

S'entraîner avec numpy

Jouez les détectives

Vous avez dorénavant les connaissances nécessaires pour enquêter sur vos données.

Pour ce faire, des questions vous seront demandées et reprendront tous les points vus en cours.

Vous devez utiliser le fichier de données qui se trouvent sur Connected. Celui-ci porte le nom de « Base de données sur le playstore ». Pour en connaître son formatage, un fichier « README_PLAYSTORE_NUMPY » se trouve également sur le Connected.

1. Chargez votre fichier de données et nommez-le « google ».
2. Enquêtez sur sa structure. Regardez sa taille (shape), sa dimension (ndim), le nombre total d'informations (size).
3. Sélectionnez les 10 premières lignes de votre fichier de données.
4. Sélectionnez les 10 dernières lignes de votre fichier de données.
5. Sélectionnez la première et la dernière colonne de deux manières différentes.
6. Trouvez la valeur maximale de chaque colonne.
7. Triez votre tableau par le nombre de « reviews » (par ordre décroissant) et sauvegardez le résultat dans « google_sorted_reviews ».
8. Réalisez une moyenne sur les 100 premières lignes de vos données précédemment triées. Interprétez les résultats.
9. Obtenez le nombre d'applications qui ont un rating plus grand que la moyenne. Calculez le % qui se trouvent au-dessus et en dessous de la moyenne. Pour ce faire, utilisez une moyenne arithmétique et une moyenne pondérée. Interprétez les résultats.
10. « Il y a plus d'applications gratuites que d'applications payantes ». Est-ce vrai ? Pour répondre à cette question, comptez toutes les valeurs possibles. Par exemple : il y a 300 applications gratuites, il y a 50 applications payantes à 2.99€, ...
11. « Les applications gratuites reçoivent plus d'avis que les applications payantes ». Est-ce vrai ? Pour répondre à cette question, utilisez les conditions pour filtrer votre matrice de données.