

RC1. Data Explorable

Domaines	Data science, Data Mining, Big Data...
Librairies (proposées)	scikit-learn, seaborn, bokeh, d3js, chart.js, etc.
Langages (proposé)	Python, javascript, html/css, etc.
Responsable du sujet	Rémy Cazabet
Informations à voir	http://cazabetremy.fr/Teaching/LIFPROJET.html

Le principe de ce projet est de choisir une source de données accessible (open data, data crawling, etc), de collecter ces données, les analyser, et construire un “explorable” permettant à n'importe qui d'explorer ces données de manière interactive, typiquement sous la forme d'un site web.

Vous êtes libre de choisir les données qui vous intéresse, et nous discuterons des analyses que nous pouvons en faire (restitution sous forme de cartes, de réseaux/graphes, de graphiques interactifs, sous forme de site-web ou non, etc.).

Quelques exemples classiques de données: Données météorologiques, Données de santé (Covid...), Données sportives (résultats sportifs par saison ou par match...), Culturelles (données de films, de musique...), Restaurants/Musées... Mais aussi: (politique, réseaux sociaux, économie, environnement, sécurité, etc.)

Si vous ne savez pas quelles données existent ou non, partez du principe que tout existe, et nous chercherons ensemble ce qui est effectivement accessible. Si vous ne savez pas où trouver des données, quelques points de départ: <https://data.grandlyon.com> , <https://www.data.gouv.fr/fr/> , <https://datasetsearch.research.google.com> .

Vous pouvez consulter quelques projets précédents sur ma page <http://cazabetremy.fr/Teaching/StudentsProduction.html> .

