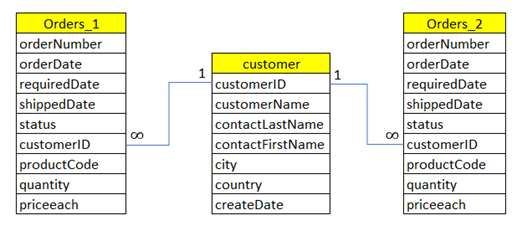
**Project Data Analysis for B2B Retail : Customer Analytics Report**

**Latar Belakang**

xyz.com adalah perusahaan rintisan B2B yang menjual berbagai produk tidak langsung kepada end user tetapi ke bisnis atau perusahaan lainnya. Sebagai data- driven company, maka setiap pengambilan keputusan di xyz.com selalu berdasarkan data. Setiap quarter xyz.com akan mengadakan townhall dimana seluruh atau perwakilan devisi akan berkumpul untuk me-review peformance perusahaan selama quarter terakhir.

Tabel yang digunakan

Dalam project ini tabel yang digunakan sebagai berikut.



1. Tabel orders\_1 : berisi data terkait transaksi penjualan periode quarter 1 ( Jan – Mar 2004)
2. Tbel orders\_2 : berisi data terkait transaksi penjualan periode quarter 2 (Apr – Jun 2004)
3. Tabel Customer : berisi data profil customer yang mendaftar menjadi customer xyz.com

Tugas dan Langkah

Sebagai seorang data analyst, kamu dimintai untuk menyediakan data dan analisa mengenai kondisi perusahaan bulan terakhir untuk dipresentasikan di townhall tersebut. (Asumsikan tahun yang sedang berjalan adalah tahun 2004).

Adapun hal yang akan direview adalah :

1. Bagaimana pertumbuhan penjualan saat ini?
2. Apakah jumlah customers xyz.com semakin bertambah ?
3. Dan seberapa banyak customers tersebut yang sudah melakukan transaksi?
4. Category produk apa saja yang paling banyak dibeli oleh customers?
5. Seberapa banyak customers yang tetap aktif bertransaksi?

Langkah yang akan dilakukan :

1. Menggunakan klausa “Select … From …” untuk mengambil data di database
2. Menggunakan klausa Where dan Operator untuk menfilter data
3. Menggunakan “group by”dan fungsi aggregat untuk aggregasi penjualan dan revenue
4. Menggunakan “order by” untuk mengurutkan data
5. Menggunakan “union” untuk menggabungkan tabel data penjualan
6. Menggunakan “date and time function” dan fungsi text untuk data manipulation
7. Menggunakan subquery untuk menyimpan hasil sementara untuk digunakan kembali dalam query.

Next >

**Memahami tabel**

Sebelum memulai membuat query SQL dan menganalisa dari hasil query yang kita buat, hal pertama yang perlu dilakukan adalah memahami tabel yang akan digunakan. Hal ini akan berguna dalam menentukan kolom mana yang berkaitan dengan problem yang akan dianalisa, dan proses manipulasi data apa yang perlu dilakukan untuk kolom – kolom tersebut.

Mengecek tabel orders\_1 : **SELECT\*FROM orders\_1 limit 5;**

Output dihasilkan:

+-------------+------------+--------------+-------------+---------+------------+-------------+----------+-----------+

| orderNumber | orderDate | requiredDate | shippedDate | status | customerID | productCode | quantity | priceeach |

+-------------+------------+--------------+-------------+---------+------------+-------------+----------+-----------+

| 10234 | 2004-03-30 | 2004-04-05 | 2004-04-02 | Shipped | 412 | S72\_1253 | 40 | 45690 |

| 10234 | 2004-03-30 | 2004-04-05 | 2004-04-02 | Shipped | 412 | S700\_2047 | 29 | 83280 |

| 10234 | 2004-03-30 | 2004-04-05 | 2004-04-02 | Shipped | 412 | S24\_3816 | 31 | 78830 |

| 10234 | 2004-03-30 | 2004-04-05 | 2004-04-02 | Shipped | 412 | S24\_3420 | 25 | 65090 |

| 10234 | 2004-03-30 | 2004-04-05 | 2004-04-02 | Shipped | 412 | S24\_2841 | 44 | 67140 |

+-------------+------------+--------------+-------------+---------+------------+-------------+----------+-----------+

Tabel orders\_1 . Transaksi penjualan periode quarter 1 ( Jan – Mar 2004)

Mengecek tabel orders-2 : **SELECT\*FROM orders\_2 limit 5;**

Output dihasilkan :

+-------------+------------+--------------+-------------+---------+------------+-------------+----------+-----------+

| orderNumber | orderDate | requiredDate | shippedDate | status | customerID | productCode | quantity | priceeach |

