TOP SPOOKY MOVIES
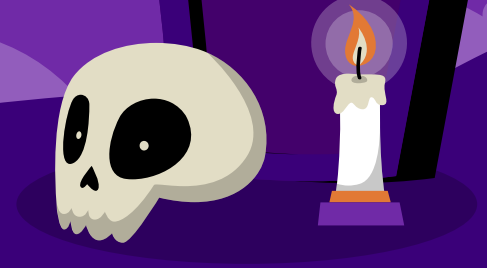
By Marcelino Gamino

My LinkedIn

SPOOKY MOVIE ANALYSIS

Are you having a hard time trying to decide where to begin your spooky movie binge this October? If this sounds like you, this analysis will help point you in the right direction by listing the top movies categorized under the "Horror" genre.

The data used in this analysis is from the IMDB Movies Dataset. Though it includes the top 1000 movies based on IMDB rating, only the movies that fall under the "Horror" genre will be considered.

*NOTE
- This dataset only contains movies
released between 1920-2020

TABLE OF CONTENTS

01

EXCEL ANALYSIS

THE DATA

Here is a snip of what the data we'll be working with looks like.

We will only be focusing on the columns highlighted in blue throughout this analysis.

- Series_Title - Name of the movie
- Released_Year - Year at which that movie released
- Genre - Genre of the movie
- IMDB_Rating - Rating of the movie at IMDB site
- Meta_score - Score earned by the movie
- No_of_votes - Total number of votes

imdb_top_1000

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
1	Poster_Link	Series_Title	Released_Year	Certificate	Runtime	Genre	IMDB_Rating	Overview	Meta_score	Director	Star1	Star2	Star3	Star4	No_of_Votes	Gross
2	https://m.med	The Shawshank Redemption	1994	A	142 min	Drama	9.3	Two imprison	80	Frank Darabont	Tim Robbins	Morgan Fr	Bob Gunte	William Sa	2343110	28,341,469
3	https://m.med	The Godfather	1972	A	175 min	Crime, Drama	9.2	An organized	100	Francis Ford Coppol	Marlon Brando	Al Pacino	James Caa	Diane Kea	1620367	134,966,411
4	https://m.med	The Dark Knight	2008	UA	152 min	Action, Crime, Drama	9	When the me	84	Christopher Nolan	Christian Bale	Heath Led	Aaron Eck	Michael C	2303232	534,858,444
5	https://m.med	The Godfather: Part II	1974	A	202 min	Crime, Drama	9	The early life	90	Francis Ford Coppol	Al Pacino	Robert De	Robert Du	Diane Kea	1129952	57,300,000
6	https://m.med	12 Angry Men	1957	U	96 min	Crime, Drama	9	A jury holdou	96	Sidney Lumet	Henry Fonda	Lee J. Cob	Martin Ba	John Fiedl	689845	4,360,000
7	https://m.med	The Lord of the Rings: The Return of the Kin	2003	U	201 min	Action, Adventure, Drama	8.9	Gandalf and	94	Peter Jackson	Elijah Wood	Viggo Mor	Ian McKell	Orlando B	1642758	377,845,905
8	https://m.med	Pulp Fiction	1994	A	154 min	Crime, Drama	8.9	The lives of tw	94	Quentin Tarantino	John Travolta	Uma Thur	Samuel L.	Bruce Will	1826188	107,928,762
9	https://m.med	Schindler's List	1993	A	195 min	Biography, Drama, History	8.9	In German-oc	94	Steven Spielberg	Liam Neeson	Ralph Fien	Ben Kingsl	Caroline G	1213505	96,898,818
10	https://m.med	Inception	2010	UA	148 min	Action, Adventure, Sci-Fi	8.8	A thief who s	74	Christopher Nolan	Leonardo DiCaprio	Joseph Go	Elliot Page	Ken Watar	2067042	292,576,195
11	https://m.med	Fight Club	1999	A	139 min	Drama	8.8	An insomniac	66	David Fincher	Brad Pitt	Edward N	Meat Loaf	Zach Gren	1854740	37,030,102

HORROR MOVIES

The first thing we want to do is to make a new worksheet with data relevant only to movies that fall under the "Horror" genre.

To do this, we want to focus on the "F" column since it contains the genres we want to go through and the "B" column since it contains the movie titles we want returned when conditions are met.

F
Genre
Drama
Crime, Drama
Action, Crime, Drama
Crime, Drama
Crime, Drama
Action, Adventure, Drama
Crime, Drama
Biography, Drama, History
Action, Adventure, Sci-Fi
Drama

B
Series_Title
The Shawshank Redemption
The Godfather
The Dark Knight
The Godfather: Part II
12 Angry Men
The Lord of the Rings: The Return of the Kin
Pulp Fiction
Schindler's List
Inception
Fight Club

=IF(ISNUMBER(SEARCH("*Horror*", imdb_top_1000!F:F) = TRUE, imdb_top_1000!B:B, "")

HORROR MOVIES CONTINUED

Since it is likely that "Horror" will not be the only genre to describe a movie, the SEARCH() function is used because it can handle wildcards. The function will return a number that signifies the starting location of the word "Horror" if found within the cell.

We then use the ISNUMBER() function which returns "TRUE" if a number is returned to it, which will be dependent on the SEARCH() function.

If "TRUE" is returned, the IF() function is designed to return the matching movie title.

`=IF(ISNUMBER(SEARCH("*Horror*", imdb_top_1000!F:F)) = TRUE, imdb_top_1000!B:B, "")`

	A
	TITLE
2	Psycho
3	The Shining
4	Alien
5	Tumbbad
6	The Thing
7	What Ever Happened to Baby Jane?
8	Das Cabinet des Dr. Caligari
9	The Exorcist
10	Rosemary's Baby

MOVIE DETAILS 1/4

With our new worksheet, which we'll call "Horror_Movies_Detailed", narrowed down to only "Horror" movies, let's look up the rest of the information associated with them.

