Data Engineer

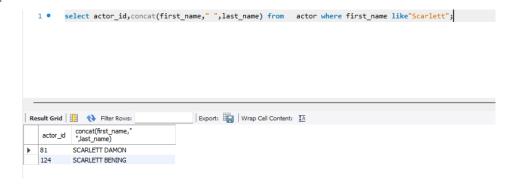## Assignment 2

Under supervisor

ENG: Hazem Omar

Name: Abdullah Ibrahim Mahmoud

Q1) Which actors have the first name 'Scarlett'

Solution

```sql
select actor_id,concat(first_name," ",last_name) from   actor where first_name like"Scarlett";
```

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

actor_id	concat(first_name," ",last_name)
81	SCARLETT DAMON
124	SCARLETT BENING

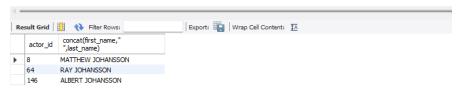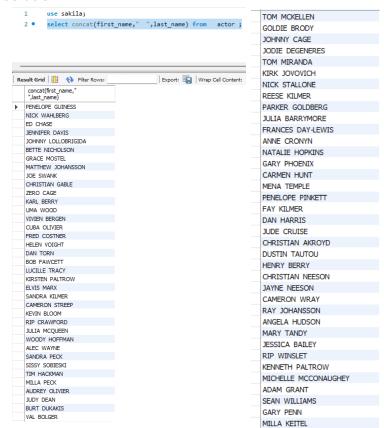## Q2) Which actors have the last name 'Johansson

Solution

```sql
select actor_id,concat(first_name," ",last_name) from   actor where last_name like"Johansson";
```
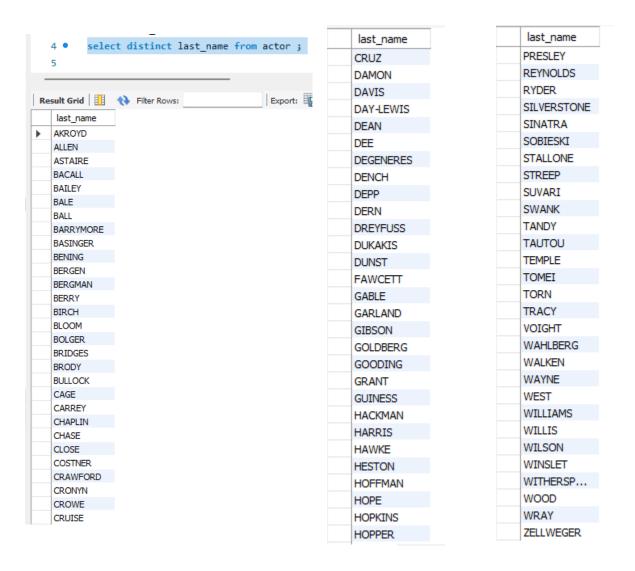
Result Grid | Filter Rows: | Export: | Wrap Cell Content:

actor_id	concat(first_name," ",last_name)
8	MATTHEW JOHANSSON
64	RAY JOHANSSON
146	ALBERT JOHANSSON

Q3) Display the first and last names of all actors from the table actor

Solution

```sql
use sakila;
select concat(first_name,"  ",last_name) from   actor ;
```

