

Assignment 1

Note: Must change Date Datatype to get accurate answers

```
1 • use netflix;
2
3 • select * from netflix_data;
4
5 • desc netflix_data;
6 • ALTER TABLE netflix_data
7   MODIFY Premiere_Date date;
8
9   /*changed Premiere_Date datatype */
10 • update netflix_data
11   SET Premiere_Date=str_to_date(Premiere_Date,"%d-%m-%Y");
```

Task 1:

Retrieve all Netflix Originals with an IMDb score greater than 7, runtime greater than 100 minutes, and the language is either English or Spanish.

The screenshot shows a SQL IDE interface. At the top, there's a toolbar with icons for file operations, search, and execution. Below the toolbar, a SQL query is entered in a text area. The query is as follows:

```
1 /* Q1.
2 Retrieve all Netflix Originals with an IMDb score greater than 7, runtime greater than 100 minutes, and the
3 language is either English or Spanish */
4
5
6 select * from netflix_data;
7
8 select *
9 from netflix_data
10 where IMDBScore > 7 AND Runtime > 100 AND Language IN ('English','Spanish');
```

Below the query editor, there's a "Result Grid" section. It contains a table with the following data:

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

Below the result grid, there's an "Output" section. It shows the execution of the query and the number of rows returned. The output is as follows:

#	Time	Action	Message
1	15:38:46	select * from netflix_data LIMIT 0, 1000	584 row(s) returned
2	15:38:51	select * from netflix_data where IMDBScore > 7 AND Runtime > 100 AND Language IN ('English','Spanish') ...	42 row(s) returned

Task 2:

Find the total number of Netflix Originals in each language, but only show those languages that have more than 5 titles.

```
1  /* Find the total number of Netflix Originals in each language, but only show those languages that have more
2  than 5 titles.*/
3
4
5  • select * from netflix_data;
6
7  • select language, count(*) 'Netflix Originals'
8  from netflix_data
9  Group by Language
10 Having count(*) > 5;
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content: |

	Language	Netflix Originals
▶	English	420
	Spanish	34
	Italian	14
	Hindi	33
	Korean	6
	Indonesian	9
	French	20
	Portuguese	12
	Japanese	6

Task 3:

Get the top 3 longest-running movies in Hindi language sorted by IMDb score in descending order.

```
1  /* Get the top 3 longest-running movies in Hindi language sorted by IMDb score in descending order */
2
3
4  • select * from netflix_data;
5
6  • select Title, Language, Runtime, IMDBScore
7  from netflix_data
8  where Language='Hindi'
9  Order by Runtime DESC, IMDBScore DESC
10 Limit 3;
```





Result Grid | Filter Rows: | Export: | Wrap Cell Content: | Fetch rows: |

	Title	Language	Runtime	IMDBScore
▶	Ludo	Hindi	149	7.6
	Raat Akeli Hai	Hindi	149	7.3
	Drive	Hindi	147	3.5

Task 4:

Retrieve all titles that contain the word "House" in their name and have an IMDb score greater than 6.

```
1  /* Retrieve all titles that contain the word "House" in their name and have an IMDb score greater than 6 */
2
3
4
5
6  • select * from netflix_data;
7
8  • select Title,IMDBScore
9    from netflix_data
10   where Title like '%House%' AND IMDBScore > 6;
```

Result Grid   Filter Rows: Export:  Wrap Cell Content: 

Title	IMDBScore
His House	6.5
Knock Down the House	7.1

Task 5:

Find all Netflix Originals released between the years 2018 and 2020 that are in either English, Spanish, or Hindi.

