**Tp2 en équipe de 2**

On vous demande d’exploiter des données en format « csv » pour répondre à différentes questions statistiques ( le nombre de, la moyenne de, les x meilleurs (les pires), les plus (les moins),…)

Il y a différents ensembles de données, de type CSV, disponibles sur le site <https://www.kaggle.com/datasets>

<http://donnees.ville.montreal.qc.ca/>

On vous demande de :

* Choisir un data set et le télécharger
* Analyser la structure du dataSet :
  + **Structurellement** : nombre de colonnes, types de données de chacune d’elles (potentiellement car pour un fichier csv, tout est considéré de type texte).
  + La qualité des données : valeurs nulles pour chacune des colonnes, forme (casse, contenu,…)
* Charger les données dans une table Oracle :
  + Il faut procéder à des **conversions** de type ou à une certaine mise en forme(casse par exemple), si nécessaire.
* Utiliser la table obtenue pour répondre à différentes questions « **INTERESSANTES** » en exploitant les données(min 12 requêtes).

**Consignes :**

* Le chargement des données peut se faire de différentes manières :
* Outils visuels ([Sql Developer](https://docs.oracle.com/database/121/ADMQS/GUID-7068681A-DC4C-4E09-AC95-6A5590203818.htm" \l "ADMQS0826))
* [Tables externes](https://oracle-base.com/articles/9i/external-tables-9i)
* [SqlLoader](https://oracle-base.com/articles/10g/load-lob-data-using-sql-loader)

À remettre :

* Le lien du fichier
* Le contenu du fichier
* Le script sql (Vous devez utiliser un commentaire avant chaque requête pour formuler le besoin auquel vous répondez)
* La table.