Industrial Project Report

*Submitted in partial fulfillment of the degree*

**B-tech in Computer Science Engineering**

PROJECT ON :- MOVIE RECOMMENDATION SYSTEM USING TENSOR FLOW

**Second-year student of**

SILIGURI INSTITUTE OF TECHNOLOGY

THIS IS SUMBITTED IN FULFILLMENT OF THE REQUIREMENTS FO THE DEGREE OF

**AFFLIATED BY**

MAULANA ABDUL KALAM AZAD UNIVERSITY OF TECHNOLOGY

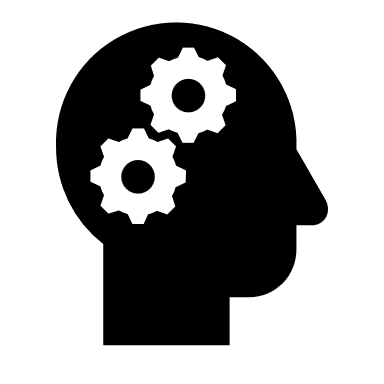


**UNDER THE SUPERVISION OF :- MR. RIPAM KUNDU**

**SIKHARTHY INFOTECH Pvt. Ltd.**

Siliguri Institute Of Technology Hill Cart Road, Salbari, Sukna, West Bengal 734009

SUBMITTED BY-

****

**TEAM LEADER:**

***Kamalika Saha-11900121057***

**TEAM MEMBERS:**

***Sweety Nag-11900121055***

***Monami Ghosh-11900121045***

***Annesha Basu-11900121004***

***Greeny Kundu-11900121024***

***Priyadarshini Sen-11900121049***

DEPARTMENT OF COMPUTER SCIENCE ENGINEERING

I hereby forward the documentation prepared under my supervision by Mr. **Ripam Kundu Sir** entitled Siliguri Institute Of Technology to be accepted as fulfillment of the requirement for the Degree of Bachelor of Technology in Electrical Engineering, **Siliguri Institute Of Technology** affiliated to **Maulana Abul Kalam Azad University of Technology** (**MAKAUT**).

|  |  |
| --- | --- |
| \_\_  \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_    **Mr.Ripam Kundu**  **(Software Developer)**  **Project Guide**  **Sikharthy Infotech Pvt. Ltd.**  **Shilpi Ghosal**  **(Director)**  **Sikharthy Infotech Pvt. Ltd.** | \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_    **HOD**  **Department Of Electrical Engineering, SIT**  **TPO**  **Siliguri Institute of Technology** |

**Certificate of Approval**

The foregoing project is hereby approved as a creditable study for the B.Tech in Computer Science Engineering presented in a manner of satisfactory to warrant its acceptance as a prerequisite to the degree for which it has been submitted. It is understood that by this approval the undersigned do not necessarily endorsed or approved any statement made, opinion expressed or conclusion therein but approve this project only for the purpose for which it is submitted.

Final Examination for

Evaluation of the Project ----------------------------------------

-----------------------------------------

------------------------------------------

**Signatures of Examiners**

**ABSTRACT**

A movie recommendation is important in our social life due to its strength in providing enhanced entertainment. Such a system can suggest a set of movies to users based on their interest, or the popularities of the movies. Although, a set of movie recommendation systems have been proposed, most of these either cannot recommend a movie to the existing users efficiently or to a new user by any means. In this paper we propose a movie recommendation system that has the ability to recommend movies to a new user as well as the others. It mines movie databases to collect all the important information, such as, popularity and attractiveness, required for recommendation. It generates movie swarms not only convenient for movie producer to plan a new movie but also useful for movie recommendation. Experimental studies on the real data reveal the efficiency and effectiveness of the proposed system.

**ACKNOWLEDGMENT**

It is a great pleasure for me to acknowledge the assistance and participation of a large number of individuals in this attempt. Our project report has been structured under the valued suggestion, support, and guidance of **Mr. Ripam Kundu**. Under his guidance, we have accomplished the challenging task in a very short time.

Finally, we express our sincere thankfulness to our family members for inspiring me all throughout and always encouraging us.

**Group Members’ Signature**

**TABLE OF CONTENTS**

1. **Libraries Used**
2. **Importing Dataset**
3. **Exploratory Data Analysis**
4. **Data Visualisation**
5. **Building Reccomendation System**
6. **Most Rated Movies**
7. **User-based Collaborative Filtering**
8. **Create a Copy of train and test Dataset**
9. **Predicting the user ratings on the movies**
10. **Functional Requirements of the System**
11. **Reference**
12. **Conclusion**
13. **Contribution**

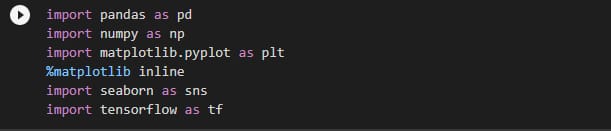
**LIBRARIES USED**

## **Importing Libraries**

The analysis will be done using the following libraries :

* Pandas : This library helps to load the data frame in a 2D array format and has multiple functions to perform analysis tasks in one go.
* NumPy: NumPy arrays are very fast and can perform large computations in a very short time.
* Matplotlib/Seaborn : This library is used to draw visualizations.
* Tensorflow: This library is a free and open-source software library for machine learning and artificial intelligence. It can be used across a range of tasks but has a particular focus on training and inference of deep neural networks.