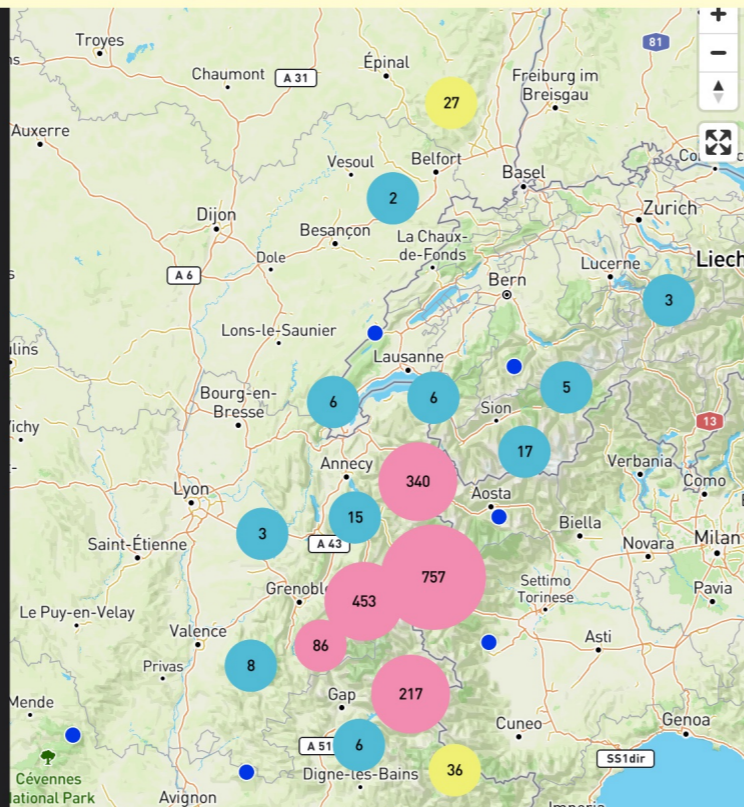


AVALANCHES - Data Mining

Infos sur l'avalanche

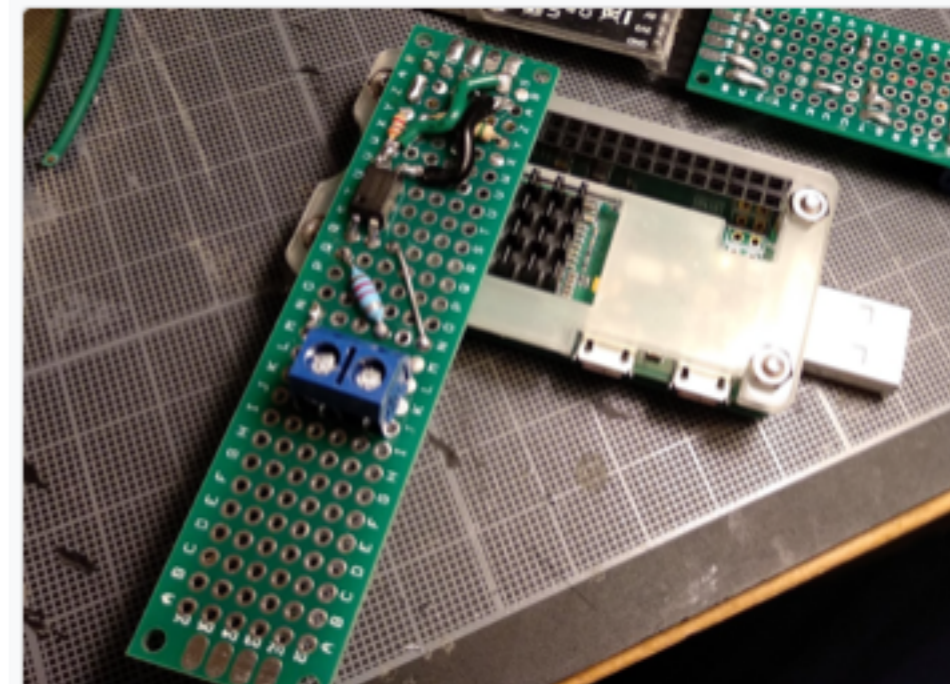
Date : Origine :
 Massif : Décédés :
 Site : Neige :
 Orientation : Activité :
 Description :
 Source :

Tri par année



Restaurants à Lyon

Ce projet permet d'explorer l'offre de restaurants de Lyon à travers une visualisation cartographique. Les données proviennent de Google



LIGHT-DASH

Ce projet permet de récupérer les données de consommation électrique à partir du compteur Linky, ainsi qu'un serveur web informations collectées.



Clique sur un nœud du graphe !

Daenerys Targaryen

Taille des nœuds par :

Couleurs des nœuds par :

Visualisation Réseaux Dynamiques V2

Deuxième version de l'outil de visualisation de réseaux dynamique avec des données de séries télévisés.

RC2. Kaggle Challenge

Domaines	Machine Learning (“Intelligence Artificielle”)
Librairies (proposées)	scikit-learn, keras, tensor-flow...
Langages (proposé)	Python, autre
Responsable du sujet	Rémy Cazabet
Informations à voir	🌐 Challenges Kaggle en cours

[🌐 Kaggle](#) est le site web le plus connu de challenges en machine learning.

En 2 mots: des entreprises proposent des challenges auxquels n'importe qui peut tenter de proposer une solution. La meilleure solution gagne. Mais là où Kaggle est particulièrement intéressant, c'est qu'il propose des tutoriels pour résoudre certains problèmes classiques, en utilisant différentes méthodes. Votre objectif: 1) Suivre un/des tutoriels pour résoudre un problème tout fait, 2) Choisir un challenge en cours et appliquer ce que vous avez appris. Vous pouvez consulter quelques projets précédents sur ma page [🌐 http://cazabetremy.fr/Teaching/StudentsProduction.html](http://cazabetremy.fr/Teaching/StudentsProduction.html).

