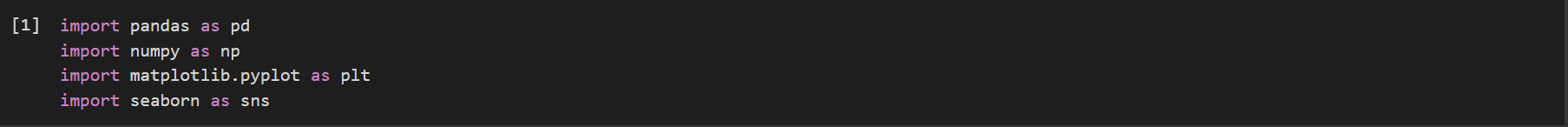
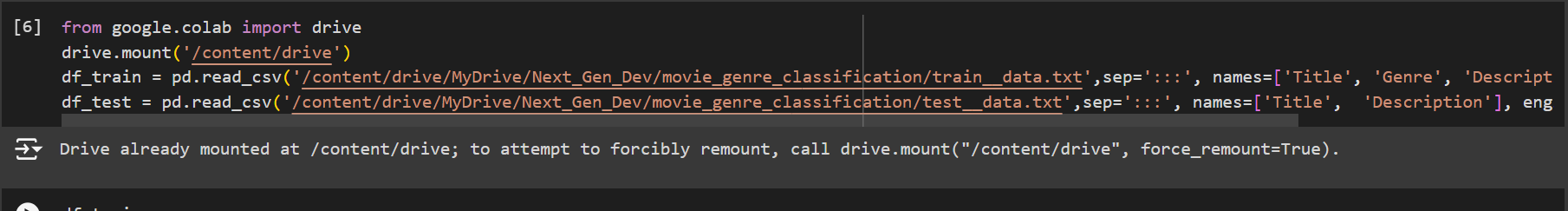
**Movie Genre Classification**

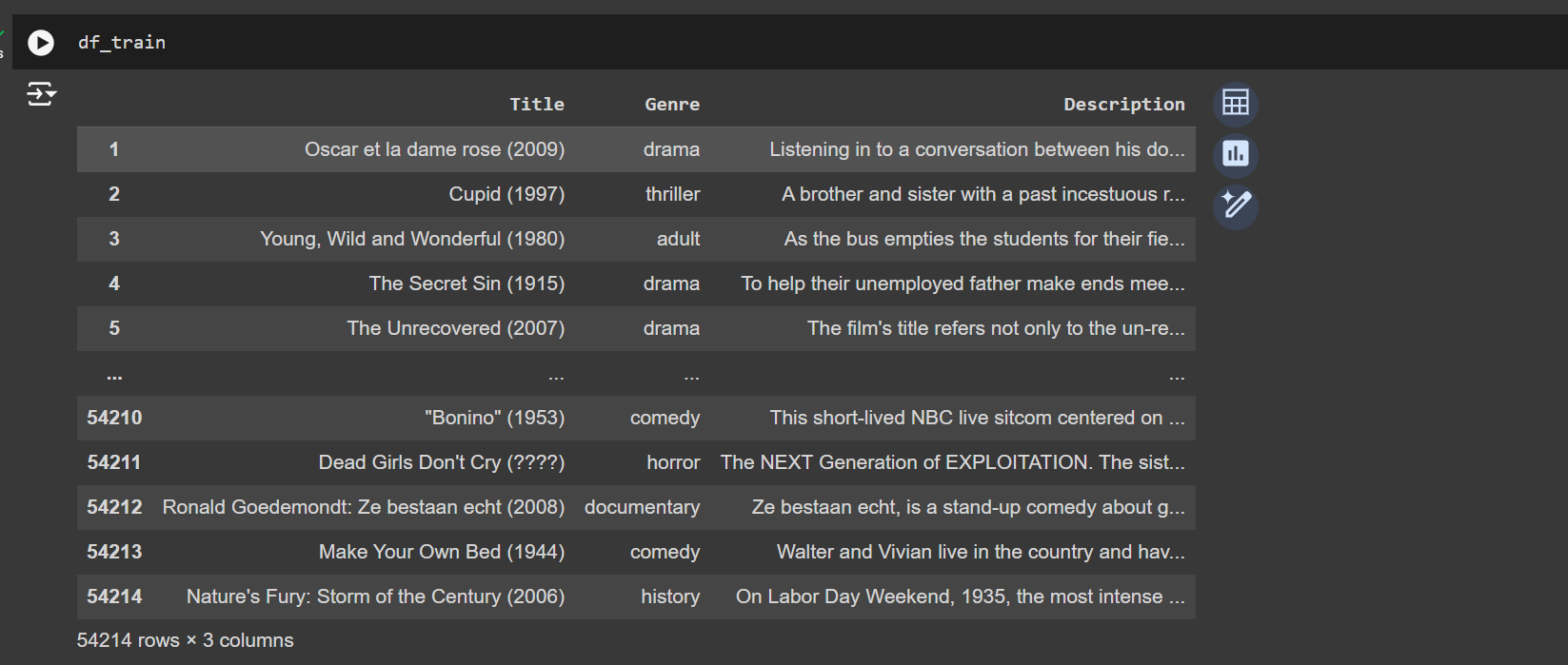
**Importing required modules and libraries: Here we imported numpy, pandas, matplotlib, and seaborn.**