We do this with the VLOOKUP() function, looking up "Horror" movie titles from our "Horror_Movies_Detailed" worksheet within the original dataset (imdb_top_1000).

Horror_Movies_Detailed

	A
	TITLE
1	
2	Psycho
3	The Shining
4	Alien
5	Tumbbad
6	The Thing
7	What Ever Happened to Baby Jane?
8	Das Cabinet des Dr. Caligari
9	The Exorcist
10	Rosemary's Baby

=VLOOKUP(A2,imdb_top_1000!B:P,2,FALSE)

imdb_top_1000

MOVIE DETAILS 2/4

In this example, the desired column to be returned to our "Horror_Movies_Detailed" worksheet is the "C" column which is represented by the number "2" in the formula.

Once completed, we should have the "Released year" associated with each of our "Horror" movies.

Horror_Movies_Detailed

=VLOOKUP(A2,imdb_top_1000!B:P,2,FALSE)

VLOOKUP() RETURNS

	A	B
1	TITLE	RELEASED
2	Psycho	1960
3	The Shining	1980
4	Alien	1979
5	Tumbbad	2018

imdb_top_1000

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
1	Poster_Link	Series_Title	Released_Year	Certificate	Runtime	Genre	IMDB_Rating	Overview	Meta_score	Director	Star1	Star2	Star3	Star4	No_of_Votes	Gross
2	https://m.med	The Shawshank Redemption	1994	A	142 min	Drama	9.3	Two imprison	80	Frank Darabont	Tim Robbins	Morgan F	Bob Gunt	William Sa	2343110	28,341,469
3	https://m.med	The Godfather	1972	A	175 min	Crime, Drama	9.2	An organized	100	Francis Ford Coppol	Marlon Brando	Al Pacino	James Caa	Diane Kea	1620367	134,966,411
4	https://m.med	The Dark Knight	2008	UA	152 min	Action, Crime, Drama	9	When the me	84	Christopher Nolan	Christian Bale	Heath Led	Aaron Eck	Michael C	2303232	534,858,444
5	https://m.med	The Godfather: Part II	1974	A	202 min	Crime, Drama	9	The early life	90	Francis Ford Coppol	Al Pacino	Robert De	Robert Du	Diane Kea	1129952	57,300,000
6	https://m.med	12 Angry Men	1957	U	96 min	Crime, Drama	9	A jury holdou	96	Sidney Lumet	Henry Fonda	Lee J. Cob	Martin Ba	John Fiedl	689845	4,360,000
7	https://m.med	The Lord of the Rings: The Return of the	in 2003	U	201 min	Action, Adventure, Drama	8.9	Gandalf and /	94	Peter Jackson	Elijah Wood	Viggo Mor	Ian McKell	Orlando B	1642758	377,845,905
8	https://m.med	Pulp Fiction	1994	A	154 min	Crime, Drama	8.9	The lives of t	94	Quentin Tarantino	John Travolta	Uma Thur	Samuel L.	Bruce Will	1826188	107,928,762
9	https://m.med	Schindler's List	1993	A	195 min	Biography, Drama, History	8.9	In German-oc	94	Steven Spielberg	Liam Neeson	Ralph Fien	Ben Kingsl	Caroline G	1213505	96,898,818
10	https://m.med	Inception	2010	UA	148 min	Action, Adventure, Sci-Fi	8.8	A thief who s	74	Christopher Nolan	Leonardo DiCaprio	Joseph Go	Elliot Page	Ken Wata	2067042	292,576,195
11	https://m.med	Fight Club	1999	A	139 min	Drama	8.8	An insomniac	66	David Fincher	Brad Pitt	Edward N	Meat Loaf	Zach Gren	1854740	37,030,102

Since our range is from "B:P", the number "2" represented column "C" in the last example because it is the second column according to that order range.

In this example, we want column "F" returned so the number to represent that column is "5", which returns "Genre."

`=VLOOKUP($A2,imdb_top_1000!$B:$P,5,FALSE)`

imdb_top_1000

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
1	Poster_Link	Series_Title	Released_Year	Certificate	Runtime	Genre	IMDB_Rating	Overview	Meta_score	Director	Star1	Star2	Star3	Star4	No_of_Votes	Gross
2	https://m.med	The Shawshank Redemption	1994	A	142 min	Drama	9.3	Two imprison	80	Frank Darabont	Tim Robbins	Morgan Fr	Bob Gunt	William Sa	2343110	28,341,469
3	https://m.med	The Godfather	1972	A	175 min	Crime, Drama	9.2	An organized	100	Francis Ford Coppol	Marlon Brando	Al Pacino	James Caa	Diane Kea	1620367	134,966,411
4	https://m.med	The Dark Knight	2008	UA	152 min	Action, Crime, Drama	9	When the me	84	Christopher Nolan	Christian Bale	Heath Led	Aaron Eck	Michael C.	2303232	534,858,444
5	https://m.med	The Godfather: Part II	1974	A	202 min	Crime, Drama	9	The early life	90	Francis Ford Coppol	Al Pacino	Robert De	Robert Du	Diane Kea	1129952	57,300,000
6	https://m.med	12 Angry Men	1957	U	96 min	Crime, Drama	9	A jury holdou	96	Sidney Lumet	Henry Fonda	Lee J. Cob	Martin Ba	John Fiedl	689845	4,360,000
7	https://m.med	The Lord of the Rings: The Return of the Kin	2003	U	201 min	Action, Adventure, Drama	8.9	Gandalf and A	94	Peter Jackson	Elijah Wood	Viggo Mor	Ian McKell	Orlando B	1642758	377,845,905
8	https://m.med	Pulp Fiction	1994	A	154 min	Crime, Drama	8.9	The lives of t	94	Quentin Tarantino	John Travolta	Uma Thur	Samuel L.	Bruce Will	1826188	107,928,762
9	https://m.med	Schindler's List	1993	A	195 min	Biography, Drama, History	8.9	In German-oc	94	Steven Spielberg	Liam Neeson	Ralph Fien	Ben Kingsl	Caroline G	1213505	96,898,818
10	https://m.med	Inception	2010	UA	148 min	Action, Adventure, Sci-Fi	8.8	A thief who s	74	Christopher Nolan	Leonardo DiCaprio	Joseph Go	Elliot Page	Ken Watar	2067042	292,576,195
11	https://m.med	Fight Club	1999	A	139 min	Drama	8.8	An insomniac	66	David Fincher	Brad Pitt	Edward N	Meat Loaf	Zach Gren	1854740	37,030,102

MOVIE DETAILS 4/4

After repeating this process for all the desired columns, our worksheet we will be working with from now on should look like this.

