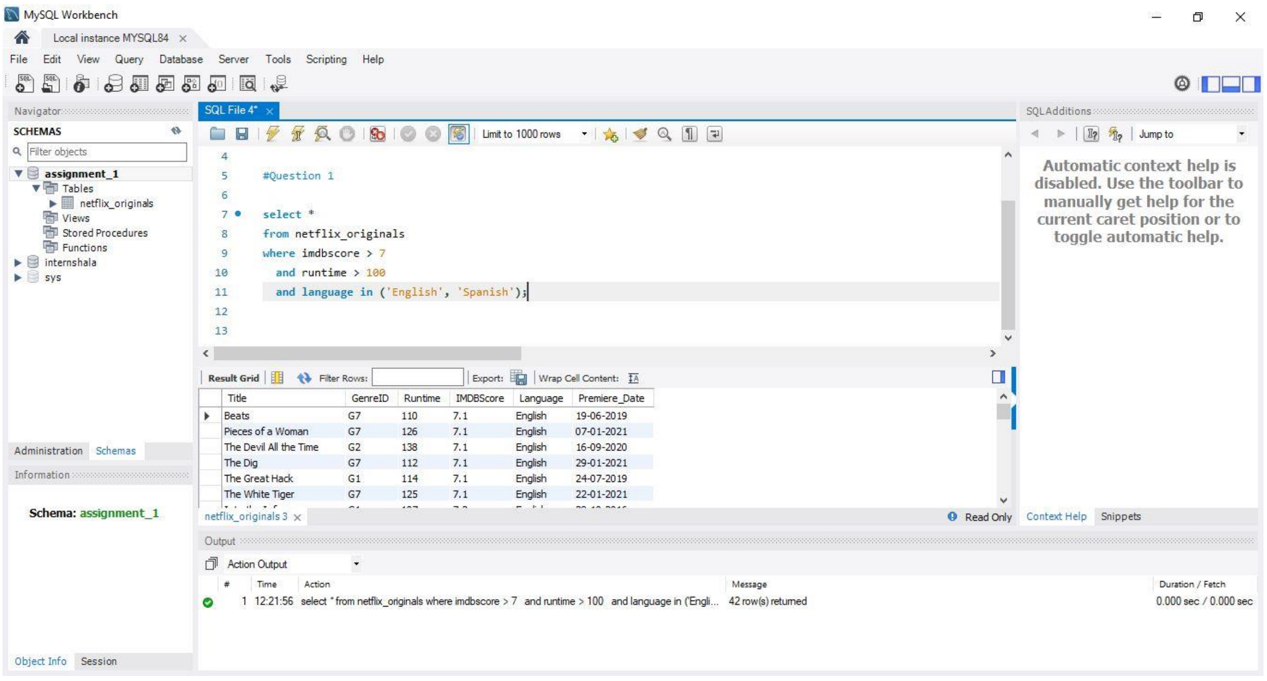
# MySQL

# ASSIGNMENT 1ST

**Problem statement** : The task involves analyzing the Netflix Originals dataset to derive insights about movie genres, runtime, IMDb scores, and premiere dates. The dataset holds valuable information regarding the content Netflix produces, and understanding these attributes can help identify trends and patterns. By using SQL queries, the aim is to perform complex filtering, aggregation, and sorting operations, providing meaningful insights for business decisions.

**Task 1st:** To retrieve all Netflix Originals with an IMDb score greater than 7, runtime greater than 100 minutes, and the language being either English or Spanish, I have used the following SQL query:

```
select *
from netflix_originals
where imdbscore > 7
  and runtime > 100
  and language in ('English', 'Spanish');
```



```
#Question 1

select *
from netflix_originals
where imdbscore > 7
  and runtime > 100
  and language in ('English', 'Spanish');
```

| Title | GenreID | Runtime | IMDBScore | Language | Premiere_Date |
|---|---|---|---|---|---|
| Beats | G7 | 110 | 7.1 | English | 19-06-2019 |
| Pieces of a Woman | G7 | 126 | 7.1 | English | 07-01-2021 |
| The Devil All the Time | G2 | 138 | 7.1 | English | 16-09-2020 |
| The Dig | G7 | 112 | 7.1 | English | 29-01-2021 |
| The Great Hack | G1 | 114 | 7.1 | English | 24-07-2019 |
| The White Tiger | G7 | 125 | 7.1 | English | 22-01-2021 |

```
#  Time      Action                                                                              Message                Duration / Fetch
1  12:21:56  select * from netflix_originals where imdbscore > 7  and runtime > 100  and language in ('Engli...  42 row(s) returned   0.000 sec / 0.000 sec
```

