



Guide préparatif « Automatiser sa vigie sur Twitter avec R »¹

Date : Jeudi 25 juin 2015

Horaire : 9h à 16h

Local : C-2059

Adresse :

Université de Montréal
Pavillon Lionel Groulx
3150, rue Jean-Brillant
Montréal QC H3T 1N8

Contact (Informations) :

Francis Fortin
Francis.Fortin@umontreal.ca

Bonjour à toutes et à tous,

Ce document constitue un guide préparatif en vue de l'atelier de formation « Automatiser sa vigie sur Twitter avec R » auquel vous êtes inscrit(e). Il vise à vous fournir un certain nombre d'informations préparatoires en vue de cette journée. **Veillez prendre connaissance de l'intégralité du document afin d'assurer le bon déroulement de cet atelier.** Le document abordera notamment les éléments suivants :

- I. Installation des logiciels R et RStudio (à effectuer en amont de l'atelier);
- II. Quelques préalables et autres exercices introductifs (à effectuer en amont de l'atelier);
- III. Ressources complémentaires;
- IV Agenda final de l'atelier de formation.

I. Installation des logiciels et configurations (à effectuer en amont de l'atelier) :

L'atelier de formation nécessite d'avoir recours à quelques logiciels. Ces derniers seront utilisés à différentes étapes lors de la journée de formation. Afin de faciliter le déroulement, mais surtout de maximiser le temps que nous passerons à effectuer directement les manipulations de collecte de données, il vous est demandé d'installer en amont de l'atelier de formation une série de logiciels et d'effectuer des configurations préalables. En effet, il est important, en particulier pour les participants disposant d'ordinateurs institutionnels d'être en mesure d'installer en amont de l'atelier de formation les logiciels et les modules indiqués afin de prévenir tout blocage potentiel en raison de droits d'accès informatiques restreints. Les modalités d'installation pour les différents logiciels sont décrites ci-dessus.

A. Installation de R et RStudio

La première étape est d'installer le logiciel R. Ce logiciel est le moteur des opérations. Par la suite, vous devez installer RStudio qui est l'interface de R. Il est la carrosserie apposée sur le moteur. Toutes vos interactions seront effectuées grâce à RStudio.

¹ Merci à Benjamin Ducol pour m'avoir permis d'utiliser plusieurs sections dans ce document.



Installer R : <http://cran.r-project.org/>

Installer R-Studio : <http://www.rstudio.com/products/rstudio/download/>

Pour ceux qui veulent en savoir plus :

Installation version longue (Français)

<https://www.youtube.com/watch?v=z59abRrXDw>

*Introduction à R et RStudio (Anglais) **Je vous conseille celle-ci :***

<https://www.youtube.com/watch?v=1VKMsawju8w>

Sur YouTube, vous pouvez chercher R tutorial.

B. Création d'un compte sur Amazon (facultatif voir dans le texte)

Plusieurs d'entre vous connaissent Amazon pour la vente de livre. Amazon est toutefois devenu un géant dans les services infonuagiques (le cloud). Pour les besoins de la formation et pour ceux qui sont intéressés, nous allons créer un espace dans les nuages pour héberger RStudio Server. Essentiellement, on demande à Amazon de nous donner un ordinateur virtuel avec lequel on va interagir. Les avantages sont nombreux : 1) l'ordinateur est toujours en ligne, 2) on peut le contrôler à partir de n'importe où, 3) le système est très sécuritaire, etc.

Dans le cadre cette formation, il sera possible d'utiliser deux formes d'installation du logiciel R / RStudio : 1) le serveur qui fonctionne sur Amazon (RStudio Server) et 2) le logiciel qui fonctionne localement (RStudio client).

Comment choisir mon type d'installation ?

1. Amazon fournit un espace d'hébergement gratuit pour 1 an. Pour obtenir cet accès, il faut donner sa carte de crédit. Si ceci pose un problème, vous pourrez utiliser l'installation locale (celle décrite au point « A »).
2. Si vous choisissez d'utiliser Amazon Cloud Services, il vous suffit de donner vos informations ici : <https://aws-portal.amazon.com/gp/aws/developer/registration/index.html>
3. Nous procéderons à l'installation et à la configuration de votre compte Amazon lors de la formation. À noter que cette installation est très rapide et efficace. Aucun besoin d'avoir un niveau technique avancé.

