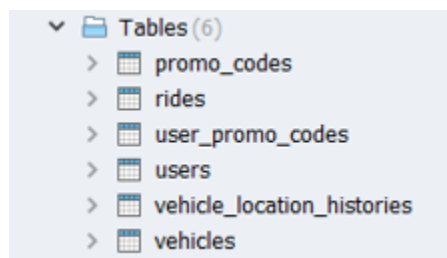# SQL PROJECT- SAGOber Data Breach Handling Case

The following are the screenshots for the SQL Data Analyst Project. The SQL files and installation instructions are present in the SQL Project Folder

The following shows the list of database tables:

Checking vehicle location history:

```
1   --Checking Vehicle loation table
2   SELECT * FROM vehicle_location_histories;
3
```

	city	ride_id	timestamp	lat	long
1	amsterdam	d2bf77dd-7af2-43fa-86a0-3c035c7474ac	2020-06-23 12:53:23.09373	5.802552	53.188701
2	amsterdam	d2bf77dd-7af2-43fa-86a0-3c035c7474ac	2020-06-23 12:53:23.093734	5.878306	53.124095
3	amsterdam	d2bf77dd-7af2-43fa-86a0-3c035c7474ac	2020-06-23 12:53:23.093771	5.891314	53.13176
4	amsterdam	d2bf77dd-7af2-43fa-86a0-3c035c7474ac	2020-06-23 12:53:23.093782	5.018579	53.248133
5	amsterdam	d2bf77dd-7af2-43fa-86a0-3c035c7474ac	2020-06-23 12:53:23.093803	5.800994	53.261533
6	amsterdam	d2bf77dd-7af2-43fa-86a0-3c035c7474ac	2020-06-23 12:53:23.093821	5.230389	53.224581
7	amsterdam	d2bf77dd-7af2-43fa-86a0-3c035c7474ac	2020-06-23 12:53:23.093867	5.117866	53.012253
8	amsterdam	d2bf77dd-7af2-43fa-86a0-3c035c7474ac	2020-06-23 12:53:23.09387	5.201437	53.26801
9	amsterdam	d2bf77dd-7af2-43fa-86a0-3c035c7474ac	2020-06-23 12:53:23.093899	5.744734	53.174305
10	amsterdam	d2bf77dd-7af2-43fa-86a0-3c035c7474ac	2020-06-23 12:53:23.093903	5.721832	53.292598
11	los angeles	fee1db35-fb44-4f09-94ef-b85851705208	2020-06-23 12:53:27.01199	119.197573	35.037608
12	los angeles	fee1db35-fb44-4f09-94ef-b85851705208	2020-06-23 12:53:27.011996	119.007412	35.03341
13	los angeles	fee1db35-fb44-4f09-94ef-b85851705208	2020-06-23 12:53:27.012027	119.053337	35.031508
14	los angeles	fee1db35-fb44-4f09-94ef-b85851705208	2020-06-23 12:53:27.012056	119.23191	35.007428
15	boston	01aa7790-9803-4ce2-9768-479f87c8a9a	2020-06-23 13:22:22.569293	71.932466	42.02335
16	boston	01aa7790-9803-4ce2-9768-479f87c8a9a	2020-06-23 13:22:22.62732	71.010282	43.160937
17	boston	01aa7790-9803-4ce2-9768-479f87c8a9a	2020-06-23 13:22:22.691515	71.002725	42.571925

Number of records: 175293   Number of fields: 5   Query time: 3.802 second(s)   ☑ Read-Only

Query 1: Checking the above table and using available information to narrow down the query

```
3
4   --Clue 1:
5   SELECT * FROM vehicle_location_histories AS vlh WHERE
6           city = 'new york' AND long BETWEEN 40.5 AND 40.6 AND lat BETWEEN -74.997 AND -74.9968   AND
7           vlh.timestamp::date =  '2020-06-23'::date
8   ORDER BY long;
```