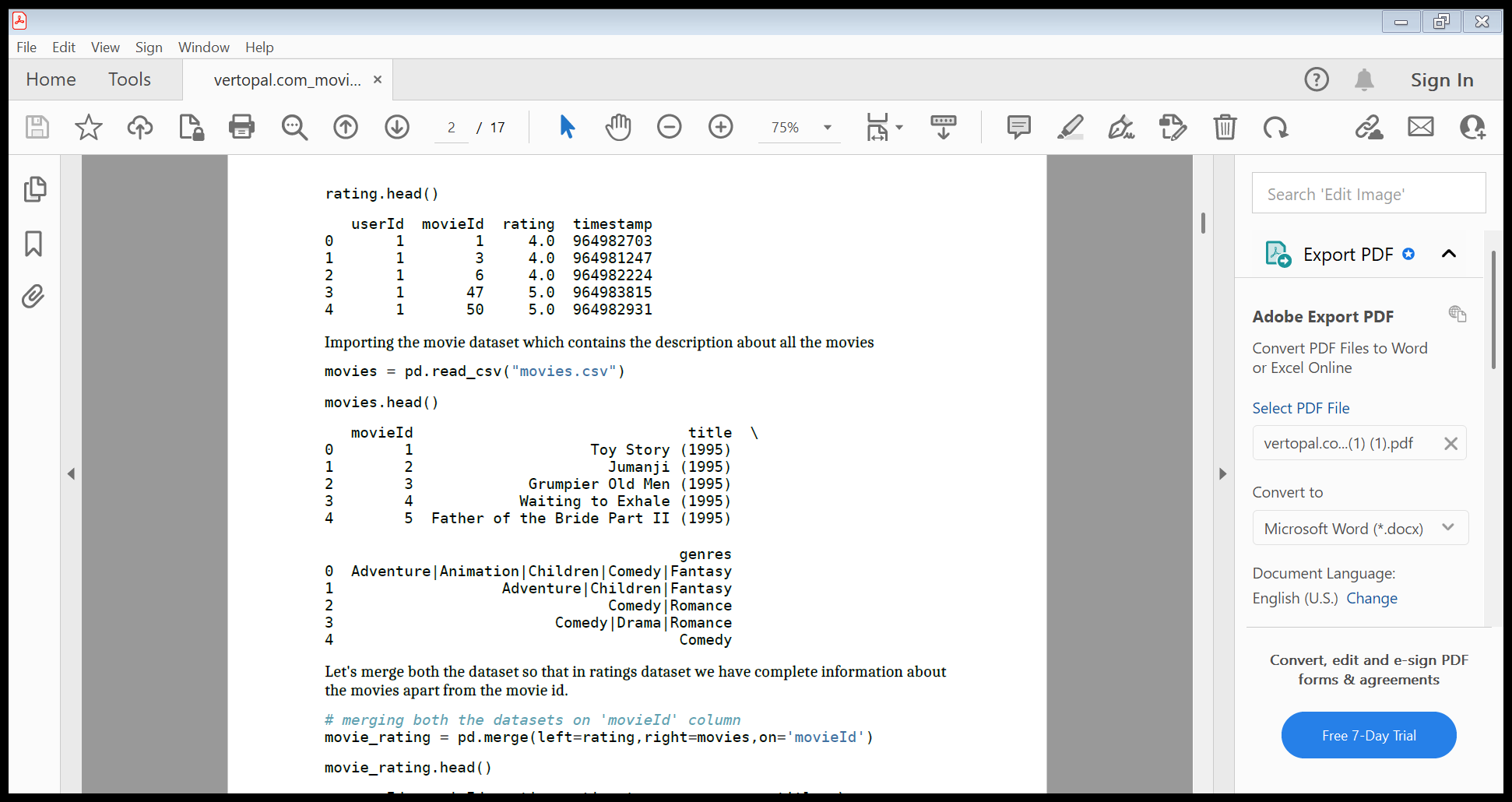
To import all these libraries, we can use the code below:

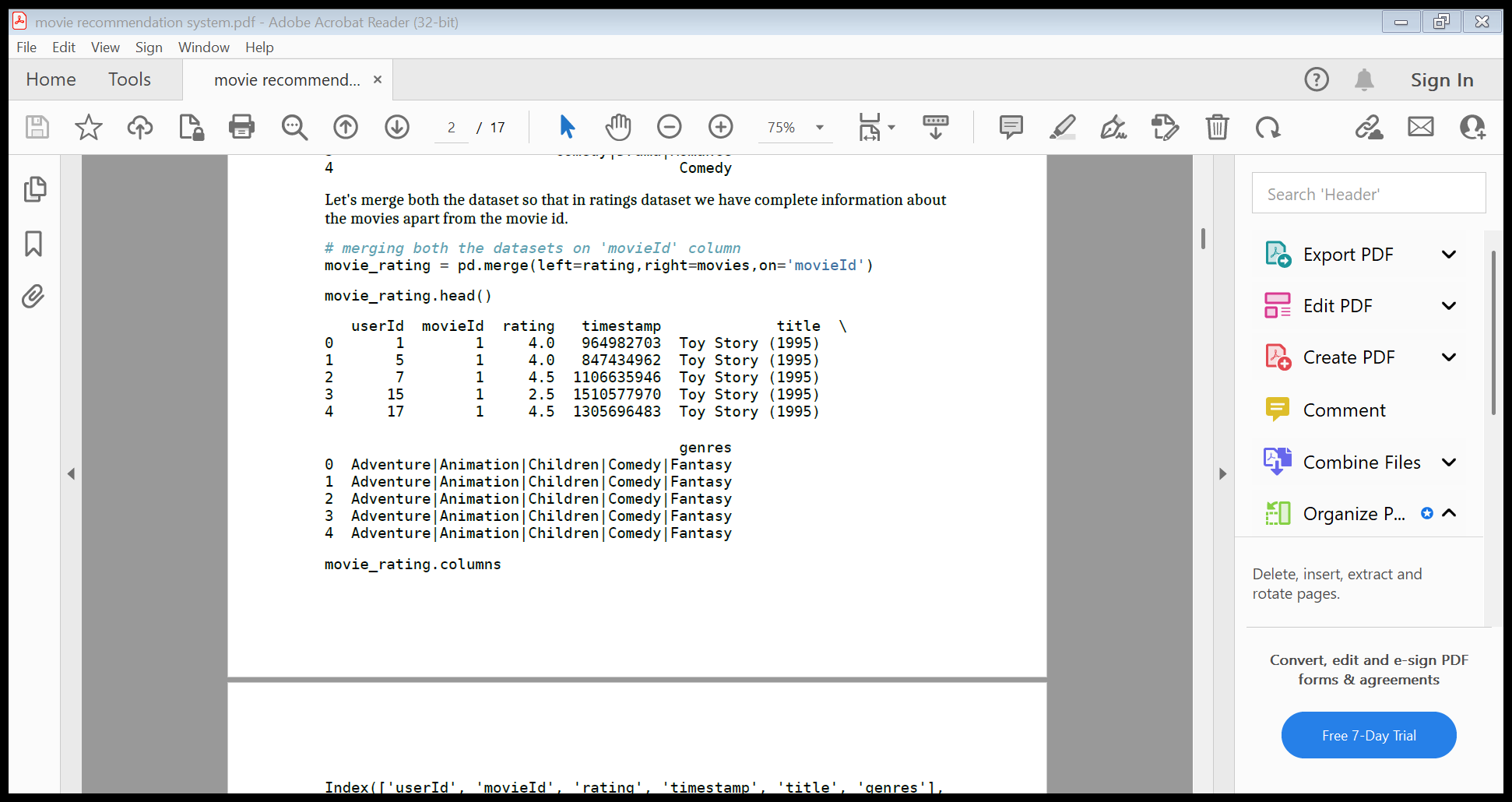


**IMPORTING DATASET**

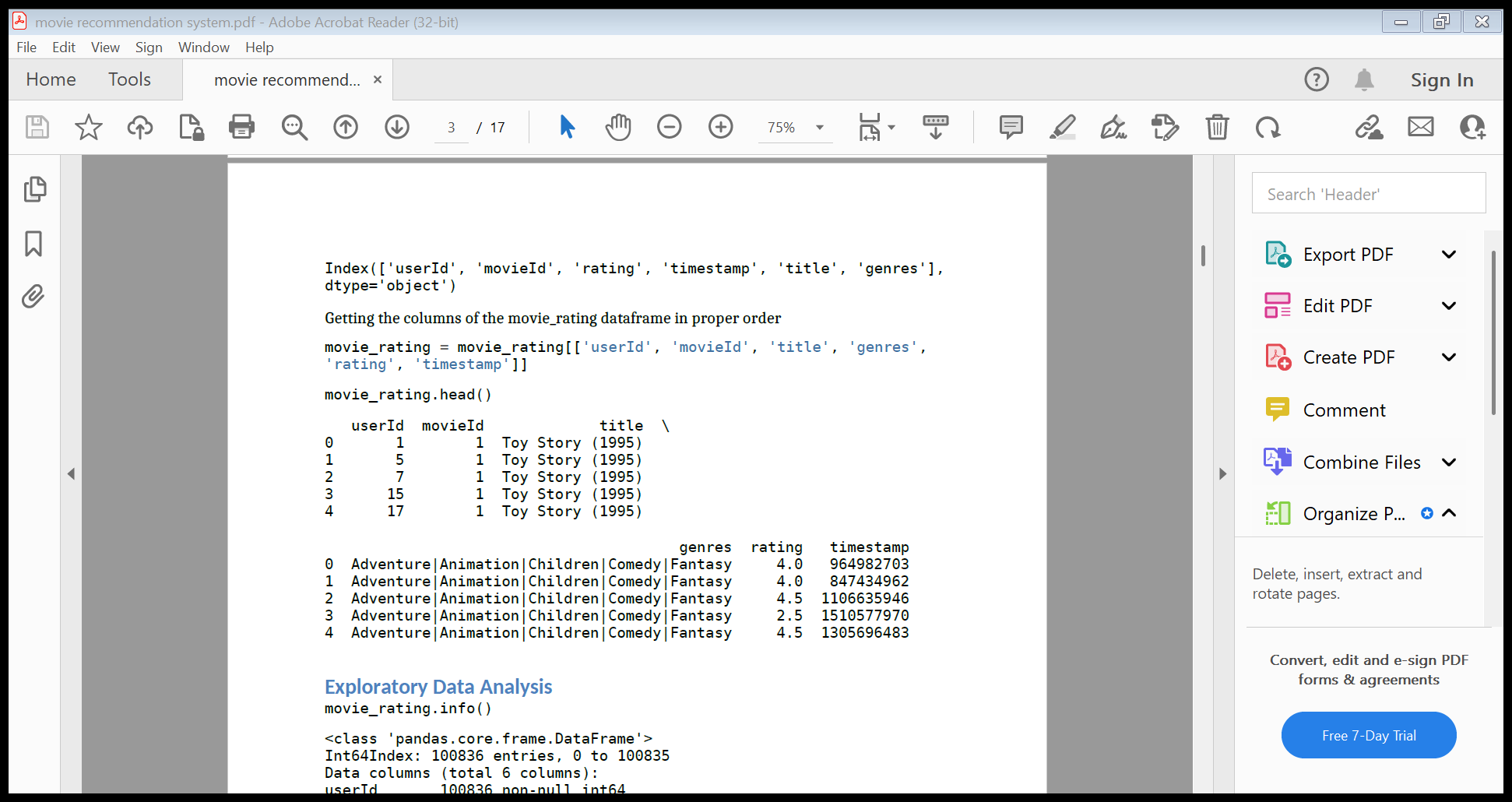
Importing the rating data set which contains ratings given by the users to the movies they

watched.

