UNDERSTANDING CUSTOMER BEHAVIOUR ON ZOMATO

INTERNSHIP MINOR PROJECT AT ACADEMOR

AUTHOR/ANALYST: SAGNIK ADAK

January 11, 2023

DATA IMPORTATION

According to instruction I imported the data in excel. I have created four tables
Table A (columns): *userid , gold_signup_date*
Table B (columns): *Product_id, product_name, price*
Table C (columns): *Userid, created_date, product_id*
Table D (columns): *Userid, signup_date*

	1	2
1	userid	gold_signup_dat
2	1	2017-09-22
3	3	2017-04-21

(TABLE A)

	1	2	3
1	product_id	product_name	price
2	1	p1	980
3	2	p2	870
4	3	p3	330
5			
6			

(TABLE B)

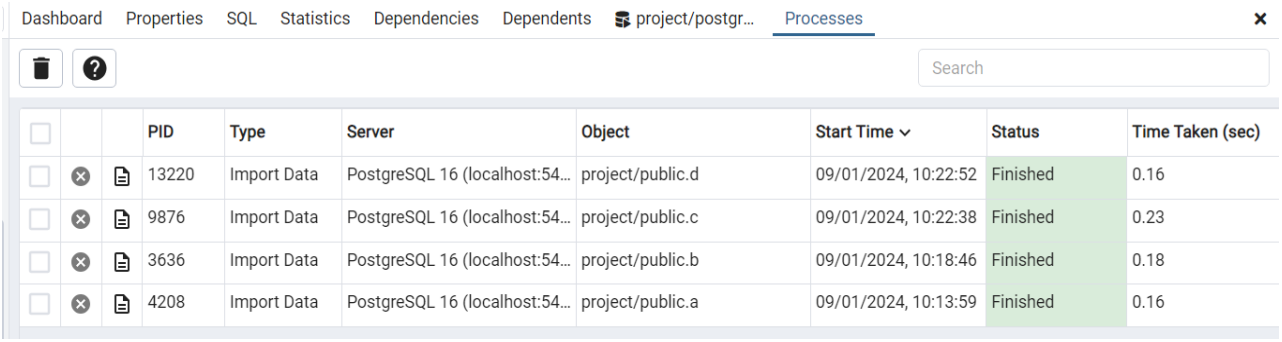
	1	2	3
1	userid	created	product_id
2	1	2017-04-19	2
3	3	2019-12-18	1
4	2	2020-07-20	3
5	1	2019-10-23	2
6	1	2018-03-19	3
7	3	2016-12-20	2
8	1	2016-11-09	3
9	1	2016-05-20	3
10	2	2017-09-24	1
11	1	2016-03-11	1
12	3	2016-11-10	1
13	3	2017-12-07	2
14	3	2016-12-15	2
15	2	2017-11-08	2
16	2	2018-09-10	3
17			

(TABLE C)

	1	2
1	userid	signup_date
2	1	2014-09-02
3	2	2015-01-15
4	3	2014-04-11
5		

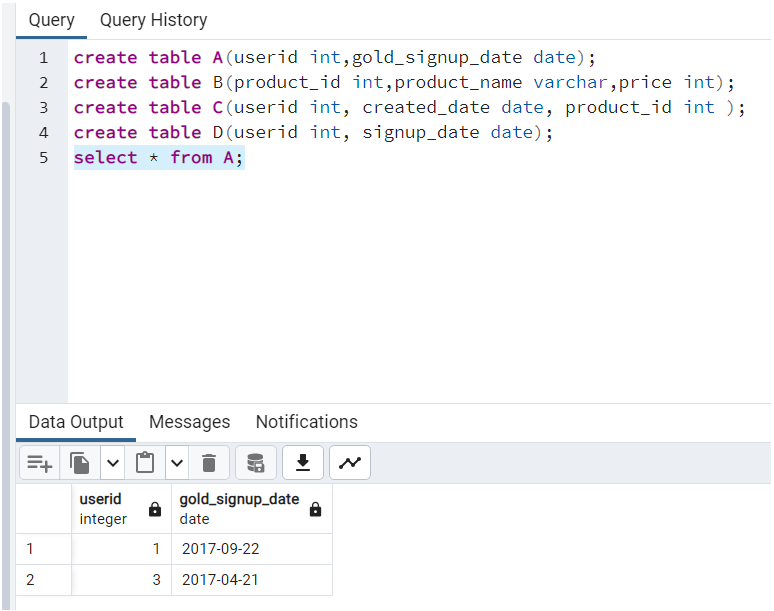
(TABLE D)