	city	ride_id	timestamp	lat	long
1	new york	a666f366-e734-495e-bc5c-dc68c47898f8	2020-06-23 12:58:23.238631	-74.99691	40.500439
2	new york	5c7ca8a2-dc65-4df4-9aa0-01a7ba448f3c	2020-06-23 13:17:23.494156	-74.996874	40.500661
3	new york	3c0a5961-b87e-4ec9-8cc2-9b27c720052e	2020-06-23 13:13:30.185195	-74.996883	40.500788
4	new york	d9623cd9-2ea5-4fab-a53a-48e846a07391	2020-06-23 13:04:28.999567	-74.996862	40.500817
5	new york	7f8b2cb4-af47-4c84-8edc-66aac14bc32d	2020-06-23 13:01:01.348577	-74.996823	40.5018
6	new york	5c7ca8a2-dc65-4df4-9aa0-01a7ba448f3c	2020-06-23 13:17:55.686569	-74.996916	40.501812
7	new york	dae0e43a-5368-4c78-a4c2-	2020-06-23 13:27:55.654841	-74.996888	40.502037
8	new york	a55f7789-8734-4322-9e9b-bb8b75240ff2	2020-06-23 13:07:04.739516	-74.996991	40.502291
9	new york	dae0e43a-5368-4c78-a4c2-	2020-06-23 13:28:07.691506	-74.996972	40.502498
10	new york	a8597aa2-ea10-4545-a8f0-786a2d910791	2020-06-23 13:23:41.901785	-74.996829	40.502733
11	new york	d9623cd9-2ea5-4fab-a53a-48e846a07391	2020-06-23 13:03:53.94456	-74.996872	40.502933
12	new york	eb3f36bd-fd19-4b08-be45-c39326efe8cc	2020-06-23 13:10:07.559108	-74.996912	40.503477

Number of records: 428   Number of fields: 5   Query time: 116 millisecond(s)   ☑ Read-Only

Let's Create a View (stored procedure) for our query:

```
10 ⊟  --Thousands of poeple come to Sagouber so narrowing it down to 428 is a blessing.
11 └  --We can interrogate the drivers. save the above query as View
12 ⊟  CREATE VIEW suspected_rides AS
13     SELECT * FROM vehicle_location_histories AS vlh WHERE
14         city = 'new york' AND long BETWEEN 40.5 AND 40.6 AND lat BETWEEN -74.997 AND -74.9968  AND
15         vlh.timestamp::date = '2020-06-23'::date
16  └ ORDER BY long;--avoiding writting the same query, one step closer to finding the mystery
17
```

	city	ride_id	timestamp	lat	long
1	new york	a666f366-e734-495e-bc5c-dc68c47898f8	2020-06-23 12:58:23.238631	-74.99691	40.500439
2	new york	5c7ca8a2-dc65-4df4-9aa0-01a7ba448f3c	2020-06-23 13:17:23.494156	-74.996874	40.500661
3	new york	3c0a5961-b87e-4ec9-8cc2-9b27c720052e	2020-06-23 13:13:30.185195	-74.996883	40.500788
4	new york	d9623cd9-2ea5-4fab-a53a-48e846a07391	2020-06-23 13:04:28.999567	-74.996862	40.500817
5	new york	7f8b2cb4-af47-4c84-8edc-66aac14bc32d	2020-06-23 13:01:01.348577	-74.996823	40.5018
6	new york	5c7ca8a2-dc65-4df4-9aa0-01a7ba448f3c	2020-06-23 13:17:55.686569	-74.996916	40.501812
7	new york	dae0e43a-5368-4c78-a4c2-	2020-06-23 13:27:55.654841	-74.996888	40.502037
8	new york	a55f7789-8734-4322-9e9b-bb8b75240ff2	2020-06-23 13:07:04.739516	-74.996991	40.502291
9	new york	dae0e43a-5368-4c78-a4c2-	2020-06-23 13:28:07.691506	-74.996972	40.502498
10	new york	a8597aa2-ea10-4545-a8f0-786a2d910791	2020-06-23 13:23:41.901785	-74.996829	40.502733
11	new york	d9623cd9-2ea5-4fab-a53a-48e846a07391	2020-06-23 13:03:53.94456	-74.996872	40.502933
12	new york	eb3f36bd-fd19-4b08-be45-c39326efe8cc	2020-06-23 13:10:07.559108	-74.996912	40.503477

Number of records: 428   Number of fields: 5   Query time: 84 millisecond(s)   ☑ Read-Only

Query 2: Using Multijoin to get vehicle and owner details

```
28
29    -----------------------------------------multijoin--specific drivers-----------------------------------------
30    SELECT DISTINCT r.vehicle_id, u.name AS "Owner",u.address, v.status, v.current_location
31    FROM suspected_rides AS sr
32    INNER JOIN rides AS r ON r.id = sr.ride_id
33    INNER JOIN vehicles AS v ON v.id = r.vehicle_id
34    INNER JOIN users AS u ON u.id = v.owner_id;
35
```