```
1  /* Find all Netflix Originals released between the years 2018 and 2020 that are in either English, Spanish, or Hindi. */
2
3
4  • select * from netflix_data;
5
6  select *
7    from netflix_data
8   where Premiere_Date between '2018-01-01' and '2020-12-31' and Language IN ('English','Spanish','Hindi');
```

Result Grid   Filter Rows: Export:  Wrap Cell Content: 

Title	GenreID	Runtime	IMDBScore	Language	Premiere_Date
Enter the Anime	G1	58	2.5	English	2019-08-05
Dark Forces	G2	81	2.6	Spanish	2020-08-21
The Open House	G9	94	3.2	English	2018-01-19
Kaali Khuli	G4	90	3.4	Hindi	2020-10-30
Drive	G5	147	3.5	Hindi	2019-11-01
The Last Days of American Crime	G2	149	3.7	English	2020-06-05
Paradox	G8	73	3.9	English	2018-03-23
After the Raid	G1	25	4.3	Spanish	2019-12-19
Ghost Stories	G9	144	4.3	Hindi	2020-01-01
The Last Thing He Wanted	G2	115	4.3	English	2020-02-21
Hello Privilege. It's Me, Chelsea	G1	64	4.4	English	2019-09-13
Secret Obsession	G2	97	4.4	English	2019-07-18

Task 6:

Find all movies that either have a runtime less than 60 minutes or an IMDb score less than 5, sorted by Premiere Date.

```
1  /* Find all movies that either have a runtime less than 60 minutes or an IMDb score less than 5, sorted by
2  Premiere Date */
3
4  •  select * from netflix_data;
5
6  •  select Title, Runtime, IMDBScore, Premiere_Date
7  from netflix_data
8  where Runtime < 60 OR IMDBScore < 5
9  order by Premiere_Date;
```

Result Grid				
Filter Rows: <input type="text"/> Export: Wrap Cell Content:				
	Title	Runtime	IMDBScore	Premiere_Date
▶	A Very Murray Christmas	56	5.5	2015-12-04
	Extremis	24	7.3	2016-09-13
	The White Helmets	40	7.5	2016-09-16
	I Am the Pretty Thing That Lives in the House	89	4.6	2016-10-28
	Mercy	90	4.2	2016-11-22
	Coin Heist	97	4.8	2017-01-06
	Take the 10	80	4.8	2017-01-20
	13th: A Conversation with Oprah Winfrey & Av...	36	7.1	2017-01-26
	Michael Bolton's Big, Sexy, Valentine's Day Special	54	6.7	2017-02-07
	Rodney King	52	5.8	2017-04-28
	Death Note	100	4.4	2017-08-25
	Resurface	27	7	2017-09-01

Task 7:

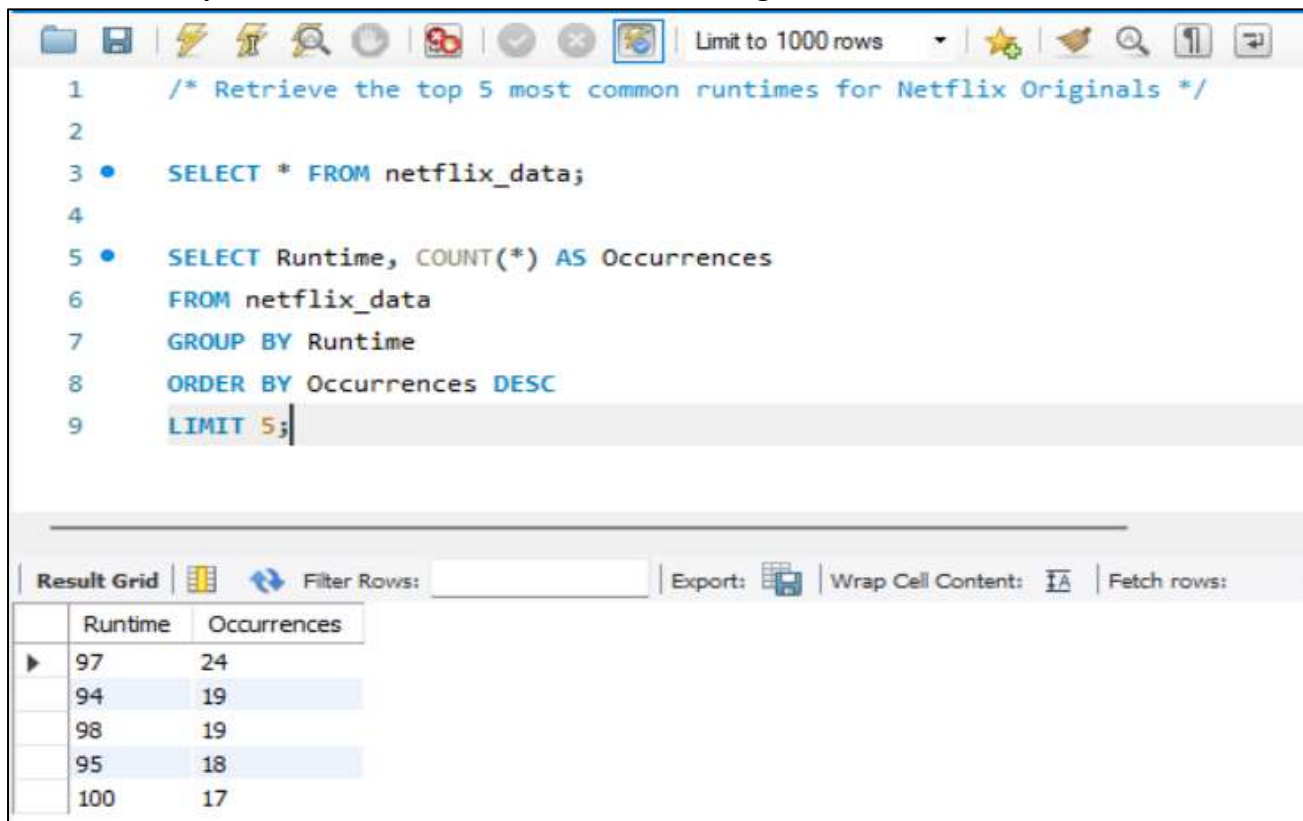
Get the average IMDb score for each genre where the genre has at least 10 movies.