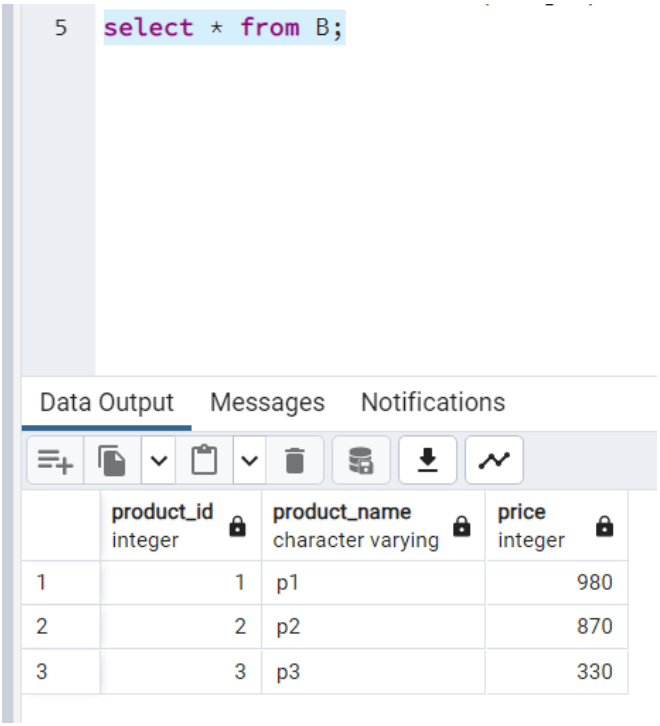
These tables are created in excel and I converted them into CSV files. Then I imported the data for SQL operation (I used postgresql for SQL).



(data-importation is finished)

After creating the tables I checked them in SQL





```
4 •  select * from C
```

Result Grid | Filter Rows:

	userid	created_date	product_id
▶	1	2017-04-19	2
	3	2019-12-18	1
	2	2020-07-20	3
	1	2019-10-23	2
	1	2018-03-19	3
	3	2016-12-20	2
	1	2016-11-09	3
	1	2016-05-20	3
	2	2017-09-24	1
	1	2016-03-11	1
	3	2016-11-10	1
	3	2017-12-07	2
	3	2016-12-15	2
	2	2017-11-08	2
	2	2018-09-10	3

```
5  select * from D;
```

Data Output | Messages | Notificat

	userid integer	signup_date date
1	1	2014-09-02
2	2	2015-01-15
3	3	2014-04-11

Q1) What is the total amount each customer spent on Zomato?

```
28
29  select C.userid as customer, sum(price) as total_price
30  from B inner join C
31  on
32  B.product_id = C.product_id
33  group by userid
34  order by userid
35
36
37
38
39
40
41
42
43
44
45
46
```

Data Output | Messages | Notifications

	customer integer	total_price bigint
1	1	3710
2	2	2510
3	3	4570

Q2) What is the total amount each customer spent on each product on Zomato?

```
60
61  select  C.userid, b.product_name,sum(price) as total_indivisual_product_price
62
63  from B full outer join C
64  on
65  B.product_id = C.product_id
66
67  group by userid, product_name
68  order by userid
69
70
71
72
73
74
75
```

Data Output | Messages | Notifications

	userid integer	product_name character varying	total_indivisual_product_price bigint
1	1	p1	980
2	1	p2	1740
3	1	p3	990
4	2	p1	980
5	2	p2	870
6	2	p3	660
7	3	p1	1960
8	3	p2	2610

Q3) How many days has each customer visited Zomato?

```
select userid , count(distinct created_date) as visited_date
from c
group by userid
```

Data Output | Messages | Notifications

	userid integer	visited_date bigint
1	1	6
2	2	4
3	3	5

Q4) What was the first product purchased by each customer?

```
select c.userid, min (b.product_name) as first_product
from B full outer join C
on B.product_id = C.product_id
group by c.userid

--5th
```

Data Output | Messages | Notifications

	userid integer	first_product text
1	3	p1
2	2	p1
3	1	p1

Q5) What is the most purchased item on menu & how many times was it puchased by all customers?

```
--5th

select b.product_name, count(C.product_id) as purchase_count
from B inner join C
on B.product_id = C.product_id
group by b.product_name
order by purchase_count desc
limit 1
```

Data Output | Messages | Notifications

	product_name character varying	purchase_count bigint
1	p2	6

Q6) Which item was most popular for each customer?

```
select userid, product_name
from
(
  select c.userid, b.product_name, row_number() over(partition by c.userid order by count(c.product_id)desc) as rank
    from B inner join C
    on B.product_id = C.product_id
    group by c.userid, b.product_name
) as ranked_products
where rank=1;
```

Data Output | Messages | Notifications

	userid integer	product_name character varying
1	1	p3
2	2	p3
3	3	p2

Q7) Which item was purchased first after they became a gold member?

```
--7th

select a.userid, b.product_name, min(c.created_date) as first_date_purchase
from A inner join c
on a.userid= c.userid
inner join b on c.product_id = b.product_id
where c.created_date > a.gold_signup_date
group by a.userid, b.product_id, b.product_name
order by a.userid
```

