**HACKATHON**

**HORROR MOVIE RATING**

**TASK 1:-**

1. By webscraping getting the horror movie details for year 2015.
2. By webscraping getting the horror movie details for year 2016.
3. With the help of Beautiful Soup
4. Doing the webscraping in jupyter notebook

**TASK 2:-**

1. Converting all the scraped data into dataframes for both years separately
2. And converting dataframes into a single CSV file named Gross.csv

**TASK 3:-**

The data is already scraped and converted into dataframes

**TASK 4:-**

1. The dataframes are converted into a single CSV file named as Gross.csv
2. New Database called Movies is created with Horror Collection
3. And it is imported to the MongoDB
4. The data types are set accordingly
5. The below columns are used in Horror Collection

**Movie** String

**Category** String

**Director** String

**Cast** String

**Country** String

**TheatricalDistributor** String

**TASK 5:-**

**#Movie with highest gross**

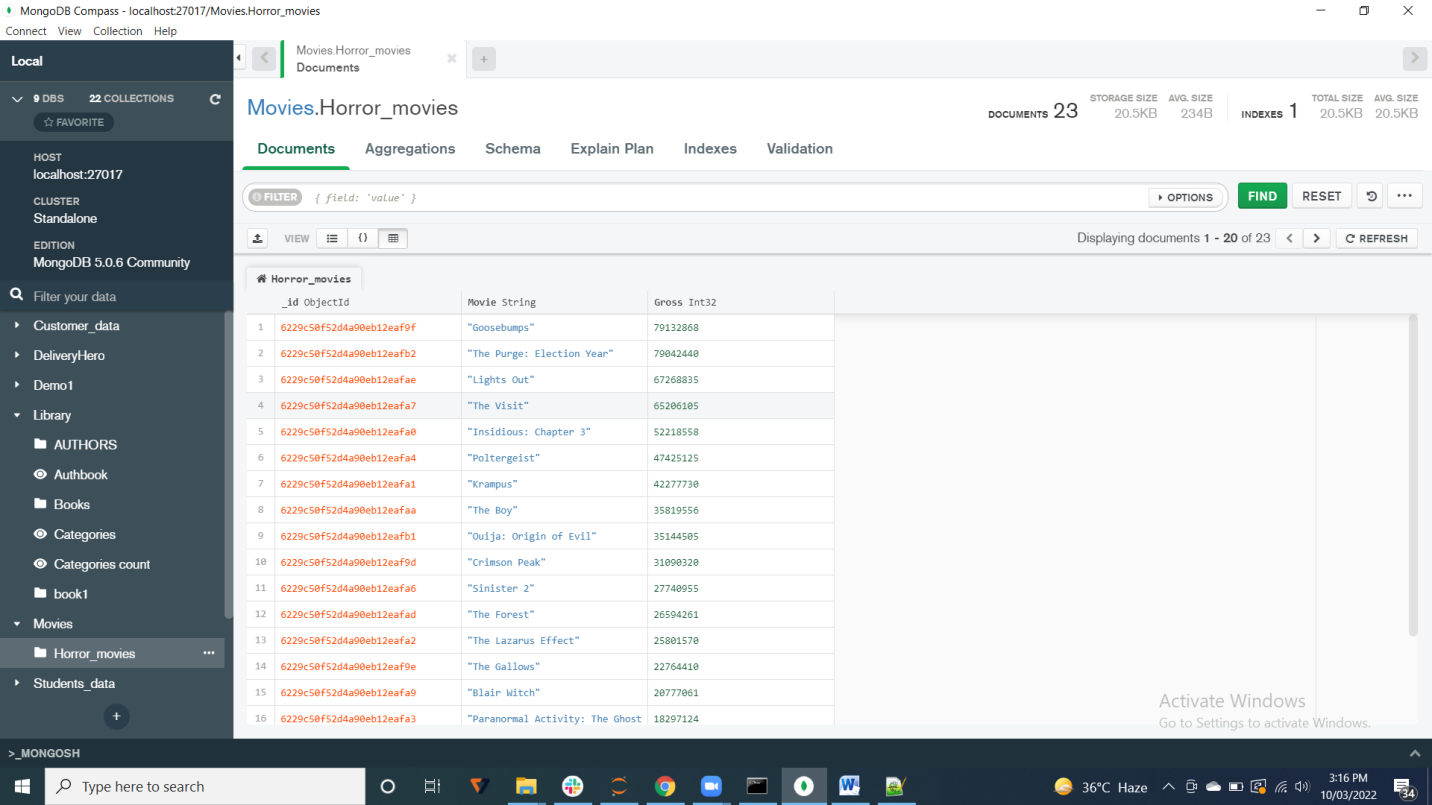
PROJECT:-

*{Gross:1,* *Movie:1}*

FILTER:-

*{Gross:-1}*

Goosebumps is the movie with highest cost =$7,91,32,868



**TASK 6:-**

**#Countries with highest count of horror movies**

Here a view called Cat\_count is created.