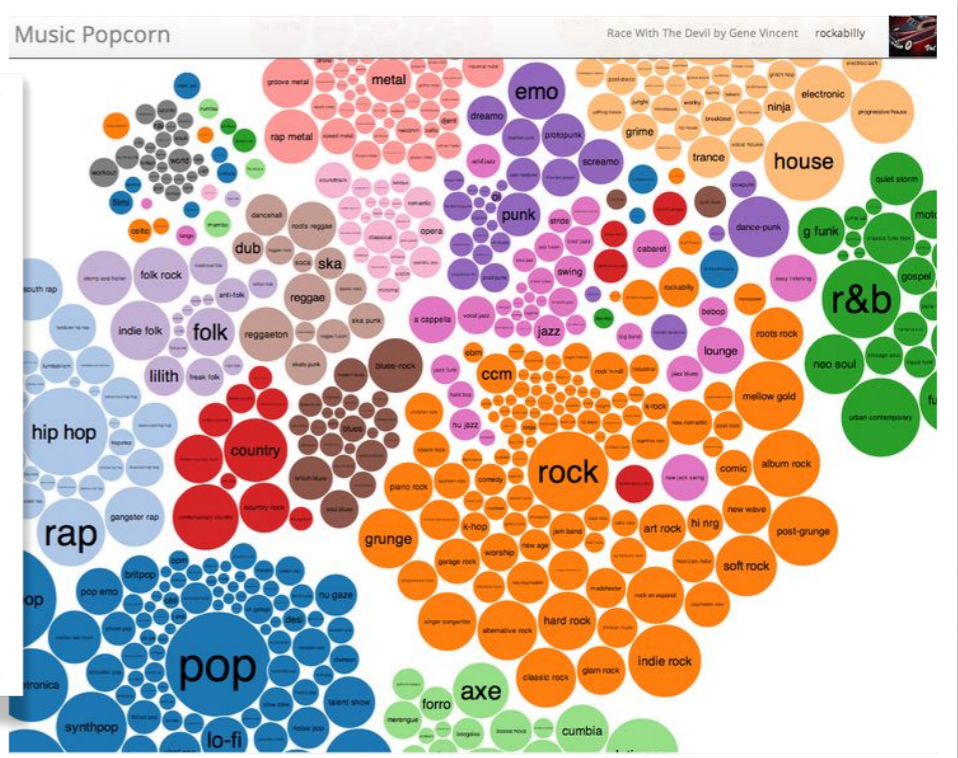
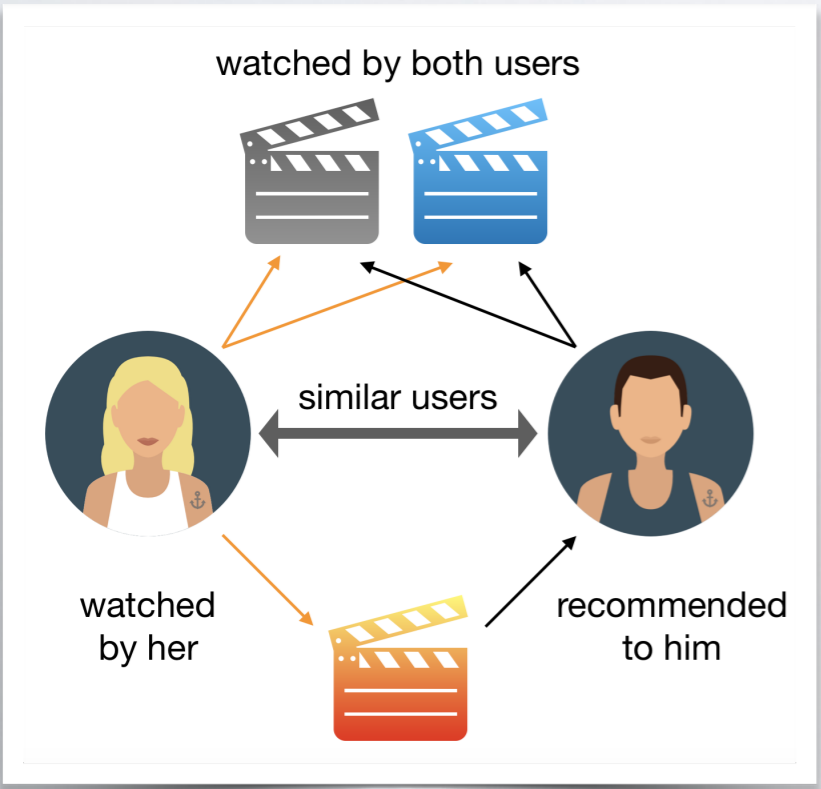
[✎ Modifier](#)