****

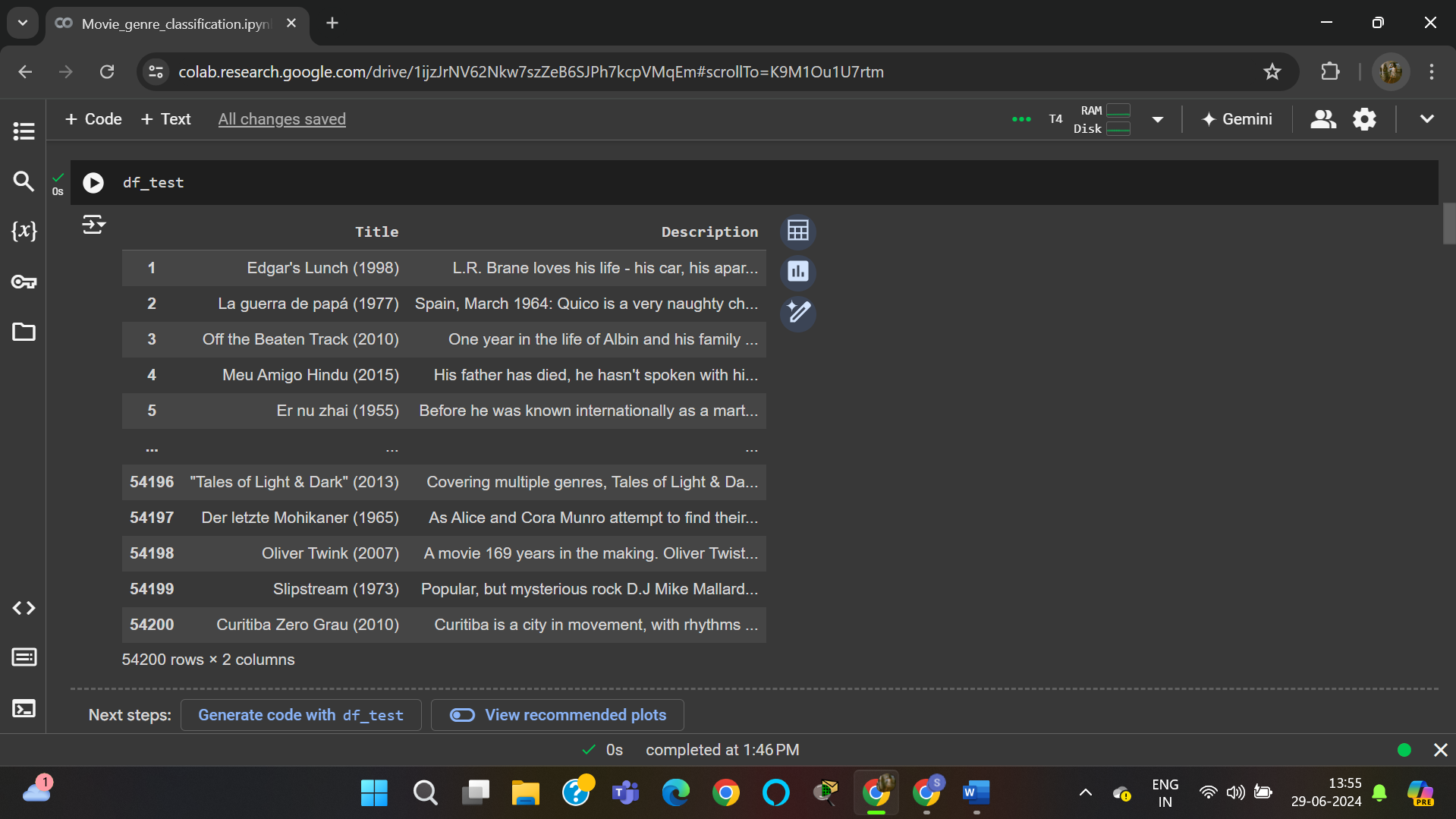
**Dataset Loading**: The data set is divided into 2 parts training data set and testing data set. Here we read both csv format data files and separate the title, genre and the description of the movie.

