

Analyse fréquentielle de texte

Présentation

L'analyse fréquentielle, ou analyse de fréquences consiste à examiner la fréquence des lettres employées dans un texte.

C'est une méthode qui est utilisée aussi en **cryptanalyse** pour décoder des messages chiffrés par substitution, dont un exemple très simple est le **chiffre de César**. La description la plus ancienne de cette méthode est réalisée par **Al-Kindi** au IXe siècle.

- [Exemple d'utilisation de l'analyse fréquentielle](#)

Principe

Source Wikipedia : https://fr.wikipedia.org/wiki/Analyse_fr%C3%A9quentielle Selon la langue, un texte comportera une répartition particulière des fréquences de lettres. Par exemple en français, les lettres les plus fréquentes, c'est-à-dire les lettres que l'on retrouve le plus souvent, sont le E, suivi du A, du I et du S ... On obtient ainsi la répartition de fréquences des lettres suivante (en %) :

	A	B	C	D	E	F	G	H	I	J	K	L	M
Français	9,42	1,02	2,64	3,39	15,87	0,95	1,04	0,77	8,41	0,89	0,00	5,34	3,24
	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
	7,15	5,14	2,86	1,06	6,46	7,90	7,26	6,24	2,15	0,00	0,30	0,24	0,32
	A	B	C	D	E	F	G	H	I	J	K	L	M
Anglais	8,08	1,67	3,18	3,99	12,56	2,17	1,80	5,27	7,24	0,14	0,63	4,04	2,60
	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
	7,38	7,47	1,91	0,09	6,42	6,59	9,15	2,79	1,00	1,89	0,21	1,65	0,07

Ce qui donne l'ordre suivant pour la langue française :

E A I | S T N R U L O D M P C V Q G B F J H Z X Y K W

Cela est cependant **approximatif** et dépend de nombreux paramètres tels que le niveau de langue du texte, le style d'écriture. On peut aussi analyser la fréquence dans un texte des **digrammes**, c'est-à-dire des groupes de deux lettres ce qui amène des indices importants pour décrypter un texte chiffré car on sait que l'on ne pourra trouver des digrammes tels que **XK** ou **WX** dans le texte clair.

Les activités ...

[Je reviens à la liste des activités.](#)

From:
[/ - Les cours du BTS SIO](#)

Permanent link:
[/doku.php/icn/facultatif/pythonanalysefrequentielle](https://doku.php/icn/facultatif/pythonanalysefrequentielle)

Last update: 2017/12/21 23:25

