

IMDB Hollywood queries

1. The number of science fiction movies over the decades.

- Code

```
select (premiered/10)*10 as decade, count(*) as scifi
from titles
where premiered and genres is not null
and genres like "%Sci-fi%"
group by premiered/10
```

- Result

decade	scifi
1890	1
1910	23
1920	30
1930	69
1940	51
1950	255
1960	353
1970	421
1980	728
1990	791
2000	1149
2010	3424
2020	600

- Analysis

Overall, the number of sci-fi movies increases rapidly over the decades, especially in the 50s, 80s, and 00s. From 1950 to 1960, the end of the WWII gave rise to the birth of apocalyptic outlook of the future, as well as the fear of the alienated communism. Therefore, the number of science fiction movies were five times higher than the previous decade. During the 80s, CGI came out and comic book adaptations became popular in the industries. It was not until 2000 that the technological development remarkably enhanced the production of science fiction movies, especially the application of the 3D technology. In 2020, there are much fewer movies due to the pandemic.

However, regarding the development of the economics over the decade, the number of science fiction movies is also affected by the total number of movies.

- Code: The percentage of science fiction movies over the decade.

```

select sf.decade, scifi, round((scifi*1.0)/(film*1.0)*100, 2) as scifipercen
from (select (premiered/10)*10 as decade, count(*) as scifi
      from titles
      where premiered and genres is not null
      and genres like "%Sci-fi%"
      group by premiered/10) as sf
inner join
      (select (premiered/10)*10 as decade, count(*) as film
      from titles
      where premiered is not null
      group by premiered/10) as film
on sf.decade = film.decade

```

- Result

decade	scifi	scifipercen
1890	1	2.63
1910	23	0.18
1920	30	0.14
1930	69	0.34
1940	51	0.35
1950	255	1.17
1960	353	1.16
1970	421	1.08
1980	728	1.72
1990	791	1.8
2000	1149	1.59
2010	3424	2.22
2020	600	3.11

- Analysis

1890 is the outlier due to a significantly small number of movies produced that year. In general, the percentage of sci-fi movies increased until 1950. It dropped from 1950 to 1970 before notably rose until now. Thus, during the 50s and 60s, although the total number of sci-fi movies increased, the percentage of which decreased. This can be explained by the shift in the movie industry when the New Hollywood era was born. The more realistic world view of the young generation aligned with and captured the interest of the mass audience. The production of non-sci fi movies became more popular. On the other hand, even though the number of sci-fi movies decreased to only 600 in 2020, its percentage rose also due to the pandemic.

2. The average runtime associated to the cinematographer's age.

- Code

```
select (p.died - p.born) as age, round(avg(runtime_minutes), 1) as runtime
from titles t, crew c, people p
where t.title_id = c.title_id
and c.person_id = p.person_id
and c.category = "cinematographer"
and t.runtime_minutes is not null
and p.died > p.born
group by age
```

- Result

age	runtime
21	90.0
23	58.9
24	81.4
25	60.2
26	90.0
27	72.0
28	62.4
29	70.0
30	71.3
31	82.8
32	70.6
33	70.8
34	82.3
35	85.9
36	105.2
37	73.5
38	60.3
39	76.3
40	83.1
41	82.9
42	76.1
43	76.6
44	91.4
45	79.7
46	83.4
47	79.3
48	80.1
49	92.3
50	80.6
51	75.1
52	78.8

53	85.5
54	81.7
55	82.3
56	77.0
57	81.8
58	79.7
59	89.2
60	90.0
61	88.3
62	88.3
63	85.2
64	86.6
65	81.2
66	87.7
67	76.6
68	82.1
69	90.2
70	85.3
71	85.7
72	86.0
73	90.7
74	86.9
75	82.5
76	85.2
77	85.9
78	90.5
79	85.8
80	81.1
81	90.1
82	90.9
83	86.3
84	86.8
85	88.5
86	85.1
87	79.7
88	84.3
89	91.6
90	86.4
91	90.6
92	88.1
93	88.4
94	86.9
95	85.1
96	92.6

97	81.5
98	89.2
99	95.2
100	94.6
101	82.9
102	75.5
103	99.3
104	92.8

- Analysis

There seems to be no relations between the average runtime and the cinematographers' age: as the age of the cinematographer increases, the average runtime scatters randomly. However, the age group with the highest and lowest runtimes are noticeable.

- Code

- Highest runtimes age groups

```
select (p.died - p.born) as age, round(avg(runtime_minutes), 1) as
runtime
from titles t, crew c, people p
where t.title_id = c.title_id
and c.person_id = p.person_id
and c.category = "cinematographer"
and t.runtime_minutes is not null
and p.died > p.born
group by age
order by runtime desc
limit 10
```

- Lowest runtimes age groups

```
select (p.died - p.born) as age, round(avg(runtime_minutes), 1) as
runtime
from titles t, crew c, people p
where t.title_id = c.title_id
and c.person_id = p.person_id
and c.category = "cinematographer"
and t.runtime_minutes is not null
and p.died > p.born
group by age
order by runtime
limit 10
```

- Result

- Highest runtimes age groups

age	runtime
36	105.2
103	99.3
99	95.2
100	94.6
104	92.8
96	92.6
49	92.3
89	91.6
44	91.4
82	90.9

- Lowest runtimes age groups

age	runtime
23	58.9
25	60.2
38	60.3
28	62.4
29	70.0
32	70.6
33	70.8
30	71.3
27	72.0
37	73.5

- Analysis

Unexpectedly, the cinematographers who produce the longest movies are more likely to be in their 90s and 100s. On the other hand, the cinematographers who produce the shortest movies are more likely to be in their 30s. Younger cinematographers nowadays are provided with a large number of streaming services that allow the production of shorter films to reach the mass audience. The older cinematographer's longer movies might be due to the runtimes becoming longer over the decades.

- Code

```
select (premiered/10)*10 as decade,
round(avg(runtime_minutes),1) as runtime
from titles
where runtime_minutes and decade is not null
group by premiered/10
```

- Result

decade runtime

1890	60.3
1900	72.6
1910	64.5
1920	72.2
1930	81.1
1940	86.1
1950	91.6
1960	93.5
1970	90.5
1980	91.9
1990	92.3
2000	89.4
2010	90.0
2020	90.2

- Analysis

As can be seen, movies' runtime gradually increases until 1960 where it stabilizes at about 90 minutes. To explain, the development of the economics helps movie industry produce on a larger scale with smaller budget, aiding by the technological advances.

Therefore, cinematographers who are in their 90s and 100s contribute to the shortest runtimes when they are in their 30s.