1971, spark@exmaple.com
Mary, Sam, 40 West Ave, Anycity, 30/01/1990,  msam@example.com
Kumar, Anil, Gachibowli 500032, Hyd, 11/12/1989, kanil@example.com
```

Write `sed` commands for the following.

1. Who lives in a city named `Anycity`? Store their details in `output.txt`.
2. Remove all lines that begin with a vowel.
3. How many people are born in the 80s? Their names?
4. Replace all numeric values by `'?'` and all punctuation marks by `'*'`.
5. Replace all months in DoB (may be prefixed by 0) by their string equivalents. (e.g. `12/05/1980 or 12/5/1980` → `12/May/1980`)
6. Reverse the order of the Given name and Surname.
   (e.g. `"Mary, Sam"` → `"Sam, Mary"`).
7. Swap the first and last characters of a line.
   (e.g. `"Alice ...  example.com"` → `"mlice...example.coA"`).

3 awk

You are given the marks of students taking a course named ABC.

Name	Gender	Mid1 (25%)	Mid2 (25%)	Endsem (50%)
AA	M	20	15	35
BB	F	22	17	44
CC	F	19	14	25
DD	M	15	20	42
EE	F	18	22	30
FF	M	0	20	45

Now, write an awk command to achieve the following:

1. Display just the names, genders and endsem marks of:
   (a) all students (b) the first 3 students only (c) the last student
2. Store the male students in `males.txt` and the females in `females.txt`.
3. Compute the total mark for each student and display it along with their names.
4. Who is the topper? Which students scored above the class average?
5. Generate a grade report for each student based on the following marking scheme.

   `A = [95-100], A- = [90-95), B = [85-90), B- = [80-85)`

   `C = [75-80), C- = [70-75), D = [60-70), F` $<$ `60`

   Your report must contain a header and a footer/end sections as well. The header be entitled, "*** Grade Report for the ABC course ***". At the end of the report, display the total number of students, the highest, lowest *&* average marks as well as the message, "*** End of Grade Report ***" .