Horror_Movies_Detailed

	A	B	C	D	E	F
1	TITLE	RELEASED	GENRE	IMDB_rating	Meta_score	Total Votes
2	Psycho	1960	Horror, Mystery, Thriller	8.5	97	604211
3	The Shining	1980	Drama, Horror	8.4	66	898237
4	Alien	1979	Horror, Sci-Fi	8.4	89	787806
5	Tumbbad	2018	Drama, Fantasy, Horror	8.3	0	27793
6	The Thing	1982	Horror, Mystery, Sci-Fi	8.1	57	371271

*NOTE
- The total count is 32
- "TITLE" = "Series_Title"
- "RELEASED" = "Released_Year"
- "Total Votes" = "No_of_Votes"

IMDB RATING, META SCORE, AND TOTAL VOTES

So, what do we do with our data of "Horror" movies? We could create a list out of the highest rated movies according to their IMDB rating. But lets also not forget to consider their meta scores and total votes, categories some may prefer over a IMDB rating.

After using conditional formatting to highlight the top 6 items for each of these columns, you'll notice that there are differences that may generate a different list of movies depending on the column of interest.

IMDB_rating	Meta_score	Total Votes
8.5	97	604211
8.4	66	898237
8.4	89	787806
8.3	0	27793
8.1	57	371271
8.1	75	50058
8.1	0	57428
8	81	362393
8	96	193674
8	0	61503
7.9	76	512249
7.9	71	111512
7.9	89	116557
7.9	90	78991
7.9	80	42117
7.9	0	88794
7.8	88	27007
7.8	95	43542
7.8	91	65341
7.7	85	492851
7.7	76	157498
7.7	87	233106
7.7	91	48883
7.7	90	171739
7.7	0	31354
7.7	90	27620
7.7	92	44839
7.7	87	30683
7.6	70	138959
7.6	46	379020
7.6	73	376853
7.6	74	337651

TOP HORROR MOVIES

Since the three columns may produce lists that differ from one another, lets go ahead and consider them all.

With a Pivot Table we can determine the top 6–7 movies according to the "IMDB rating", "Meta score", and "Total votes" columns.

IMDB Rating

Top Movies	IMDB Rating
Psycho	8.5
Alien	8.4
The Shining	8.4
Tumbbad	8.3
What Ever Happened to Baby Jane?	8.1
Das Cabinet des Dr. Caligari	8.1
The Thing	8.1

Meta Score

Top Movies	Meta Score
Psycho	97
Rosemary's Baby	96
Bride of Frankenstein	95
Invasion of the Body Snatchers	92
Frankenstein	91
Repulsion	91

Total Votes

Top Movies	Total Votes
The Shining	898237
Alien	787806
Psycho	604211
Shaun of the Dead	512249
Get Out	492851
Saw	379020

GENRES IN DETAIL

At this point, it's possible that you can decide on what to watch from those three lists, but let's see if we can make things a little more interesting.

So, let's go ahead and point our attention towards the other genres that are associated with these "Horror" movies.

With the TEXTSPLIT() function, we can separate the genres within a single cell into new columns with the "," delimiter.

C
GENRE
Horror, Mystery, Thriller
Drama, Horror
Horror, Sci-Fi
Drama, Fantasy, Horror
Horror, Mystery, Sci-Fi
Drama, Horror, Thriller
Fantasy, Horror, Mystery
Horror
Drama, Horror

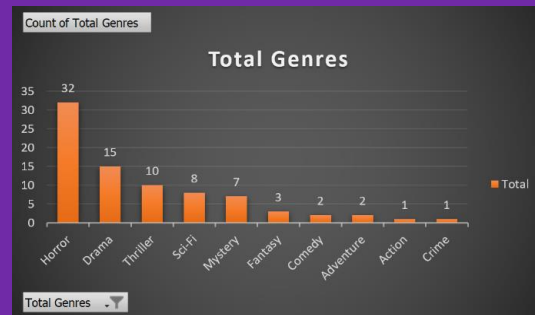| C	D	E	F
GENRE	GENRE_1	GENRE_2	GENRE_3
Horror, Mystery, Thriller	Horror	Mystery	Thriller
Drama, Horror	Drama	Horror	
Horror, Sci-Fi	Horror	Sci-Fi	
Drama, Fantasy, Horror	Drama	Fantasy	Horror
Horror, Mystery, Sci-Fi	Horror	Mystery	Sci-Fi
Drama, Horror, Thriller	Drama	Horror	Thriller
Fantasy, Horror, Mystery	Fantasy	Horror	Mystery
Horror	Horror		
Drama, Horror	Drama	Horror	

GENRES IN DETAIL CONTINUED

We can then copy and paste these separated genres into a single column.

We do this so that we can count all the genres and see which ones associate themselves the most with these "Horror" movies.

We can accomplish this with a PivotTable, totaling the genre count and revealing the top three associated genres.

	A
1	Total Genres
2	Horror
3	Drama
4	Horror
5	Drama
6	Horror
7	Drama
8	Fantasy

Genres	Count of Total Genres
Horror	32
Drama	15
Thriller	10
Sci-Fi	8
Mystery	7
Fantasy	3
Comedy	2
Adventure	2
Action	1
Crime	1

HORROR/DRAMA

We find that "Drama" is one of the genres highly associated with "Horror" when it comes to this dataset of movies.