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

concat(first_name," ",last_name)
PENELOPE GUINESS
NICK WAHLBERG
ED CHASE
JENNIFER DAVIS
JOHNNY LOLLOBRIGIDA
BETTE NICHOLSON
GRACE MOSTEL
MATTHEW JOHANSSON
JOE SWANK
CHRISTIAN GABLE
ZERO CAGE
KARL BERRY
UMA WOOD
VIVIEN BERGEN
CUBA OLIVIER
FRED COSTNER
HELEN VOIGHT
DAN TORN
BOB FAWCETT
LUCILLE TRACY
KIRSTEN PALTROW
ELVIS MARX
SANDRA KILMER
CAMERON STREEP
KEVIN BLOOM
RIP CRAWFORD
JULIA MCQUEEN
WOODY HOFFMAN
ALEC WAYNE
SANDRA PECK
SISSY SOBIESKI
TIM HACKMAN
MILLA PECK
AUDREY OLIVIER
JUDY DEAN
BURT DUKAKIS
VAL BOLGER
TOM MCKELLEN
GOLDIE BRODY
JOHNNY CAGE
JODIE DEGENERES
TOM MIRANDA
KIRK JOVOVICH
NICK STALLONE
REESE KILMER
PARKER GOLDBERG
JULIA BARRYMORE
FRANCES DAY-LEWIS
ANNE CRONYN
NATALIE HOPKINS
GARY PHOENIX
CARMEN HUNT
MENA TEMPLE
PENELOPE PINKETT
FAY KILMER
DAN HARRIS
JUDE CRUISE
CHRISTIAN AKROYD
DUSTIN TAUTOU
HENRY BERRY
CHRISTIAN NEESON
JAYNE NEESON
CAMERON WRAY
RAY JOHANSSON
ANGELA HUDSON
MARY TANDY
JESSICA BAILEY
RIP WINSLET
KENNETH PALTROW
MICHELLE MCCONAUGHEY
ADAM GRANT
SEAN WILLIAMS
GARY PENN
MILLA KEITEL

Q4) How many distinct actors last names are there?

Solution

```
1    use sakila;
2 •  SELECT COUNT(DISTINCT last_name) AS distinct_last_names FROM actor;
3    |
4
```

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

distinct_last_names
▶ 121

Q5) Which last names are not repeated?

Solution

```
4 •  select distinct last_name from actor ;
5
```

Result Grid | ↻ Filter Rows: | Export:

last_name
▶ AKROYD
ALLEN
ASTAIRE
BACALL
BAILEY
BALE
BALL
BARRYMORE
BASINGER
BENING
BERGEN
BERGMAN
BERRY
BIRCH
BLOOM
BOLGER
BRIDGES
BRODY
BULLOCK
CAGE
CARREY
CHAPLIN
CHASE
CLOSE
COSTNER
CRAWFORD
CRONYN
CROWE
CRUISE

last_name
CRUZ
DAMON
DAVIS
DAY-LEWIS
DEAN
DEE
DEGENERES
DENCH
DEPP
DERN
DREYFUSS
DUKAKIS
DUNST
FAWCETT
GABLE
GARLAND
GIBSON
GOLDBERG
GOODING
GRANT
GUINESS
HACKMAN
HARRIS
HAWKE
HESTON
HOFFMAN
HOPE
HOPKINS
HOPPER

last_name
PRESLEY
REYNOLDS
RYDER
SILVERSTONE
SINATRA
SOBIESKI
STALLONE
STREEP
SUVARI
SWANK
TANDY
TAUTOU
TEMPLE
TOMEI
TORN
TRACY
VOIGHT
WAHLBERG
WALKEN
WAYNE
WEST
WILLIAMS
WILLIS
WILSON
WINSLET
WITHERSP...
WOOD
WRAY
ZELLWEGER

Q6) Write a query to find the full name of the actor who has acted in the maximum number of movies.?
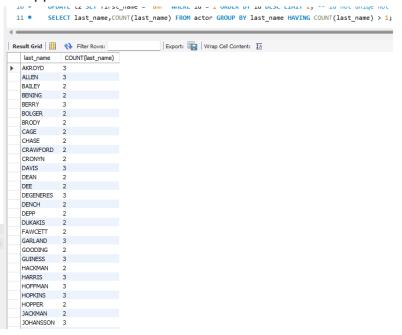
Solution

```
12 •   select concat(a.first_name," " ,a.last_name )as "full name ",COUNT(fc.film_id)from actor AS a
13     INNER join film_actor AS fc
14     using (actor_id )
15     group by a.actor_id
16     order by COUNT(fc.film_id) desc
17     limit 1 ;
```