RC3. Recommendation de contenus

Domaines	Machine Learning
Librairies (proposées)	scikit-learn, keras, tensor-flow...
Langages (proposé)	Python, autre
Responsable du sujet	Rémy Cazabet
Informations à voir	Notebook pour commencer

L'objectif de ce projet est de créer un outil de recommandation de contenus: pour un utilisateur qui indique ses préférences (notes ou like) sur des contenus (films, musiques, jeux vidéos...), vous lui recommanderez des contenus similaires, découverts automatiquement à partir des préférences des autres utilisateurs (en se basant sur des bases de données existantes). Vous pouvez travailler uniquement sur l'algorithme lui-même (python uniquement donc), ou faire un site web complet de recommandation. Un exemple de méthode classique pour le faire est le collaborative filtering: https://en.wikipedia.org/wiki/Collaborative_filtering, mais d'autres méthodes plus avancées seront possible pour les groupes qui avancent rapidement.

Modifier



RC4. Web crawling et data analysis de données de publications scientifiques

Domaines	Web crawling, base de données, analyse de données
Librairies	beautifulsup, languages web...
Langages (proposé)	python, javascript
Responsable du sujet	Rémy Cazabet
Informations à voir	https://taylorandfrancis.com/partnership/commercial/accelerated-publication/

Dans le cadre d'un projet de recherche, nous souhaitons étudier le temps écoulé entre la soumission d'un article à un journal scientifique et sa publication. Plus particulièrement, nous nous intéressons à un nouveau modèle économique pour lequel les chercheurs peuvent payer pour accélérer le temps d'évaluation de leurs articles: nous voulons voir les effets sur la vitesse d'acceptation. L'idée est d'écrire des scripts de web-crawling pour récupérer automatiquement ces dates dans des journaux d'intérêts. En particulier ceux proposant cette nouvelle solution (<https://taylorandfrancis.com/partnership/commercial/accelerated-publication/>), mais aussi d'autres journaux typiques pour avoir un point de comparaison. Ces données devront être stockées dans un format standard dans une base de données, et mis à jour automatiquement à intervalle fixe, par exemple chaque semaine. Vous disposerez d'un serveur web pour faire tourner vos scripts. Si vous le souhaitez, vous développerez un site web pour permettre aux chercheurs de voir le temps moyen d'acceptation d'un article dans chaque journal

Modifier

Taylor & Francis
an informa business

Search by subject, title, author, keyword

- Biomarkers*
- Blood Pressure*
- Cancer Investigation*
- Current Eye Research*
- Current Medical Research and Opinion
- Drug Development and Industrial Pharmacy*
- Expert Opinion on Biological Therapy
- Expert Opinion on Drug Delivery
- Expert Opinion on Drug Discovery
- Expert Opinion on Drug Metabolism & Toxicology
- Expert Opinion on Drug Safety
- Expert Opinion on Emerging Drugs
- Expert Opinion on Investigational Drugs
- Expert Opinion on Orphan Drugs
- Expert Opinion on Pharmacotherapy
- Expert Opinion on Therapeutic Patents
- Expert Opinion on Therapeutic Targets
- Expert Review of Anticancer Therapy
- Expert Review of Anti-infective Therapy
- Expert Review of Cardiovascular Therapy
- Expert Review of Clinical Immunology
- Expert Review of Clinical Pharmacology
- Expert Review of Endocrinology & Metabolism
- Expert Review of Gastroenterology & Hepatology
- Expert Review of Hematology
- Expert Review of Medical Devices
- Expert Review of Molecular Diagnostics
- Expert Review of Neurotherapeutics
- Expert Review of Ophthalmology
- Expert Review of Precision Medicine and Drug Development
- Expert Review of Pharmacoeconomics & Outcomes Research
- Expert Review of Proteomics
- Expert Review of Respiratory Medicine
- Expert Review of Vaccines
- Hospital Practice
- International Journal of Neuroscience*
- Journal of Asthma*
- Journal of Dermatological Treatment*
- Journal of Drug Targeting*
- Journal of Medical Economics**
- Minimally Invasive Therapy & Allied Technologies*
- Nanotoxicology*
- Postgraduate Medicine
- Pharmaceutical Development and Technology*
- Scandinavian Journal of Gastroenterology*
- The Physician and Sportsmedicine

