

Assessment Requirement

SQL Foundation

- Peserta mampu melakukan dan menerapkan query yang lebih kompleks.
- Peserta dapat menemukan minimal 10 insight menarik dari data dan menggunakannya untuk membuat keputusan bisnis yang tepat.
- Peserta memiliki kemampuan untuk mentransformasikan data menjadi narasi yang mudah dipahami, sehingga orang-orang tanpa latar belakang data pun dapat memperoleh pemahaman dari informasi tersebut.

SC61015_La Ode Muhammad Sampaga_Foundation for Data Science

PRODUCT KNOWLEDGE

CASE 1: Perusahaan Menginginkan Brand dan Kategori Produk apa saja yang Menghasilkan Penjualan Tertinggi.

a. Brand Product dengan Penjualan Tertinggi

	product_brand	Jumlah_Penjualan ▾
1.	Allegra K	6.242
2.	Calvin Klein	3.199
3.	Carhartt	2.592
4.	Hanes	1.974
5.	Volcom	1.859

Allegra K memiliki perbedaan yang signifikan dalam Jumlah Penjualan dibandingkan dengan Brand Lainnya.

b. Product Category dengan Penjualan Tertinggi

	product_category	Jumlah_Penjualan ▾
1.	Intimates	13.344
2.	Jeans	12.627
3.	Tops & Tees	11.951
4.	Fashion Hoodies & Sweatshirts	11.736
5.	Sleep & Lounge	11.288

Kategori **Intimates** memiliki jumlah Penjualan Terbesar namun Perbedaan dalam **Product Category** Tidak terlalu Signifikan.

c. Brand Product dengan Penghasilan Tertinggi

	Product_Brand	revenue
1.	Calvin Klein	300.242
2.	Diesel	278.405,09
3.	7 For All Mankind	271.241,87
4.	Carhartt	264.929,07
5.	True Religion	245.114,91
6.	Tommy Hilfiger	164.518,74
7.	Volcom	159.883,28
8.	Joe's Jeans	155.796,08
9.	Quiksilver	152.794,26
10.	Columbia	151.043,68

Calvin Klein menempati posisi pertama dengan Penjualan Tertinggi

INSIGHT:

Setelah mendapatkan Data Kategori dan Brand Produk dengan Penjualan Tertinggi Kita Dapat Melakukan:

- **Optimalkan Potensi Penjualan:** Menggunakan data produk dengan jumlah pembelian tertinggi sebagai sumber kekuatan untuk penjualan selanjutnya.
- **Pemasaran yang Terarah dan Efektif:** Dengan memahami preferensi pelanggan, kita dapat merancang strategi pemasaran yang lebih terarah dan efektif.
- **Memperkuat Merek dan Memperkenalkan Produk Baru:** Pemahaman yang mendalam tentang preferensi pelanggan memungkinkan kita untuk memperkuat merek-merek yang kuat (Seperti *Allegria K*) dan memperkenalkan produk-produk baru yang menarik, Ini akan memperluas basis pelanggan dan meningkatkan loyalitas.
- **Kolaborasi dengan Pemasok dan Distributor:** Berkolaborasi dengan pemasok dan distributor akan memungkinkan kita untuk memenuhi kebutuhan pasar dengan lebih efisien, mengoptimalkan rantai pasokan dan meningkatkan ketersediaan produk.

QUERY:

```
4  --JUMLAH PEMBELIAN TERBANYAK UNTUK PRODUCT CATEGORY
5  SELECT
6  |   product_category as product_category,
7  |   COUNT(product_category) as Jumlah
8  FROM `bigquery-public-data.thelook_ecommerce.inventory_items`
9  WHERE sold_at IS NOT NULL
10 GROUP BY product_category
11 ORDER BY Jumlah DESC
12 LIMIT 5;
13
14 --JUMLAH PEMBELIAN TERBANYAK UNTUK PRODUCT BRAND
15 SELECT
16 |   product_brand as product_brand,
17 |   COUNT(product_brand) as Jumlah
18 FROM `bigquery-public-data.thelook_ecommerce.inventory_items`
19 WHERE sold_at IS NOT NULL
20 GROUP BY product_brand
21 ORDER BY Jumlah DESC
22 LIMIT 5;
23
24 -- BRAND PRODUCT DENGAN PENGHASILAN TERTINGGI
25 SELECT
26 |   p.brand AS Product_Brand,
27 |   ROUND(SUM(oi.sale_price * o.num_of_item), 2) AS revenue,
28 |   SUM(o.num_of_item) AS Jumlah_Penjualan
29 FROM
30 |   bigquery-public-data.thelook_ecommerce.order_items as oi,
31 |   bigquery-public-data.thelook_ecommerce.orders as o,|
32 |   bigquery-public-data.thelook_ecommerce.products as p
33 WHERE
34 |   oi.order_id = o.order_id
35 |   AND oi.product_id = p.id
36 |   AND oi.status NOT IN ('Cancelled', 'Returned')
37 GROUP BY
38 |   p.brand
39 ORDER BY
40 |   Jumlah_Penjualan DESC;
41
```

