

## TP n°5 – filtres sur les fichiers m[*o*é]té?o

Ce TP ne donne aucune piste pour les réponses. A vous de piocher dans le catalogue des filtres et leurs options. Ce TP est également très long, vous pourrez essayer d'avancer sur la partie 2 en dehors des créneaux. Ce TP constitue le programme du contrôle de TP.

### 1) Partie 1 : la météo US

Recopier le fichier `/usr/local/anonymous/SYS/1A/tp5/meteoUS` chez vous. Ce fichier est un bulletin météo américain (qui date de 2001, mais les données sont intéressantes pour faire des recherches).

Chaque ligne contient 7 champs séparés par des espaces. Le premier champ est le nom de la ville, ensuite les températures de la journée min et max (en °F, les températures négatives, s'il y en a, sont indiquées par un B comme *below* précédant le nombre, ex : B52 signifie -52°F), ensuite on trouve les prévisions : il y a la couleur du ciel (partially, moderately, very cloudy, rain, snow, cindy, storms, showers...) et les températures min/max prévues, et ce pour les deux jours suivants.

NB : la liste est déjà triée dans l'ordre des villes.

#### a) Facile

1. Afficher (seulement) la ligne concernant BATON\_ROUGE.
2. Afficher page par page les lignes triées sur les températures maximales du jour. Remarquer que les données sont mélangées aux lignes de titre. Voir la question pour éviter cela.
3. Afficher les lignes des villes dont le nom commence par A (11 réponses).
4. Compter le nombre de villes où le temps est FAIR (*correct* en français). Attention, la bonne réponse est 11, pas 12, pourquoi ?! Il faut prendre l'habitude de toujours vérifier les résultats quand c'est possible.
5. Afficher les lignes dont le nom de ville contient la lettre X (6 réponses).

#### b) Moyen

6. Avec la commande `egrep`, créer un second fichier (`meteo`) ne contenant pas les lignes d'en-tête, uniquement ce qui concerne les villes. On remarquera que ces lignes commencent toutes par une lettre et pas par un espace (2 solutions).

Travaillez maintenant uniquement avec le fichier `meteo` filtré.

7. Afficher les lignes où au moins l'une des températures est située entre 20 et 29 (11 réponses).
8. Afficher les lignes concernant les villes dont le nom fait exactement 4 lettres (4 réponses).
9. Afficher les lignes dont le nom de ville contient la lettre Y (19 réponses). NB : il faut garder les lignes entières.
10. Afficher les lignes où la température maximale du jour est comprise entre 90°F et 99°F (4 réponses). Gardez les lignes entières.
11. Afficher les lignes (entières) où la température minimale du jour est comprise entre 70°F et 79°F.

#### c) Difficile

12. Afficher la température maximale la plus fréquente (11 villes ont 76°F).
13. Afficher les lignes où le nom de la ville contient CITY et le temps est SUNNY.
14. Afficher les lignes où le nom de la ville contient CITY ou le temps est WINDY (ou non exclusif).
15. Afficher les lignes où soit le nom de la ville contient CITY, soit le temps est WINDY mais pas les deux à la fois (ou exclusif).
16. Afficher les 5 villes les plus froides (plusieurs sens à cette phrase, choisissez le votre).

#### d) Très difficile

17. Afficher les villes dont le nom comprend un caractère non alphabétique quelconque (un tiret ou autre). Vous pouvez vérifier votre réponse en relançant la commande avec l'option `-v`, ça ne doit afficher que les villes dont le nom n'est composé que de lettres.

18. Certaines villes ont un nom composé qui « commence pareil ». Par exemple, on trouve FORT\_SMITH et FORT\_WAYNE, SAN\_FRANCISCO et SAN\_JOSE. On désire afficher la liste des débuts identiques.

Pour commencer, remarquer que le nom est composé de mots séparés par un `_`. Il existe une commande capable de couper les lignes sur un certain caractère, l'utiliser pour garder le début du nom. Ensuite afficher les noms en double ou plus.

## 2) Partie 2 : championnat de vitesse moto

Recopier le fichier `/usr/local/anonymous/SYS/1A/tp5/moto2011` chez vous. Ce fichier donne le classement 2011 des coureurs en moto de vitesse (comme pour la météo, les données sont obsolètes et en plus, elles sont fausses : elles ont été retouchées pour être intéressantes pour ce TP). Il y a les trois séries dans le fichier : `motogp` (800 cm<sup>3</sup>), `moto2` (250 cm<sup>3</sup>) et `125` cm<sup>3</sup>. C'est le premier champ qui indique ça pour chaque coureur. Ensuite, on trouve le classement dans la série, le nom du coureur, son pays, l'équipe ou la marque de la moto et enfin la somme des points obtenus dans les différentes courses de la saison.

1. Compter le nombre de coureurs inscrits en 250 cm<sup>3</sup>.
2. Compter le nombre de coureurs français.
3. Compter le nombre de coureurs allemands (pays = GER). Appelez ensuite l'enseignant pour qu'il vienne vous taper sur les doigts si vous n'avez pas spontanément pensé à vérifier le résultat.
4. Afficher le nombre d'équipes (marques) différentes présentes dans le fichier (la plupart des équipes ont plusieurs coureurs, éventuellement dans plusieurs séries).
5. Afficher le nom de la série (800, 250 ou 125) qui possède le plus d'inscrits.
6. Afficher les noms des coureurs qui sont inscrits dans plusieurs séries. (En fait, ça a été le cas il y a quelques années, mais le coureur en question n'est plus dans cette situation).
7. Afficher le nom du coureur qui a le plus de points dans l'absolu. (C'est le pilote qui a gagné le plus de courses toutes séries confondues car le nombre de points est obtenu de la même manière, 25 pts pour le premier, 20 pour le second, 16 pour le troisième...)
8. Afficher le nom de l'équipe qui a le plus de coureurs inscrits toutes séries confondues.
9. Afficher les pays qui ne comptent pas plus d'un seul inscrit par série. ex: FIN car en 125, il n'y a que Niklas AJO...

Ça y est, ça commence à rentrer ?.. sinon, les enseignants sont là pour vous aider.