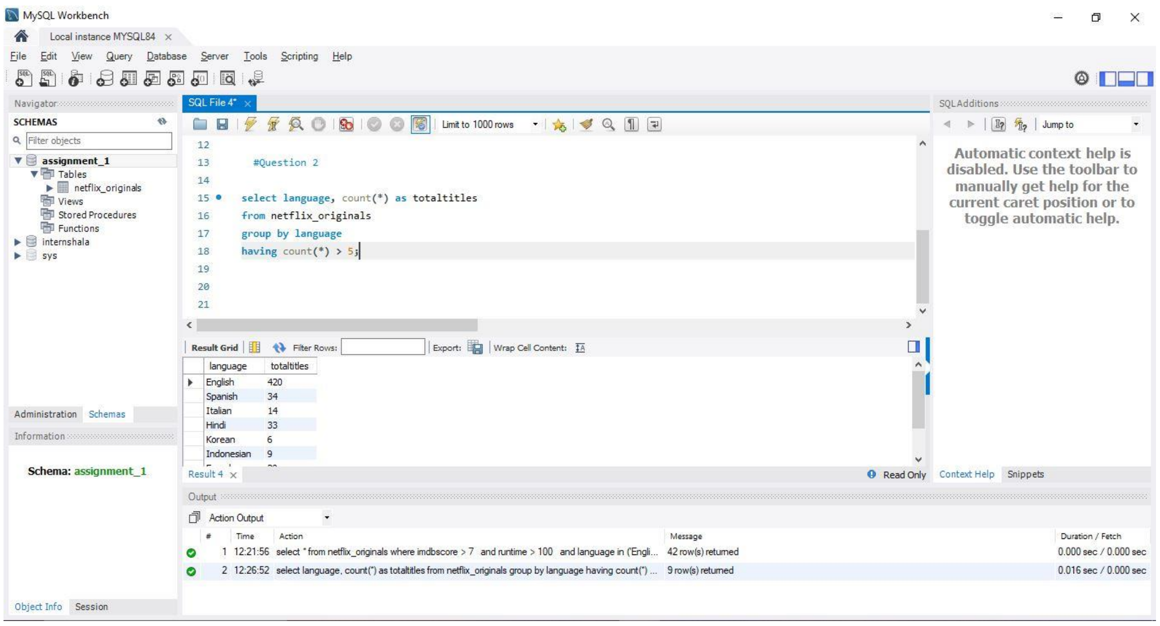
**Task 2nd:** Finding the total number of Netflix Originals in each language, but only show those languages that have more than 5 titles.

**select language, count(*) as totaltitles**

**from netflix_originals**

**group by language**

**having count(*) > 5;**



```
#Question 2

select language, count(*) as totaltitles
from netflix_originals
group by language
having count(*) > 5;
```

| language | totaltitles |
|---|---|
| English | 420 |
| Spanish | 34 |
| Italian | 14 |
| Hindi | 33 |
| Korean | 6 |
| Indonesian | 9 |

| # | Time | Action | Message | Duration / Fetch |
|---|---|---|---|---|
| 1 | 12:21:56 | select * from netflix_originals where imdbscore > 7 and runtime > 100 and language in ('Engli... | 42 row(s) returned | 0.000 sec / 0.000 sec |
| 2 | 12:26:52 | select language, count(*) as totaltitles from netflix_originals group by language having count(*) ... | 9 row(s) returned | 0.016 sec / 0.000 sec |

Task 3rd: To get the top 3 longest-running movies in Hindi language sorted by IMDb score in descending order

