

CIRCLE 97

Fintech Events - Cards and Users.

__altschoolafrica
DATA ENGINEERING

DEEP DIVE INTO OUR DATA

- Overview of the cards and users tables.

select * from events.cards Enter a SQL expression to filter results (use Ctrl+Space)										
	123 id	123 user_id	ABC created_by_name	updated_at	created_at	active	ABC type	event_at	ABC event_id	
1	4,940	545	Joshua Graves	2023-10-23 23:30:30.000	2023-10-23 23:26:25.000	[]	card	2023-10-23 23:30:30.000	00017a9e-67b3-4124-8a9e-c58c05680369	
2	303	235	Victoria Thomas	2023-10-23 23:14:01.000	2023-10-23 23:09:56.000	[v]	card	2023-10-23 23:14:01.000	00101c06-31ac-452a-b3ab-973bf1456869	
3	2,240	608	Glenn Cook	2023-10-23 23:22:33.000	2023-10-23 23:16:52.000	[]	card	2023-10-23 23:22:33.000	00165065-1f6b-45e0-8ff0-e69422ee1319	
4	2,433	483	Kimberly Harris	2023-10-23 23:28:43.000	2023-10-23 23:17:34.000	[]	card	2023-10-23 23:28:43.000	002eae69-2bce-4491-8d04-eab12de23932	
5	1,134	7	Karina Kramer	2023-10-23 23:21:02.000	2023-10-23 23:12:58.000	[]	card	2023-10-23 23:21:02.000	0033ce8a-996c-4067-b5d6-9bdbee3ec4df	

select * from events.users Enter a SQL expression to filter results (use Ctrl+Space)									
	ABC type	event_at	ABC event_id	123 id	ABC name	ABC address	ABC job	123 score	
1	user	2023-10-23 22:55:01.000	006759bd-9a00-47c4-8a35-f7cf9877f9fa	507	Jacob Wise	474 Jennifer UnionfReyesburgh, PW	Engineer, electrical	0.5269983	
2	user	2023-10-23 22:55:01.000	00a3a97f-33b9-434e-a83d-a35c8daf31c1	978	Kelly Nguyen	96204 Elliott PathfJessicamouth, VA	Retail manager	0.49702907	
3	user	2023-10-23 22:55:01.000	01310be4-a3f0-4b48-ba86-767c9ae5afb1	264	Jody Hoffman	2619 Jeremy Junctions Suite 700fSo	Ranger/warden	0.18239447	
4	user	2023-10-23 22:55:01.000	01644360-a490-4f22-ada6-bb5ac13f0afb	50	Katherine Chang	504 Miller Vista Apt. 801fWest Frank	Recycling officer	0.03797031	
5	user	2023-10-23 22:55:01.000	0176dd66-10e7-4ba3-a7f8-b045aec8fad6	291	Amy Carter	16662 Walker ManorfBrycefort, MT	Multimedia programmer	0.04850558	

DEEP DIVE INTO OUR DATA Contd...

- Analyzing the data from the event.users and events.cards table yielded no missing values and duplicate users.

```
SELECT COUNT(*) AS missing_count
FROM events.users
WHERE id IS NULL OR name IS NULL OR address IS NULL OR job IS NULL OR score IS NULL;
```

Results 1

missing_count
123

```
SELECT id, COUNT(*) AS users_duplicate
FROM events.users
GROUP BY id
HAVING COUNT(*) > 1;
```