REVIEW
OF CLINICAL PHARMACOLOGY

Volume 15, 2022 - Issue 7

Submit an article | Journal homepage

243 Views


0 CrossRef citations to date

0 Altmetric

Listen

Editorial

Biosimilar and generic formulations of novel antidiabetic drugs: the role of liraglutide in clinical pharmacology of type 2 diabetes

Manfredi Rizzo , Francesco Cosentino & Christos Mantzoros

Pages 795-797 | Received 09 Feb 2022, Accepted 15 Jun 2022, Published online: 04 Aug 2022

Download citation | <https://doi.org/10.1080/17512433.2022.2108400> | Check for updates

Full Article | Figures & data | References | Citations | Metrics | Reprints & Permissions | PDF | EPUB

KEYWORDS: Biosimilar | diabetes | generics | GLP-1 | liraglutide | treatment

Previous article | View issue table of contents | Next article

Related research

People also read

An affinity thres efficacy in anti-F

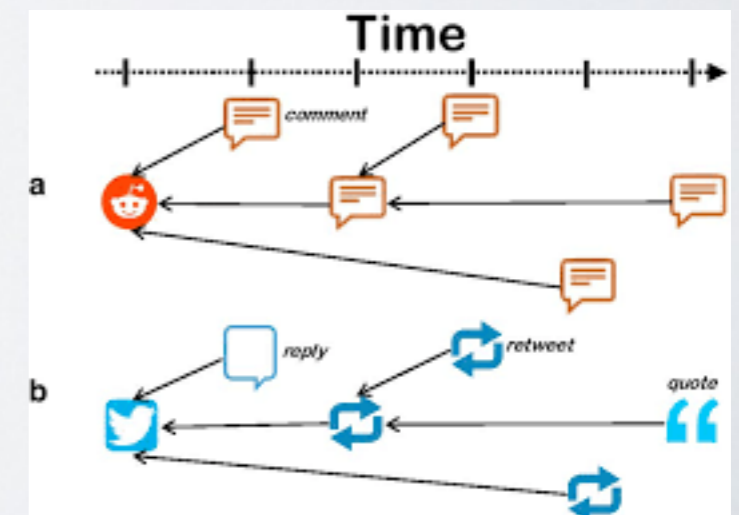
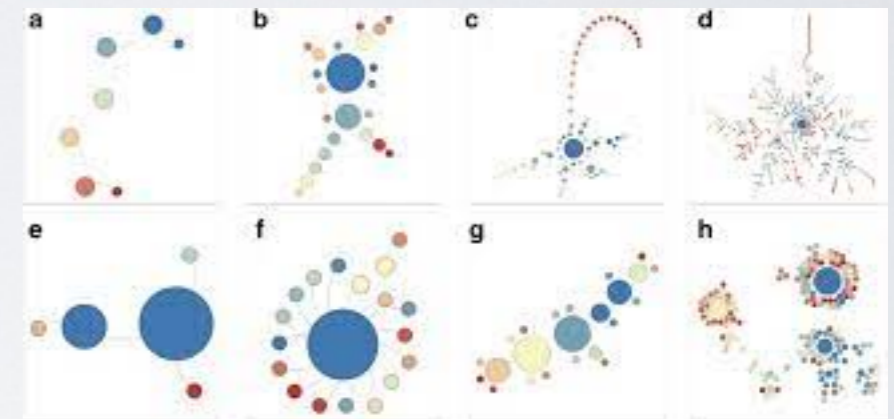
RC5. Outil de visualisation d'un arbre de tweet

Domaines	data viz, analyse de réseaux sociaux, etc.
Librairies	API twitter, javascript, (python..)
Langages (proposé)	python, javascript
Responsable du sujet	Rémy Cazabet