+-------------+------------+--------------+-------------+---------+------------+-------------+----------+-----------+

| 10235 | 2004-04-02 | 2004-04-12 | 2004-04-06 | Shipped | 260 | S18\_2581 | 24 | 81950 |

| 10235 | 2004-04-02 | 2004-04-12 | 2004-04-06 | Shipped | 260 | S24\_1785 | 23 | 89720 |

| 10235 | 2004-04-02 | 2004-04-12 | 2004-04-06 | Shipped | 260 | S24\_3949 | 33 | 55270 |

| 10235 | 2004-04-02 | 2004-04-12 | 2004-04-06 | Shipped | 260 | S24\_4278 | 40 | 63030 |

| 10235 | 2004-04-02 | 2004-04-12 | 2004-04-06 | Shipped | 260 | S32\_1374 | 41 | 90900 |

+-------------+------------+--------------+-------------+---------+------------+-------------+----------+-----------+

Tabel orders\_2 : Transaksi penjualan periode quarter 2 ( Apr – Jun 2004)

Mengecek tabel customer : **SELECT\*FROM customer limit 5;**

Output dihasilkan:

+------------+----------------------------+-----------------+------------------+-----------+-----------+------------+

| customerID | customerName | contactLastName | contactFirstName | city | country | createDate |

+------------+----------------------------+-----------------+------------------+-----------+-----------+------------+

| 103 | Atelier graphique | Schmitt | Carine | Nantes | France | 2004-02-05 |

| 112 | Signal Gift Stores | King | Jean | Las Vegas | USA | 2004-02-05 |

| 114 | Australian Collectors, Co. | Ferguson | Peter | Melbourne | Australia | 2004-02-20 |

| 119 | La Rochelle Gifts | Labrune | Janine | Nantes | France | 2004-02-05 |

| 121 | Baane Mini Imports | Bergulfsen | Jonas | Stavern | Norway | 2004-02-05 |

+------------+----------------------------+-----------------+------------------+-----------+-----------+------------+

Tabel customer xyz.com

Query yang dibuat ini memunculkan data 5 baris pertama, sehingga kita dapat melihat dan memahami masing-masing tabel yang akan digunakan.

**Total Penjualan dan Revenue pada Quarter -1(Jan, Feb, Mar) dan Quarter- 2 (Apr,Mei,Jun)**

Dalam menentukan total penjualan dan revenue pada quarter-1 dan quarter-2 dilihat pada tabel orders\_1 dan tabelk orders\_2. Hal yang kita lakukan sebagai berikut:

1. Kita tentukan dari tabel orders\_1 dengan melakukan penjumlahan pada kolom quantity dengan fungsi aggregat SUM() yang akan diberi nama sebagai “total\_penjualan”, dengan mengalikan quantitiy dengan kolom priceEach dan kemudian menjumlahkan hasil perkalian kedua kolom tersebut dan diberi nama sebagai “revenue”.
2. Perusahaan hanya ingin menghitung penjualan dari produk yang terkirim saja, sehingga kita perlu mem-filter kolom “status” yang hanya menampilkan order dengan statur “Shipped”.

Menentukan total penjulan dan revenue pada tabel orders\_1:

**SELECT SUM(quantity) as total\_penjualan, SUM(quantity\*priceeach) as revenue FROM orders\_1;**

Output :

+-----------------+-----------+

| total\_penjualan | revenue |

+-----------------+-----------+

| 8694 | 799579310 |

+-----------------+-----------+

Tabel total penjualan dan revenue pada tabel orders\_1 ( Quarter-1)

Menghitung penjulan produk yang hanya terkirim saja :

**SELECT SUM(quantity) as total\_penjualan, SUM(quantity\*priceeach) as revenue FROM orders\_1**

**WHERE status = 'Shipped';**

Output :

+-----------------+-----------+

| total\_penjualan | revenue |

+-----------------+-----------+

| 8694 | 799579310 |

+-----------------+-----------+

Tabel total penjualan dan revenue pada Quarter-1 dengan menghitung produk yang terkirim saja

Dari kedua output diatas dapat disimpulkan hasil total penjualan (8694) dan revenue (799579310) pada tabel order\_ 1 sama dengan hasil total penjualan dan revenue yang tekirim.

