

Messieurs Mopolo, Pasquier, Simonian et Winkler,

Afin de faciliter la gestion de notre projet TPA, nous avons créé un dépôt sur GitHub qui regroupe l'ensemble des scripts, codes et visualisations associés à ce projet.

Ce dépôt est accessible à l'adresse suivante : <https://github.com/adamspd/TPA>. Il comprend :

- un dossier intitulé "Analyse" contenant l'analyse effectuée avec R,
- un dossier "DataViz" qui regroupe les visualisations,
- un dossier "Mongo" qui décrit la procédure d'installation et d'ajout de données dans MongoDB,
- un dossier "Tables_externes" qui contient les scripts utilisés pour créer des tables externes dans Hive,
- un dossier "nosql" qui regroupe le script permettant l'ajout de données dans Oracle NoSQL,
- un dossier "src/main/java" qui contient un projet Maven pour faciliter la compilation des fichiers Java et la création des jars pour l'exécution dans Hadoop,
- un dossier "vps" qui regroupe les scripts et configurations utilisés pour notre serveur dédié.

Le rapport complet est également disponible sur ce dépôt, ainsi que sur ce lien :

https://docs.google.com/document/d/1QV3StyaYxqpyu_HhiBd4IDKOL3fLPjr8qGP_EsysuSM/edit?usp=sharing.

Nous vous invitons à consulter ces éléments pour une meilleure compréhension de notre travail.