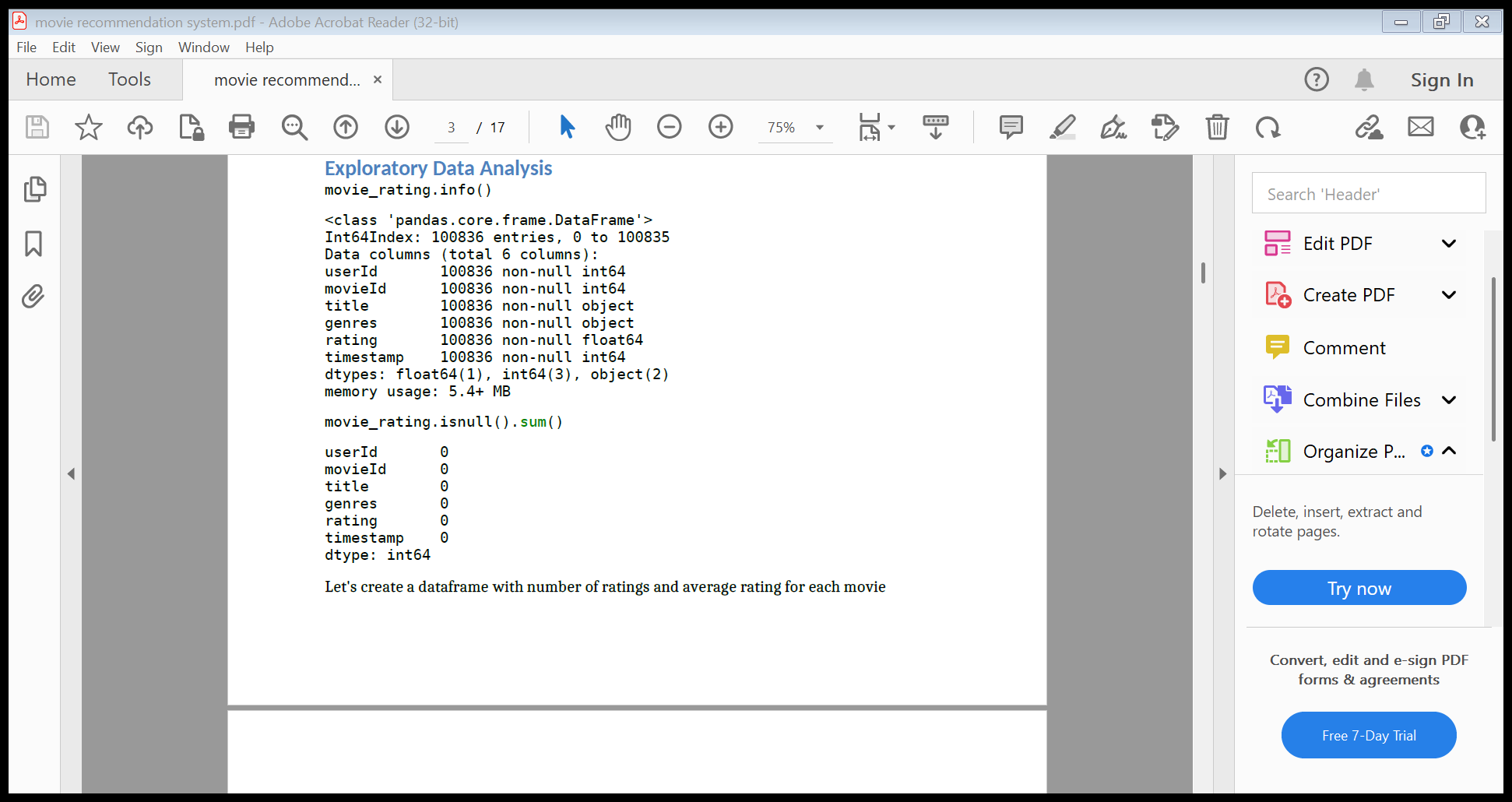
rating = pd.read\_csv ( ' ratings.csv ' )

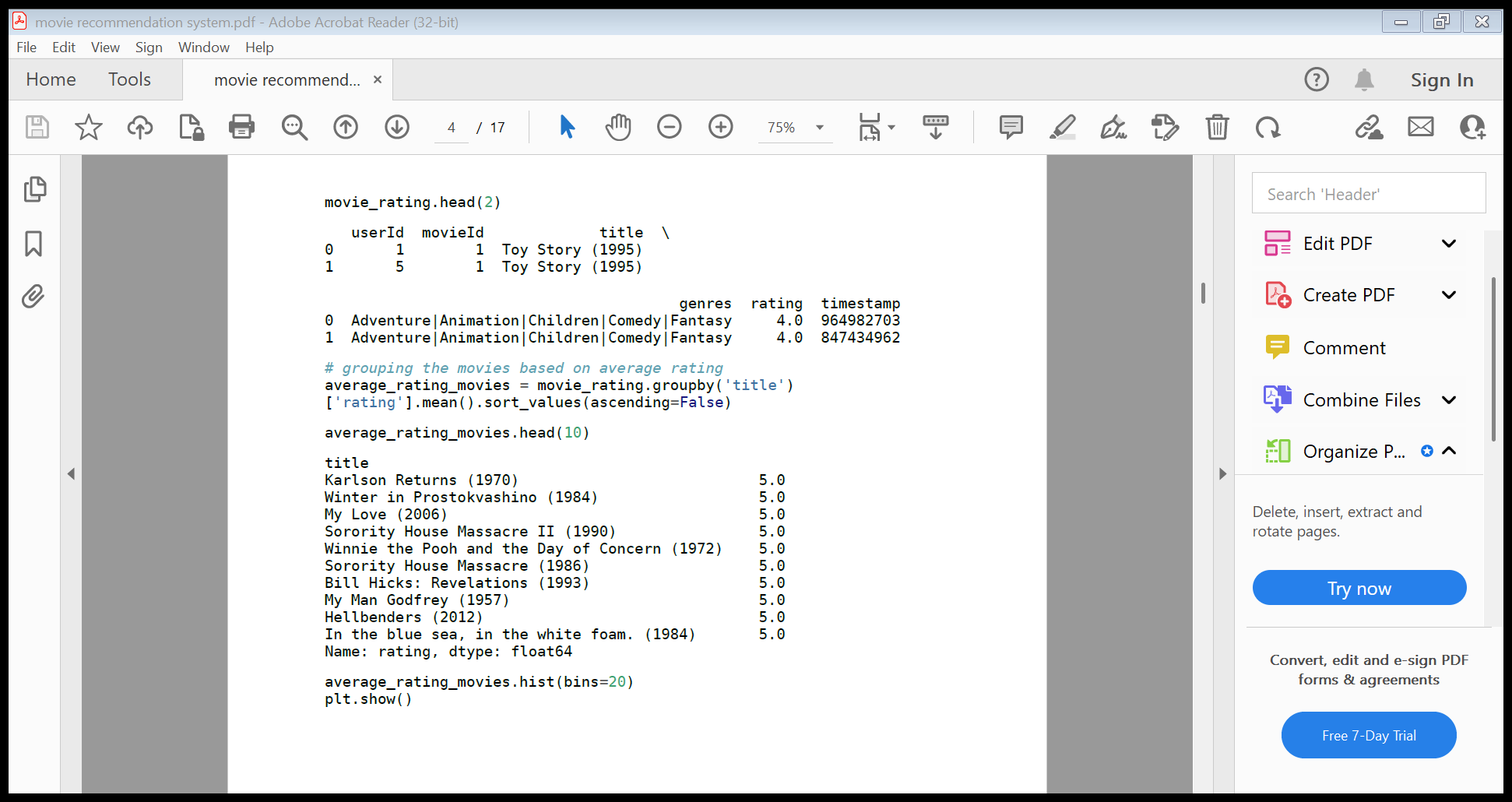


movie\_rating.columns



**Exploratory Data Analysis**



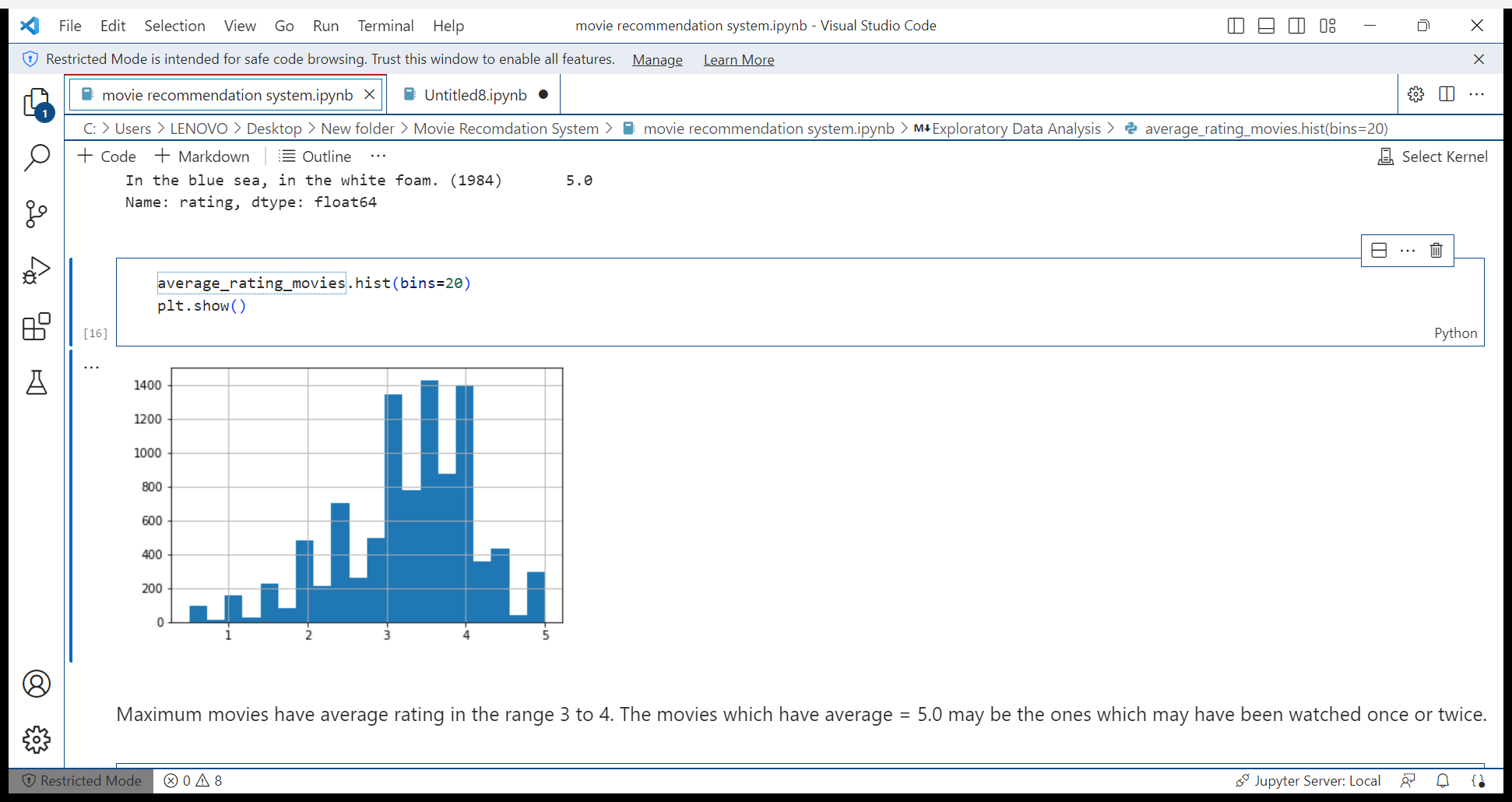


**DATA VIZUALISATION :-**

**Data visualization is the graphical representation of information and data. By using visual elements like charts, graphs and maps, data visualization tools provide an accessible way to see and understand trends, outliers, and patterns in data. Additionally, it provides an excellent way for employees or business owners to present data to non-technical audiences without confusion.**

average\_rating\_movies.hist(bins=20)

plt.show()