CASE 2: Perusahaan ingin tau Brand apa saja yang menghasilkan Penjualan

Terendah

Row	Product_Brand	revenue	Jumlah_Penjualan
1	Snob	5.9	1
2	July Women	8.95	1
3	MMG	8.99	1
4	Custom '77	15.0	1
5	Kikit	17.99	1
6	LLange Maternity	18.99	1
7	Ms Maggie	19.49	1
8	Slimming Tees	19.99	1
9	Zicac	19.99	1
10	Estee Lauder	21.99	1

INSIGHT:

Pada E-Commerce Kita masih memiliki banyak Brand dengan Penjualan dibawah 10.

Oleh Karena itu,Kita dapat melakukan pendekatan kepada brand terkait untuk menawarkan Strategi Pemasaran yang akan menguntungkan kedua belah pihak.

Seperti Flash Sale, Diskon, DLL.

QUERY:

```
24 -- BRAND PRODUCT DENGAN PENGHASILAN TERENDAH
25 SELECT
26     p.brand AS Product_Brand,
27     ROUND(SUM(oi.sale_price * o.num_of_item), 2) AS revenue,
28     SUM(o.num_of_item) AS Jumlah_Penjualan
29 FROM
30     bigquery-public-data.thelook_ecommerce.order_items as oi,
31     bigquery-public-data.thelook_ecommerce.orders as o,
32     bigquery-public-data.thelook_ecommerce.products as p
33 WHERE
34     oi.order_id = o.order_id
35     AND oi.product_id = p.id
36     AND oi.status NOT IN ('Cancelled', 'Returned')
37 GROUP BY
38     p.brand
39 ORDER BY
40     jumlah_penjualan ASC
41     LIMIT 10;
42
```

CASE 3: Perusahaan Ingin Melihat Status Pengiriman terhadap Product Mereka

Row	status ▼	jumlah ▼
1	Shipped	53653
2	Complete	45606
3	Processing	36670
4	Cancelled	27304
5	Returned	18216

INSIGHT:

Dalam analisis ini, terlihat bahwa status pengiriman terbanyak adalah "**Shipped**" dan "**Complete**", namun terdapat sebanyak **25%** pengiriman yang mengalami **pembatalan** dan **pengembalian**. Ketidakmampuan menangani masalah ini dapat berdampak negatif pada bisnis kita.

Oleh karena itu, kami perlu mengidentifikasi produk-produk apa saja yang paling sering mengalami pembatalan dan pengembalian.

QUERY:

```
3 -- TOTAL STATUS PENGIRIMAN BARANG
4 SELECT status, COUNT(status) as jumlah
5 FROM `bigquery-public-data.thelook_ecommerce.order_items`
6 GROUP BY status
7 ORDER BY jumlah DESC;
```

CASE 4: Product Apa saja yang paling sering di Returned dan Cancelled

	brand	cancelled_count ▾	returned_count	Percentage
1.	Allegra K	923	642	0,25
2.	Calvin Klein	467	317	0,25
3.	Carhartt	367	286	0,25
4.	Quiksilver	288	193	0,27
5.	Hanes	282	181	0,23

INSIGHT:

Berikut Daftar produk dengan jumlah pengembalian dan pembatalan tertinggi yang telah disusun.

Diberikan Juga Persentase Pengembalian dan Pembatalan agar tidak terjadi kesalah pahaman karena jumlah Pengembalian dan Pembatalan dapat dipengaruhi oleh Jumlah Penjualan yang tinggi.

Salah satu strategi penyelesaian yang disarankan adalah berkomunikasi dan melakukan investigasi dengan Brand terkait untuk memahami permasalahan pada produk.

QUERY:

```
61 -- Barang yang sering di return dan cancel|
62 WITH cancelled_returned_counts AS (
63     SELECT
64         p.brand AS brand,
65         SUM(CASE WHEN oi.status = 'Cancelled' THEN 1 ELSE 0 END) AS cancelled_count,
66         SUM(CASE WHEN oi.status = 'Returned' THEN 1 ELSE 0 END) AS returned_count,
67         COUNT(*) AS total_count
68     FROM bigquery-public-data.thelook_ecommerce.order_items oi
69     INNER JOIN bigquery-public-data.thelook_ecommerce.products p
70     ON oi.product_id = p.id
71     GROUP BY brand
72 )
73
74 SELECT
75     brand,
76     cancelled_count,
77     returned_count,
78     ROUND(((returned_count + cancelled_count) / total_count),2) AS Percentage,
79 FROM
80     cancelled_returned_counts
81 ORDER BY
82     cancelled_count DESC;
```

CASE 5: Perusahaan Ingin Tahu Siapa Saja User yang Paling Sering Men-Cancel dan Return Product

Row	user_id	Total_Cancel_Return	gender	email
1	27568	4	F	lauriethomas@example.com
2	60911	4	M	keiththompson@example.com
3	28285	4	M	ronniestewart@example.com
4	85816	4	F	mariapalmer@example.net
5	89751	4	M	luisstanley@example.org

INSIGHT:

Kita juga tidak bisa mengenyampingkan kemungkinan User-User yang tidak bertanggung jawab dengan mengembalikan dan membatalkan pesanan dengan sengaja.