Data Output | Messages | Notifications

	userid integer	product_name character varying	first_date_purchase date
1	1	p2	2019-10-23
2	1	p3	2018-03-19
3	3	p2	2017-12-07
4	3	p1	2019-12-18

Q8) Which item was purchased just before they became a gold member?

```
select a.userid, b.product_name, max(c.created_date) as last_date_purchase
from A inner join c
on a.userid= c.userid
inner join b on c.product_id = b.product_id
where c.created_date < a.gold_signup_date
group by a.userid, b.product_id, b.product_name
order by a.userid
```

Data Output | Messages | Notifications

	userid integer	product_name character varying	last_date_purchase date
1	1	p2	2017-04-19
2	1	p1	2016-03-11
3	1	p3	2016-11-09
4	3	p2	2016-12-20
5	3	p1	2016-11-10

Q9) What is total orders and amount spent for each member before they became a gold member?

```
96
97  select a.userid, b.product_name, max(c.created_date) as last_date_purchase
98  from A inner join c
99  on a.userid= c.userid
100 inner join b on c.product_id = b.product_id
101 where c.created_date < a.gold_signup_date
102 group by a.userid, b.product_id, b.product_name
103 order by a.userid
104
105 --9th
106
107
108
109
110
111
112
113
114
115
116
117
118
```

Data Output | Messages | Notifications

	userid integer	product_name character varying	last_date_purchase date
1	1	p2	2017-04-19
2	1	p1	2016-03-11
3	1	p3	2016-11-09
4	3	p2	2016-12-20
5	3	p1	2016-11-10

Q10) rank all transaction of the customers?

```
119
120 select c.userid, b.product_name, c.created_date,
121  rank() over(partition by c.userid  order by c.created_date) as Transaction_rank
122 from c inner join b
123 on c.product_id = b.product_id
124 order by Transaction_rank, c.userid
125
126
127
```

Data Output | Messages | Notifications

	userid integer	product_name character varying	created_date date	transaction_rank bigint
1	1	p1	2016-03-11	1
2	2	p1	2017-09-24	1
3	3	p1	2016-11-10	1
4	1	p3	2016-05-20	2
5	2	p2	2017-11-08	2
6	3	p2	2016-12-15	2
7	1	p3	2016-11-09	3
8	2	p3	2018-09-10	3
9	3	p2	2016-12-20	3
10	1	p2	2017-04-19	4
11	2	p3	2020-07-20	4
12	3	p2	2017-12-07	4
13	1	p3	2018-03-19	5
14	3	p1	2019-12-18	5
15	1	p2	2019-10-23	6

Q11) rank all transaction for each member whenever they are zomato gold member. For every non goldmember transaction mark as na

```
select c.userid, b.product_name, c.created_date,
case
 when a.userid is not null then rank() over(partition by c.userid order by c.created_date)

  end as Transaction_rank
from c inner join b
on c.product_id = b.product_id left join a on c.userid= a.userid
order by c.userid, c.created_date
```

Data Output | Messages | Notifications

	userid integer	product_name character varying	created_date date	transaction_rank bigint
1	1	p1	2016-03-11	1
2	1	p3	2016-05-20	2
3	1	p3	2016-11-09	3
4	1	p2	2017-04-19	4
5	1	p3	2018-03-19	5
6	1	p2	2019-10-23	6
7	2	p1	2017-09-24	[null]
8	2	p2	2017-11-08	[null]
9	2	p3	2018-09-10	[null]
10	2	p3	2020-07-20	[null]
11	3	p1	2016-11-10	1
12	3	p2	2016-12-15	2
13	3	p2	2016-12-20	3
14	3	p2	2017-12-07	4
15	3	p1	2019-12-18	5

Q12) In the first one year after customer joins the gold program (including the join date) irrespective of what customer has purchased

a) earn 5 zomato points for every 10rs spent. who earned more? 1 or 3

```
select a.userid,sum(case when created_date between gold_signup_date and date_add(gold_signup_date, interval '1 year')
                  then price/10 * 5 else 0 end ) as earned_points
                  from c inner join b
on c.product_id = b.product_id inner join a on c.userid=a.userid
where a.userid in(1,3)
group by a.userid
```

Data Output | Messages | Notifications

	userid integer	earned_points bigint
1	3	435
2	1	165

b) what is the earning in first year? 1zp = 2rs

```
160
161 select a.userid,sum(case when created_date between gold_signup_date and date_add(gold_signup_date, interval '1 year')
162                   then price/2 else 0 end ) as earned_points
163                   from c inner join b
164 on c.product_id = b.product_id inner join a on c.userid=a.userid
165
166 group by a.userid
167
168
169
170
171
172
173
174
175
```

Data Output | Messages | Notifications

	userid integer	earned_points bigint
1	1	165
2	3	435