1)$unwind

{

path: '$Category',

preserveNullAndEmptyArrays: true

}

2)$group

{

\_id: '$Category',

Count: {

$sum:1

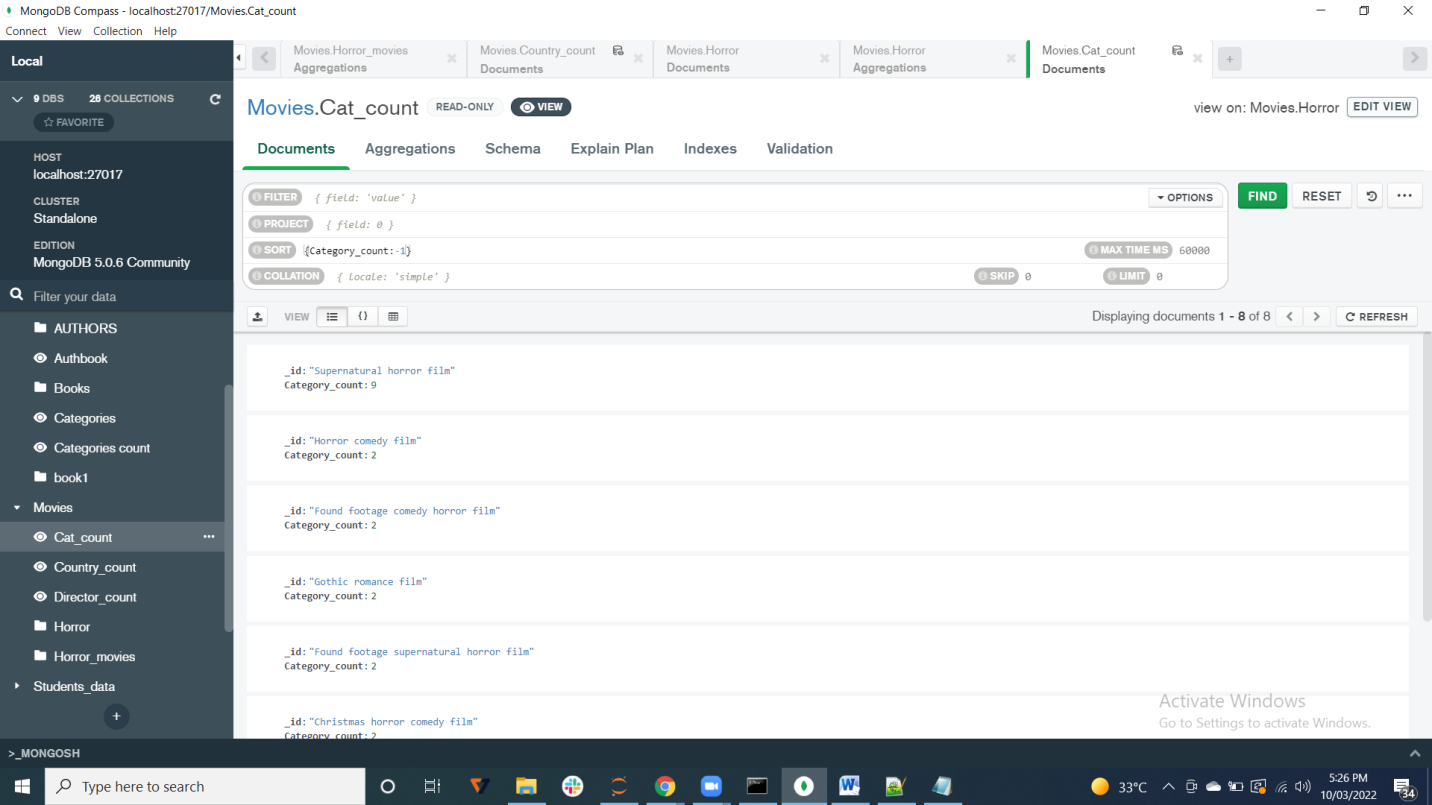
}

}

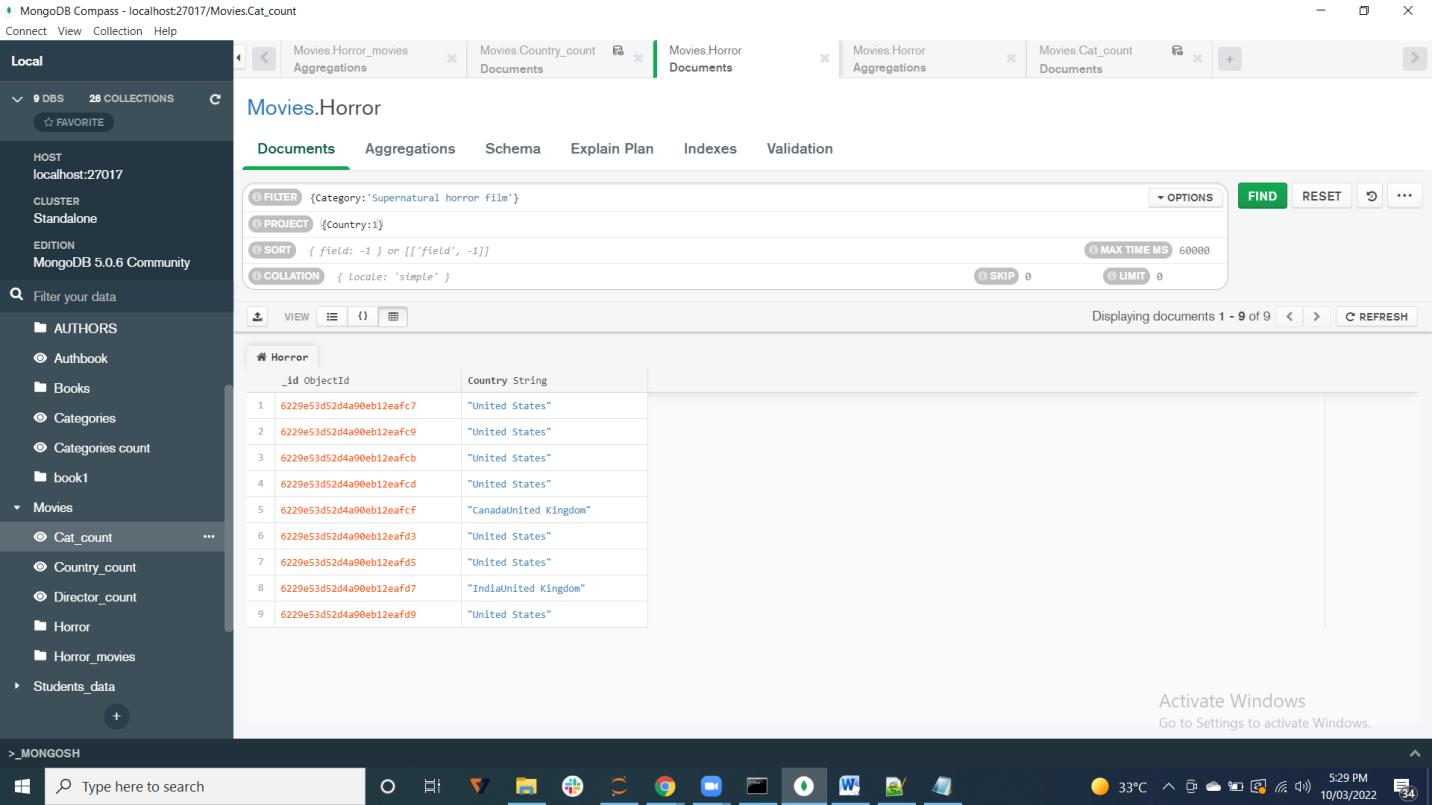
SORT:-

*{Category\_count:-1}*

Supernatural horror film is the most frequent category of horror movie



United States is the country.



**TASK 7:-**

**#Most horror directors**

PROJECT:-

*{Director:1,Gross:1}*

SORT:-

*{Gross:-1}*

Rob Letterman is the most horror director.

