Murder Mystery using SQL

Libraries import

In [2]:
```
import pandas as pd
import sqlite3 as sql
import  numpy as np
pd.set_option('display.max_columns', None)

```

Loading Database & Tables

In [3]:
```
database = r"sql-murder-mystery.db"
connection = sql.connect(database)

```

In [7]:
```
query_1 = 'SELECT * FROM crime_scene_report'
crime_scene_report =  pd.read_sql_query(query_1, connection)
crime_scene_report.head()
```

Out[7]:

	date	type	description	city
0	20180115	robbery	A Man Dressed as Spider-Man Is on a Robbery Spree	NYC
1	20180115	murder	Life? Dont talk to me about life.	Albany
2	20180115	murder	Mama, I killed a man, put a gun against his he...	Reno
3	20180215	murder	REDACTED REDACTED REDACTED	SQL City
4	20180215	murder	Someone killed the guard! He took an arrow to ...	SQL City

In [5]:
```
query_2 =  'SELECT * FROM drivers_license'
drivers_license =  pd.read_sql_query(query_2, connection)
drivers_license.head()
```

Out[5]:

	id	age	height	eye_color	hair_color	gender	plate_number	car_make	car_model
0	100280	72	57	brown	red	male	P24L4U	Acura	MDX
1	100460	63	72	brown	brown	female	XF02T6	Cadillac	SRX
2	101029	62	74	green	green	female	VKY5KR	Scion	xB
3	101198	43	54	amber	brown	female	Y5NZ08	Nissan	Rogue
4	101255	18	79	blue	grey	female	5162Z1	Lexus	GS

In [6]:
```
query_3 = 'SELECT * FROM facebook_event_checkin'
facebook_event_checkin = pd.read_sql_query(query_3, connection)
facebook_event_checkin.head()
```

Out[6]:

	person_id	event_id	event_name	date
0	28508	5880	Nudists are people who wear one-button suits.\n	20170913
1	63713	3865	but that's because it's the best book on anyth...	20171009
2	63713	3999	If Murphy's Law can go wrong, it will.\n	20170502
3	63713	6436	Old programmers never die. They just branch t...	20170926
4	82998	4470	Help a swallow land at Capistrano.\n	20171022

In [8]:
```
query_4 = 'SELECT * FROM get_fit_now_check_in'
get_fit_now_check_in = pd.read_sql_query(query_4, connection)
get_fit_now_check_in.head()
```

Out[8]:

	membership_id	check_in_date	check_in_time	check_out_time
0	NL318	20180212	329	365
1	NL318	20170811	469	920
2	NL318	20180429	506	554
3	NL318	20180128	124	759
4	NL318	20171027	418	1019

In [9]:
```
query_5 = 'SELECT * FROM get_fit_now_member'
get_fit_now_member = pd.read_sql_query(query_5, connection)
get_fit_now_member.head()
```

Out[9]:

	id	person_id	name	membership_start_date	membership_status
0	NL318	65076	Everette Koepke	20170926	gold
1	AOE21	39426	Noe Locascio	20171005	regular
2	2PN28	63823	Jeromy Heitschmidt	20180215	silver
3	0YJ24	80651	Waneta Wellard	20171206	gold
4	3A08L	32858	Mei Bianchin	20170401	silver

In [11]:
```
query_6 = 'SELECT * FROM income '
income = pd.read_sql_query(query_6, connection)
income.head()

```

Out[11]:

	ssn	annual_income
0	100009868	52200
1	100169584	64500
2	100300433	74400
3	100355733	35900
4	100366269	73000

In [13]:
```
query_7 = 'SELECT * FROM interview'
interview = pd.read_sql_query(query_7, connection)
interview.head()

```

Out[13]:

	person_id	transcript
0	28508	‘I deny it!’ said the March Hare.\n
1	63713	\n
2	86208	way, and the whole party swam to the shore.\n
3	35267	lessons in here? Why, there’s hardly room for ...
4	33856	\n

In [14]:
```
query_8 = 'SELECT * FROM person'
person = pd.read_sql_query(query_8, connection)
person.head()
```

Out[14]:

	id	name	license_id	address_number	address_street_name	ssn
0	10000	Christoper Peteuil	993845	624	Bankhall Ave	747714076
1	10007	Kourtney Calderwood	861794	2791	Gustavus Blvd	477972044
2	10010	Muoi Cary	385336	741	Northwestern Dr	828638512
3	10016	Era Moselle	431897	1987	Wood Glade St	614621061
4	10025	Trena Hornby	550890	276	Daws Hill Way	223877684

1. Retrieve Crime Scene Report:

Task: Run a query to retrieve the crime scene report for the murder that occurred on Jan.15, 2018, in SQL City. Gather all available details from the report.