To make our analysis a little more interesting, let's consider those who may want a "Horror/Drama" paired movie.

Like our very first step of finding all the "Horror" movies from the original dataset (imdb_top_1000), we will use our IF(ISNUMBER(SEARCH())) formula. Except this time, we will look for the "Drama" genre within our "Horror_Movies_Detailed" worksheet instead since it's already filtered for only "Horror" movies.

Horror_Movies_Detailed

	A	B	C	D	E	F
1	TITLE	RELEASED	GENRE	IMDB_rating	Meta_score	Total Votes
2	Psycho	1960	Horror, Mystery, Thriller	8.5	97	604211
3	The Shining	1980	Drama, Horror	8.4	66	898237
4	Alien	1979	Horror, Sci-Fi	8.4	89	787806
5	Tumbbad	2018	Drama, Fantasy, Horror	8.3	0	27793
6	The Thing	1982	Horror, Mystery, Sci-Fi	8.1	57	371271

=IF(ISNUMBER(SEARCH("*Drama*", Horror_Movies_Detailed C2:C33)) = TRUE, Horror_Movies_Detailed A2:A33, "")

HORROR/DRAMA CONTINUED

Once this is completed along with our VLOOKUP() steps, this new worksheet, which we will call "Drama_Movies", should look like this.

We can repeat this for the other top genres, "Sci-Fi" and "Thriller."

Drama_Movies

	A	B	C	D	E
1	Drama_Movies	RELEASED	IMDB_rating	Meta_score	Total Votes
3	The Shining	1980	8.4	66	898237
5	Tumbbad	2018	8.3	0	27793
7	What Ever Happened to Baby Jane?	1962	8.1	75	50058
10	Rosemary's Baby	1968	8	96	193674
11	Les diaboliques	1955	8	0	61503
16	Freaks	1932	7.9	80	42117

*NOTE
- We failed to include the "Genre" column in this worksheet to verify whether all these titles are truly characterized by both "Horror" and "Drama" but we will confirm this later with SQL.

TOP HORROR/DRAMA MOVIES

Just like we did to find the "Top Horror Movies" within our "Horror_Movies_Detailed" worksheet, let's do the same for our "Drama_Movies" worksheet.

IMDB Rating

Top Movies	IMDB Rating
The Shining	8.4
Tumbbad	8.3
What Ever Happened to Baby Jane?	8.1
Les diaboliques	8
Rosemary's Baby	8
Freaks	7.9

Meta Score

Top Movies	Meta Score
Rosemary's Baby	96
Bride of Frankenstein	95
Invasion of the Body Snatchers	92
Frankenstein	91
Repulsion	91
The Birds	90
Les yeux sans visage	90

Total Votes

Top Movies	Total Votes
The Shining	898237
28 Days Later...	376853
Rosemary's Baby	193674
The Birds	171739
La piel que habito	138959
Frankenstein	65341

TOP HORROR/SCI-FI MOVIES

IMDB Rating

Top Movies	IMDB Rating
Alien	8.4
The Thing	8.1
King Kong	7.9
Bride of Frankenstein	7.8
Frankenstein	7.8
The Invisible Man	7.7
Invasion of the Body Snatchers	7.7

Meta Score

Top Movies	Meta Score
Bride of Frankenstein	95
Invasion of the Body Snatchers	92
Frankenstein	91
King Kong	90
Alien	89
The Invisible Man	87

Total Votes

Top Movies	Total Votes
Alien	787806
28 Days Later...	376853
The Thing	371271
King Kong	78991
Frankenstein	65341
Invasion of the Body Snatchers	44839

TOP HORROR/THRILLER MOVIES

IMDB Rating

Top Movies	IMDB Rating
Psycho	8.5
What Ever Happened to Baby Jane?	8.1
Night of the Living Dead	7.9
Get Out	7.7
Repulsion	7.7
Halloween	7.7
Peeping Tom	7.7

Meta Score

Top Movies	Meta Score
Psycho	97
Repulsion	91
Night of the Living Dead	89
Halloween	87
Get Out	85
What Ever Happened to Baby Jane?	75

Total Votes

Top Movies	Total Votes
Psycho	604211
Get Out	492851
Saw	379020
The Others	337651
Halloween	233106
La piel que habito	138959

VERIFICATION WITH SQL

02

DOUBLE CHECKING OUR WORK

So, we went ahead and created a total of 12 different lists for you to choose from. Depending on what columns and genres are most important to you, whether that be preferring "IMDB rating" over "Total votes" or a "Horror/Sci-FI" movie over a general "Horror" movie, you should be able to choose something right for you.

Still, that is a lot of lists to consider and movies to choose from. This analysis is supposed to make it nice and simple for you to decide. To make this happen, we will narrow down our choices even more with SQL but before we do that, let's verify and confirm our Excel work by replicating it with SQL.

OUR SETUP

For this part of our analysis, a *Cleaned version of the original "imdb_top_1000" dataset is used, containing only the columns used in our Excel analysis.

We are also running our queries with PostgreSQL in pgAdmin.

	A	B	C	D	E	F
1	Series_Title	Released_Year	Genre	IMDB_Rating	Meta_score	No_of_Votes
2	The Shawshank R	1994	Drama	9.3	80	2343110
3	The Godfather	1972	Crime, Drama	9.2	100	1620367
4	The Dark Knight	2008	Action, Crime, Drama	9	84	2303232
5	The Godfather: Pa	1974	Crime, Drama	9	90	1129952
6	12 Angry Men	1957	Crime, Drama	9	96	689845
7	The Lord of the Ri	2003	Action, Adventure, Dram	8.9	94	1642758
8	Pulp Fiction	1994	Crime, Drama	8.9	94	1826188
9	Schindler's List	1993	Biography, Drama, Histor	8.9	94	1213505
10	Inception	2010	Action, Adventure, Sci-Fi	8.8	74	2067042
11	Fight Club	1999	Drama	8.8	66	1854740

FIRST CONFIRMATION

If you recall, we had to rely on a few fancy formulas and jumped back and forth between worksheets during our Excel analysis to produce our first list of "Top Horror Movies" that considered the IMDB rating column.