C. Création des droits d'accès aux APIs de Twitter (pour ceux n'ayant pas déjà leurs codes de développeurs)

Dernière étape dans nos installations, la création des droits d'accès aux APIs – Application Programming Interface ou Interface de Programmation – de Twitter. Avant toute chose, si vous n'avez pas de compte Twitter, il sera nécessaire de vous en créer un. En effet, Twitter contrôle l'accès à son interface de programmation et aux données disponibles du réseau en ne donnant accès qu'aux utilisateurs dûment enregistrés et ayant obtenu les droits d'accès nécessaires. Si vous ne disposez pas déjà d'un compte Twitter personnel, c'est le moment d'en créer un. Pour cela, il vous suffit de suivre les instructions à partir de la page suivante : <https://twitter.com/?lang=fr>

Une fois votre compte personnel Twitter crée, il convient obtenir vos droits d'accès personnels afin de pouvoir vous connecter aux APIs de Twitter. De nombreux tutoriels en ligne expliquent comment obtenir ces droits et les identifiants qui y sont associés. Vous retrouverez sur cette page les principales étapes à suivre pour obtenir ces droits et ces identifiants personnels vous permettant de connecter aux APIs de Twitter : <http://www.sketchthemes.com/tutorials/getting-new-twitter-api-consumer-and-secret-keys/>. Je vous conseille de suivre les étapes évoquées ci-dessus en vous aidant néanmoins des captures d'écran proposées dans la page Web ci-dessus :



- Connectez-vous sur votre compte *Twitter* avec votre identifiant et votre mot de passe habituel.
- Allez ensuite sur l'URL : <https://apps.twitter.com>.
- Remplissez les champs suivants :

Name => mettez un nom unique pour votre, par exemple « Scaper Twitter » ou « Collecteur de données sur Twitter »

Description => mettez une brève description de l'outil tel que « Cet outil vise à collecter des données sur Twitter »

Website => mettez « <http://outil.local/> »

Callback URL=> mettez « <http://outils.local/> »

- Après avoir lu et validé les conditions, cliquez sur créer votre application. Vous arrivez à cette étape sur la page d'accueil de votre
- Dans l'onglet *Details* et allez à la ligne *Access* afin de choisir l'option « Read, Write and Access direct messages » puis validez ce choix.
- Dans l'onglet *Settings* et cochez « Allow this application to be used to Sign in with Twitter »
- Dans l'onglet *Keys and Token*, allez à la ligne *Access* afin de choisir l'option « Read, Write and Access direct messages » puis validez ce choix. Toujours dans cette page *Keys and Token*, cliquez sur l'onglet « Regenerate Consumer Key and Secret », puis cliquez sur l'onglet « Regenerate Access Token and Token Secret ».
- Créer un fichier *Word* dans lequel vous allez sauvegarder les informations disponibles sur la page *Keys and Token* :

Consumer Key (API Key) : XXXXXXXXXXXXXXXXXXXXXXXXXXXX

Consumer Secret (API Secret) : XXXXXXXXXXXXXXXXXXXXXXXXXXXX

Access Token : XXXXXXXXXXXXXXXXXXXXXXXXXXXX

Access Token : XXXXXXXXXXXXXXXXXXXXXXXXXXXX

À cette étape vous avez créé les droits d'accès et vos identifiants pour accéder aux APIs de *Twitter*. Félicitations...

D) Créer un compte de courriel pour l'occasion

Il est conseillé de créer un compte Hotmail, Gmail ou Yahoo pour l'occasion.



D. Accès au contenu

Dans le cadre de l'atelier, je serai amené à partager avec vous des documents et des scripts R. Tous les contenus seront disponibles sur le site <http://www.formationcyber.com> . Si on vous demande le mot de passe, veuillez indiquer ceci : gardeso2015 .

II. Quelques préalables et autres exercices introductifs (À effectuer en amont de l'atelier) :

L'atelier de formation constitue une introduction à la collecte de données sur Twitter Elle ne requiert pas de connaissances préalables importantes en matière de programmation ou en informatique, même si une connaissance minimum dans ces domaines constitue un atout certain. En vue de faciliter la compréhension des concepts et des dimensions abordés lors de l'atelier de formation, je vous propose quelques lectures, vidéos et exercices introductifs. S'ils ne sont pas obligatoires pour suivre la journée de formation, ces préalables et exercices introductifs vous permettront de comprendre plus facilement le contenu de celle-ci.

- 1) *Introduction à R et RStudio (Anglais) **Je vous conseille celle-ci** :*
<https://www.youtube.com/watch?v=IVKMsawju8w>
- 2) *Un excellent tutoriel pour R (comprendre les fonctions statistiques)*
<http://tryr.codeschool.com/levels/1/challenges/7>

IV. Agenda final de l'atelier de formation

9h-9h45	Présentation
	Présentation des objectifs Présentation du déroulement de l'atelier Installation (ou vérification d'installation) des logiciels nécessaires Explication des outils et des techniques utilisés
9h45-10 h30	Créer son espace sur Amazon EC2
	Qu'est-ce qu'un service infonuagique? Les avantages et inconvénients. Configuration de notre espace. Familiarisation de l'interface R-Studio
10h30 - 10h45	Pause
10h45-12h	R, RStudio et Twitter : un mariage heureux
	Présentation des concepts de bases dans R / RStudio Présentation des fonctionnalités de base de R / RStudio
12h - 13h15	Repas
13h15-14 h30	Analyser les données de Twitter
	Expérimentation du script CriminiTwit v1.5 <i>Mots clés, profils et tendances.</i>
14h30-14 h45	Pause
14h45 - 16h	Visualiser ses données
	Nuages de mots Analyse de sentiments Diagrammes de relation (d3Network) Évaluation