****

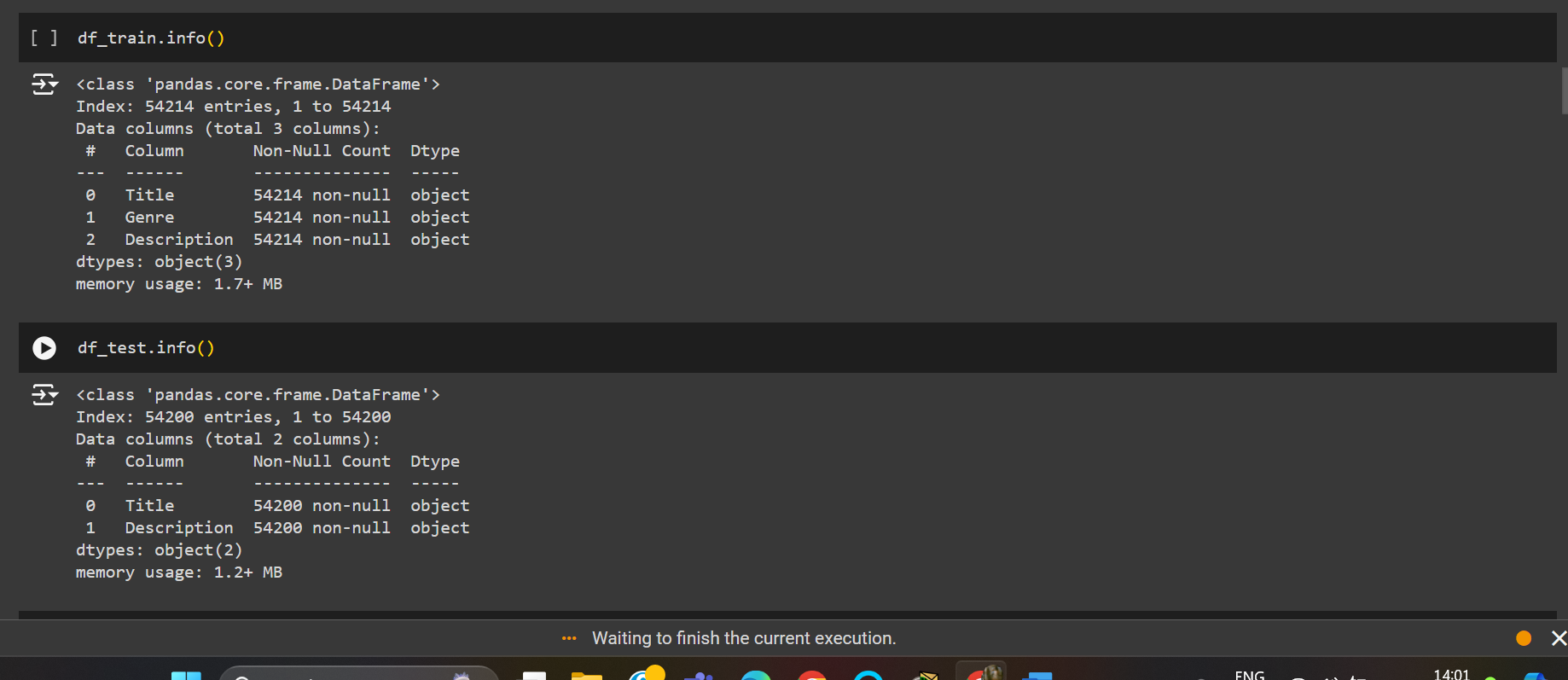
**Training dataset stored in data frame:** This command shows us all the data stored in the training file according to the format we read it with titles, genre and description separated.