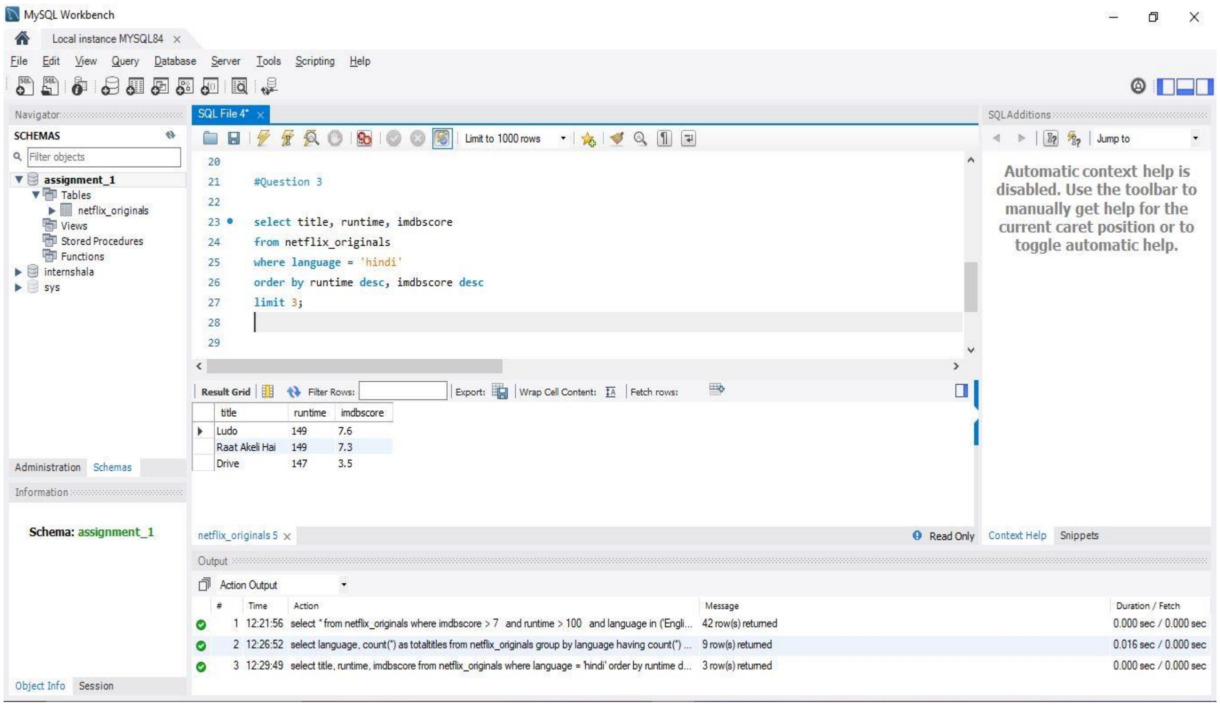
**select title, runtime, imdbscore**

**from netflix_originals**

**where language = 'hindi'**

**order by runtime desc, imdbscore desc**

**limit 3;**



```
#Question 3

select title, runtime, imdbscore
from netflix_originals
where language = 'hindi'
order by runtime desc, imdbscore desc
limit 3;
```

| title | runtime | imdbscore |
|---|---|---|
| Ludo | 149 | 7.6 |
| Raat Akeli Hai | 149 | 7.3 |
| Drive | 147 | 3.5 |

Action Output

| # | Time | Action | Message | Duration / Fetch |
|---|---|---|---|---|
| 1 | 12:21:56 | select * from netflix_originals where imdbscore > 7  and runtime > 100  and language in ('Engli... | 42 row(s) returned | 0.000 sec / 0.000 sec |
| 2 | 12:26:52 | select language, count(*) as totaltitles from netflix_originals group by language having count(*) ... | 9 row(s) returned | 0.016 sec / 0.000 sec |
| 3 | 12:29:49 | select title, runtime, imdbscore from netflix_originals where language = 'hindi' order by runtime d... | 3 row(s) returned | 0.000 sec / 0.000 sec |

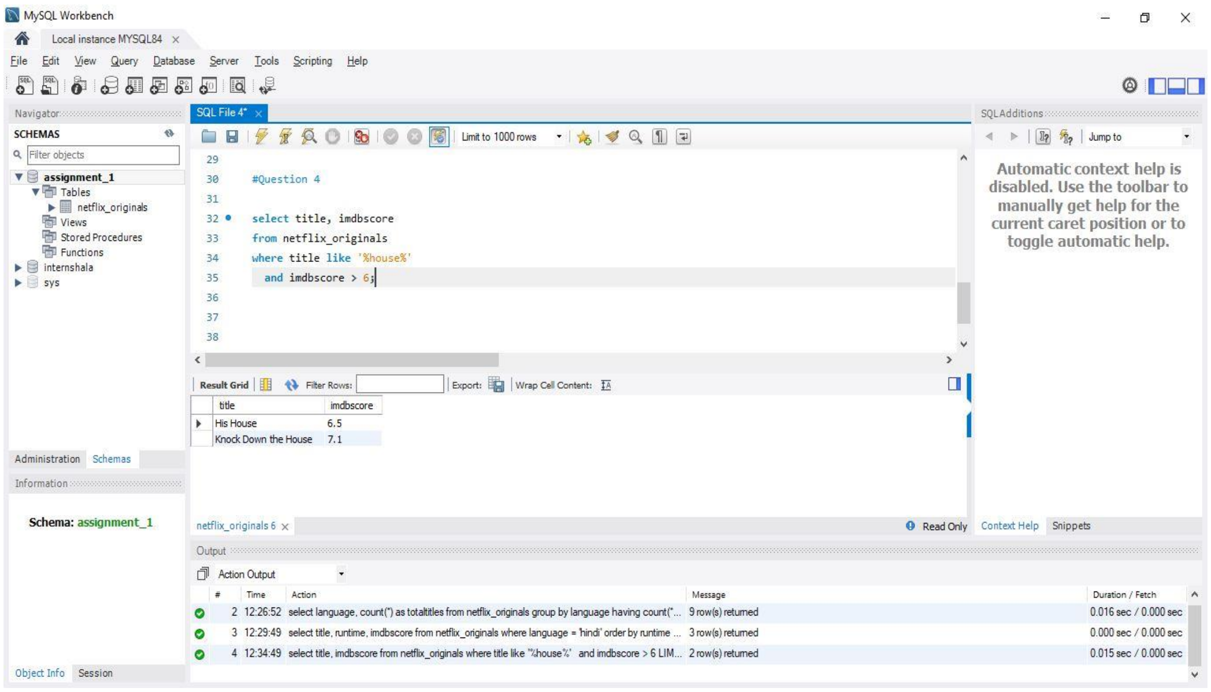
**Task 4th :** To retrieve all titles that contain the word "House" in their name and have an IMDb score greater than 6.