In [15]:
```
task_1 = "SELECT * FROM crime_scene_report WHERE date = 20180115 AND ty
murder_on_jan15 = pd.read_sql_query(task_1,connection)
pd.set_option('display.max_colwidth', None)
murder_on_jan15
```

Out[15]:

	date	type	description	city
0	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

2. Witness Personal Details:

Task: Check the personal details of witnesses involved in the case. Retrieve their names, addresses, and any other relevant information.

In [16]:
```
task_2 = "SELECT * FROM person where address_street_name=='Northwestern
location = pd.read_sql_query(task_2, connection)
max(location["address_number"])
```

Out[16]: 4919

In [17]:
```
task_2_1 = "SELECT * FROM person where address_number=='4919'"
first_witness = pd.read_sql_query(task_2_1, connection)
first_witness
```

Out[17]:

	id	name	license_id	address_number	address_street_name	ssn
0	14887	Morty Schapiro	118009	4919	Northwestern Dr	111564949

In [18]:
```
task_2_2 = "SELECT * FROM person where address_street_name = 'Franklin
second_witness = pd.read_sql_query(task_2_2, connection)
second_witness
```

Out[18]:

	id	name	license_id	address_number	address_street_name	ssn
0	16371	Annabel Miller	490173	103	Franklin Ave	318771143

3. View Witness Interviews:

Task: Access the recorded interviews of witnesses conducted after the murder. Gather insights into their statements and potential clues.

In [19]:
```
interview.head()
```

Out[19]:

	person_id	transcript
0	28508	'I deny it!' said the March Hare.\n
1	63713	\n
2	86208	way, and the whole party swam to the shore.\n
3	35267	lessons in here? Why, there's hardly room for YOU, and no room at all\n
4	33856	\n

In [20]:
```
task_3 = "SELECT * FROM interview WHERE person_id =16371"
interview_of_witnesses_1 = pd.read_sql_query(task_3, connection)
interview_of_witnesses_1
```

Out[20]:

	person_id	transcript
0	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.

In [21]:
```
task_3_1 = "SELECT * FROM interview WHERE person_id =14887"
interview_of_witnesses_2 = pd.read_sql_query(task_3_1, connection)
interview_of_witnesses_2
```

Out[21]:

	person_id	transcript
0	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".

Witnesses have said that the murderer is a MALE with a Gold membership.

4. Check Gym Database:

Task: Investigate the gym database using details obtained from the crime scene report and witness interviews. Look for any gym-related information that might be relevant.

In [22]:
```
get_fit_now_member.head()
```

Out[22]:

	id	person_id	name	membership_start_date	membership_status
0	NL318	65076	Everette Koepke	20170926	gold
1	AOE21	39426	Noe Locascio	20171005	regular
2	2PN28	63823	Jeromy Heitschmidt	20180215	silver
3	0YJ24	80651	Waneta Wellard	20171206	gold
4	3A08L	32858	Mei Bianchin	20170401	silver

Looking for somebody with gold membership in get_fit_now gym and has a car with plate "48z"

In [24]:
```
task_4 = "SELECT * FROM get_fit_now_member WHERE membership_status= 'go
suspects = pd.read_sql_query(task_4, connection)
suspects
```

Out[24]:

	id	person_id	name	membership_start_date	membership_status
0	48Z7A	28819	Joe Germuska	20160305	gold
1	48Z55	67318	Jeremy Bowers	20160101	gold

According to witness 2, there are two suspects who have gold memberships and gym IDs that resemble 48Z. We must now examine the car's details.

5. Check Car Details:

Task: Examine the car details associated with the crime scene. Retrieve information about the vehicles present during the incident.

In [25]:
```
task_5 = "SELECT * FROM drivers_license WHERE plate_number LIKE '%H42W%
car_details =pd.read_sql_query(task_5, connection)
car_details
```

Out[25]:

	id	age	height	eye_color	hair_color	gender	plate_number	car_make	car_model
0	423327	30	70	brown	brown	male	0H42W2	Chevrolet	Spark LS
1	664760	21	71	black	black	male	4H42WR	Nissan	Altima

6. Personal Details:

Task: Identify and collect personal details mentioned in the previous query. This includes names, addresses, and any additional details.

so we have 2 MALES who are having a car plate similar to H42W

In [26]:
```
task_6 = "SELECT * FROM  drivers_license WHERE id= 423327"
suspect_car_details = pd.read_sql_query(task_6,connection)
suspect_car_details
```

Out[26]:

	id	age	height	eye_color	hair_color	gender	plate_number	car_make	car_model
0	423327	30	70	brown	brown	male	0H42W2	Chevrolet	Spark LS

In [24]:
```
person.head()
```

Out[24]:

	id	name	license_id	address_number	address_street_name	ssn
0	10000	Christoper Peteuil	993845	624	Bankhall Ave	747714076
1	10007	Kourtney Calderwood	861794	2791	Gustavus Blvd	477972044
2	10010	Muoi Cary	385336	741	Northwestern Dr	828638512
3	10016	Era Moselle	431897	1987	Wood Glade St	614621061
4	10025	Trena Hornby	550890	276	Daws Hill Way	223877684