```
1  /* Get the average IMDb score for each genre where the genre has at least 10 movies */
2
3
4  select * from netflix_data;
5
6  select GenreID, round(AVG(IMDBScore),2) as 'AVG IMDBScore'
7  from netflix_data
8  group by GenreID
9  Having Count(*) >= 10;
```

Result Grid		
Filter Rows: <input type="text"/> Export: Wrap Cell Content:		
	GenreID	AVG IMDBScore
▶	G1	6.94
	G2	5.7
	G3	5.61
	G9	5.36
	G5	5.74
	G6	5.74
	G7	6.38
	G17	5.84
	G10	6.09
	G11	6.67
	G15	6.44

Task 8:

Retrieve the top 5 most common runtimes for Netflix Originals

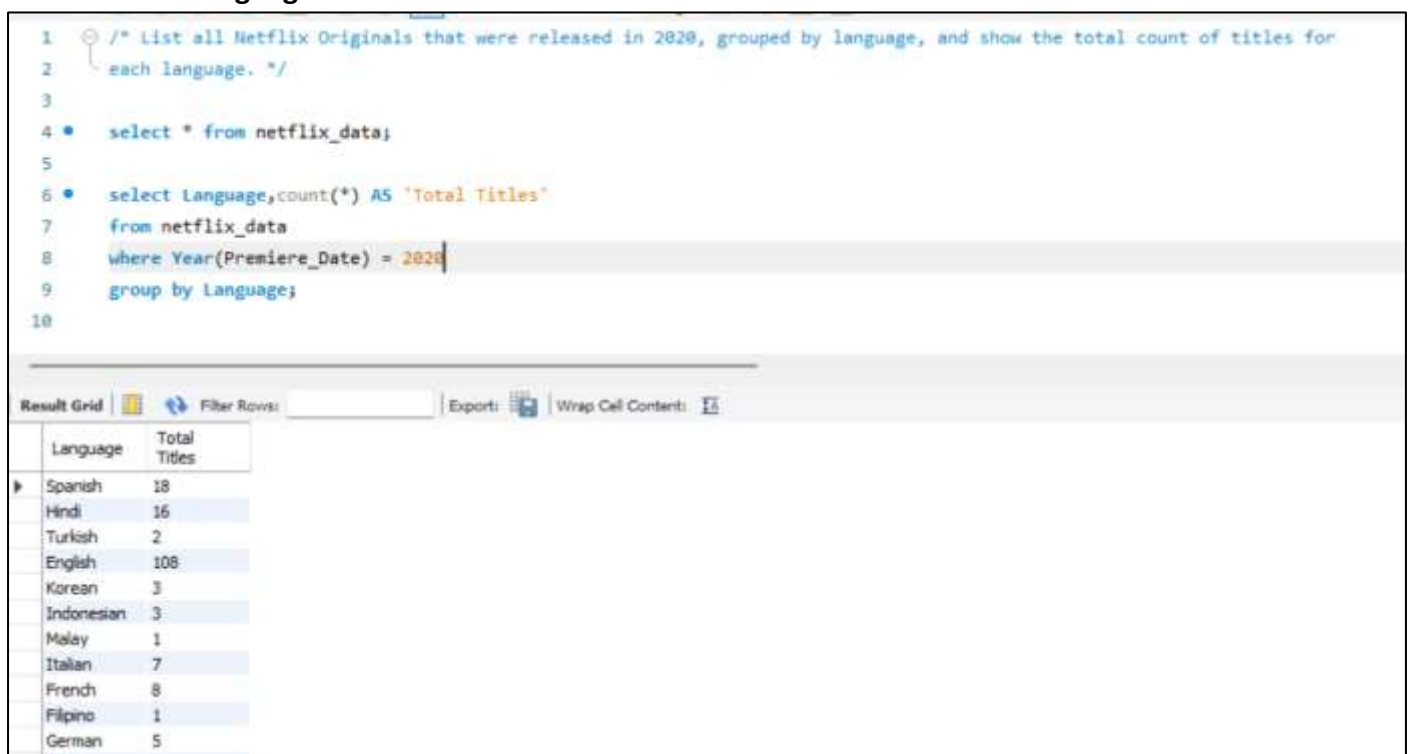


