

TP DataScience 1 : Introduction à Pandas

TP Data Science

Ilyas Bambrik

Table des matières



I - Exercice : Lecture d'un DataFrame	3
II - Exercice : Sélection	4
III - Exercice : Exploration du DataFrame	5

Exercice : Lecture d'un DataFrame



Le DataFrame utilisé dans ce TP est similaire au DataFrame du cours mais propose plus d'informations. Le *Times Higher Education World University Rankings 2024* comporte 1 904 universités dans 108 pays et régions. *THE (Times Higher Education)* fournit depuis 2004 des données fiables sur les performances des universités aux étudiants, leurs familles, aux dirigeants universitaires, aux gouvernements et à l'industrie.

Question

- Lisez le DataFrame à partir du fichier *timeswr.csv*. Combien de colonnes contient ce DataFrame ? Utilisez l'attribut `.columns` du DataFrame.
- Quelles sont les cinq meilleurs universités selon ce classement ?
- Quelle est le nom de la colonne qui désigne le pays de l'université ?
- Affichez les dix premières lignes du DataFrame.
- Affichez les dix dernières lignes du DataFrame.
- Changez l'index du DataFrame tel que la colonne *name* devient l'index.

Exercice : Exploration du DataFrame



Le DataSet laptops.csv comporte des informations concernant les PC portables disponibles sur le marché.

Question

- Quel sont les types de RAM (*RAM_TYPE*) adoptées par les laptop dans ce DataSet ?
- Quelle est le fabricant de processeur (*Processor_Brand*) ayant le plus large nombre de modèles présent dans le DataSet ?
- Donnez les noms des cinq fabricants (*Brand*) dominant ce DataSet ?
- Donnez le nom du modele (*Name*) et fabricant (*Brand*) du laptop le plus coûteux.