With this query, the same results are produced with only the order varying a bit.

```sql
1  SELECT * FROM movies
2  WHERE "Genre" LIKE '%Horror%'
3  AND "IMDB_Rating" >= (SELECT "IMDB_Rating" FROM movies
4  WHERE "Genre" LIKE '%Horror%'
5  AND "IMDB_Rating" IS NOT NULL
6  ORDER BY "IMDB_Rating" DESC
7  LIMIT 1 OFFSET 5)
8  ORDER BY "IMDB_Rating" DESC, "Series_Title"
```

	Series_Title text	Released_Year text	Genre text	IMDB_Rating numeric	Meta_score integer	No_of_Votes integer
1	Psycho	1960	Horror, Mystery, Thriller	8.5	97	604211
2	Alien	1979	Horror, Sci-Fi	8.4	89	787806
3	The Shining	1980	Drama, Horror	8.4	66	898237
4	Tumbbad	2018	Drama, Fantasy, Horror	8.3	[null]	27793
5	Das Cabinet des Dr. Caligari	1920	Fantasy, Horror, Myste…	8.1	[null]	57428
6	The Thing	1982	Horror, Mystery, Sci-Fi	8.1	57	371271
7	What Ever Happened to Baby Jane?	1962	Drama, Horror, Thriller	8.1	75	50058

Top Movies	IMDB Rating
Psycho	8.5
Alien	8.4
The Shining	8.4
Tumbbad	8.3
What Ever Happened to Baby Jane?	8.1
Das Cabinet des Dr. Caligari	8.1
The Thing	8.1

THE FIRST QUERY

The key to this query is to order our subquery results from highest to lowest and to find the value of the item that is in the 6th position of that order.

That value is then returned to the main query so that it can return values that are greater or equal to that 6th position value.

```
1  SELECT * FROM movies
2  WHERE "Genre" LIKE '%Horror%'
3  AND "IMDB_Rating" >= (SELECT "IMDB_Rating" FROM movies
4  WHERE "Genre" LIKE '%Horror%'
5  AND "IMDB_Rating" IS NOT NULL
6  ORDER BY "IMDB_Rating" DESC
7  LIMIT 1 OFFSET 5)
8  ORDER BY "IMDB_Rating" DESC, "Series_Title"
```

THE FIRST QUERY CONTINUED

The reason this is a key part to the query is because we originally intended on creating a "Top 6" movie list, but what if the movie in the 6th position of that list shared the same value with the movie in the 7th position?

This query takes that into consideration, returning all the movie titles that share that 6th position value if present, along with everything greater than it.

Excel performs this logic automatically when we used conditional formatting.

	Series_Title text	Released_Year text	Genre text	IMDB_Rating numeric
1	Psycho	1960	Horror, Mystery, Thriller	8.5
2	Alien	1979	Horror, Sci-Fi	8.4
3	The Shining	1980	Drama, Horror	8.4
4	Tumbbad	2018	Drama, Fantasy, Horror	8.3
5	Das Cabinet des Dr. Caligari	1920	Fantasy, Horror, Myste…	8.1
6	The Thing	1982	Horror, Mystery, Sci-Fi	8.1
7	What Ever Happened to Baby Jane?	1962	Drama, Horror, Thriller	8.1

GENRE CONFIRMATION

Like the TEXTSPLIT() function we used in our Excel analysis to separate the genres, the STRING_TO_ARRAY() function will make an array out of all the genres in a single field and then the unnest() function will transfer those individual values into a single row.

From there, we can group all the genres to count and total them.

Remember that we only want the top three genres other than "Horror", so we'll skip the highest on the list, since it is "Horror."

```
1  SELECT unnest(STRING_TO_ARRAY("Genre", ', ')) AS genres,
2  COUNT(*) AS total
3  FROM movies
4  WHERE "Genre" LIKE '%Horror%'
5  GROUP BY genres
6  ORDER BY total DESC
7  LIMIT 3 OFFSET 1
```

	genres text	total bigint
1	Drama	15
2	Thriller	10
3	Sci-Fi	8

Genres	Count of Total Genres
Horror	32
Drama	15
Thriller	10
Sci-Fi	8
Mystery	7
Fantasy	3
Comedy	2
Adventure	2
Action	1
Crime	1

HORROR/GENRE/COLUMN QUERY

Now that we have confirmed the top genres associated with "Horror", lets include them in our query. Essentially, the same query is used throughout the verification. It all depends on what columns and genres you're interested in returning.

"Horror/Sci-Fi/No_of_Votes" query

```
1   SELECT * FROM movies
2   WHERE "Genre" LIKE '%Horror%'
3   AND "Genre" LIKE '%Sci-Fi%'
4   AND "No_of_Votes" >= (SELECT "No_of_Votes" FROM movies
5   WHERE "Genre" LIKE '%Horror%'
6   AND "Genre" LIKE '%Sci-Fi%'
7   AND "No_of_Votes" IS NOT NULL
8   ORDER BY "No_of_Votes" DESC
9   LIMIT 1 OFFSET 5)
10  ORDER BY "No_of_Votes" DESC, "Series_Title"
```

"Horror/Thriller/Meta_score" query

```
1   SELECT * FROM movies
2   WHERE "Genre" LIKE '%Horror%'
3   AND "Genre" LIKE '%Thriller%'
4   AND "Meta_score" >= (SELECT "Meta_score" FROM movies
5   WHERE "Genre" LIKE '%Horror%'
6   AND "Genre" LIKE '%Thriller%'
7   AND "Meta_score" IS NOT NULL
8   ORDER BY "Meta_score" DESC
9   LIMIT 1 OFFSET 5)
10  ORDER BY "Meta_score" DESC, "Series_Title"
```

