**Atelier Données – Formation SNT**

**Titanic « Les femmes et les enfants d’abord ? »**

**Partie A – Avec le tableur.**

A l’aide du fichier CSV proposé, répondre à l’ensemble des questions suivantes :

1. Identifier les différents descripteurs de la collection.

Préciser le type et le format de donnée d’un objet.

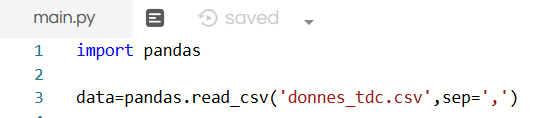
1. Combien de personnes étaient âgées de 10 ans ?
2. Quel est le prix du billet le plus élevé ?
3. Combien de femmes en 1ère classe voyageaient sur le Titanic ? Quel âge pour la plus jeune ? pour la plus âgée ?
4. « Les femmes et les enfants d’abord ? »

En considérant que l'on est enfant lorsqu'on a moins de 18 ans, déterminer le nombre de femmes et d'enfants qui ont survécu.

Que penser alors de l'adage "les femmes et les enfants d'abord" ?

1. Quel est le tarif moyen du billet des personnes qui voyageaient en 1ere classe ?
2. Quelle est la moyenne d’âge des personnes qui voyageaient en 3ème classe ?
3. Indiquer selon le genre des personnes de chaque classe, la fréquence du descripteur « Survivant. »

**Partie B : Approfondissement – Avec Python et le module PANDAS**

Pour cela, on utilise la bibliothèque **pandas** (Python Data ANalysiS Library). *Voir « Du côté de la technique ».*

Exécuter les lignes suivantes :

**L’objet data contient ainsi l’ensemble des données.**

1. Exécuter successivement dans la console, les instructions suivantes et commenter les résultats :

data.head() ; data.shape ; data.dtypes.

1. Exécuter la commande data.describe()

Quel est le prix moyen du billet ?

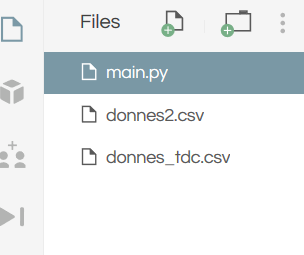
1. Voici un panel de fonctions de la librairie pandas :

data[‘nom du descripteur’].mean()

groupe1=data.groupby([‘nom du descripteur 1’,’nom du descripteur 2’]).mean()

groupe2=data.dropna()

groupe3=data.query(‘sexe==2 or age<18’)

Reprendre l’ensemble des questions de la partie A.

**Du côté de la technique**

Avec Edupython :

Installer la librairie pandas. Pour cela, menu Outils/Outils/Installer un nouveau module. Choisir conda de préférence.

Ouvrir un nouveau programme, le sauver.

Déposer le fichier csv contenant les données à traiter dans le répertoire où   
a été sauvé le programme.

Avec Repl.it, Rien à installer. Glisser ou télécharger le fichier des données   
dans la liste des fichiers.