SQL Murder Mystery

Module 3: Practice SQL command

ITC 6000

Prepared by: Heejae Roh
Presented to: Professor, Venkata Duvvuri

Date: Dec 9th, 2022

INTRODUCTION

A crime has occurred. We have large amounts of data through SQL. By examining data with the evidence, I will gradually get closer to the answer.

SQL has the advantage of being able to store and share data with others. And the biggest advantage of SQL is that the data can be stored in large quantities. As a result of examining the Murder mystery data, the amount of data is quite large, and it took pretty long time to run it. Let's get used to SQL commands with relatively larger data than the ones we've been dealing with so far.

FAMILIARIZE

Aggregation 1: Count and Simple description

```
SELECT count(*) FROM person;
SELECT count(*) FROM interview;
SELECT count(*) FROM crime_scene_report;
```

	Person	Interview	Crime scene report
counts	10006	4991	1228

```
SELECT membership_status, COUNT(*) FROM get_fit_now_member GROUP BY membership_status;
```

	Gold	Regular	Silver
Membership status	67	53	63

Aggregation 2: What's in the data

```
SELECT * FROM person LIMIT 3;
```

id	name	License_id	Address_number	Address_street_name	ssn
10000	Christoper Peteuil	993845	624	Bankhall Ave	747714076
10007	Kourtney Calderwood	861794	2791	Gustavus Blvd	477972044
10010	Muoi Cary	385336	741	Northwestern Dr	828638512

```
SELECT * FROM crime_scene_report WHERE type="robbery" LIMIT 3;
```

Data	Type	description	city
20180115	robbery	A Man Dressed as Spider-Man Is on a Robbery Spree	NYC
20171110	robbery	The Gjallarhorn shoulder-mounted rocket system was forged from …	SQL City
20170105	robbery	uglifying!' it exclaimed. 'You know what to beautify is, I suppose?'	Seaside

Join and filter tables Example

```
SELECT name, annual_income as income, gender, eye_color as eyes, hair_color as hair
FROM income i JOIN person p ON i.ssn = p.ssn
JOIN drivers_license d ON p.license_id = d.id WHERE annual_income > 450000;
```

Name	Income	Gender	Eyes	hair
Claudio Carlan	473100	male	black	brown
Felice Prudden	486600	female	green	green

SOLVE THE MYSTERY

Start from type, data and city of the murder with codes

SELECT * FROM crime_scene_report WHERE type='murder' and city='SQL city' AND date = 20180115;

Date	Type	Description	City
20180115	murder	Security footage shows that there were 2 witnesses. The first witness lives at the last house on `Northwestern Dr`. The second witness, named Annabel, lives somewhere on `Franklin Ave`.	SQL City

Clue 1: Finding Witness #1 Using security footage

SELECT * FROM person WHERE address_street_name = "Northwestern Dr"
ORDER BY address_number DESC;

Id	Name	License id	Address_number	Address_street_name	SSN
14887	Morty Schapiro	118009	4919	Northwestern Dr	111564949

Finding Witness #2 Using security footage

SELECT * FROM person WHERE name like "Annabel%" AND address_street_name="Franklin Ave";

Id	Name	License id	Address_number	Address_street_name	SSN
16371	Annabel Miller	490173	103	Franklin Ave	318771143

Clue 2: Interview & Result: male, gym bag, membership '48Z', plate 'H42W', and Jan 9th

SELECT * FROM interview WHERE (person_id = 14887 OR person_id = 16371);

Person_id	transcript
14887	I heard a gunshot and then saw a **man** run out. He had a `Get Fit Now **Gym` bag**. The membership number on the bag started with **`48Z`**. Only **gold** members have those bags. The man got into a car with a plate that included **`H42W`.**
16371	I saw the murder happen, and I recognized the killer from my gym when I was working out last week on **January the 9th**.

Identify the perpetrator: by witness 14887(person_id)

SELECT * FROM get_fit_now_member WHERE id LIKE '48Z%' AND membership_status = 'gold';

Id	Person_id	Name	Membership_start_date	Membership_status
48Z55	67318	Jeremy Bowers	20160101	gold

SELECT p.*, d.* from person p JOIN drivers_license d ON
p.license_id=d.id WHERE gender = 'male' AND plate_number LIKE '%H42W%';

matched 'Jeremy Bowers' in 2 results

Id	Name	License_id	Address_street_name	ssn	age	Plate_number	Car_model
67318	Jeremy Bowers	423327	Washington Pl, Apt 3A	871539279	30	0H42W2	Spark LS
51739	Tushar Chandra	664760	Phi St	137882671	21	4H42WR	Altima

Identify the perpetrator: by witness 16371(person_id)

```
SELECT * FROM get_fit_now_check_in c
JOIN get_fit_now_member m ON m.id = c.membership_id
WHERE membership_id = "48Z55" AND check_in_date = "20180109";
```

Membership_id	Check_in_date	Check_in_time	Check_out_time	Id	Person_id	name
48Z55	20180109	1530	1700	48Z55	67318	**Jeremy Bowers**

Identify the perpetrator: Jeremy Bowers

```
SELECT * FROM get_fit_now_member m
JOIN get_fit_now_check_in c On m.id = c.membership_id
 JOIN person p ON p.id = m.person_id
JOIN drivers_license d ON d.id=p.license_id
WHERE check_in_date= '20180109' AND m.id LIKE '48Z%'
AND membership_status = 'gold' AND d.plate_number LIKE '%H42W%';
```

Id	Person_id	Name	Membership_start_date	Membership_status	Membership_id	Check_in_date
48Z55	67318	**Jeremy Bowers**	20160101	gold	48Z55	20180109

CONCLUSION

I have become familiar with SQL statements and learned how to utilize data. The way to utilize the data in SQL is quite intuitive. A person who seeing the statement for the first time could guess what the result could be.

Aliases are the one thing I learned this project. An aliase simplifies the codes. SQL aliases are used to give a table, or a column in a table, a temporary name (3schools, n.d). The JOIN procedure was also important. Intuitive words of codes helped the meaning. I think SQL is a programming language with lower hurdles than I thought. I will be closer through other examples in the future.

REFERENCE

MAILCHIMP. (2022). COUNT() with GROUP by. Retrieved from https://www.w3resource.com/sql/aggregate-functions/count-with-group-by.php

3schools. (n.d.). SQL aliases. Retrieved from https://www.w3schools.com/sql/sql_alias.asp