MORE CONFIRMATIONS

```
1  SELECT * FROM movies
2  WHERE "Genre" LIKE '%Horror%'
3  AND "Genre" LIKE '%Sci-Fi%'
4  AND "No_of_Votes" >= (SELECT "No_of_Votes" FROM movies
5  WHERE "Genre" LIKE '%Horror%'
6  AND "Genre" LIKE '%Sci-Fi%'
7  AND "No_of_Votes" IS NOT NULL
8  ORDER BY "No_of_Votes" DESC
9  LIMIT 1 OFFSET 5)
10 ORDER BY "No_of_Votes" DESC, "Series_Title"
```

	Series_Title text	Released_Year text	Genre text	IMDB_Rating numeric	Meta_score integer	No_of_Votes integer
1	Alien	1979	Horror, Sci-Fi	8.4	89	787806
2	28 Days Later...	2002	Drama, Horror, Sci-Fi	7.6	73	376853
3	The Thing	1982	Horror, Mystery, Sci-Fi	8.1	57	371271
4	King Kong	1933	Adventure, Horror, Sci-Fi	7.9	90	78991
5	Frankenstein	1931	Drama, Horror, Sci-Fi	7.8	91	65341
6	Invasion of the Body Snatchers	1956	Drama, Horror, Sci-Fi	7.7	92	44839

Top Movies	Total Votes
Alien	787806
28 Days Later...	376853
The Thing	371271
King Kong	78991
Frankenstein	65341
Invasion of the Body Snatchers	44839

"Horror/Sci-Fi/No_of_Votes" query

```
1  SELECT * FROM movies
2  WHERE "Genre" LIKE '%Horror%'
3  AND "Genre" LIKE '%Thriller%'
4  AND "Meta_score" >= (SELECT "Meta_score" FROM movies
5  WHERE "Genre" LIKE '%Horror%'
6  AND "Genre" LIKE '%Thriller%'
7  AND "Meta_score" IS NOT NULL
8  ORDER BY "Meta_score" DESC
9  LIMIT 1 OFFSET 5)
10 ORDER BY "Meta_score" DESC, "Series_Title"
```

	Series_Title text	Released_Year text	Genre text	IMDB_Rating numeric	Meta_score integer	No_of_Votes integer
1	Psycho	1960	Horror, Mystery, Thriller	8.5	97	604211
2	Repulsion	1965	Drama, Horror, Thriller	7.7	91	48883
3	Night of the Living Dead	1968	Horror, Thriller	7.9	89	116557
4	Halloween	1978	Horror, Thriller	7.7	87	233106
5	Get Out	2017	Horror, Mystery, Thriller	7.7	85	492851
6	What Ever Happened to Baby Jane?	1962	Drama, Horror, Thriller	8.1	75	50058

Top Movies	Meta Score
Psycho	97
Repulsion	91
Night of the Living Dead	89
Halloween	87
Get Out	85
What Ever Happened to Baby Jane?	75

"Horror/Thriller/Meta_score" query

03
FURTHER ANALYSIS WITH SQL

FINAL STAGE

Now that we have verified our work with SQL, we can finally take our analysis to the next level and narrow down our choices to find the perfect spooky movie.

Top Movies	Meta Score
Bride of Frankenstein	95
Invasion of the Body Snatchers	92
Frankenstein	91
King Kong	90
Alien	89
...le Man	87

Movies	Meta Score
...o	97
...nary's Baby	96
...of Frankenstein	95

Top Movies	Total Votes
...e Alien	787806
	376853
	371271
	78991
	65341
...S	44839

Top Movies	
Rosemary's Baby	
Bride of Frankenstein	
Invasion of the Body Snatchers	

Top Movies	IMDB Rating
The Shining	8.4
Tumbbad	8.3
What Ever Happened to Baby Jane?	8.1
Les diaboliques	8
Rosemary's Baby	8
Freaks	7.9

Movies	Total Votes
...o	
...sion	
...of the Living Dead	
...ween	
...ut	
...Ever Happened to Baby	

Top Movies	Total Votes
Psycho	604211
Get Out	492851
Saw	379020
The Others	337651
Halloween	233106
La piel que habito	138959

Top	IMDB Rating
Psyc...	
What...	
Night...	
Get Out	

IMDB Rating	
8.4	
8.1	
7.9	
7.8	
7.8	
7.7	
7.7	

7.7	
7.7	
7.7	

Top Movies	IMDB Rating
Psycho	8.5
Alien	8.4
The Shining	8.4
Tumbbad	8.3
What Ever Happened to Baby Jane?	8.1
Das Cabinet des Dr. Caligari	8.1
The Thing	8.1

...ankenstein	
...ein	
...le Man	
...the Body Snatchers	

I Votes	
898237	
787806	
604211	
512249	
492851	
379020	

Votes	
898237	
376853	
...aby 193674	
171739	
...abito 138959	
65341	

COMMONALITIES

How will we find this perfect spooky movie?

You may or may have not noticed but the lists we've created share some common movie titles. We will narrow our results to these common titles that the lists share and determine them as the best choices for your next spooky movie.

Top Horror Movies

IMDB Rating

Top Movies	IMDB Rating
Psycho	8.5
Alien	8.4
The Shining	8.4
Tumbbad	8.3
What Ever Happened to Baby Jane?	8.1
Das Cabinet des Dr. Caligari	8.1
The Thing	8.1

Meta Score

Top Movies	Meta Score
Psycho	97
Rosemary's Baby	96
Bride of Frankenstein	95
Invasion of the Body Snatchers	92
Frankenstein	91
Repulsion	91

Total Votes

Top Movies	Total Votes
The Shining	898237
Alien	787806
Psycho	604211
Shaun of the Dead	512249
Get Out	492851
Saw	379020

COMMON "TOP HORROR MOVIES" QUERY

To make things a little more organized, three CTEs were created from the queries used to verify our Excel work.