Untuk menentukan total penjulan dan revenue pada Quarter-2 dilakukan langkah yang sama, sehingga query nya sebagai berikut:

Menentukan jumlah total penjulan dan revenue pad atabel orders\_2 :

**SELECT SUM(quantitiy) as total\_penjualan, SUM(quantity\*priceeach) as revenue FROM orders\_2;**

Output :

+-----------------+-----------+

| total\_penjualan | revenue |

+-----------------+-----------+

| 8513 | 779271810 |

+-----------------+-----------+

Tabel total penjualan dan revenue pada tabel orders\_1 ( Quarter-1)

**SELECT SUM(quantitiy) as total\_penjualan, SUM(quantity\*priceeach) as revenue FROM orders\_2 WHERE status = ’Shipped’;**

Output :

+-----------------+-----------+

| total\_penjualan | revenue |

+-----------------+-----------+

| 6717 | 607548320 |

+-----------------+-----------+

Dari kedua tabel output Quarter – 2 sebelum di filter status “ Shipped” terdapat perbedaan total penjualan dan pendapatkan. Dapat disimpulkan custumer melakukan pembatalan pembelian sehingga perusahaan mengalami penurunan revenue sebesar 171723490.

**Menghitung Presentasi Keseluruhan Penjualan**

Untuk menghitung presentasi keseluruhan penjualan dilakukan penggabungan tabel orders\_1 dan orders\_2 dengan status “Shipped”.

Query: SELECT quarter , SUM ( quantity) AS total\_penjualan , SUM (quantity\*priceeach) AS revenue

FROM (SELECT orderNumber, status,quantity, priceeach, ‘1’ AS quarter FROM orders\_1

UNION

SELECT orderNumber, status, quantity, priceeach, ‘2’ AS quarter FROM orders\_2) AS table\_a

WHERE status= “Shipped’

GROUP BY quarter ;

Ouput :

+---------+-----------------+-----------+

| quarter | total\_penjualan | revenue |

+---------+-----------------+-----------+

| 1 | 8694 | 799579310 |

| 2 | 6717 | 607548320 |

+---------+-----------------+-----------+

Tabel total penjualan dan revenue pada Quarter – 1 dan Quarter-2 dengan menghitung produk yang terkirim saja

Untuk menghitung pertumbuhan keseluruhan penjualan dapat dilakukan dengan menghitung presentase penjumlahan bulan lalu dengan jumlah penjumlahan sebelumnya dan selisih pendapatan bulan terakhir dengan pendapatan periode sebelumnya. Yang diformulasikan sebagai berikut:

**%Y (t) = jumlah ( jumlah (t)) – jumlah (jumlah (t-1) )/ jumlah (jumlah (t-1))**

**%X(t) = jumlah (harga (t) – jumlah(harga(t-1))/ jumlah (harga (t-1))**

Sehingga presentase penjualan dan pendapatan quarter 1 dan quarter sebagai berikut :

% Growth Penjualan = (Total penjualan (quarter-2) – Total penjulan (quarter-1) / total penjualan quarter-1 x 100%

= (6717 – 8694) / 8694 x 100%

= - 22%

% Growth Revenue =(Revenue (quarter-2) – Revenue (quarter-1) )/ Revenue quarter-1 x 100%

= (607,548,320 – 799,579,310)/799,579,310 x 100%

= -24%

Disimpulkan nilai penjualan mengalami penurunan sebesar 22% dan pendapatan mengalami penurunan 24%.

**Apakah jumlah customer xyz.com semakin bertambah?**

Penambahan jumlah customer dapat diukur dengan membandingkan total jumlah customers yang terdaftar di periode saat ini dengan total jumlah customers yang terdaftar diakhir periode selanjutnya. Untuk mengetahui total customer pada quarter-1 dan quarter-2 pada periode antara 1 Januari 2004 dan 30 Juni 2004, digunakan query berikut:

SELECT quarter, COUNT(DISTINCT customerID) AS total\_customers FROM(SELECT customerID,createDate,quarter(createDate) AS quarter FROM customer WHERE createDate BETWEEN ‘2004-01-01’ AND ‘2004-06-01’ ) AS tabel\_b

GROUP BY quarter;

Output :

+---------+-----------------+

| quarter | total\_customers |

+---------+-----------------+

| 1 | 43 |

| 2 | 35 |

+---------+-----------------+

Tabel total customers pada periode 1 januari 2004 dan 30 juni 2004

Dengan periode waktu yang sama kita melihat seberapa banyak customer yang terdaftar di periode quarter-1 dan quarter- 2 yang telah melakukan transaksi, dapat dibuat query sebagai berikut:

**SELECT QUARTER(createDate) AS quarter, COUNT(DISTINCT customerID) AS total\_customers FROM customer**

**WHERE customerID IN (SELECT customerID from ORDERS\_1 UNION SELECT customerID FROM orders\_2) AND createDate BETWEEN ‘2004-01-01 AND ‘2004-06-30’**

**GROUP BY quarter;**

Output :

+---------+-----------------+

| quarter | total\_customers |

+---------+-----------------+

| 1 | 25 |

| 2 | 19 |

+---------+-----------------+

Tabel. Total customer yang melakukan transaksi pada periode 1 Januari 2004 dan 30 Juni 2004

Ketertarikan cutomer untuk berbelanja di xyz.com dapat ditentukan presentase nya dengan menghitung jumlah customer berbelanja dibagi dengan total customer, sehingga:

= = 56.4%

Presentase ketertarikan customers baru berbelanja rata-rata 56.4%.

