Université USTHB – Bab-Ezzouar Faculté d'Informatique, Département de l'Informatique, Module : Data Mining Bab-Ezzouar, 2022 / 2023 2<sup>ème</sup> année Master Informatique Système Informatique Intelligent Semestre 1

# Exercice 1 Analyse des Données

#### Boîte à outils JAVA:

### Importation du contenu d'un fichier :

BufferedReader < IDF\_READER> = new BufferedReader(new FileReader(< PATH\_TO\_FILE>));

## Lecture du contenu du fichier ligne par ligne :

String <IDF\_STR> = <IDF\_READER>.readLine();

Retourne null SI la fin du flux est atteinte, SINON une chaîne de caractères.

# Séparation des attributs de chaque ligne :

La méthode split() permet de scinder une chaîne de caractères :

String[] <IDF\_STR\_ARRAY> = <IDF\_STR>.split(<SEPARATEUR>);

Retourne un tableau de chaînes de caractères après découpage de la chaîne globale selon le séparateur choisie.

## Conversion d'une chaîne de caractère selon le type désiré :

Double <IDF\_DOUBLE> = Double.parseDouble(<IDF\_STR>);
int <IDF\_INTEGER> = Integer.parseInteger(<IDF\_STR>);

#### **Ouestions:**

- 1- Écrire une fonction Java permettant de charger le dataset.
- 2- Écrire une fonction Java permettant d'afficher quelques informations de base sur le dataset.
- 3- Écrire une fonction Java permettant de calculer les tendances centrales d'un attribut.
- 4- Écrire une fonction Java permettant de calculer les quartiles  $(Q_0,Q_1,Q_2,Q_3,Q_4)$  d'un attribut.
- 5- Écrire une fonction Java permettant d'afficher le nombre et pourcentage de valeurs manquantes d'un attribut.

Dataset: Dataset-Exos.txt disponible sur ce lien .

Have fun!

HOUACINE Naila Aziza: n.a.houacine@gmail.com