

TD 4.3 - Big Data

Ex 1 - Reprises des notions

- 1- Faire un schéma qui explicite la démarche de Big Data.
- 2- Expliquer comment une radio privée peut mettre en place une démarche Big Data intéressante pour elle ?
- 3- Expliquer comment une entreprise comme Carrefour peut mettre en place une démarche Big Data intéressante pour elle ?

Ex2 - Messageries et réseaux sociaux

Facebook, Messenger, Telegram, Gmail, Whatsapp, Gchat, Skype, LinkedIn, twitter, Signal ...

- 1) A qui appartiennent ces différents réseaux ? Décrire leur fonctionnement en quelques lignes.
- 2) Quel est l'intérêt pour une entreprise d'animer une messagerie gratuite ?
- 3) Comment pourrait-on s'y prendre pour valoriser les données accumulées par ces entreprises (plusieurs idées attendues)

Ex3 - Les Tableaux croisés dynamiques, une première démarche de big data

De plus en plus sites proposent l'accès à des données publiques. C'est le cas par exemple du site data.gouv.fr. On trouve sur ce site des données très diverses issues d'organismes publics ou de ministères.

- 1) Explorer le site et trouver un jeu de données qu'on téléchargera.
- 2) Analyser la donnée à l'aide d'un TCD pour faire apparaître un fait intéressant.
- 3) Envoyer un mail à l'enseignant pour faire un résumé d'activité indiquant la donnée retenue et la conclusion tirée.

A noter un autre site permettant de récupérer des données libres : <http://data.assemblee-nationale.fr/> ou le site <https://fr.statista.com/>

Ex4

Une organisation humanitaire spécialisée dans la formation sanitaire aux populations en difficulté a décidé de mener une campagne pour obtenir de nouveaux dons. Cette campagne a deux aspects :

A- identifier ceux qui pourraient donner plus parmi les donateurs déjà inscrits

B- trouver de nouveaux donateurs à partir de fichiers de prospectus achetés en lignes.

Pour le premier aspect, l'organisation dispose de plusieurs informations sur les donateurs :

- montant donné dans le passé
- type des dons : réguliers / ponctuels.
- durée des dons
- CSP
- Age
- Genre

Pour le second aspect, l'organisation dispose d'un fichier d'individus obtenu auprès d'un institut de sondage. Ce fichier indique pour chaque individu sa CSP, son age, les associations auxquelles il donne, le montant mensuel de ses dons,

Il est coûteux d'envoyer un message à chaque donateur, l'organisation veut être la plus efficace possible pour cibler les donateurs les plus susceptibles de contribuer.

Quelle démarche pourrait-on lui conseiller ?

Pour aller plus loin :

<https://www.espace-sciences.org/conferences/big-data-en-sante-quels-usages-quels-benefices>

<https://www.parismatch.com/Actu/Politique/L-arme-secrete-du-geek-Fillon-1133708>

<https://www.nouvelobs.com/sciences/20180308.OBS3317/sur-twitter-les-fake-news-se-propagent-beaucoup-plus-vite-que-la-verite.html>