Testing data set stored in data frame: This command shows us all the data stored in the testing file according to the format we read it with titles, genre and description separated.



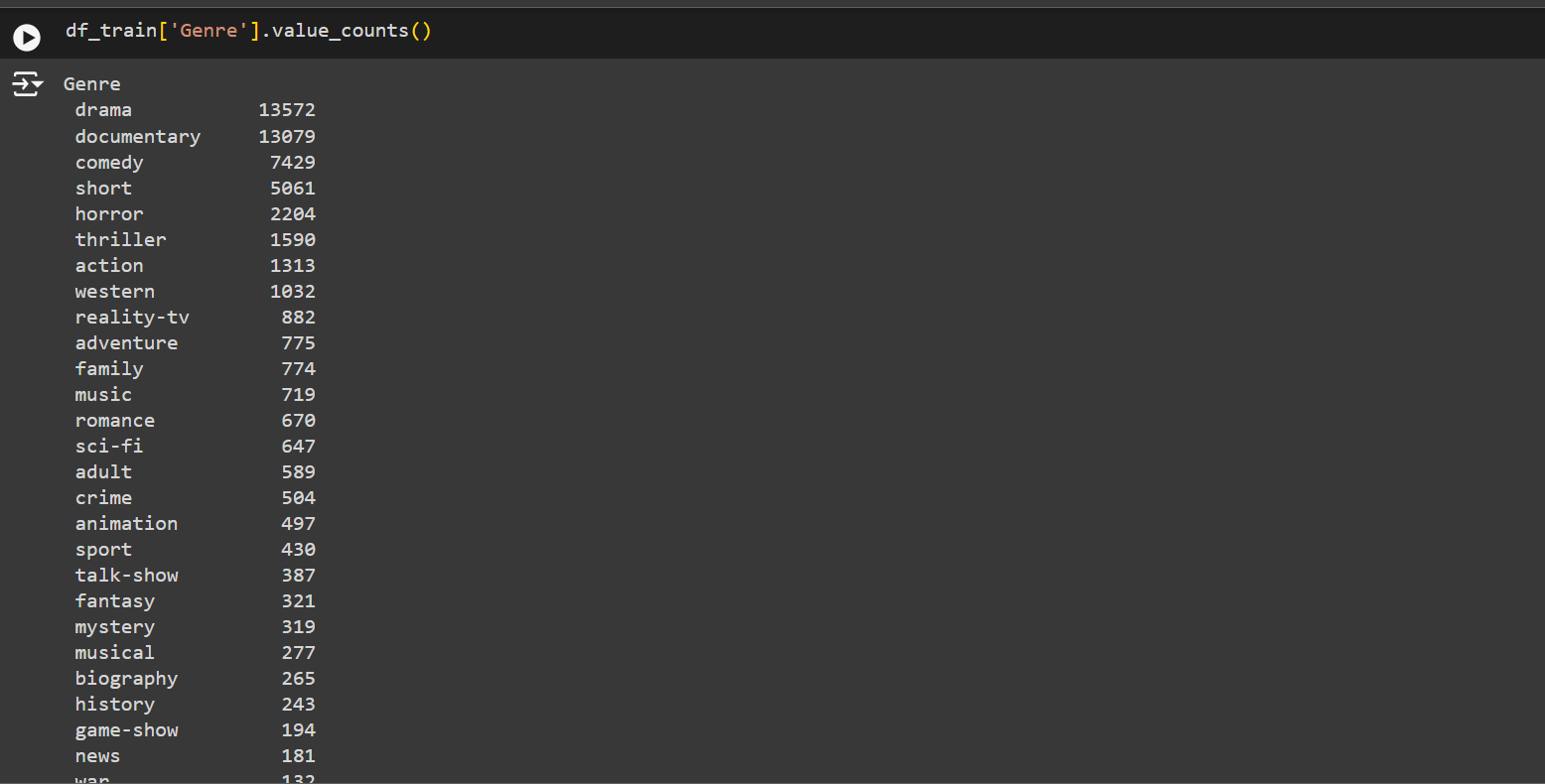
Displaying info about training and testing dataset:



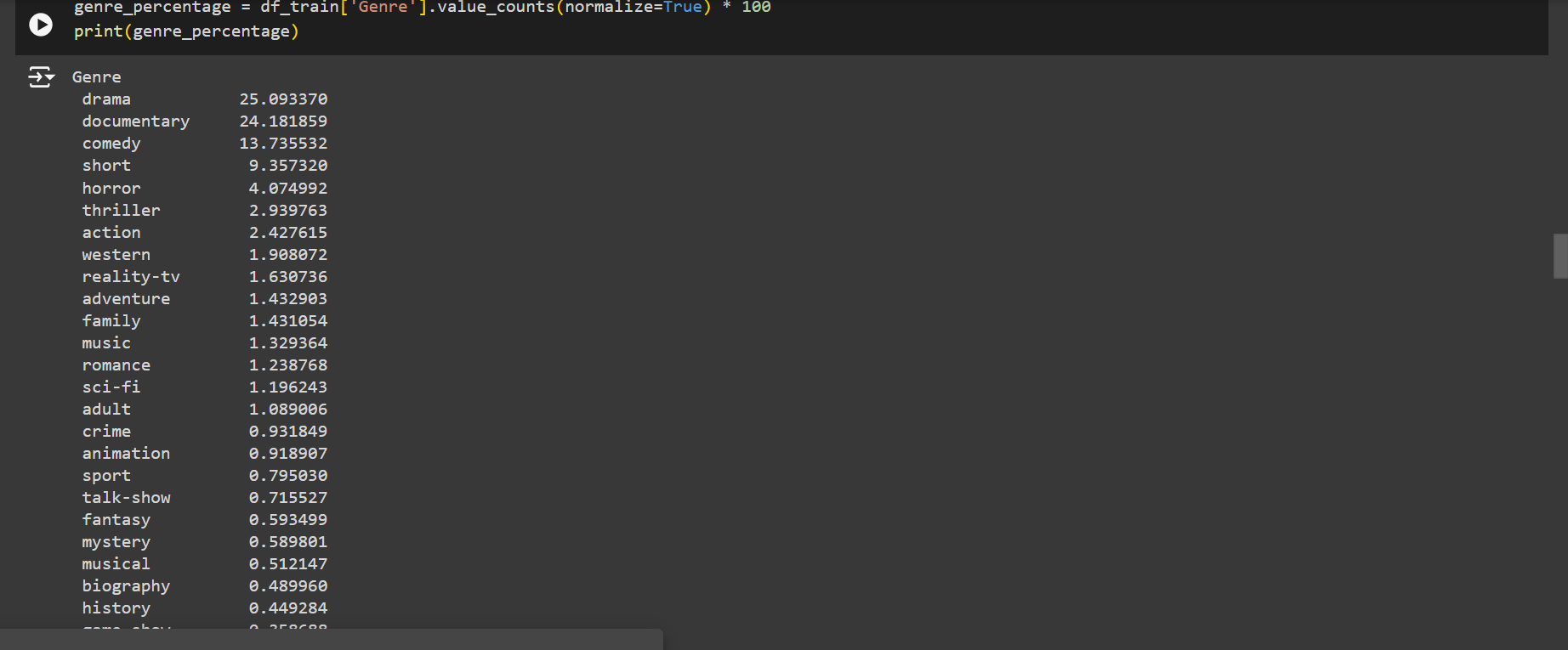
Removing duplicate values and displaying Training info without duplicate values:



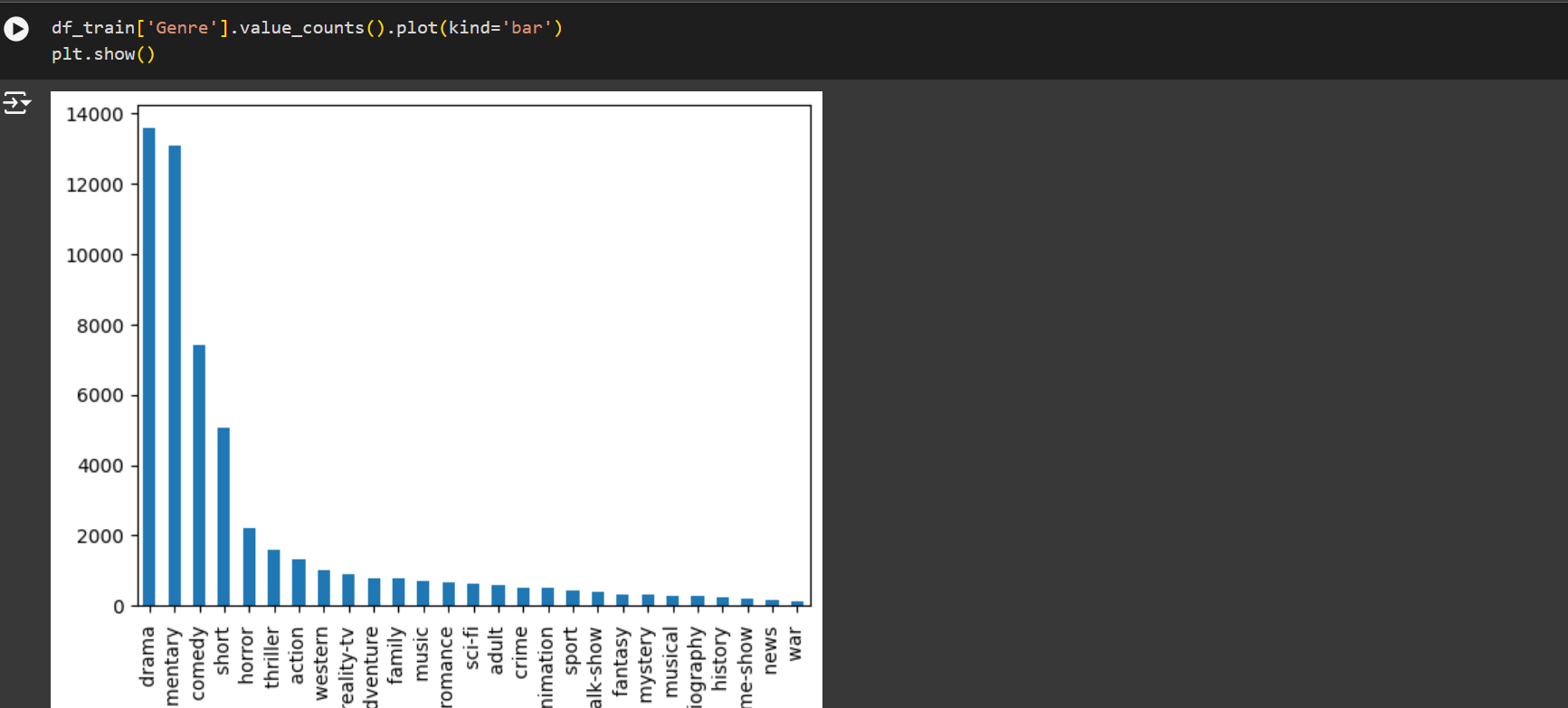
Total types and number of each genre in training set:



Showing percentages of different genres in the training data set (like drama genre is of 25% in whole data):



Displaying genre and the number of times the genre is found in training data set in a bar graph:





Preprocessing the data:

