# Customer Analysis

**Data Analysis Case Study**

This analysis provides insights of the customer's spending, subscriptions, items purchased & points earned to track the overall customer behavior & market trend.

-Created by
Abhishek S. Gharat

# Understanding Data

**There are four tables overall including sales data, product data, users data & gold membership users data.**

- **sales** table includes: User ID as **userid**, Date of Sales as **created_at** & item purchased as **product_id**
- **product** table includes: Product ID as **product_id**, Name of Product as **product_name** & Product Price as **price**
- **users** table includes: User ID as **userid,** User Signup Date as **signup_date**
- **goldusers_signup** table includes: User ID as **userid,** Membership Signup Date as **gold_signup_date**

`select * from sales;`

| 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 | 1 |
| 1 | 2016-05-20 | 3 |

`select * from product;`

| product_id | product_name | price |
|------------|--------------|-------|
| 1 | p1 | 980 |
| 2 | p2 | 870 |
| 3 | p3 | 330 |

`select * from users;`

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

`select * from goldusers_signup;`

| userid | gold_signup_date |
|--------|------------------|
| 1 | 2017-09-22 |
| 3 | 2017-04-21 |

## What is the total amount spent by each customer ?

```sql
SELECT
    s.userid, SUM(p.price) AS Total_amount_spent
FROM
    sales s
        INNER JOIN
    product p ON s.product_id = p.product_id
GROUP BY s.userid
ORDER BY s.userid;
```

| userid | Total_amount_spent |
|--------|--------------------|
| 1 | 5230 |
| 2 | 2510 |
| 3 | 4570 |

**Insights / Findings:**

- **User with ID 1 has spent the highest amount ( about 42.5% of total sales).**
- **User with ID 2 has spent the lowest amount (about 20.4% of total sales).**
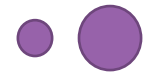
# How many days has each customer visited the store?

```sql
SELECT
    userid, COUNT(DISTINCT created_date) AS No_of_visits
FROM
    sales
GROUP BY userid
ORDER BY userid;
```

| userid | No_of_visits |
|--------|--------------|
| 1      | 7            |
| 2      | 4            |
| 3      | 5            |

**Insights / Findings:**

- **User with ID 1 has visited the most.**
- **User with ID 2 has visited the least.**

# What was the first product purchased by each customer?

```sql
Select userid, created_date,product_id from (
SELECT *, rank() over(partition by userid order by created_date) rnk from sales) a
where rnk=1;
```

| userid | created_date | product_id |
|--------|--------------|------------|
| 1 | 2016-03-11 | 1 |
| 2 | 2017-09-24 | 1 |
| 3 | 2016-11-10 | 1 |

**Insights / Findings:**

- **Product with ID 1 has been the first product ever purchased by all users.**

# What is the most purchased product on menu and how many times was it purchased by all customers?

```sql
SELECT
    userid,product_id,COUNT(product_id) AS total_purchase
FROM
    sales
WHERE
    product_id = (SELECT
            product_id
        FROM
            sales
        GROUP BY product_id
        ORDER BY COUNT(product_id) DESC
        LIMIT 1)
GROUP BY userid, product_id
ORDER BY userid;
```

| userid | product_id | total_purchase |
|--------|------------|----------------|
| 1 | 2 | 3 |
| 2 | 2 | 1 |
| 3 | 2 | 3 |

**Insights / Findings:**

- **Product with ID 2 has the highest number of sales.**
- **User with ID 1 & 3 purchased product 3 times and User with ID 2 purchased same 1 time.**

## Which item was the most popular for each customer?

```
select userid,product_id,cnt from
(select *, rank() over(partition by userid order by cnt desc) rnk from
(select userid, product_id, count(product_id) as cnt from sales group by userid, product_id) a) b
where rnk=1;
```

| userid | product_id | cnt |
|--------|-----------|-----|
| 1 | 2 | 3 |
| 2 | 3 | 2 |
| 3 | 2 | 3 |