**select title, imdbscore**

**from netflix_originals**

**where title like '%house%'**

  **and imdbscore > 6;**



```
#Question 4

select title, imdbscore
from netflix_originals
where title like '%house%'
  and imdbscore > 6;
```

| | title | imdbscore |
|---|---|---|
| ▶ | His House | 6.5 |
| | Knock Down the House | 7.1 |

```
     #  Time      Action                                                                                      Message            Duration / Fetch
 ✓   2  12:26:52  select language, count(*) as totaltitles from netflix_originals group by language having count(*...  9 row(s) returned  0.016 sec / 0.000 sec
 ✓   3  12:29:49  select title, runtime, imdbscore from netflix_originals where language = 'hindi' order by runtime ...  3 row(s) returned  0.000 sec / 0.000 sec
 ✓   4  12:34:49  select title, imdbscore from netflix_originals where title like '%house%'  and imdbscore > 6 LIM...  2 row(s) returned  0.015 sec / 0.000 sec
```

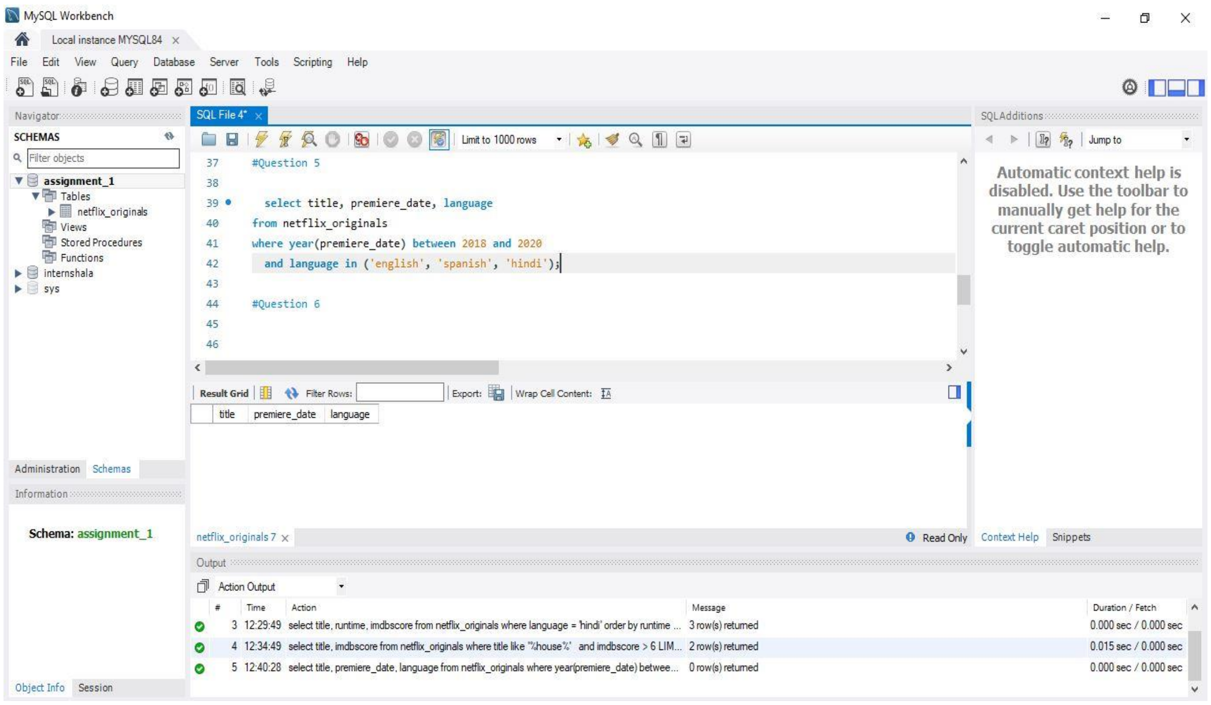
**Task 5th:** Finding all Netflix Originals released between the years 2018 and 2020 that are in either English, Spanish, or Hindi.

**select title, premiere_date, language**

**from netflix_originals**

**where year(premiere_date) between 2018 and 2020**

**and language in ('english', 'spanish', 'hindi');**



```
#Question 5


  select title, premiere_date, language
from netflix_originals
where year(premiere_date) between 2018 and 2020
  and language in ('english', 'spanish', 'hindi');


#Question 6
```

Result Grid: title | premiere_date | language

Schema: assignment_1

Automatic context help is disabled. Use the toolbar to manually get help for the current caret position or to toggle automatic help.

```
#   Time      Action                                                                               Message           Duration / Fetch
3   12:29:49  select title, runtime, imdbscore from netflix_originals where language = 'hindi' order by runtime ...   3 row(s) returned   0.000 sec / 0.000 sec
4   12:34:49  select title, imdbscore from netflix_originals where title like "%house%"  and imdbscore > 6 LIM...   2 row(s) returned   0.015 sec / 0.000 sec
5   12:40:28  select title, premiere_date, language from netflix_originals where year(premiere_date) betwee...   0 row(s) returned   0.000 sec / 0.000 sec
```