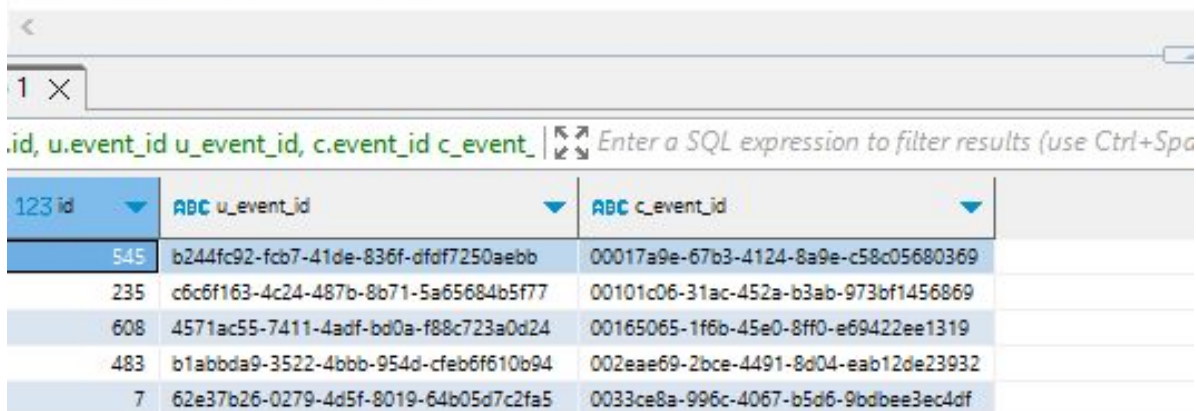
Results 1

id	users_duplicate
123	123

Deep Dive Contd.

- Event_id on both tables for the same user are different

```
select u.id, u.event_id u_event_id, c.event_id c_event_id
from events.users u
join events.cards c
on c.user_id = u.id;
```

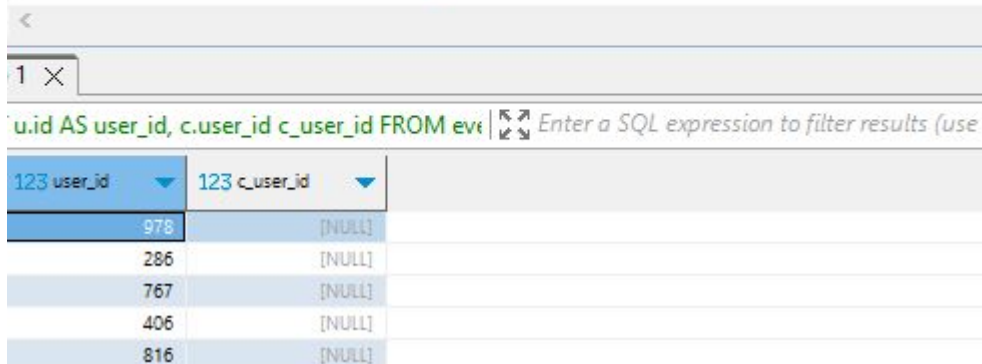


123 id	ABC u_event_id	ABC c_event_id	
545	b244fc92-fcb7-41de-836f-dfdf7250aebb	00017a9e-67b3-4124-8a9e-c58c05680369	
235	c6c6f163-4c24-487b-8b71-5a65684b5f77	00101c06-31ac-452a-b3ab-973bf1456869	
608	4571ac55-7411-4adf-bd0a-f88c723a0d24	00165065-1f6b-45e0-8ff0-e69422ee1319	
483	b1abbd9-3522-4bbb-954d-cfeb6f610b94	002eae69-2bce-4491-8d04-eab12de23932	
7	62e37b26-0279-4d5f-8019-64b05d7c2fa5	0033ce8a-996c-4067-b5d6-9bdbee3ec4df	

Deep Dive Contd...

- Checking for users without corresponding cards returns a null, which implies that we have users without cards.

```
-- Check for users without corresponding cards
SELECT u.id AS user_id, c.user_id c_user_id
--select u.event_id u_event_id, c.event_id c_event_id
FROM events.users u
left JOIN events.cards c ON u.id = c.user_id
WHERE c.user_id IS NULL;
```



123 user_id	123 c_user_id
978	[NULL]
286	[NULL]
767	[NULL]
406	[NULL]
816	[NULL]

Conclusion and Recommendation

- From the analysis done, the data collected by the engineering team needs a fine-tuning and cannot be used for industrial and analytical purposes.
- Attention needs to be paid to the users id, event_id, event_at columns for accurate and precise data collection.
-
- For further reference on this project please refer to the Circle 97 github repository with the link below

[oakerekan/Circle97 \(github.com\)](https://github.com/oakerekan/Circle97)