

Projet de la semaine data science

Introduction

Vous allez devoir trouver des données intéressantes dans un ensemble de messages (de votre choix) et les représenter graphiquement.

Export

Vous pouvez sur Gmail, Messenger, WhatsApp et bien d'autres services de mails/messagerie instantanée, exporter vos données (Utiliser votre moteur de recherche préféré pour trouver comment).

Manipulation

Vous allez devoir ouvrir le fichier exporté, interpréter son contenu correctement, éventuellement le nettoyer.

Vous allez devoir en extraire les données que vous jugerez intéressantes, comme par exemple (mais pas exclusivement) :

- Quel contact répond le plus
- Qui écrit avec le moins de fautes d'orthographe
- Qui écrit les plus gros pavés
- Courbe de fréquence de messages
- Courbe de temps entre chaque message

Exporter les données

Vous allez ensuite devoir générer un fichier HTML contenant vos diverses données. L'esthétisme, bien qu'apprécié, ne sera pas évalué.

Modalités de rendu

Vous devrez rendre votre projet par mail à alexandre.th.manuel@gmail.com

Objet : [IIM][DSC] NOM prénom

Pièce jointe : nom.prenom.zip

Vous mettrez dans nom.prenom.zip :

- le fichier Python contenant votre code (extension .py)
- un fichier texte/word/pdf expliquant ce que vous avez fait comme features, sur quelles données vous vous êtes basé, ainsi qu'un exemple du format des données d'entrée.
- Le fichier HTML qu'aura généré votre programme, bien que facultatif, sera apprécié