**Insights / Findings:**

- **Product with ID 2 is the most purchased item by users with ID 1 & 3.**
- **Product with ID 3 is the most purchased item by user with ID 2.**

# Which item was purchased first by customer after they became gold member?

```sql
Select userid,product_id,created_date,gold_signup_date from
(Select *, rank() over(partition by userid order by created_date) rnk from
(Select s.userid, s.product_id, s.created_date, g.gold_signup_date from sales s inner join goldusers_signup g
on s.userid=g.userid  where s.created_date >= g.gold_signup_date) a) b where rnk=1;
```

| userid | product_id | created_date | gold_signup_date |
|--------|-----------|--------------|------------------|
| 1 | 3 | 2018-03-19 | 2017-09-22 |
| 3 | 2 | 2017-12-07 | 2017-04-21 |

**Insights / Findings:**

- Product with ID 3 is the first purchased item by user with ID 1 after they subscribe for gold membership.
- Product with ID 2 is the first purchased item by user with ID 3 after they subscribe for gold membership.

# Which item was purchased just before the customer became gold member?

```sql
Select userid,product_id,created_date,gold_signup_date from
(Select *, rank() over(partition by userid order by created_date desc) rnk from
(Select s.userid, s.product_id, s.created_date, g.gold_signup_date from sales s inner join goldusers_signup g
on s.userid=g.userid  where s.created_date <= g.gold_signup_date) a) b where rnk=1;
```

| userid | product_id | created_date | gold_signup_date |
|--------|-----------|--------------|------------------|
| 1 | 2 | 2017-04-19 | 2017-09-22 |
| 3 | 2 | 2016-12-20 | 2017-04-21 |

**Insights / Findings:**

- **Product with ID 3 is the first purchased item by user with ID 1 after they subscribe for gold membership.**
- **Product with ID 2 is the first purchased item by user with ID 3 after they subscribe for gold membership.**
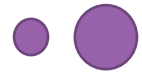
# What is the total orders and amount spent by each customer before they became member?

```sql
select userid, count(created_date) as total_orders, sum(price) as total_sales from
(Select s.userid, s.product_id, s.created_date, g.gold_signup_date, p.price
from sales s inner join goldusers_signup g on s.userid=g.userid inner join product p on s.product_id= p.product_id
where s.created_date <= g.gold_signup_date) a group by userid order by userid;
```

| userid | total_orders | total_sales |
|--------|--------------|-------------|
| 1      | 5            | 4030        |
| 3      | 3            | 2720        |

**Insights / Findings:**

- User with ID 1 has total sales of Rs. 4030 with total orders of 5 Whereas User with ID 3 has total sales of Rs. 2720 with total orders of 3 before they became member.
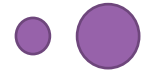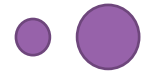
If buying each product generates points, e.g. 5Rs= 2 Points and each product has different purchasing points, for e.g for p1 5RS=1 Point, for p2 10Rs= 5 Point and p3 5Rs= 1 Point

Calculate total points collected by each customer against each product.

```sql
select *, round(total_amount/each_point_money,0) as total_points from (select *, case when product_id=1 then 5
when product_id=2 then 2 when product_id=3 then 5 else 0 end as each_point_money
from (Select s.userid, s.product_id, sum(p.price) as total_amount from sales s inner join product p
on s.product_id=p.product_id group by s.userid, s.product_id) a) b order by userid, total_points desc;
```

| userid | product_id | total_amount | each_point_money | total_points |
|--------|-----------|--------------|------------------|--------------|
| 1 | 2 | 2610 | 2 | 1305 |
| 1 | 1 | 1960 | 5 | 392 |
| 1 | 3 | 660 | 5 | 132 |
| 2 | 2 | 870 | 2 | 435 |
| 2 | 1 | 980 | 5 | 196 |
| 2 | 3 | 660 | 5 | 132 |
| 3 | 2 | 2610 | 2 | 1305 |
| 3 | 1 | 1960 | 5 | 392 |

If buying each product generates points, e.g. 5Rs= 2 Points and each product has different purchasing points, for e.g for p1 5RS=1 Point, for p2 10Rs= 5 Point and p3 5Rs= 1 Point
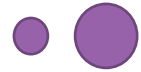
Calculate total points collected by each customer.

```sql
select userid, sum(total_points) as total_points_earned from (select *, round(total_amount/each_point_money,0) as total_points from
(select *, case when product_id=1 then 5 when product_id=2 then 2 when product_id=3 then 5 else 0 end as each_point_money
from (Select s.userid, s.product_id, sum(p.price) as total_amount from sales s inner join product p
on s.product_id=p.product_id group by s.userid, s.product_id) a) b) c group by userid order by userid;
```

| userid | total_points_earned |
|--------|---------------------|
| 1      | 1829                |
| 2      | 763                 |
| 3      | 1697                |

**Insights / Findings:**

- **User with ID 1 has earned highest points (1829) while User with ID 2 has earned highest points (763)**

If buying each product generates points, e.g. 5Rs= 2 Points and each product has different purchasing points, for e.g for p1 5RS=1 Point, for p2 10Rs= 5 Point and p3 5Rs= 1 Point
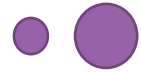
Calculate total money earned from total points.

```
select userid, sum(total_points)* 2.5 as total_money_earned from (select *, round(total_amount/each_point_money,0) as total_points
from (select *, case when product_id=1 then 5 when product_id=2 then 2 when product_id=3 then 5 else 0 end as each_point_money
from (Select s.userid, s.product_id, sum(p.price) as total_amount from sales s inner join product p
on s.product_id=p.product_id group by s.userid, s.product_id) a) b) c group by userid order by userid;
```

| userid | total_money_earned |
|--------|--------------------|
| 1      | 4572.5             |
| 2      | 1907.5             |
| 3      | 4242.5             |

**Insights / Findings:**

- **User with ID 1 has higher earnings as compared to others.**

If buying each product generates points, e.g. 5Rs= 2 Points and each product has different purchasing points, for e.g for p1 5RS=1 Point, for p2 10Rs= 5 Point and p3 5Rs= 1 Point

Calculate product with most points earned.

```sql
select * from (select *, rank() over(order by total_points_earned desc) rnk
from (select product_id, sum(total_points) as total_points_earned
from (select *, round(total_amount/each_point_money,0) as total_points
from (select *, case when product_id=1 then 5 when product_id=2 then 2 when product_id=3 then 5 else 0 end as each_point_money
from (Select s.userid, s.product_id, sum(p.price) as total_amount from sales s inner join product p
on s.product_id=p.product_id group by s.userid, s.product_id) a) b) c group by product_id) d) e where rnk=1;
```

| product_id | total_points_earned |
|------------|---------------------|
| 2          | 3045                |

**Insights / Findings:**

• **Product with ID 2 has resulted into higher number of earned points from its sale.**
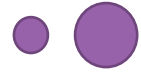
In the first one year after a customer joins the gold program (including joining date) irrespective of what customer has purchased they earn 5 points for every 10Rs spent. What was the points earnings of users in their first year?

```sql
select a.*, p.price* 0.5 as total_points from (Select s.userid, s.product_id, s.created_date, g.gold_signup_date
from sales s inner join goldusers_signup g on s.userid=g.userid
where s.created_date >= g.gold_signup_date and s.created_date <= date_add(g.gold_signup_date, interval 1 year)) a
inner join product p on a.product_id=p.product_id ;
```

| userid | product_id | created_date | gold_signup_date | total_points |
|--------|-----------|--------------|------------------|--------------|
| 3 | 2 | 2017-12-07 | 2017-04-21 | 435.0 |
| 1 | 3 | 2018-03-19 | 2017-09-22 | 165.0 |

**Insights / Findings:**

- User with ID 3 has earned 435 points while User with ID 1 has earned 165 points.