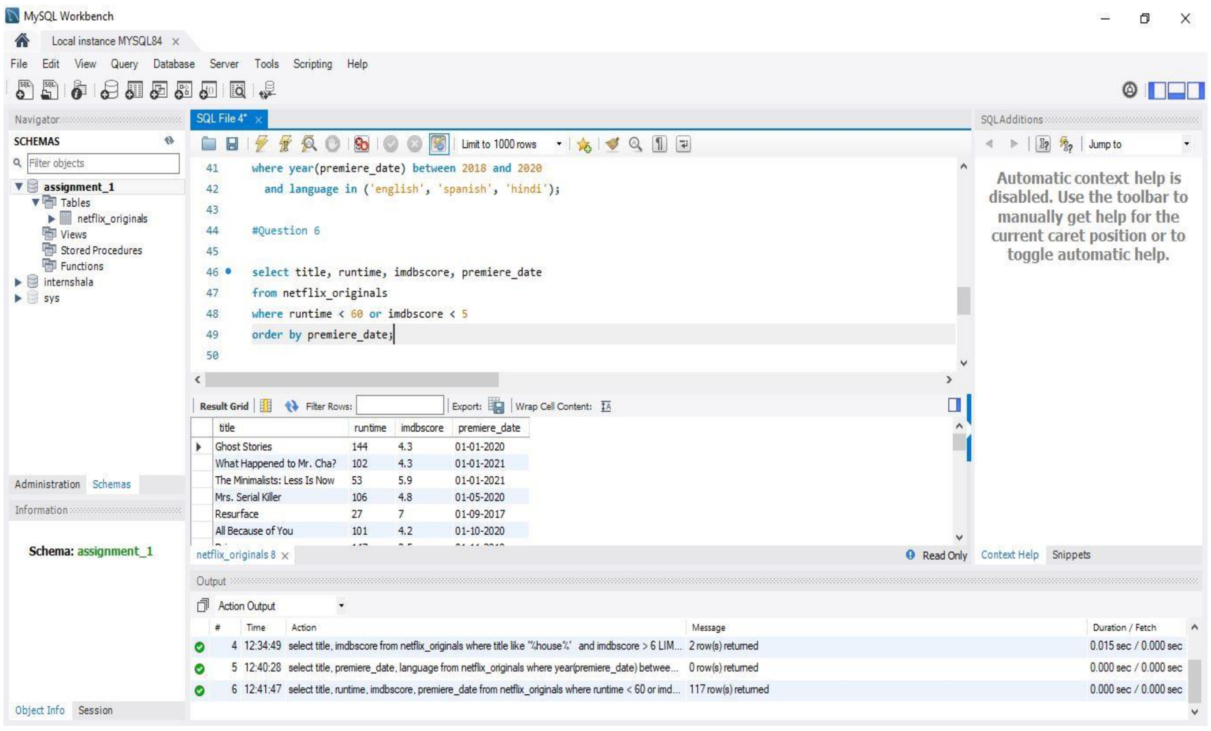
**Task 6th :** Finding all movies that either have a runtime less than 60 minutes or an IMDb score less than 5, sorted by Premiere Date.

**select title, runtime, imdbscore, premiere_date**

**from netflix_originals**

**where runtime < 60 or imdbscore < 5**

**order by premiere_date;**



```
where year(premiere_date) between 2018 and 2020
  and language in ('english', 'spanish', 'hindi');

#Question 6

select title, runtime, imdbscore, premiere_date
from netflix_originals
where runtime < 60 or imdbscore < 5
order by premiere_date;
```

| title | runtime | imdbscore | premiere_date |
|---|---|---|---|
| Ghost Stories | 144 | 4.3 | 01-01-2020 |
| What Happened to Mr. Cha? | 102 | 4.3 | 01-01-2021 |
| The Minimalists: Less Is Now | 53 | 5.9 | 01-01-2021 |
| Mrs. Serial Killer | 106 | 4.8 | 01-05-2020 |
| Resurface | 27 | 7 | 01-09-2017 |
| All Because of You | 101 | 4.2 | 01-10-2020 |

| # | Time | Action | Message | Duration / Fetch |
|---|---|---|---|---|
| 4 | 12:34:49 | select title, imdbscore from netflix_originals where title like '%house%'  and imdbscore > 6 LIM... | 2 row(s) returned | 0.015 sec / 0.000 sec |
| 5 | 12:40:28 | select title, premiere_date, language from netflix_originals where year(premiere_date) betwee... | 0 row(s) returned | 0.000 sec / 0.000 sec |
| 6 | 12:41:47 | select title, runtime, imdbscore, premiere_date from netflix_originals where runtime < 60 or imd... | 117 row(s) returned | 0.000 sec / 0.000 sec |