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

full name	COUNT(fc.film_id)
GINA DEGENERES	42

Q7) Write a query to find the full name of the actor who has acted in the third most number of movies.
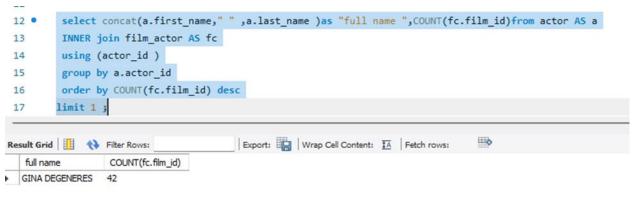
Solution

```
12 •   select concat(a.first_name," " ,a.last_name )as "full name ",COUNT(fc.film_id)from actor AS a
13     INNER join film_actor AS fc
14     using (actor_id )
15     group by a.actor_id
16     order by COUNT(fc.film_id) desc
17     limit 1 offset 2 ;
18
```

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

full name	COUNT(fc.film_id)
MARY KEITEL	40

Q8) . Which last names appear more than once?

Solution

```
10 •   UPDATE t2 SET first_name = aim  WHERE id = 1 ORDER BY id DESC LIMIT 2, -- id not unique not
11 •   SELECT last_name,COUNT(last_name) FROM actor GROUP BY last_name HAVING COUNT(last_name) > 1;
```

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

last_name	COUNT(last_name)
AKROYD	3
ALLEN	3
BAILEY	2
BENING	2
BERRY	3
BOLGER	2
BRODY	2
CAGE	2
CHASE	2
CRAWFORD	2
CRONYN	2
DAVIS	3
DEAN	2
DEE	2
DEGENERES	3
DENCH	2
DEPP	2
DUKAKIS	2
FAWCETT	2
GARLAND	3
GOODING	2
GUINESS	3
HACKMAN	2
HARRIS	3
HOFFMAN	3
HOPKINS	3
HOPPER	2
JACKMAN	2
JOHANSSON	3

Q9). Which actor has appeared in the most films?

Solution

```sql
12 •   select concat(a.first_name," " ,a.last_name )as "full name ",COUNT(fc.film_id)from actor AS a
13     INNER join film_actor AS fc
14     using (actor_id )
15     group by a.actor_id
16     order by COUNT(fc.film_id) desc
17     limit 1 ;
```

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

full name	COUNT(fc.film_id)
GINA DEGENERES	42

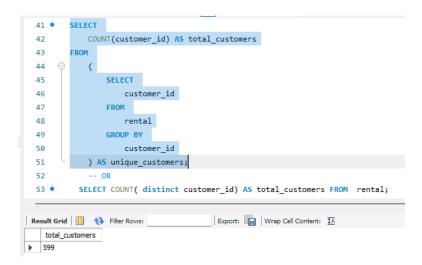## Q10). Write a query to find the full names of those customers who have rented at least one movie and belong to the city Arlington.

Solution

```sql
1 ▾   use sakila;
2 •   SELECT DISTINCT CONCAT(c.first_name, ' ', c.last_name) AS "full_name", count(r.rental_id),ci.city
3     FROM customer as c
4     INNER JOIN address as a
5      ON c.address_id = a.address_id
6     INNER JOIN city as ci
7     ON a.city_id = ci.city_id
8     INNER JOIN rental as r
9     ON c.customer_id = r.customer_id
10    WHERE ci.city = "Arlington"
11    group by c.customer_id
12    having count(r.rental_id)> 1 ;
```

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

full_name	count(r.rental_id)	city
RICHARD MCCRARY	25	Arlington

Q11). Use group by statement to calculate the total customers who have rented in Sakila DB

Solution

```sql
41 •   SELECT
42         COUNT(customer_id) AS total_customers
43     FROM
44     (
45             SELECT
46                 customer_id
47             FROM
48                 rental
49             GROUP BY
50                 customer_id
51     ) AS unique_customers;
52         -- OR
53 •   SELECT COUNT( distinct customer_id) AS total_customers FROM  rental;
```

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

total_customers
599

Q12).. Select all the information from the customer entity.

