

Énoncé des TD, TD1

[Énoncé provisoire -- dernière mise à jour : 2017.01.24 15h45]

TD1 : analyse de texte

Le premier TD consiste à écrire un programme simple en Python tout en pratiquant le traitement des chaînes de caractères pour lequel Python est particulièrement utile.

Le travail consiste à calculer [l'indice de lisibilité Flesch \(Flesch Reading Ease\)](#) d'un extrait du texte *Colonel Chabert* de Balzac :

"Ces diverses exclamations partirent à la fois au moment où le vieux plaideur ferma la porte avec cette sorte d'humilité qui dénature les mouvements de l'homme malheureux. L'inconnu essaya de sourire, mais les muscles de son visage se détendirent quand il eut vainement cherché quelques symptômes d'aménité sur les visages inexorablement insouciantes des six clercs. Accoutumé sans doute à juger les hommes, il s'adressa fort poliment au saute-ruisseau, en espérant que ce pâtre lui répondrait avec douceur. "

Le programme doit afficher l'indice de lisibilité, ainsi que l'ensemble des valeurs utilisées dans la formule de calcul (nombre de phrases, mots, etc...) en utilisant les expressions régulières pour déterminer ces valeurs.

Le fichier doit pouvoir être exécuter en ligne de commande avec `python td1.py` ou être importé avec `import td1`. Dans le premier cas, l'ensemble des valeurs doit être affichée. Dans le second cas, une fonction `td1.calculerFlesh(texte)`, qui prend en argument une chaîne de caractère et renvoie l'indice de Flesh, doit être disponible.

Les façons habituelles de spécifier des mots et des phrases doivent être reconnues.

Le code doit être documenté.

Documentation

- [Tutorial Python](#)
- [Documentation sur les expressions régulières](#)

Lecture d'un fichier

Pour lire le contenu d'un fichier, vous pouvez exécuter les commandes suivantes:

```
>>> with open("col-chabert.txt", "r", encoding="utf8") as f:
...     data = f.read()
```

Accueil

Séances

Plan

Descriptions

Évaluation

Références

TD

TD1

TD2

TD3

X

Fichiers

- [col-chabert.txt](#) : texte à analyser dans lequel les syllabes d'un mot sont séparées par le caractère barre oblique ('/'). L'encodage est UTF-8.

Exigences et dépôt du travail

- Vous devez remettre le travail en format électronique sur le site Moodle du cours.
- Remettez un fichier "td1.py" qui contient les commentaires et tout le code. Vous n'avez pas à écrire de rapport.
- Remettez un fichier `instructions.txt` contenant les noms et matricules des coéquipiers et autres informations s'il y a lieu.
- Le programme doit pouvoir rouler avec la ligne de commande suivante :

```
python td1.py
```


Il doit rapporter l'indice Flesch-kincaid et la valeur des différents termes de la formule.
- Le code doit contenir des tests unitaires

Barème TD1

- Exigences satisfaites
- Utilisation des structures de données et de contrôle appropriées
- L'architecture est modulaire et se conforme aux principes présentés
- Démonstration de la connaissance de Python
- Commentaires et/ou documentation adéquats

Accueil
Séances
Plan
Descriptions
Évaluation
Références
TD
TD1
TD2
TD3