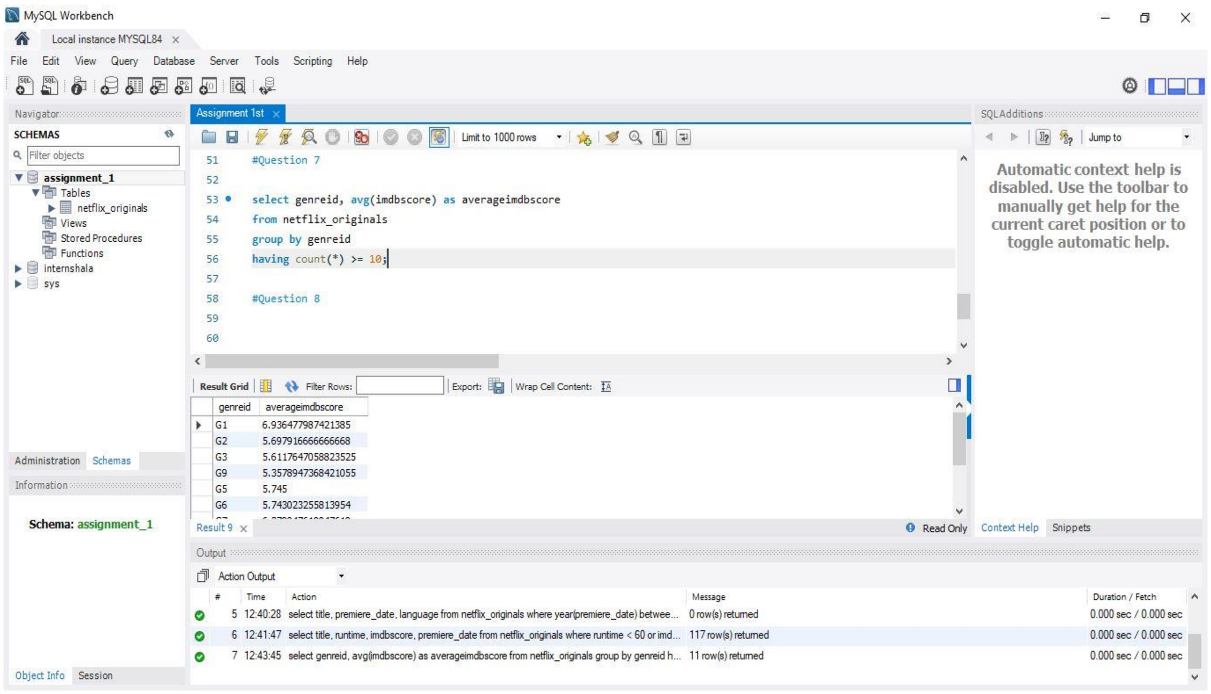
**Task 7th:** To get the average IMDb score for each genre where the genre has at least 10 movies.

**select genreid, avg(imdbscore) as averageimdbscore**

**from netflix_originals**

**group by genreid**

**having count(*) >= 10;**



Schema: assignment_1

```
#Question 7

select genreid, avg(imdbscore) as averageimdbscore
from netflix_originals
group by genreid
having count(*) >= 10;

#Question 8
```

Automatic context help is disabled. Use the toolbar to manually get help for the current caret position or to toggle automatic help.

| genreid | averageimdbscore |
|---|---|
| G1 | 6.936477987421385 |
| G2 | 5.697916666666668 |
| G3 | 5.6117647058823525 |
| G9 | 5.3578947368421055 |
| G5 | 5.745 |
| G6 | 5.743023255813954 |

```
#  Time      Action                                                                                   Message            Duration / Fetch
5  12:40:28  select title, premiere_date, language from netflix_originals where year(premiere_date) betwee...  0 row(s) returned    0.000 sec / 0.000 sec
6  12:41:47  select title, runtime, imdbscore, premiere_date from netflix_originals where runtime < 60 or imd...  117 row(s) returned  0.000 sec / 0.000 sec
7  12:43:45  select genreid, avg(imdbscore) as averageimdbscore from netflix_originals group by genreid h...  11 row(s) returned   0.000 sec / 0.000 sec
```