	vehicle_id	Owner	address	status	current_location
1	01d60181-5762-4603-831e-08a29d3ebb5	Tyler Evans	0031 Teresa Path Apt. 577 New	available	6595 Rice Hill Lawrencemouth
2	03002163-c9d9-42e9-81c9-a9ca83476d5c	Jeffery Garcia	63238 Sharon Field Apt. 382 Pitt	available	49059 Jose Route Suite 770 Te
3	06145a48-2fde-47da-8214-22e48b8af3b2	Joshua Sherman	4776 Lambert Gateway Suite 07...	available	08835 Burgess Groves Apt. 91...
4	18c9ef45-b7fe-41f5-a2be-fdb4e404861e	Crystal Lynch	98230 Melendez Avenue Suite 0...	available	43933 Long Fields Suite 080 W
5	19f0933c-abec-487d-a928-624089147d1d	Chad Scott	PSC 4484, Box 9126 APO AA 931	available	50985 Becker Island Suite 164
6	1a27e91f-64b3-4240-a4b3-347d65bde039	James Smith	6594 Barry Shoal Connierort, TN	available	322 Jose Cape Apt. 612 East Br
7	1a55e793-6b5a-4975-9bb9-e7614cbb6184	Stacy Shah	015 Manuel Dale North Jeffrey S	available	20982 Pollard Stream West Ba
8	1db131ad-f05a-4f84-9ae9-c1e4ffe21dd1	Juan Richards	978 Romero Crescent Suite 494	available	9177 Jones Mountain Apt. 987
9	20764109-e583-4a41-86d3-5244ef812008	Morgan Morris	030 Ho Cliffs Yorkmouth, TN 66922	available	155 Hobbs Street East Don OK
10	208bdd75-0bf7-40cb-a349-6daf73ec6f35	Michael Matthews	Unit 0165 Box 5973 DPO AA 50237	available	2221 Brown Crest Waberton C
11	229edecf-522e-4b6b-90a7-c485e1adc752	Randy Wong	5623 Meyers Plain Apt. 158 Dan	available	8346/ John Center Suite 014
12	28735478-ea77-4dc9-b1a0-d3b47c80bf3f	Anthony Dickson	0371 Pearson Springs Suite 321 227 Brian Corner	available	6682 Hamilton Cliffs Apt. 4/4 E USS Allen

Number of records: 89   Number of fields: 5   Query time: 668 millisecond(s)   ☑ Read-Only

Query 3: Using Joins to connect vehicle and passenger

```
35
36    --Interrogate all drivers but nothing. its one of riders(passengers)
37
38    SELECT DISTINCT r.vehicle_id, u.name AS "Rider Name",u.address
39    FROM suspected_rides AS sr
40    INNER JOIN rides AS r ON r.id = sr.ride_id
41    INNER JOIN users AS u ON u.id = r.rider_id;
42
```

	vehicle_id	Rider Name	address
1	01d60181-5762-4603-831e-08a29d3ebb5	Antonio Ruiz	16683 Jacqueline Camp Brownf
2	01d60181-5762-4603-831e-08a29d3ebb5	Christie Delgado	269 Watson Stravenue Suite 4...
3	03002163-c9d9-42e9-81c9-a9ca83476d5c	Carla Hill	72035 Joseph Curve Lake Iofft
4	06145a48-2fde-47da-8214-22e48b8af3b2	Candice Kenne...	5535 Scott Lane New Sethbury
5	18c9ef45-b7fe-41f5-a2be-fdb4e404861e	Anna Garza	2487 Francisco Roads Apt. 43...
6	18c9ef45-b7fe-41f5-a2be-fdb4e404861e	Carla Hill	72035 Joseph Curve Lake Iofft
7	19f0933c-abec-487d-a928-624089147d1d	Veronica Lewis	2147 Duran Fords Brucemouth
8	1a27e91f-64b3-4240-a4b3-347d65bde039	Lisa Hayes	5924 Williams Estate South Ma
9	1a55e793-6b5a-4975-9bb9-e7614cbb6184	Joshua Baxter	9435 Lisa Cove New Nicholefur
10	1db131ad-f05a-4f84-9ae9-c1e4ffe21dd1	Marc Rogers	2016 Hailey Mountain Apt. 925
11	20764109-e583-4a41-86d3-5244ef812008	Amanda Farrell	077 Allen Stravenue New Trici
12	20764109-e583-4a41-86d3-5244ef812008	Kathy Martinez	842 Odonnell Street Suite 319 Unit 2110 Box 1697

Number of records: 109 Number of fields: 3 Query time: 2.234 second(s) ☑ Read-Only