**Category produk apa saja yang paling banyak di order oleh customers di Quarter-2?**

Katagori produk yang paling banyak dibeli dapat dilakukan dengan menghitung total order dan jumlah penjualan dari setiap produk.

Kita akan melihat katagori produk apa saja yang paling banyak di order pada quarter-2 dengan query :

**SELECT \* FROM (SELECT categoryID , COUNT(DISTINCT orderNumber) , SUM(quantity) AS total\_penjualan**

**FROM (SELECT productCode, orderNumber, quantity, status,LEFT(productCode,3) AS categoryID**

**FROM orders\_2**

**WHERE status =’Shipped’) AS tabel\_c**

**GROUP BY categoryID) a ORDER BY total\_order DESC;**

Output:

+------------+-------------+-----------------+

| categoryID | total\_order | total\_penjualan |

+------------+-------------+-----------------+

| S18 | 25 | 2264 |

| S24 | 21 | 1826 |

| S32 | 11 | 616 |

| S12 | 10 | 491 |

| S10 | 8 | 492 |

| S50 | 8 | 292 |

| S70 | 7 | 675 |

| S72 | 2 | 61 |

+------------+-------------+-----------------+

Presentase total order dan total penjulan paling tinggi pada categoryID S18 dan S24. CategoryID S18 dengan presentase total order 27.17 % sedangakan S24 dengan presentase 22.83%. sehingga produk dengan categoryID S18 dan S24 berkontribusi total order sebesar 50% dengan presentase total penjualan S18 dan S24 sebesar 33.71 % dan 27.18% dengan memberi kontribusi 60% dari total penjualan.

**Seberapa banyak customers yang tetap aktif bertransaksi setelah transaksi pertamanya?**

Mengetahui seberapa banyak customers yang tetap aktif menunjukkan apakah xyz.com digemari oleh customers untuk memesan kebutuhan bisnis mereka. Hal ini juga dapat menjadi dasar bagi tim product dan business untuk pengembangan product dan business kedepannya. Oleh karena itu dilakukan analisa retention rate customer yang dihitung dari customers yang berbelanja di Quarter 1 dan kembali berbelanja di Quarter 2. Hal pertama kita menghitung total unik customers yang melakukan transaksi di Quarter-1, query nya :

**SELECT COUNT(DISTINCT customerID) AS total\_customers FROM orders\_1;**

Output :

+-----------------+

| total\_customers |

+-----------------+

| 25 |

+-----------------+

**SELECT 1 AS quarter, (COUNT(DISTINCT customersID)\*100)/25 AS Q2 FROM orders\_1 WHERE customerID IN (SELECT DISTINCT**

**customerID FROM orders\_2);**

output :

+---------+---------+

| quarter | Q2 |

+---------+---------+

| 1 | 24.0000 |

+---------+---------+

Kesimpulan berdasarkan data yang diperoleh melalui query SQL, yaitu:

1. Performance xyz.com menurun signifikan di quarter ke -2 , terlihat nilai penjualan dan revenue yang drop hingga 22 % dan 24%.
2. Perolehan customer baru juga tidak terlalu baik , dan sedikit menurun dibandingkan quarter sebelumnya.
3. Ketertarikan cusomer baru untuk berbelanja di xyz.com masih berkurang , hanya sekitar 56% yang sudah bertransaksi . Disarankan tim Produk untuk mempelajari behaviour customer dan melakukan product improvement , sehingga conversion rate (register to transaction) dapat memingkat.
4. Produk kategori S18 dan S24 berkontribusi 50% dari total order dan 60% dari total penjualan , sehingga xyz.com sebaiknya fokus untuk pengembangan category S18 dan S24
5. Retention rate customer xyz.com juga sangat rendah yaitu hanya 24% , artinya banyak customer yang sudah bertransaksi di quarter – 1 tidak kembali melakukan order di quarter ke-2 atau tidak melakukan repeat order.
6. Xyz.com mengalami pertimbihan negatif di quarter-2 dan perlu melakukan banyak improvement baik itu di sisi produk dan bisnis marketing , jika ingin mencncapai target positif growth di quarter ke-3. Rendahnya retention rate dan conversation rate biasa menjadi diagnosa awal bahwa customer tidak tertarik atau kurang puas ataupun kecewa berbelanja di xyz.com.