**Task 8th:** To retrieve the top 5 most common runtimes for Netflix Originals.

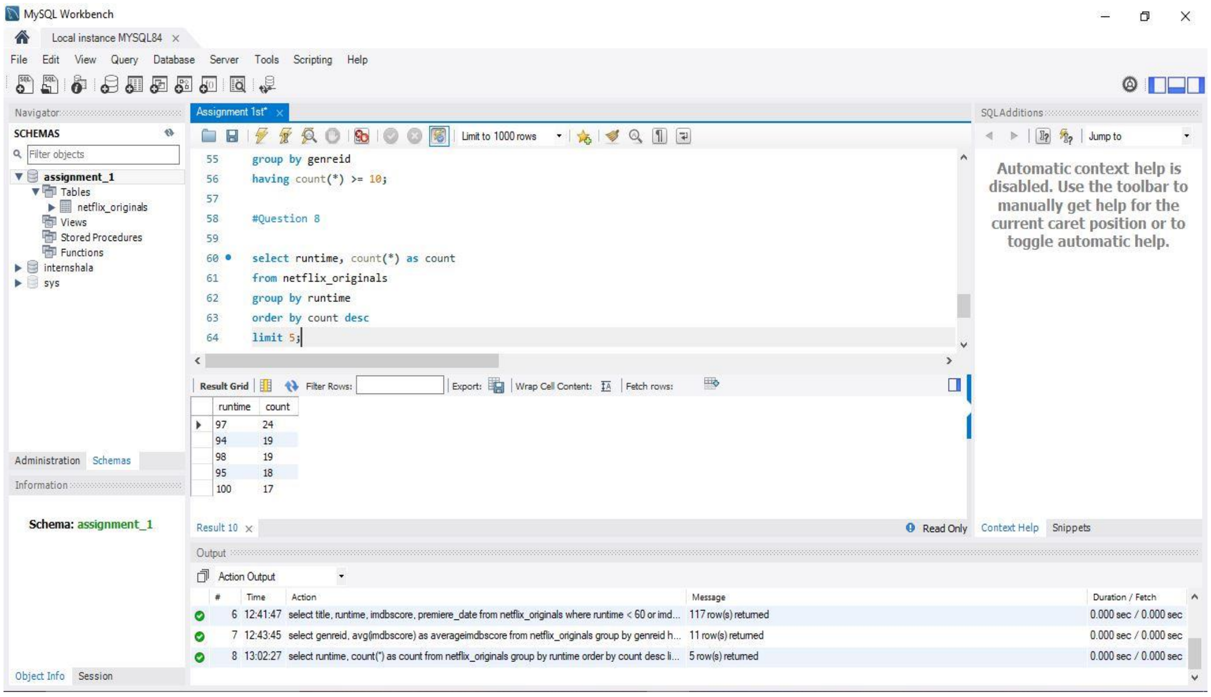
**select runtime, count(*) as count**

**from netflix_originals**

**group by runtime**

**order by count desc**

**limit 5;**



```
group by genreid
having count(*) >= 10;

#Question 8

select runtime, count(*) as count
from netflix_originals
group by runtime
order by count desc
limit 5;
```

| runtime | count |
|---|---|
| 97 | 24 |
| 94 | 19 |
| 98 | 19 |
| 95 | 18 |
| 100 | 17 |

```
#  Time      Action                                                                                      Message             Duration / Fetch
6  12:41:47  select title, runtime, imdbscore, premiere_date from netflix_originals where runtime < 60 or imd...  117 row(s) returned  0.000 sec / 0.000 sec
7  12:43:45  select genreid, avg(imdbscore) as averageimdbscore from netflix_originals group by genreid h...      11 row(s) returned   0.000 sec / 0.000 sec
8  13:02:27  select runtime, count(*) as count from netflix_originals group by runtime order by count desc li...   5 row(s) returned    0.000 sec / 0.000 sec
```

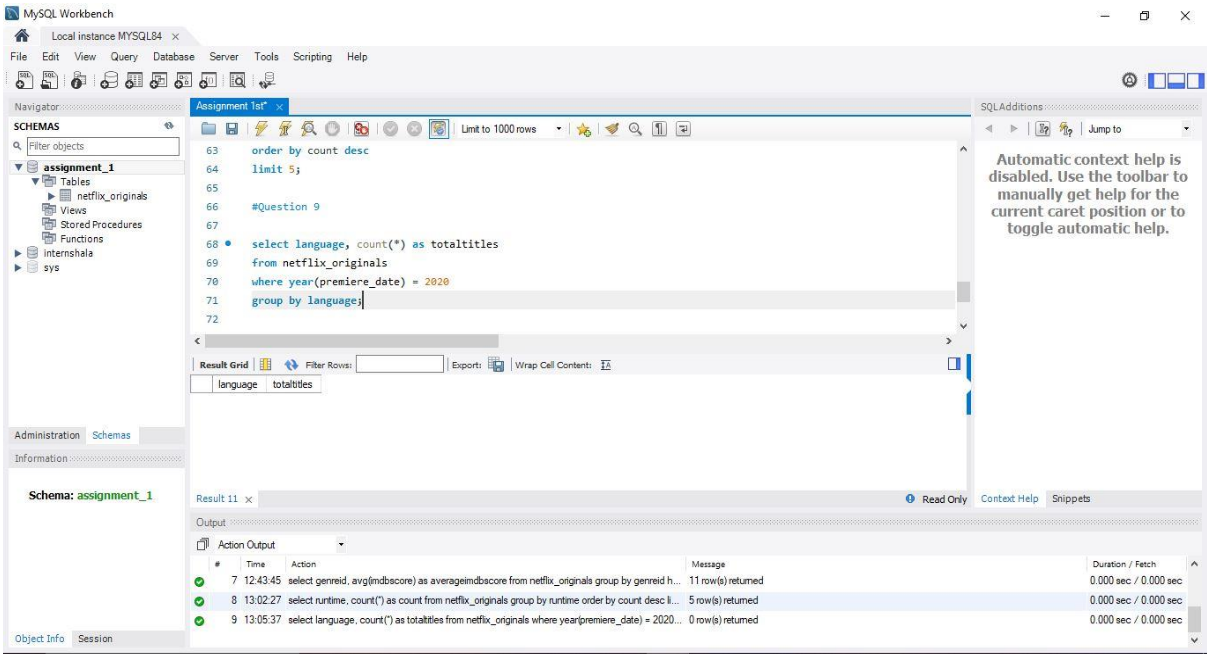
**Task 9th :** To listing all Netflix Originals that were released in 2020, grouped by language, and show the total count of titles for each language.

