

Code de partage avec Capytale : 1680-1272508

Préparation

Dans un premier temps, nous allons à nouveau importer le jeu de données. Sur Capytale ou dans le fichier Python, on écrira et on exécutera :

```
import pandas
films=pandas.read_csv('films.csv', delimiter=';', encoding='utf8')
```

Représenter des données avec des diagrammes en bâtons

1 Préambule : vers la représentation graphique

Pour représenter des données sous la forme de diagramme en bâtons, on utilisera la commande `plt.bar` de la bibliothèque `matplotlib`

Comme pour la représentation de fonctions, une particularité de cette commande est qu'elle nécessite deux listes : une pour les abscisses et une pour les ordonnées. On pourra tester avec le programme ci-dessous.

```
import matplotlib.pyplot as plt
valeurs = [7, 58, 49, 32]
abscisses = [1, 2, 3, 4]
plt.bar(abscisses, valeurs)
plt.show()
```

On peut aussi définir des abscisses qui ne sont pas des nombres :

```
valeurs = [7, 58, 49, 32]
abscisses = [A, B, C, D]
plt.bar(abscisses, valeurs)
plt.show()
```

2 Sélectionner les données

Avec le même jeu de données (films) :

1. compter le nombre de films pour une année donnée (par exemple 2004) ;
2. calculer le revenu moyen des films sur une année (par exemple 2004).

Nota bene : on pourra utiliser les commandes `.count()` et `.mean()`

3 Représenter les données choisies

Représenter à l'aide de diagrammes en bâtons :

1. l'évolution du nombre de films par année ;
2. l'évolution du revenu moyen des films par année.

Pour améliorer un peu, on pourra enlever les films dont le revenu est 0 (l'information n'est sans doute pas renseignée dans ce cas).

Pour aller plus loin, on pourra tester les paramètres suivants :

```
plt.xticks(x) # à adapter
plt.xlabel("Années")
plt.ylabel("Revenu moyen")
plt.title("Revenu moyen des films par année de 1917 à 2017")
```