

Bing

بالطبع، سأشرح لك كيف يعمل هذا الكود بالتفصيل:

1. `import pandas as pd, import numpy as np, import matplotlib.pyplot as plt, import seaborn as sns, %matplotlib inline`: هذه الأوامر تقوم للرسومات البيانية، و `matplotlib` و `seaborn`، للتعامل مع العمليات الرياضية `numpy`، تستخدم للتعامل مع البيانات `pandas` بتحميل المكتبات المطلوبة لتنفيذ الكود. المكتبة `%matplotlib inline` لعرض الرسومات البيانية مباشرة في الدفتر.
2. `movie_titles = pd.read_csv("movie_titles.csv")`: هذا السطر يقرأ ملف CSV هذا الأفلام ويحفظه في `DataFrame` يدعى `movie_titles`.
3. `movie_ratings = pd.read_csv('movie_ratings.data', sep='\t', names=['user_id', 'item_id', 'rating', 'date'])`: هذا السطر يقرأ ملف يحتوي: `['user_id', 'item_id', 'rating', 'date']` يتم تحديد أسماء الأعمدة بواسطة القائمة `movie_ratings` يدعى `DataFrame` على تقييمات الأفلام ويحفظه في.
4. `movie_titles.shape, movie_ratings.shape`: هذه الأوامر تعرض عدد الصفوف والأعمدة في كل `DataFrame`.
5. `movie_ratings.drop(['date'], axis = 1, inplace = True)`: لأنه ليس مطلوبًا في هذا التحليل `movie_ratings` `DataFrame` من `date` هذا السطر يقوم بإزالة عمود.
6. `ratings = pd.merge(movie_ratings, movie_titles, on = 'item_id')`: حسب `item_id` و `movie_titles` و `movie_ratings` هذا السطر يدمج البيانات من.
7. `ratings.groupby('title')['rating'].describe()`: هذا السطر يقوم بتجميع التقييمات حسب العنوان ويعطي الإحصاءات الوصفية لكل فيلم.
8. `avg_ratings = pd.concat([ratings_count, ratings_avg], axis = 1)`: جديد يحتوي على متوسط التقييم وعدد التقييمات لكل فيلم `DataFrame` هذا السطر يقوم بإنشاء.
9. `avg_ratings['average_rating'].plot(bins=100, kind='hist')` و `avg_ratings['total_reviews'].plot(bins=100, kind='hist', color = 'r')`: هذه الأوامر تقوم برسم توزيعات التقييمات وعدد التقييمات.
10. `avg_ratings[avg_ratings['average_rating'] == 5]`: هذا السطر يعرض الأفلام التي حصلت على تقييم 5.
11. `avg_ratings.sort_values('total_reviews', ascending = False).head(100)`: هذا السطر يعرض 100 فيلم حصلوا على أكبر عدد من التقييمات.

بشكل عام، هذا الكود يقوم بتحليل بيانات التقييمات ويقوم بإعداد نظام توصية بسيط يعتمد على متوسط التقييم وعدد التقييمات لكل فيلم. يمكن استخدام هذا النوع من الأنظمة لتوصية الأفلام للمستخدمين بناءً على تقييمات الأفلام من قبل المستخدمين الآخرين.