

SQL PROJECT

DATA RETRIEVAL

QUESTION 1

- --FILMS ANALYSIS
- --Q1 10 HIGHEST GROSSING FILM AND RELEASE
- --SELECT release_year,
- -- gross
- --FROM films
- --where gross is not null
- --order by gross desc
- --limit 10

Data Output Messages Graph Visualiser X Notifications

	release_year integer	gross bigint
1	2015	936627416
2	2009	760505847
3	1997	658672302
4	2015	652177271
5	2012	623279547
6	2012	623279547
7	2008	533316061
8	1999	474544677
9	1977	460935665
10	2015	458991599

QUESTION 2

- --Q2 HOW MANY FILMS IN THE DATABASE
- -- WERE RELEASE IN EACH COUNTRY,
- --AND WAHT ARE THE TOP 5 COUNTRY
- --SELECT country,
- count(release_year) AS release_count
- --FROM films
- --group by country
- --order by count(release_year) desc
- --limit 5

Data Output			Messages	Graph Visualiser X	Notifications
	country character varying	release_count bigint			
1	USA	3726			
2	UK	434			
3	France	152			
4	Canada	122			
5	Germany	97			

QUESTION 3

- --Q3 HOW MANY FILMS ARE AVAILABLE IN EACH LANGUAGE AND
- --WHAT ARE THE TOP 3 LANGUAGE PRESENTED
- --SELECT language,
- -- count(language)
- --FROM films
- --group by language
- --order by count(language) desc
- --limit 3

Data Output Messages Graph Visualiser X Notifications		
	language character varying	count bigint
1	English	4635
2	French	72
3	Spanish	40

QUESTION 4

- --Q4 WHAT IS THE AVERAGE IMDb SCORE FOR FILMS IN THE DATABASE
- --SELECT avg(imdb_score) as avg_imdb_score
- --FROM reviews

The screenshot shows the 'Data Output' tab of the AWS SageMaker console. The table displays the following data:

	avg_imdb_score double precision
1	6.42648953616907

QUESTION 5

- --Q5 WHICH COUNTRY HAS MADE THE HIGHEST PROFIT FROM MOVIE?
- --SELECT country,
- -- avg(gross)
- --FROM films
- --where gross is not null
- --group by country
- --order by avg(gross) desc
- --limit 1

Data Output

Messages

Graph Visualiser

×

Notifications

≡+

▼

▼

	country character varying	avg numeric
1	New Zealand	108180043.50000000000000

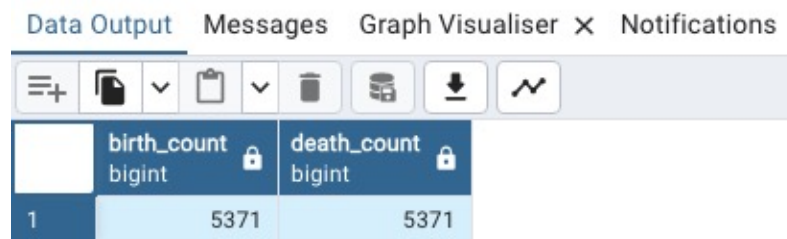
QUESTION 6

- Q6 WHICH MOVIE MADE THE HIGHEST PROFIT IN THE 21ST CENTURY
- --SELECT title, release_year, gross
- --FROM films
- --where release_year >= 2001 and
- -- gross is not null
- --order by gross desc
- --limit 1

Data Output Messages Graph Visualiser X Notifications			
	title character varying	release_year integer	gross bigint
1	Star Wars: Episode VII - The Force Awakens	2015	936627416

QUESTION 7

- --Q7 HOW MANY PEOPLE IN THE DATABASE ARE STILL ALIVE
- --(BASED ON THE BIRTHDATE AND DEATH DATE)
- --select count(birthdate) as birth_count, count(birthdate) as death_count
- --from people
- --where deathdate is null and birthdate is not null



The screenshot shows a database interface with a tab labeled 'Data Output'. Below the tab is a toolbar with icons for file operations and a query icon. The main area displays a table with the following data:

	birth_count bigint	death_count bigint
1	5371	5371

QUESTION 8

- --Q8 WHICH YEAR HAS THE HIGHEST NUMBER OF MOVIE RELEASE
- --select release_year, count(release_year) as number_of_movie
- --from films
- --where release_year is not null
- --group by release_year
- --order by count(release_year) desc
- --limit 1

Data Output Messages Graph Visualiser X Notifications		
≡+	📄	📋
▼	▼	🗑️
🗄️	📥	📈
	release_year integer	number_of_movie bigint
1	2009	260

QUESTION 9

- -- PEOPLE AND ROLE ANALYSIS --
- --Q9 DETERMINE THE TOP 10 PEOPLE WITH THE MOST ROLE IN THE DATABASE
- --select name, role
- --from people
- --full join roles
- --on people.id = roles.id
- --limit 10

Data Output Messages Graph Visualiser X Notifications

	name character varying	role character varying
1	50 Cent	director
2	A. Michael Baldwin	actor
3	A. Raven Cruz	actor
4	A.J. Buckley	actor
5	A.J. DeLucia	director
6	A.J. Langer	actor
7	Aaliyah	actor
8	Aaron Ashmore	actor
9	Aaron Hann	actor
10	Aaron Hill	actor

QUESTION 10

- --USER REVIEWS AND VOTES ANALYSIS
- --Q10 CALCULATE THE AVERAGE NUMBER OF USER AND REVIEWS FOR FILMS
- --SELECT avg(num_user) as average_number, avg (num_votes) as review_number
- --FROM REVIEWS

Data Output Messages Graph Visualiser X Notifications		
≡+	📄	📋
▼	▼	🗑️
🗄️	⬇️	📈
	average_number numeric	review_number numeric
1	275.6999393816932714	84441.991545893720

QUESTION 11

- --Q11 IDENTIFY FILMS WITH THE HIGHEST NUMBER OF USER AND CRITIC REVIEWS
- --SELECT max(num_user) as highest_user, max (num_critic) as highest_critic
- --from reviews
- --where num_user is not null
- --order by max(num_user)

Data Output Messages Graph Visualiser X Notifications			
	highest_user integer	highest_critic integer	
1	5060	828	

QUESTION 12

- --Q12 WHICH FILMS HAVE THE MOST FACEBOOK LIKES AND
- --IS THERE CORRELATION WITH THEIR IMDb SCORE

- --SELECT facebook_likes,(imdb_score)
- --FROM reviews
- --order by facebook_likes desc
- --limit 1

Data Output Messages Graph Visualiser X Notifications		
≡+	📄	📋
▼	▼	🗑️
🗄️	⬇️	📈
	facebook_likes integer	imdb_score real
1	1364147	8.1