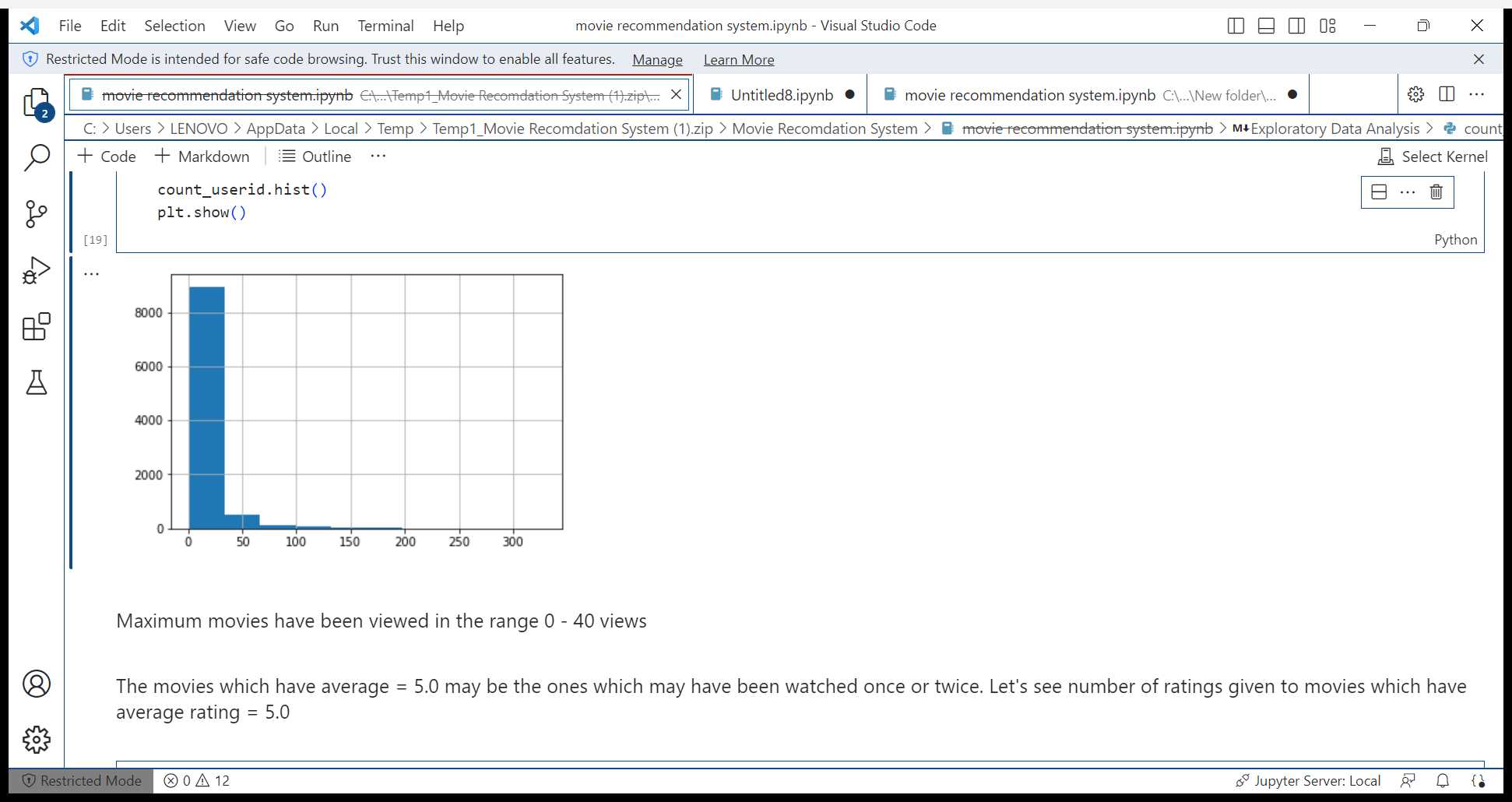


Maximum movies have average rating in the range 3 to 4. The movies which have average =

5.0 may be the ones which may have been watched once or twice.

count\_userid.hist()

plt.show()



Maximum movies have been viewed in the range 0 - 40 views

The movies which have average = 5.0 may be the ones which may have been watched once

or twice. Let's see number of ratings given to movies which have average rating = 5.0

**BUILDING RECOMMENDATION SYSTEM**

Collaborative Filtering is the most common technique used when it comes to building intelligent recommender systems that can learn to give better recommendations as more information about users is collected.

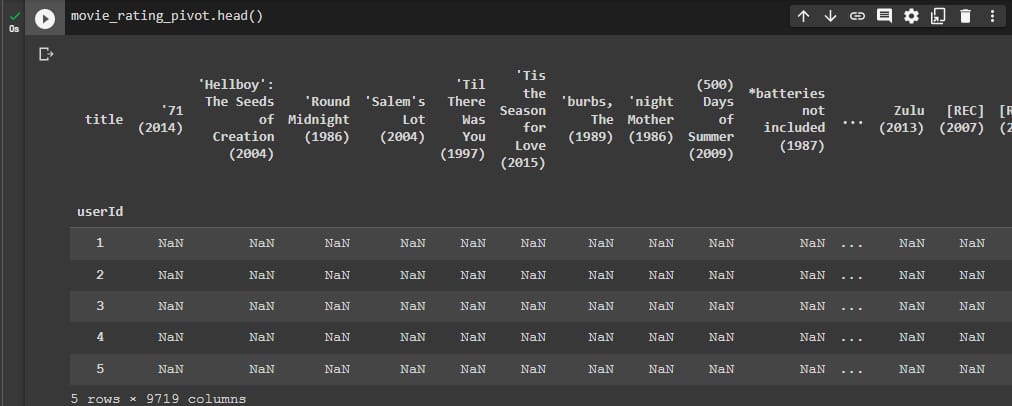
Most websites like Amazon, YouTube, and Netflix use collaborative filtering as a part of their sophisticated recommendation systems. You can use this technique to build recommenders that give suggestions to a user on the basis of the likes and dislikes of similar users.

Creating pivot table to create item by item collaborative filtering

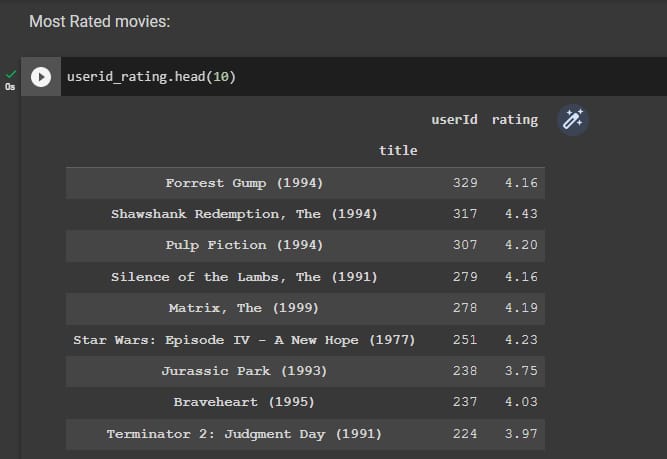
movie\_rating\_pivot = pd.pivot\_table(index='userId',columns='title',values='rating',data=movie\_rating)

There will be many Nan values because users have watched only few of the movies and given ratings only to those movies.

movie\_rating\_pivot.head()

