

TD3 Regex

OS Linux

Tom MAFILLE

Partie 1 - Pour débiter

1. Pour toutes les occurences de fraise en ignorant la casse :

```
grep -E -i '\bfraise' fruits.txt
```

Sinon, pour seulement les mot fraise ou Fraise :

```
grep -E '[Ff]raise' fruits.txt
```

2. Pour toutes les lignes dont le nom se termine par “se” :

```
grep -E 'se$' fruits.txt
```

3. Pour les lignes contenant ai :

```
grep -E 'ai' fruits.txt
```

4. Pour les lignes d’exactement 5 caractères :

```
grep -E '^.{5}$' fruits.txt
```

5. Pour les lignes contenant un chiffre :

```
grep -E '[0-9]' fruits.txt
```

Partie 2 - Pour poursuivre

1. Pour compter le nombre de lignes du fichier :

```
grep -E -c '' french.txt
```

Il y en a 348359.

2. Pour trouver le nombre de mots commençants et finissants par z :

```
grep -E -c '\bz\w*z\b' french.txt
```

Et la liste de ces mots :

```
grep -E '\bz\w*z\b' french.txt
```

3. Pour trouver le nombre de mots qui contiennent 23 lettres ou plus :

```
grep -E -c '^.{23,}$' french.txt
```

4. Pour trouver le nombre de mots qui commencent par ‘bio’ et finissent par ‘que’ :

```
grep -E -c '\bbio\w*que\b' french.txt
```

Et les afficher :

```
grep -E '\bbio\w*que\b' french.txt
```

5. Mots de 6 lettres commençant par a, suivis d’une lettre, d’un y, de deux lettres et se finissant par un e :

```
grep -E '\ba\w{1}y\w{2}e\b' french.txt
```

Partie 3- Pour aller plus loin

1. Chaîne d'ADN :

```
grep -E '^[ACGT]+$' doc.txt
```

2. URL :

```
grep -E '\bhttps?://\[a-zA-Z0-9.-]+\b' doc.txt
```

3. Couleur hexa :

```
grep -E '#[a-fA-F0-9]{6}' doc.txt
```

4. Date au format jj/mm/aaaa :

```
grep -E '\b[0-9]{2}/[0-9]{2}/[0-9]{4}\b' doc.txt
```

5. Plaque d'immatriculation

```
grep -E '\b[A-Z]{2}-[0-9]{2}-[A-Z]{2}\b' doc.txt
```

6. Numéro de sécu :

```
grep -E '\b[0-1] [0-9]{2} (0[1-9]|1[0-2]) [0-9]{2} [0-9]{3} [0-9]{3} [0-9]{2}\b' doc.txt
```

7. Adresse mail :

```
grep -E '^[a-zA-Z0-9._%+-]+@[a-zA-Z0-9.-]+\.[a-zA-Z]{2,}$' doc.txt
```