Solution

```
39 •    SELECT *FROM customer;
```

customer_id	store_id	first_name	last_name	email	address_id	active	create_date	last_update
1	1	MARY	SMITH	MARY.SMITH@sakilacustomer.org	5	1	2006-02-14 22:04:36	2006-02-15 04:57:20
2	1	PATRICIA	JOHNSON	PATRICIA.JOHNSON@sakilacustomer.org	6	1	2006-02-14 22:04:36	2006-02-15 04:57:20
3	1	LINDA	WILLIAMS	LINDA.WILLIAMS@sakilacustomer.org	7	1	2006-02-14 22:04:36	2006-02-15 04:57:20
4	2	BARBARA	JONES	BARBARA.JONES@sakilacustomer.org	8	1	2006-02-14 22:04:36	2006-02-15 04:57:20
5	1	ELIZABETH	BROWN	ELIZABETH.BROWN@sakilacustomer.org	9	1	2006-02-14 22:04:36	2006-02-15 04:57:20
6	2	JENNIFER	DAVIS	JENNIFER.DAVIS@sakilacustomer.org	10	1	2006-02-14 22:04:36	2006-02-15 04:57:20
7	1	MARIA	MILLER	MARIA.MILLER@sakilacustomer.org	11	1	2006-02-14 22:04:36	2006-02-15 04:57:20
8	2	SUSAN	WILSON	SUSAN.WILSON@sakilacustomer.org	12	1	2006-02-14 22:04:36	2006-02-15 04:57:20
9	2	MARGARET	MOORE	MARGARET.MOORE@sakilacustomer.org	13	1	2006-02-14 22:04:36	2006-02-15 04:57:20
10	1	DOROTHY	TAYLOR	DOROTHY.TAYLOR@sakilacustomer.org	14	1	2006-02-14 22:04:36	2006-02-15 04:57:20
11	2	LISA	ANDERSON	LISA.ANDERSON@sakilacustomer.org	15	1	2006-02-14 22:04:36	2006-02-15 04:57:20
12	1	NANCY	THOMAS	NANCY.THOMAS@sakilacustomer.org	16	1	2006-02-14 22:04:36	2006-02-15 04:57:20
13	2	KAREN	JACKSON	KAREN.JACKSON@sakilacustomer.org	17	1	2006-02-14 22:04:36	2006-02-15 04:57:20
14	2	BETTY	WHITE	BETTY.WHITE@sakilacustomer.org	18	1	2006-02-14 22:04:36	2006-02-15 04:57:20
15	1	HELEN	HARRIS	HELEN.HARRIS@sakilacustomer.org	19	1	2006-02-14 22:04:36	2006-02-15 04:57:20
16	2	SANDRA	MARTIN	SANDRA.MARTIN@sakilacustomer.org	20	0	2006-02-14 22:04:36	2006-02-15 04:57:20
17	1	DONNA	THOMPSON	DONNA.THOMPSON@sakilacustomer.org	21	1	2006-02-14 22:04:36	2006-02-15 04:57:20
18	2	CAROL	GARCIA	CAROL.GARCIA@sakilacustomer.org	22	1	2006-02-14 22:04:36	2006-02-15 04:57:20
19	1	RUTH	MARTINEZ	RUTH.MARTINEZ@sakilacustomer.org	23	1	2006-02-14 22:04:36	2006-02-15 04:57:20
20	2	SHARON	ROBINSON	SHARON.ROBINSON@sakilacustomer.org	24	1	2006-02-14 22:04:36	2006-02-15 04:57:20

Q13).. What is the average running time of all the films in the Sakila DB?

Solution

```
40 •    SELECT AVG(length) AS "average_running_time" FROM film;
```

average_running_time
115.2720

Q14). What is the average running time of all the films in the Sakila DB?

Solution

```
40     -- Calculate the average running time of films by category
41 •   SELECT  c.name AS category, AVG(f.length) AS average_running_time
42     FROM film AS f
43     INNER JOIN  film_category AS fc
44     USING(film_id )
45     INNER JOIN category AS c
46      USING (category_id )
47     GROUP BY  c.name;
```

category	average_running_time
Action	111.6094
Animation	111.0152
Children	109.8000
Classics	111.6667
Comedy	115.8276
Documentary	108.7500
Drama	120.8387
Family	114.7826
Foreign	121.6986
Games	127.8361
Horror	112.4821
Music	113.6471
New	111.1270
Sci-Fi	108.1967
Sports	128.2027
Travel	113.3158