These CTEs represent each list created from our "Top Horror Movies" lists and will be used in a main query to find common movie titles between them all.

IMDB Rating CTE

```
1   WITH IMDB_Rating AS (SELECT * FROM movies
2   WHERE "Genre" LIKE '%Horror%'
3   AND "IMDB_Rating" >= (SELECT "IMDB_Rating" FROM movies
4   WHERE "Genre" LIKE '%Horror%'
5   AND "IMDB_Rating" IS NOT NULL
6   ORDER BY "IMDB_Rating" DESC
7   LIMIT 1 OFFSET 5)
8   ORDER BY "IMDB_Rating" DESC, "Series_Title"
```

Meta Score CTE

```
10  Meta_Score AS (SELECT * FROM movies
11  WHERE "Genre" LIKE '%Horror%'
12  AND "Meta_score" >= (SELECT "Meta_score" FROM movies
13  WHERE "Genre" LIKE '%Horror%'
14  AND "Meta_score" IS NOT NULL
15  ORDER BY "Meta_score" DESC
16  LIMIT 1 OFFSET 5)
17  ORDER BY "Meta_score" DESC, "Series_Title"
```

Number of Votes CTE

```
19  Number_of_Votes AS (SELECT * FROM movies
20  WHERE "Genre" LIKE '%Horror%'
21  AND "No_of_Votes" >= (SELECT "No_of_Votes" FROM movies
22  WHERE "Genre" LIKE '%Horror%'
23  AND "No_of_Votes" IS NOT NULL
24  ORDER BY "No_of_Votes" DESC
25  LIMIT 1 OFFSET 5)
26  ORDER BY "No_of_Votes" DESC, "Series_Title"
27  )
```

Main Query

```
29  SELECT * FROM movies
30  WHERE "Genre" LIKE '%Horror%'
31  AND "Series_Title" IN (Select "Series_Title" FROM IMDB_Rating)
32  AND "Series_Title" IN (Select "Series_Title" FROM Meta_Score)
33  AND "Series_Title" IN (Select "Series_Title" FROM Number_of_Votes)
```

COMMON "TOP HORROR MOVIES" QUERY RESULTS

```sql
1   WITH IMDB_Rating AS (SELECT * FROM movies
2   WHERE "Genre" LIKE '%Horror%'
3   AND "IMDB_Rating" >= (SELECT "IMDB_Rating" FROM movies
4   WHERE "Genre" LIKE '%Horror%'
5   AND "IMDB_Rating" IS NOT NULL
6   ORDER BY "IMDB_Rating" DESC
7   LIMIT 1 OFFSET 5)
8   ORDER BY "IMDB_Rating" DESC, "Series_Title"
9   ),
10  Meta_Score AS (SELECT * FROM movies
11  WHERE "Genre" LIKE '%Horror%'
12  AND "Meta_score" >= (SELECT "Meta_score" FROM movies
13  WHERE "Genre" LIKE '%Horror%'
14  AND "Meta_score" IS NOT NULL
15  ORDER BY "Meta_score" DESC
16  LIMIT 1 OFFSET 5)
17  ORDER BY "Meta_score" DESC, "Series_Title"
18  ),
19  Number_of_Votes AS (SELECT * FROM movies
20  WHERE "Genre" LIKE '%Horror%'
21  AND "No_of_Votes" >= (SELECT "No_of_Votes" FROM movies
22  WHERE "Genre" LIKE '%Horror%'
23  AND "No_of_Votes" IS NOT NULL
24  ORDER BY "No_of_Votes" DESC
25  LIMIT 1 OFFSET 5)
26  ORDER BY "No_of_Votes" DESC, "Series_Title"
27  )
28
29  SELECT * FROM movies
30  WHERE "Genre" LIKE '%Horror%'
31  AND "Series_Title" IN (Select "Series_Title" FROM IMDB_Rating)
32  AND "Series_Title" IN (Select "Series_Title" FROM Meta_Score)
33  AND "Series_Title" IN (Select "Series_Title" FROM Number_of_Votes)
```

	Series_Title text	Released_Year text	Genre text	IMDB_Rating numeric	Meta_score integer	No_of_Votes integer
1	Psycho	1960	Horror, Mystery, Thriller	8.5	97	604211

IMDB Rating

Top Movies	IMDB Rating
Psycho	8.5
Alien	8.4
The Shining	8.4
Tumbbad	8.3
What Ever Happened to Baby Jane?	8.1
Das Cabinet des Dr. Caligari	8.1
The Thing	8.1

Meta Score

Top Movies	Meta Score
Psycho	97
Rosemary's Baby	96
Bride of Frankenstein	95
Invasion of the Body Snatchers	92
Frankenstein	91
Repulsion	91

Total Votes

Top Movies	Total Votes
The Shining	898237
Alien	787806
Psycho	604211
Shaun of the Dead	512249
Get Out	492851
Saw	379020

