

```
> Data Science C:\Users\scham\Desktop\From 1
> Section 3 2
> Section 4 3
> code.py 4
> world_alcohol.csv 5
> Section 5 6
> idea 7
> code.py 8
> food_info.csv 9
> Section 6 10
> code.py 11
> food_info.csv 12
> Section 7 13
> code.py 14
> titanic_survival.csv 15
> Section 8 16
> all-ages.csv 17
> code.py 18
> recent-grads.csv 19
> Section 9 20
> code.py 21
> fandango_score_comparison.csv 22
> Section 10 23
> code.py 24
> fandango_score_comparison.csv 25
> Section 11 26
> code.py 27
> thanksgiving.csv 28
> Section 12 29
> code.py 30
> unrate.csv 31
> Section 13 32
> Section 14 33
> Diplôme.jpg 34
> External Libraries 35
> Scratches and Consoles 36

import pandas as pd
import matplotlib.pyplot as plt

reviews = pd.read_csv("fandango_scores (1).csv")
cols = ["FILM", "RT_user_norm", "Metacritic_user_nom", "IMDB_norm", "Fandango_Ratingvalue"]

norm_reviews = reviews[cols]

imdb_distribution = norm_reviews["IMDB_norm"].value_counts()
imdb_distribution = imdb_distribution.sort_index()

"""fig, ax = plt.subplots()
ax.hist(norm_reviews["Fandango_Ratingvalue"], range=(0, 5))
"""

fig = plt.figure(figsize=(5, 20))
ax1 = fig.add_subplot(4, 1, 1)
ax2 = fig.add_subplot(4, 1, 2)
ax3 = fig.add_subplot(4, 1, 3)
ax4 = fig.add_subplot(4, 1, 4)

ax1.hist(norm_reviews["Fandango_Ratingvalue"], bins=20, range=(0, 5))
ax1.set_title("Distribution des notes de fandango")
ax1.set_ylim(0, 50)

ax2.hist(norm_reviews["RT_user_norm"], bins=20, range=(0, 5))
ax2.set_title("Distribution des notes de Rotten Tomatoes")
ax2.set_ylim(0, 50)

ax3.hist(norm_reviews["Metacritic_user_nom"], bins=20, range=(0, 5))
ax3.set_title("Distribution des notes de Metacritic")
ax3.set_ylim(0, 50)

ax4.hist(norm_reviews["IMDB_norm"], bins=20, range=(0, 5))
ax4.set_title("Distribution des notes de IMDB")
ax4.set_ylim(0, 50)
```

Section 3	31/01/2021 17:15	Dossier de fichiers
Section 4	31/01/2021 19:30	Dossier de fichiers
Section 5	02/02/2021 11:39	Dossier de fichiers
Section 6	09/07/2021 15:17	Dossier de fichiers
Section 7	11/07/2021 19:07	Dossier de fichiers
Section 8	14/07/2021 15:31	Dossier de fichiers
Section 9	14/07/2021 15:32	Dossier de fichiers
Section 10	14/07/2021 15:57	Dossier de fichiers
Section 11	14/07/2021 18:01	Dossier de fichiers
Section 12	26/07/2021 14:00	Dossier de fichiers
Section 13	26/07/2021 15:10	Dossier de fichiers
Section 14	26/07/2021 16:00	Dossier de fichiers



N° de certificat : UC-08364c2b-cfd5-44c8-bdc7-a8a737031be1  
Url du certificat : ude.my/UC-08364c2b-cfd5-44c8-bdc7-a8a737031be1  
Numéro de référence : 0004

CERTIFICAT DE FIN DE FORMATION

# Data Science : Analyse de données avec Python

Formateurs **Rod | Python - SQL - Data Science - Machine Learning - Deep Learning, MonCoachData | Data Science - Machine Learning - Deep Learning - Trading**

**Arthur Schamroth**

Date **26 juillet 2021**

Longueur **10 heures au total**