**select language, count(*) as totaltitles**

**from netflix_originals**

**where year(premiere_date) = 2020**

**group by language;**



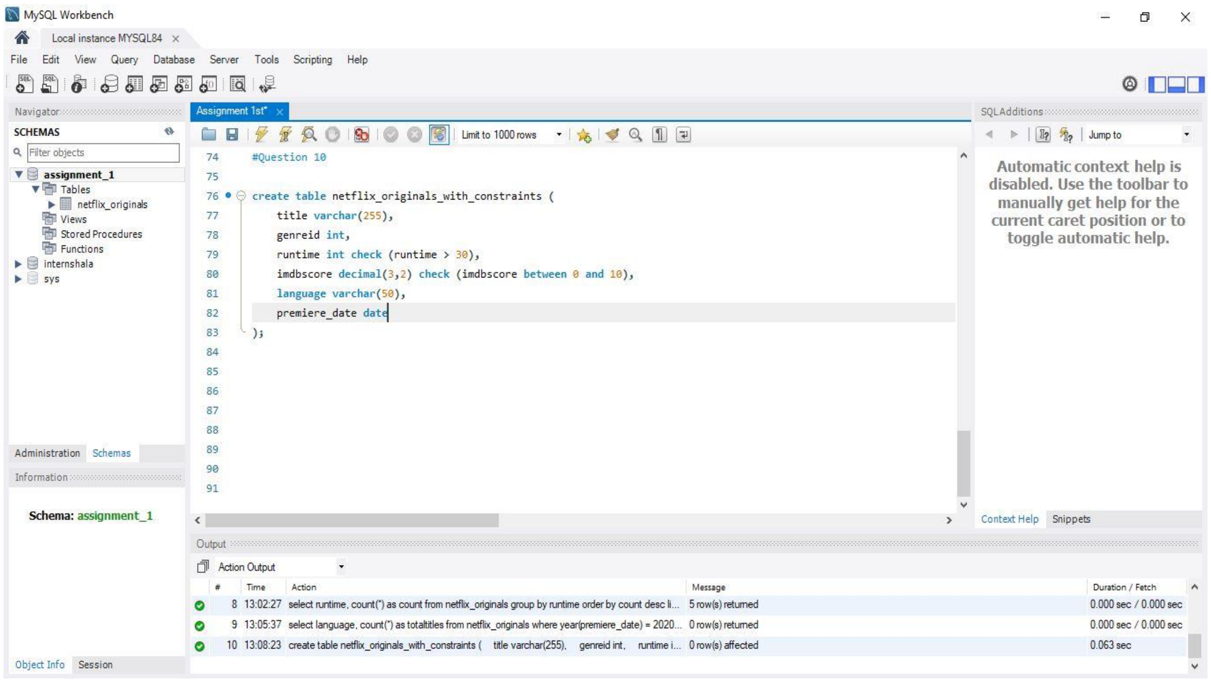
```
order by count desc
limit 5;

#Question 9

select language, count(*) as totaltitles
from netflix_originals
where year(premiere_date) = 2020
group by language;
```

```
language   totaltitles
```

```
#  Time      Action                                                                               Message            Duration / Fetch
7  12:43:45  select genreid, avg(imdbscore) as averageimdbscore from netflix_originals group by genreid h...  11 row(s) returned  0.000 sec / 0.000 sec
8  13:02:27  select runtime, count(*) as count from netflix_originals group by runtime order by count desc li...  5 row(s) returned  0.000 sec / 0.000 sec
9  13:05:37  select language, count(*) as totaltitles from netflix_originals where year(premiere_date) = 2020...  0 row(s) returned  0.000 sec / 0.000 sec
```

**Task 10th:** To creating a new table that enforces a constraint on the IMDb score to be between 0 and 10 and the runtime to be greater than 30 minutes.

```
create table netflix_originals_with_constraints (
    title varchar(255),
    genreid int,
    runtime int check (runtime > 30),
    imdbscore decimal(3,2) check (imdbscore between 0 and 10),
    language varchar(50),
    premiere_date date
);
```



```
#Question 10

create table netflix_originals_with_constraints (
    title varchar(255),
    genreid int,
    runtime int check (runtime > 30),
    imdbscore decimal(3,2) check (imdbscore between 0 and 10),
    language varchar(50),
    premiere_date date
);
```

Action Output

```
#   Time      Action                                                                                  Message              Duration / Fetch
8   13:02:27  select runtime, count(*) as count from netflix_originals group by runtime order by count desc li...  5 row(s) returned    0.000 sec / 0.000 sec
9   13:05:37  select language, count(*) as totaltitles from netflix_originals where year(premiere_date) = 2020...  0 row(s) returned    0.000 sec / 0.000 sec
10  13:08:23  create table netflix_originals_with_constraints (    title varchar(255),    genreid int,    runtime i...  0 row(s) affected    0.063 sec
```