Twitter est un media social très utilisé pour partager et suivre de l'information. Bien que donnant souvent lieu à des débats houleux entre ses membres, il n'a pas d'outil simple pour suivre l'ensemble de l'arbre de discussion autour d'un twitt initial. Vous proposerez un outil graphique, de préférence sous la forme d'un site web, pour afficher des arbres de discussion a partir d'un tweet d'origine. Vous pourrez vous concentrer sur l'aspect collecte et classement des données (quelles réponses montrer lorsqu'il y en a des milliers ?) et/ou sur l'aspect visualisation (arbre interactif avec D3.js où l'on peut plier et déplier des branches à la demande.)

Modifier

The image shows two screenshots from the Twitter interface. The left screenshot displays a thread of tweets: Rachelle O'Brien (Aug 20), Paul Kleiman (Aug 21), and Kiu Sum (Aug 20). The right screenshot shows a detailed view of a tweet by Jane Manchun Wong (@wongmjane) from Nov 8, 2019, with 3 replies. Below the tweet, there are sections for 'All replies to FS' and 'All replies to Daniel'.



RC6. Site web de jeu en ligne interactif

Domaines	Web
Librairies	Utiliser des frameworks complet (MERN, MEAN) https://www.datasciencecentral.com/profiles/blogs/mern-vs-mean-which-stack-to-use-in-2021
Langages (proposé)	Javascript, html, etc.
Responsable du sujet	Rémy Cazabet

Dans ce sujet, vous développerez un site web, en utilisant des frameworks modernes de votre choix, par exemple parmi la liste suivante :

(<https://medium.com/javarevisited/10-best-frontend-and-backend-frameworks-for-java-python-ruby-and-javascript-developers-cce3c951787a>), de préférence en les combinant (MERN, MEAN, MEVN...). Ce site web doit permettre de jouer à un jeu, quel qu'il soit, depuis un jeu que vous avez inventé à un classique de type scrabble, dames, échecs ou autre. Il doit permettre aux joueurs de jouer en temps réel les uns contre les autres (salon de jeu, codes pour rejoindre une partie...), et d'afficher des statistiques (hall of fame, meilleurs scores du jour, etc.).

RC7. Contribution à la librairie Python CDlib (Community Detection Library)

Domaines	Machine Learning, python, developpement logiciel,...
Librairies (proposées)	cdlib, networkx, scikit-learn, seaborn, etc.
Langages (proposé)	Python
Responsable du sujet	Rémy Cazabet
Informations à voir	http://cazabetremy.fr/Teaching/LIFPROJET.html

La librairie python CDlib (Community Detection library) est une librairie spécialisée dans le problème de la détection de communautés dans les graphes, un problème équivalent à celui du clustering mais pour des données de type graphe. Il s'agit de la principale librairie du domaine, avec plus de 100 téléchargements individuels par mois. Je suis l'un des 3 principaux contributeurs de cette librairie. Le sujet de ce POM consiste à contribuer à la librairie, par exemple en y ajoutant de nouveaux algorithmes, en corrigeant des bugs, en améliorant le code ou la documentation, etc. Il y a 2 aspects: 1)Aspect "machine learning", puisqu'il faudra se plonger dans domaine du machine learning, le clustering de graphe, 2)Un aspect développement logiciel, puisqu'il faudra comprendre le fonctionnement d'une librairie python, de la génération automatique de documentation, et coder avec un niveau de qualité suffisant pour intégrer du code à une librairie utilisée par des centaines de personnes.

Références: -Code de la librairie: <https://github.com/GiulioRossetti/cdlib> -Documentation: <https://cdlib.readthedocs.io/en/latest/> -Liste d'améliorations possible identifiées: <https://github.com/GiulioRossetti/cdlib/projects/2>



```
▶ from cdlib import viz  
  
pos = nx.spring_layout(g)  
viz.plot_network_clusters(g, leiden_coms, pos, figsize=(5, 5))
```

⊖ <matplotlib.collections.PathCollection at 0x7f9e05b2f160>