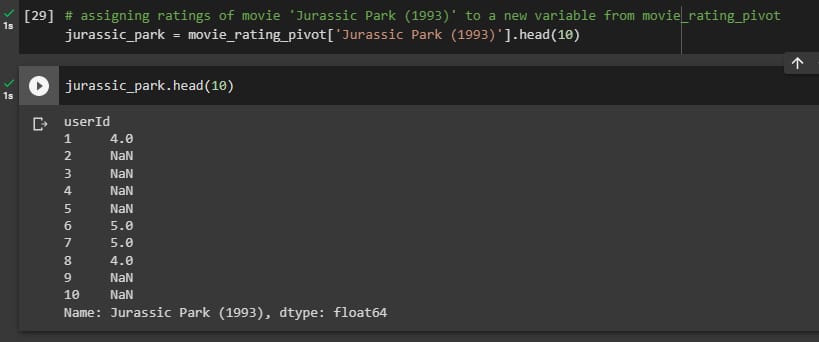


**Most Rated Movies :**

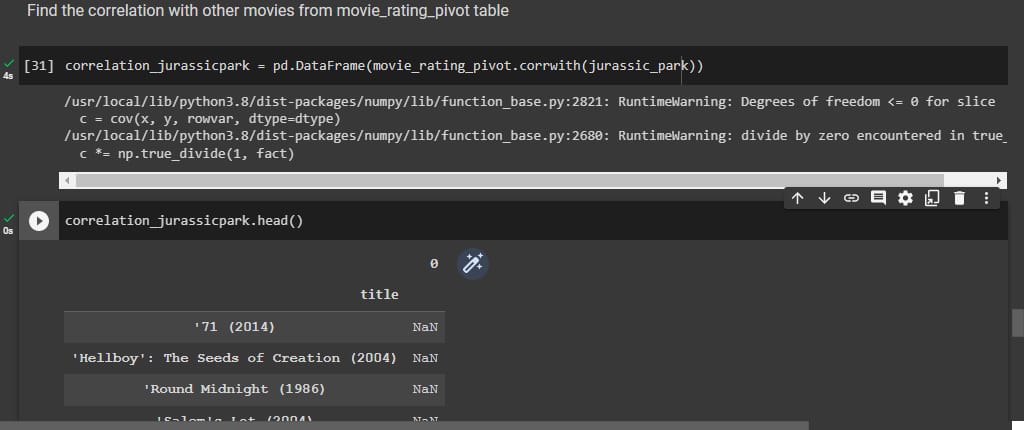


Let's find which movies to recommend to the users who have watched 'Jurassic Park (1993)'.

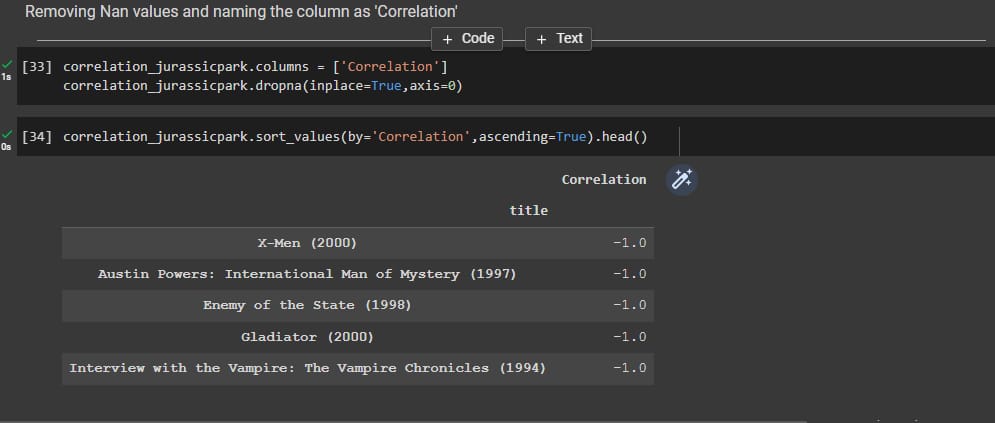
To do this we have to find correlation of 'Jurassic Park (1993)' with other movies which have been rated in a similar way by the users.



Find the correlation with other movies from movie\_rating\_pivot table

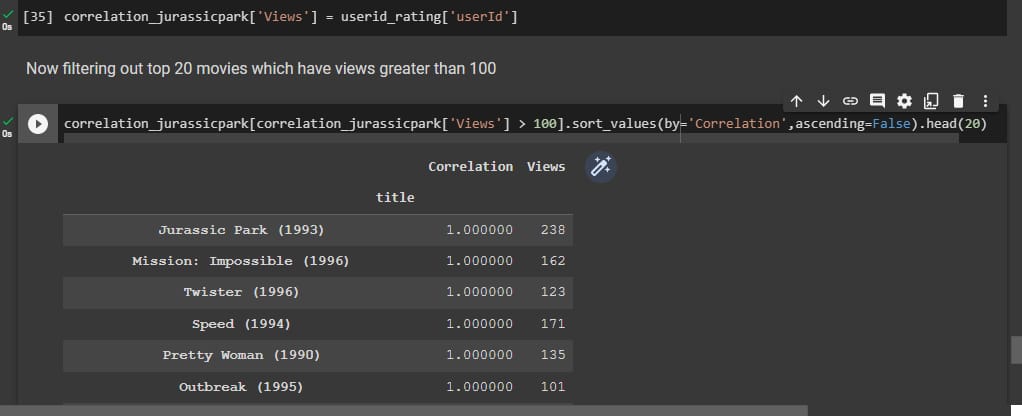


Removing Nan values and naming the column as 'Correlation'

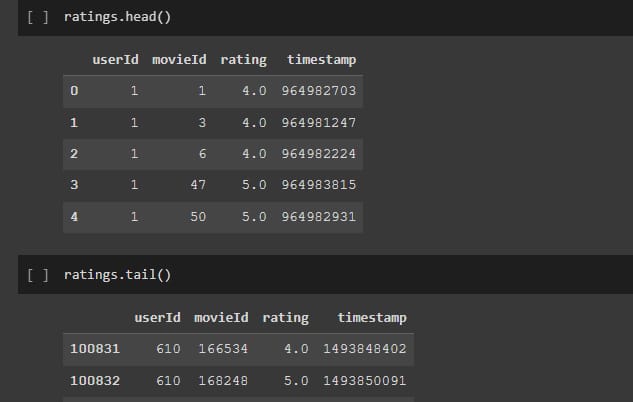


Now filtering out top 20 movies which have views greater than 100

correlation\_jurassicpark[correlation\_jurassicpark['Views'] > 100].sort\_values(by='Correlation',ascending=False).head(20)



