

Le Langage R

Librairies

library () : donne la liste des packages installés.

library (grDevices) : pour charger une librairie (si la librairie est déjà chargée, ne fait rien).

library (help = stats) : pour donner la liste des fonctions d'une librairie (ici c'est la librairie stats)

library (help = "base" ») : pour avoir de l'aide sur les fonctions de la librairie base.

Search() : voir les librairies déjà installées.

Commandes de base

- *q ()* : pour quitter.

- *help (solve) ou ?solve* : pour avoir de l'aide sur solve.

- *help ("if")* : avoir de l'aide sur if (quotes nécessaires car mot réservé).

- *help.start ()* : lance un navigateur pour l'aide en html.

- *help.search ('chi')* : pour chercher dans l'aide avec la partie de mot clef 'chi'.

- *demo (package="stats")* : liste des démos du package stats.

- *demo (nlm, package="stats")* : fait tourner la démo de nlm sur le package stats)

**** Les packages dans R :**

1- Le Package 'tm' :

« tm » est un paquet « text mining » le plus populaire de R. Il est utilisé pour l'analyse statistique de textes. La librairie aide pour tout ce qui est préparation, constitution de corpus, analyses simples (fréquences par exemple).

- Importer des fichiers avec le paquet tm

Le paquet *tm* est conçu pour marcher avec une variété de formats : textes simples, articles/papiers en PDF ou Word, documents Web (HTML, XML, SGML), etc. Il fournit entre autres les fonctionnalités suivantes :

- Un dispositif pour analyser des corpus.
- Des fonctions nouvelles pour l'analyse de textes d'un corpus.

→ Pour installer le paquet : `install.packages ('tm')`

→ Pour utiliser le paquet : `library (tm)`

- Importation de documents et corpus

Le corpus représente une collection de documents que l'on veut analyser. L'analyse de corpus doit être faite en important des textes.

Il existe 2 versions :

VCorpus : `alias>code>Corpus`. Ce type de corpus est volatile, lorsqu'on arrête R, il faut reconstruire le corpus.

PCorpus : Il est persistant, définit un endroit unique où le corpus est stocké.

2- Le package Wordcloud :

Wordcloud est un paquet qui permet de générer de word-Cloud. Autrement dit, il permet de créer de jolis nuages de mots, de visualiser les différences et similitudes entre les documents.

3- Le package SnowballC :

SnowballC est un paquet qui est utilisé pour le text stemming. Il facilite la comparaison du vocabulaire. Les langues actuellement prises en charge sont : le danois, le néerlandais, l'anglais, le finnois, le français, l'allemand, le hongrois, l'italien,

Norvégien, Portugais, roumain, russe, espagnol, suédois et turc.

4- Le package RColorBrewer :

RcolorBrewer est un paquet qui permet de faire des schémas de couleurs pour les cartes et autres graphiques.

5- Le package devtools :

Devtools est un paquet qui permet d'installer une librairie depuis Github grâce à la fonction « install_github() ».

```
Library("devtools")
```

```
Install_github("dgrtwo/broom")
```

6- Le package tidyverse :

Tidyverse est un paquet qui contient notamment :

- ***readr***: pour importer des données dans R
- ***dplyr***: pour la manipulation des données
- ***ggplot2***: pour la visualisation des données.
- ***datarium***: contient des données de démonstration pour des analyses statistiques.