```
1  /* Retrieve the top 5 most common runtimes for Netflix Originals */
2
3  •  SELECT * FROM netflix_data;
4
5  •  SELECT Runtime, COUNT(*) AS Occurrences
6     FROM netflix_data
7     GROUP BY Runtime
8     ORDER BY Occurrences DESC
9     LIMIT 5;
```

Runtime	Occurrences
97	24
94	19
98	19
95	18
100	17

Task 9:

List all Netflix Originals that were released in 2020, grouped by language, and show the total count of titles for each language.



```
1  /* List all Netflix Originals that were released in 2020, grouped by language, and show the total count of titles for
2     each language. */
3
4  •  select * from netflix_data;
5
6  •  select Language, count(*) AS 'Total Titles'
7     from netflix_data
8     where Year(Premiere_Date) = 2020
9     group by Language;
10
```

Language	Total Titles
Spanish	18
Hindi	16
Turkish	2
English	108
Korean	3
Indonesian	3
Malay	1
Italian	7
French	8
Filipino	1
German	5

Task 10:

Create 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.

```
1  /* Create a new table that enforces a constraint on the IMDb score to be between 0 and 10 and the runtime to
2  be greater than 30 minutes */
3
4  CREATE TABLE Netflix_Originals_Valid (
5      Title VARCHAR(255) NOT NULL,
6      GenreID INT NOT NULL,
7      Runtime INT NOT NULL CHECK (Runtime > 30),
8      IMDBScore DECIMAL(3,1) NOT NULL CHECK (IMDBScore BETWEEN 0 AND 10),
9      Language VARCHAR(50) NOT NULL,
10     Premiere_Date DATE NOT NULL
11 );
12
13 select * from netflix_originals_valid;
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content: |

Title	GenreID	Runtime	IMDBScore	Language	Premiere_Date
-------	---------	---------	-----------	----------	---------------