**User-based Collaborative Filtering**



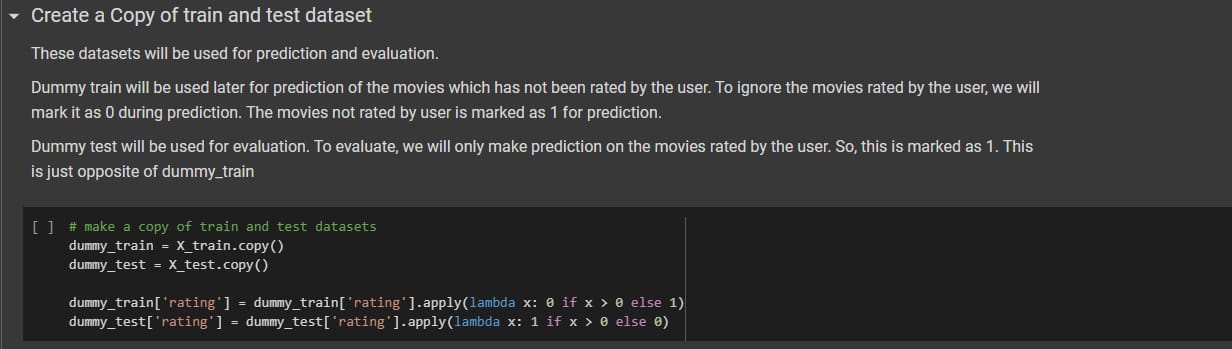
**Create a Copy of train and test Dataset**

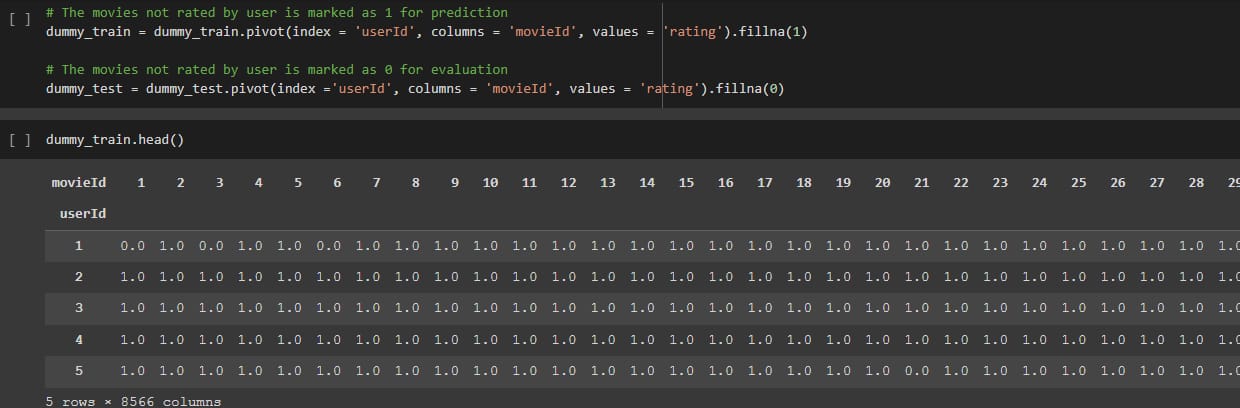
These datasets will be used for prediction and evaluation.

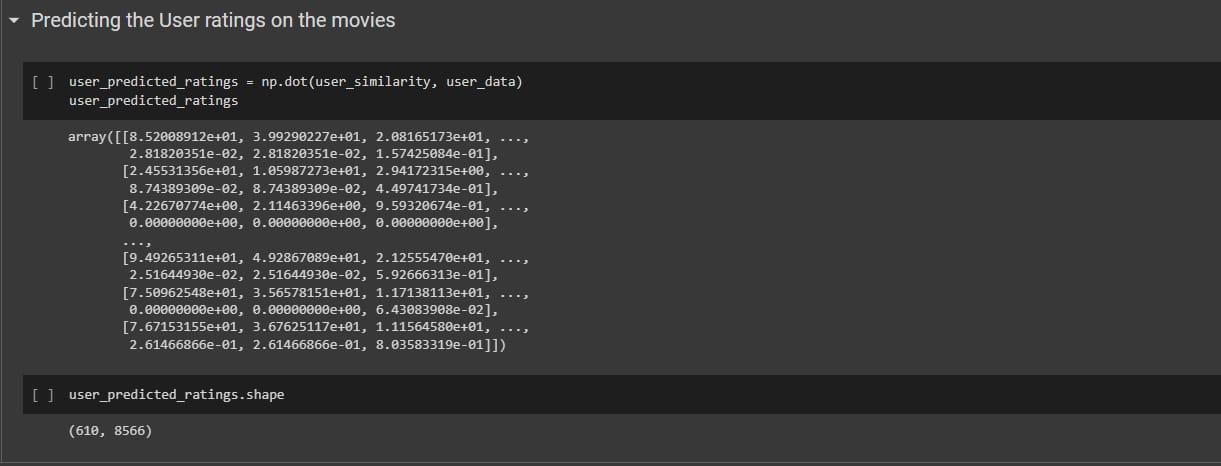
Dummy train will be used later for prediction of the movies which has not been rated by the user. To ignore the movies rated by the user, we will mark it as 0 during prediction. The movies not rated by user is marked as 1 for prediction.

Dummy test will be used for evaluation. To evaluate, we will only make prediction on the movies rated by the user. So, this is marked as 1.

This is just opposite of dummy\_train.





**Predicting the user ratings on the movies**

**FUNCTIONAL REQUIREMENTS OF THE SYSTEM**

***SOFTWARE:***

* *Operating System*
* Windows OS 11
* TensorFlow

***WEB BROWSER:***

* Internet Explorer 7
* Google Chrome

***CODING LANGUAGE :***

* Python

**REFERENCE**

[www.tensorflow.org](http://www.tensorflow.org)

[www.geeksforgeeks.org](http://www.geeksforgeeks.org)

[www.slideshare.net](http://www.slideshare.net)

**Conclusion**

Building a movie recommendation system using TensorFlow can be a great way to improve the user experience on a movie streaming platform. By utilizing machine learning algorithms, the system can analyze user data and provide personalized movie recommendations, increasing user engagement and satisfaction.

**CONTRIBUTION**

**1.** **Kamalika Saha - building recommendation system, user based and item based collaborative filtering**

**2.** **Sweety Nag - building recommendation system,data analysis**

**3.** **Monami Ghosh - ppt making**

**4.** **Priyadarshini Sen - ppt making**

**5.** **Annesha Basu - documentation**

**6.** **Greeny Kundu - documentation**

**Thank you**