

Exercice Pandas

Objectif : se familiariser avec les librairies Pandas.

Source: Frédéric Amblard, Philippe Muller, Benoit Gaudou

Note: un résumé des “fonctions essentielles de pandas” est disponible à cette adresse :

https://medium.com/@thaoly_22574/pandas-les-fonctions-essentielles-b978c6377db0#:~:text=des%20objets%20Test-,pd.,et%202020%20lignes%2C%20chiffres%20al%C3%A9atoires

Exercice - Manipulation de la librairie pandas

1. Créer un dataframe df en important le fichier pokemon.csv (disponible sur Moodle) avec la colonne Name comme index¹. Afficher le data frame. Afficher les informations générales sur ce dataframe².
2. Afficher les 10 premières lignes de df³.
3. Afficher la colonne “Type 2”.
4. Remplacer les valeurs NaN dans la colonne ‘Type 2’ par ” (i.e. le caractère vide).
5. Supprimer les colonnes “#”, “Sp. Atk” et “Sp. Def” de df. Essayer de le faire en 1 ligne.
6. Afficher les valeurs des colonnes “Type 1” et “Type 2” du Pokemon Bulbasaur.
7. Afficher les 5 dernières lignes et 2 premières colonnes de df.
8. Afficher les Pokemons légendaire de “Type 1” ‘grass’.
9. Afficher une description statistique des 4 premières colonnes (‘Total’, ‘HP’, ‘Attack’, ‘Defense’)⁴
10. Trouver le Pokemon avec la plus haute valeur de HP.
11. Trouver les 3 pokemons avec le plus grande vitesse (“speed”)⁵.
12. Compter le nombre de Pokémon légendaires.
13. Créer la variable legendary en groupant les Pokemon selon leur statuts légendaire.
14. Calculer les valeurs moyennes de ‘Total’, ‘HP’, ‘Attack’, ‘Defense’ et ‘Speed’ sur les groupes legendary.
15. [Question ouverte] Utiliser la librairie matplotlib pour des graphiques intéressants et informatifs.

¹ Utiliser le paramètre `index_col`, de la fonction `pandas.read_csv`

² cf. méthode `dataframe.info()`

³ cf. Méthode `dataframe.head(.)`

⁴ cf. Méthode `dataframe.describe()`

⁵ voir les méthodes de tri sur les dataframe et la fonction permettant d’obtenir les n dernières lignes.