In [27]:
```
task_6_1 = "SELECT * FROM person WHERE license_id=423327"
suspect_personal_detail = pd.read_sql_query(task_6_1,connection)
suspect_personal_detail
```

Out[27]:

	id	name	license_id	address_number	address_street_name	ssn
0	67318	Jeremy Bowers	423327	530	Washington Pl, Apt 3A	871539279

In [28]:
```
get_fit_now_member.head()
```

Out[28]:

	id	person_id	name	membership_start_date	membership_status
0	NL318	65076	Everette Koepke	20170926	gold
1	AOE21	39426	Noe Locascio	20171005	regular
2	2PN28	63823	Jeromy Heitschmidt	20180215	silver
3	0YJ24	80651	Waneta Wellard	20171206	gold
4	3A08L	32858	Mei Bianchin	20170401	silver

Thus, the male driver "Jeremy Bowers" who drives a Chevrolet Spark LS with H42Z license plate is the suspect.

7. Membership Status at the Gym:

Task: Determine who is identified in the previous query as a member of the gym. Utilize the gym database to confirm their membership status.

In [29]:
```
# now checking if he was in gym on 15jan or not
task_7 = "SELECT * FROM get_fit_now_member WHERE name = 'Jeremy Bowers'"
jeremy_bowers_gym = pd.read_sql_query(task_7,connection)
jeremy_bowers_gym

```

Out[29]:

	id	person_id	name	membership_start_date	membership_status
0	48Z55	67318	Jeremy Bowers	20160101	gold

Indeed, he is a Gold Member, and the bag in question was the one that the suspect saw a Male driving a car with the license plate H42z. For further information, let's now examine his interview.

In [32]:
```
task_7_1 = "SELECT * from interview WHERE person_id= 67318"
interview_of_the_murderer = pd.read_sql_query(task_7_1,connection)
interview_of_the_murderer
```

Out[32]:

	person_id	transcript
0	67318	I was hired by a woman with a lot of money. I don't know her name but I know she's around 5'5" (65") or 5'7" (67"). She has red hair and she drives a Tesla Model S. I know that she attended the SQL Symphony Concert 3 times in December 2017.\n

Considering that he was a contract murderer employed by a woman, let's identify the primary offender.

According to the transcript, the woman is wealthy, drives a Tesla Model S, has red hair, and went to three SQL Symphony concerts in December of 2017.

In [33]:
```
1 drivers_license.head()
```

Out[33]:

	id	age	height	eye_color	hair_color	gender	plate_number	car_make	car_model
0	100280	72	57	brown	red	male	P24L4U	Acura	MDX
1	100460	63	72	brown	brown	female	XF02T6	Cadillac	SRX
2	101029	62	74	green	green	female	VKY5KR	Scion	xB
3	101198	43	54	amber	brown	female	Y5NZ08	Nissan	Rogue
4	101255	18	79	blue	grey	female	5162Z1	Lexus	GS

In [34]:
```
1 task_7_2 = " SELECT * FROM drivers_license WHERE hair_color ='red' AND
2 culprit_details = pd.read_sql_query(task_7_2, connection)
3 culprit_details
```

Out[34]:

	id	age	height	eye_color	hair_color	gender	plate_number	car_make	car_model
0	202298	68	66	green	red	female	500123	Tesla	Model S
1	291182	65	66	blue	red	female	08CM64	Tesla	Model S
2	918773	48	65	black	red	female	917UU3	Tesla	Model S

We thus have three girls, and the information matches the transcript. Let's now examine which of these three went to the SQL Symphony Concert three times in December 2017.

In [35]:
```
1 task_7_3 = "SELECT *FROM person WHERE license_id = '202298' OR license_
2 culprit_1 = pd.read_sql_query(task_7_3,connection)
3 culprit_1
```

Out[35]:

	id	name	license_id	address_number	address_street_name	ssn
0	78881	Red Korb	918773	107	Camerata Dr	961388910
1	90700	Regina George	291182	332	Maple Ave	337169072
2	99716	Miranda Priestly	202298	1883	Golden Ave	987756388

In [36]:
```
1 task_7_4 = "SELECT person_id, count(*), event_name FROM facebook_event_
2 culprit_2 = pd.read_sql_query(task_7_4, connection)
3 culprit_2
```

Out[36]:

	person_id	count(*)	event_name
0	24556	3	SQL Symphony Concert
1	99716	3	SQL Symphony Concert

We therefore found that "Miranda Priestly" is the primary culprit who paid the muderer "Jeremy Bowers" to attempt the crime in SQL City on January 15, 2017

As we can see that there is just one row that is common in each of these dataframes that contains culprit details.

---------- CASE SOLVED ----------