Checking Users tables to know the Passengers details

```
43   --Breakthrough-anonymous call that somebody from inside(employees) helped someone outside. min 2 people were involved
44   --db linking is now required
45
46   SELECT * FROM users;
47   --somehow we need to connect name in users to first_name and last_name in employees;
48   SELECT DISTINCT
49          split_part( u.name, ' ', 1 ) AS "first_name",
50          split_part( u.name, ' ', 2 ) AS "last_name"
```

	id	city	name	address	credit_card
1	000934e7-725b-4f43-8f52-2789970ca086	new york	Gary Wade	8132 Washington Station So	5348345015427254
2	000ef301-f942-420f-8b1d-7642ce8daa90	detroit	Kyle Johnson	2798 Brenda Well Apt. 365 S	4624824530754341
3	0018a508-81b6-4e61-95db-842beea8b66	new york	Robert Richard	484 Patterson Village Russell	4075578018708955
4	001b48e8-98ef-40b5-aec9-10af44f1f221	boston	Lindsay Werner	111 Martin Avenue Apt. 605	47725329806550393...
5	001b7a6a-bd1e-4da2-afeb-0104a14be8b2	boston	Frederick Harris	93127 Lawrence Tunnel Wes	4630298954625570
6	003de45e-4e31-4bff-a8cd-9c8fe30415c8	new york	Kevin Rasmussen	Unit 0260 Box 8569 DPO AE 2	4985949927018472
7	00482d57-0ec8-4cc9-8401-553ad714e911	boston	Penny Frey	13859 Miller Curve Wrightsid	45326672467376837...
8	004c6a15-0276-4c21-bf20-0183ef2cf844	new york	Christina May	358 Jonathan Knoll Suite 13...	3584430741268764
9	0052bbf6-1194-4457-b4ba-8a60a592c049	boston	Jerry Mills	3002 Roberts Drives Suite 4...	4819854700900502
10	00532204-c194-4efd-8c8f-38b4390226c2	new york	Marc Mcclure	87622 Lang Place Inseland	501842032127
11	005551a5-7c50-471d-b70e-b6f472cc1aeb	boston	Mallory Lewis	PSC 3556, Box 5218 APO AE	38603768327543
12	0071a31f-a852-42d3-a46b-458f68fe4e07	boston	Abigail Lindsey	2936 George Route Hickmant	5272932796475949
				877 Thompson Mission	

Number of records: 585947 Number of fields: 5 Query time: 5.618 second(s) ☑ Read-Only

Query 4: Splitting Users name to first and last name:

```
43  --Breakthrough-anonymous call that somebody from inside(employees) helped someone outside. min 2 people were involved
44  --db linking is now required
45  --Somehow we need to connect name in users to first_name and last_name in employees;
46  --lets create a view of above----
47  CREATE VIEW suspect_riders_name AS
48  SELECT DISTINCT
49                  split_part( u.name, ' ', 1 ) AS "first_name",
50                  split_part( u.name, ' ', 2 ) AS "last_name"
51  FROM suspected_rides AS sr
52  INNER JOIN rides AS r ON r.id = sr.ride_id
53  INNER JOIN users AS u ON u.id  = r.rider_id;
54
```

	first_name	last_name
1	Alexander	Freeman
2	Alexis	Jones
3	Alicia	Jones
4	Alyssa	Carroll
5	Amanda	Farrell
6	Andrew	Owens
7	Angela	Garcia
8	Anna	Garza
9	Antonio	Ruiz

Number of records: 103 Number of fields: 2 Query time: 81 millisecond(s) ☑ Read-Only

Query 5: Connecting Users (Passengers) to our Employees database to find matching names:

```
---Now we can use it to correlate this with employees via db link
SELECT DISTINCT
                concat(t1.first_name,' ', t1.last_name) AS "Employees",
                concat(u.first_name,' ', u.last_name) AS "Riders"

FROM dblink('host=localhost user=postgres password=postgres dbname=employees',
            'select first_name, last_name from employees;')

AS t1(first_name NAME, last_name NAME)
JOIN suspect_riders_name AS u ON t1.last_name = u.last_name
ORDER BY "Riders";
```

	employee	rider
1	Xumin Peck	Jeanne Peck
2	Sreenivas Peck	Jeanne Peck
3	Barun Harllee	Danny Harllee
4	Annemarie Harllee	Danny Harllee
5	Gino Peck	Jeanne Peck
6	Deborah Peck	Jeanne Peck
7	Izaskun Harllee	Danny Harllee
8	Fumino Peck	Jeanne Peck
9	Domenico Harllee	Danny Harllee
10	Parke Morris	Morgan Morris
11	Xiaoshan Morris	Morgan Morris