Oleh karena itu saya sudah mengumpulkan 5 User yang paling sering melakukan hal tersebut.

Solusi untuk masalah ini adalah Memberikan Pendekatan kepada User melewati Email yang ada pada Database, untuk meminta Penjelasan dan Saran dari User yang Berkaitan.

QUERY:

```
23 --USER YANG SUKA MEN-CANCEL BARANG
24 WITH Cancelled_Users AS (
25     SELECT
26         user_id,
27         SUM(CASE WHEN status IN ('Cancelled', 'Returned') THEN 1 ELSE 0 END) AS Total_Cancel_Return
28     FROM
29         `bigquery-public-data.thelook_ecommerce.orders`
30     WHERE
31         status IN ('Returned', 'Cancelled')
32     GROUP BY
33         user_id
34 )
35
36 SELECT
37     cu.user_id,
38     cu.Total_Cancel_Return,
39     u.gender,
40     u.email
41 FROM
42     Cancelled_Users cu
43 JOIN
44     `bigquery-public-data.thelook_ecommerce.users` u
45 ON
46     cu.user_id = u.id
47 ORDER BY cu.Total_Cancel_Return DESC
48 LIMIT 5;
```

USER INFORMATION

CASE 6: Negara apa yang memiliki User Terbanyak dan Berapa Rata-Rata Umur di Setiap Negara?

a. Negara Dengan Jumlah User Terbanyak dan Rata-Rata Umur disetiap Negara.

	country	Jumlah_User ▾	AverageAge
1.	China	34.188	41
2.	United States	22.547	41
3.	Brasil	14.382	41
4.	South Korea	5.326	41
5.	France	4.634	41

b. Rentang Usia yang dimiliki User Ecommerce kami.

	Umur	Total_User ▾
1.	Dewasa	27.136
2.	Lansia	27.046
3.	Remaja Dewasa	13.468
4.	Remaja	8.286
5.	Anak-Anak	4.071

INSIGHT:

Negara **China** memiliki jumlah pengguna terbanyak, dengan rata-rata **usia 41 Tahun** Sama di seluruh Negara.

Dari data tersebut, juga dapat disimpulkan bahwa pengguna kami didominasi oleh kalangan **dewasa hingga lansia**.

Kita dapat meningkatkan **strategi pemasaran** dengan menargetkan pengguna di China dan mengadaptasi kebudayaan lokal.

Langkah-langkah seperti penyesuaian antarmuka pengguna (UI/UX) serta pelaksanaan promosi khusus selama perayaan **Tahun Baru Imlek** dapat menjadi pilihan. Pendekatan ini tidak hanya akan menarik perhatian pengguna di China, tetapi juga di negara-negara Asia lainnya seperti Korea, Jepang, dan sebagainya.

QUERY:

```
4  -- NEGARA DAN AVG UMUR
5  SELECT country, count(country) as Jumlah_User, ROUND(AVG(age),0) as AverageAge
6  FROM `bigquery-public-data.thelook_ecommerce.users`
7  GROUP BY country
8  ORDER BY Jumlah_User DESC
9  LIMIT 5;

11 -- JUMLAH UMUR
12 SELECT
13     CASE
14         WHEN u.age < 15 THEN 'Anak-Anak'
15         WHEN u.age BETWEEN 15 AND 20 THEN 'Remaja'
16         WHEN u.age BETWEEN 20 AND 30 THEN 'Remaja Dewasa'
17         WHEN u.age BETWEEN 30 AND 50 THEN 'Dewasa'
18         WHEN u.age > 50 THEN 'Lansia'
19     END AS Umur,
20     COUNT(DISTINCT oi.user_id) AS Total_User
21 FROM bigquery-public-data.thelook_ecommerce.order_items AS oi
22 INNER JOIN bigquery-public-data.thelook_ecommerce.users AS u
23 ON oi.user_id = u.id
24 GROUP BY Umur
25 ORDER BY Total_User DESC;
```

CASE 7: Perusahaan Ingin Mengetahui User Aplikasi Mereka didominasi oleh Gender tertentu atau Tidak

gender ▼	Jumlah ▼
F	50045
M	49955

INSIGHT:

Dalam dataset ini, gender pengguna tidak begitu berpengaruh terhadap pertimbangan bisnis karena tidak terdapat perbedaan yang signifikan, baik laki-laki maupun perempuan memiliki persentase sebesar 50%.