COMMON "TOP HORROR/SCI-FI MOVIES" QUERY RESULTS

```
1   WITH Sci_Fi_IMDB_Rating AS (SELECT * FROM movies
2   WHERE "Genre" LIKE '%Horror%'
3   AND "Genre" LIKE '%Sci-Fi%'
4   AND "IMDB_Rating" >= (SELECT "IMDB_Rating" FROM movies
5   WHERE "Genre" LIKE '%Horror%'
6   AND "Genre" LIKE '%Sci-Fi%'
7   AND "IMDB_Rating" IS NOT NULL
8   ORDER BY "IMDB_Rating" DESC
9   LIMIT 1 OFFSET 5)
10  ORDER BY "IMDB_Rating" DESC, "Series_Title"
11  ),
12  Sci_Fi_Meta_Score AS (SELECT * FROM movies
13  WHERE "Genre" LIKE '%Horror%'
14  AND "Genre" LIKE '%Sci-Fi%'
15  AND "Meta_score" >= (SELECT "Meta_score" FROM movies
16  WHERE "Genre" LIKE '%Horror%'
17  AND "Genre" LIKE '%Sci-Fi%'
18  AND "Meta_score" IS NOT NULL
19  ORDER BY "Meta_score" DESC
20  LIMIT 1 OFFSET 5)
21  ORDER BY "Meta_score" DESC, "Series_Title"
22  ),
23  Sci_Fi_Number_of_Votes AS (SELECT * FROM movies
24  WHERE "Genre" LIKE '%Horror%'
25  AND "Genre" LIKE '%Sci-Fi%'
26  AND "No_of_Votes" >= (SELECT "No_of_Votes" FROM movies
27  WHERE "Genre" LIKE '%Horror%'
28  AND "Genre" LIKE '%Sci-Fi%'
29  AND "No_of_Votes" IS NOT NULL
30  ORDER BY "No_of_Votes" DESC
31  LIMIT 1 OFFSET 5)
32  ORDER BY "No_of_Votes" DESC, "Series_Title"
33  )
34
35  SELECT * FROM movies
36  WHERE "Genre" LIKE '%Horror%'
37  AND "Series_Title" IN (Select "Series_Title" FROM Sci_Fi_IMDB_Rating)
38  AND "Series_Title" IN (Select "Series_Title" FROM Sci_Fi_Meta_Score)
39  AND "Series_Title" IN (Select "Series_Title" FROM Sci_Fi_Number_of_Votes)
```

Series_Title text	Released_Year text	Genre text	IMDB_Rating numeric	Meta_score integer	No_of_Votes integer
1	Alien	1979	Horror, Sci-Fi	8.4	89	787806
2	King Kong	1933	Adventure, Horror, Sci-Fi	7.9	90	78991
3	Frankenstein	1931	Drama, Horror, Sci-Fi	7.8	91	65341
4	Invasion of the Body Snatchers	1956	Drama, Horror, Sci-Fi	7.7	92	44839

IMDB Rating

Top Movies	IMDB Rating
Alien	8.4
The Thing	8.1
King Kong	7.9
Bride of Frankenstein	7.8
Frankenstein	7.8
The Invisible Man	7.7
Invasion of the Body Snatchers	7.7

Meta Score

Top Movies	Meta Score
Bride of Frankenstein	95
Invasion of the Body Snatchers	92
Frankenstein	91
King Kong	90
Alien	89
The Invisible Man	87

Total Votes

Top Movies	Total Votes
Alien	787806
28 Days Later...	376853
The Thing	371271
King Kong	78991
Frankenstein	65341
Invasion of the Body Snatchers	44839

OUR RESULTS

Top Horror Movie(s)

	Series_Title text	Released_Year text	Genre text	IMDB_Rating numeric	Meta_score integer	No_of_Votes integer
1	Psycho	1960	Horror, Mystery, Thriller	8.5	97	604211

Top Horror/Drama Movie(s)

	Series_Title text	Released_Year text	Genre text	IMDB_Rating numeric	Meta_score integer	No_of_Votes integer
1	Rosemary's Baby	1968	Drama, Horror	8	96	193674

Top Horror/Sci-Fi Movie(s)

	Series_Title text	Released_Year text	Genre text	IMDB_Rating numeric	Meta_score integer	No_of_Votes integer
1	Alien	1979	Horror, Sci-Fi	8.4	89	787806
2	King Kong	1933	Adventure, Horror, Sci-Fi	7.9	90	78991
3	Frankenstein	1931	Drama, Horror, Sci-Fi	7.8	91	65341
4	Invasion of the Body Snatchers	1956	Drama, Horror, Sci-Fi	7.7	92	44839

Top Horror/Thriller Movie(s)

	Series_Title text	Released_Year text	Genre text	IMDB_Rating numeric	Meta_score integer	No_of_Votes integer
1	Psycho	1960	Horror, Mystery, Thriller	8.5	97	604211
2	Get Out	2017	Horror, Mystery, Thriller	7.7	85	492851
3	Halloween	1978	Horror, Thriller	7.7	87	233106

IN CONCLUSION

So, we began by creating a worksheet of "Horror" movies from a dataset of 1000 top movies that were based on their IMDB rating.

At first, we wanted to find the top "Horror" movies according to their "IMDB rating" but to make things a little interesting, we would also consider their "Meta score" and "Total votes" for those who may value those categories over a "IMDB rating." This resulted in three lists for the "Horror" genre alone.

We then figured out the top three genres that accommodated these "Horror" movies for those who may want a more specific type of movie to watch. After taking the same steps we did for the "Horror" genre alone, we ended up with 12 total lists and many movie choices to choose from.

We did all this in Excel and then validated our work by replicating it with SQL. Once confirmed, we furthered our analysis by finding common movie titles between the lists that fell under their specific genre(s). Doing this narrowed our lists and movie choices to choose from, making it easier for you to decide on your spooky movie.

IN CONCLUSION CONTINUED

If you look at these lists/choices and still need help with your decision, "Psycho" seems to be the top general "Horror" movie according to our analysis and the only movie title shared between it's three lists. If I would have to recommend a movie from these choices, I recommend "Rosemary's Baby" or "Alien." Choose wisely and GET SPOOKED!

Top Horror Movie(s)

	Series_Title text	Released_Year text	Genre text	IMDB_Rating numeric	Meta_score integer	No_of_Votes integer
1	Psycho	1960	Horror, Mystery, Thriller	8.5	97	604211

RESOURCES/LINKS

- IMDB Movies Dataset

- PowerPoint template

- My LinkedIn

THANKS

Are there any questions?

youremail@freepik.com
+91 620 421 838
yourweb.com