QUERY:

```
27 -- JUMLAH GENDER
28 SELECT gender, count(gender) as Jumlah
29 FROM `bigquery-public-data.thelook_ecommerce.users`
30 GROUP BY gender
31 ORDER BY jumlah DESC;
32 |
```

USER BEHAVIOUR

CASE 8: Perusahaan Ingin Memberikan Reward kepada User-User Yang loyal Membeli di Ecommerce mereka.

Row	user_id	email	total_purchase
1	96463	victormiller@example.org	6204.3
2	98024	danieljohnson@example.com	5617.27
3	82782	jeffreysshort@example.net	5384.43
4	52044	christophergonzalez@example...	5126.79
5	24684	thomasshah@example.org	5116.9
6	14768	jeffreybates@example.org	5113.96
7	52474	collinstewart@example.com	4816.12
8	15659	cherylbond@example.com	4720.72
9	24158	toddgentry@example.com	4521.17
10	66132	josephaustin@example.org	4486.68

INSIGHT:

Dengan Mendapatkan Siapa Saja User yang Loyal Berbelanja di Aplikasi Ecommerce ini, Kami dapat Memberikan Reward melalui Email yang ada pada Database.

QUERY:

```
39 -- Customer Loyal
40 SELECT
41     u.id AS user_id,
42     u.email AS email,
43     ROUND(SUM(oi.sale_price * o.num_of_item), 2) AS total_purchase
44 FROM bigquery-public-data.thelook_ecommerce.order_items AS oi
45 INNER JOIN bigquery-public-data.thelook_ecommerce.orders AS o
46 ON oi.order_id = o.order_id
47 INNER JOIN bigquery-public-data.thelook_ecommerce.users AS u
48 ON o.user_id = u.id
49 GROUP BY user_id, email
50 ORDER BY total_purchase DESC
51 LIMIT 10;
```

CASE 9: Dari Platform Mana User Mengunjungi atau Mengenal Aplikasi Ecommerce Kami?

	traffic_source	Jumlah ▾
1.	Search	70.137
2.	Organic	14.982
3.	Facebook	5.950
4.	Email	4.933
5.	Display	3.998

INSIGHT:

Traffic Source yang paling banyak membawa User adalah **Search** dengan perbedaan Signifikan dari yang lainnya.

Kita dapat terus mempertahankan Teknik **SEO** kita dan Melakukan **peningkatan** untuk Traffic Source yang lainnya.

QUERY:

```
33 -- Traffic Source
34 SELECT traffic_source, count(traffic_source) as Jumlah
35 FROM `bigquery-public-data.thelook_ecommerce.users`
36 GROUP BY traffic_source
37 ORDER BY jumlah DESC;
```

CASE 10: Marketing Channel apa saja yang Paling Sering Mendatangkan Pelanggan?

traffic_source	Total_Visits	Total_Purchases	Conversion_Rate...
Adwords	725.766	54.334	7,49
Facebook	243.957	18.241	7,48
Email	1.088.876	81.344	7,47
YouTube	243.049	18.120	7,46
Organic	122.315	9.116	7,45

INSIGHT:

Dari tabel di atas, terlihat bahwa penyebaran iklan melalui **Adwords dan Email** adalah yang paling efektif dalam mendorong pengguna untuk mengklik aplikasi kami.

Rata-rata persentase pengguna yang membeli setelah melihat aplikasi e-commerce kami adalah **7,4%** untuk setiap teknik Traffic Source.

QUERY:

```
12 -- note: conversion rate
13 WITH visit AS (
14     SELECT
15         traffic_source,
16         COUNT(traffic_source) as Total_Visit
17     FROM `bigquery-public-data.thelook_ecommerce.events`
18     GROUP BY traffic_source
19 ),
20 purchase AS (
21     SELECT
22         traffic_source,
23         COUNT(traffic_source) as Total_Purchase
24     FROM `bigquery-public-data.thelook_ecommerce.events`
25     WHERE event_type = 'purchase'
26     GROUP BY traffic_source
27 )
28
29 SELECT
30     COALESCE(v.traffic_source, p.traffic_source) as traffic_source,
31     COALESCE(Total_Visit, 0) as Total_Visit,
32     COALESCE(Total_Purchase, 0) as Total_Purchase,
33     ROUND((COALESCE(Total_Purchase, 0) / COALESCE(Total_Visit, 1)) * 100, 2) AS conversion_rate
34 FROM
35     visit v
36 FULL OUTER JOIN
37     purchase p ON v.traffic_source = p.traffic_source
38 ORDER BY
39     Total_Visit DESC;
```