

SCID : SC61920  
Nama Lengkap : Calvin Adiwinata  
Nama Chapter : Foundation for Data Science

### Case Study 1

Mencari tahu produk kategori apa yang paling laku dan tidak laku di kalangan pengguna, sehingga perusahaan dapat membuat keputusan produk kategori mana yang perlu dilakukan pemasaran lebih lanjut sehingga dapat meningkatkan penjualan.

### Insight

Paling banyak disorder

Row	category	jumlah_produk_yg_diorder
1	Intimates	3323
2	Jeans	3163
3	Tops & Tees	3022
4	Fashion Hoodies & Sweatshirts	3001
5	Swim	2862
6	Sweaters	2859
7	Sleep & Lounge	2808
8	Shorts	2806
9	Accessories	2520

Dari hasil query yang dilakukan, diketahui produk yang ada di produk kategori Intimates adalah produk yang paling banyak disorder.

Sehingga perusahaan mungkin dapat menambah stok produk untuk produk kategori intimates, atau mungkin juga pengembangan produk baru, karena sudah diketahui intimates adalah produk kategori yang paling laku, sehingga kemungkinan produk baru dibeli mungkin juga tinggi.

Paling sedikit diorder

Row	category	jumlah_produk_yg_diorder
1	Clothing Sets	52
2	Jumpsuits & Rompers	218
3	Suits	292
4	Skirts	528
5	Leggings	791
6	Pants & Capris	885
7	Blazers & Jackets	890
8	Socks & Hosiery	996
9	Plus	1086
10	Suits & Sport Coats	1300

Dari hasil query yang dilakukan, diketahui produk yang ada di produk kategori clothing sets adalah produk yang paling sedikit diorder.

Sehingga perusahaan dapat melakukan analisis lebih lanjut, seperti mengapa clothing sets kurang begitu diminati oleh customer.

Perusahaan juga dapat meningkatkan pemasaran untuk produk kategori tersebut, sehingga dapat menarik perhatian calon customer.

Query

```
select
  p.category,
  count(oi.product_id) jumlah_produk_yg_diorder
from bigquery-public-data.thelook_ecommerce.order_items oi
  join bigquery-public-data.thelook_ecommerce.products p
  on oi.product_id = p.id
where status = 'Complete'
group by p.category
order by 2 desc
```

```
select
  p.category,
  count(oi.product_id) jumlah_produk_yg_diorder
from bigquery-public-data.thelook_ecommerce.order_items oi
  join bigquery-public-data.thelook_ecommerce.products p
  on oi.product_id = p.id
where status = 'Complete'
group by p.category
order by 2
```

## Case Study 2

Mencari tahu di umur berapa pengguna paling banyak melakukan order, sehingga perusahaan dapat melakukan analisis lebih lanjut untuk penargetan pasar untuk kelompok umur tertentu.

### Insight

Row	age_group	jumlah_order
1	old	44673
2	middle age	42329
3	young adult	21106
4	teenager	16671

Dari hasil query yang dilakukan, diketahui kelompok umur 'old' (diatas 50) dan 'middle age' (30-49) sangat banyak melakukan order, untuk kelompok 'old' sejumlah 44673, dan 'middle age' sejumlah '42329'.

Sehingga perusahaan mungkin dapat melakukan pemasaran produk yang lebih banyak menargetkan kategori 'old' dan 'middle age'.

### Query

```
select
  case
    when u.age <10 then 'child'
    when u.age between 10 and 19 then 'teenager'
    when u.age between 20 and 29 then 'young adult'
    when u.age between 30 and 49 then 'middle age'
    else 'old'
  end as age_group,
  count(o.order_id) jumlah_order
from bigquery-public-data.thelook_ecommerce.users u
  join bigquery-public-data.thelook_ecommerce.orders o
  on o.user_id = u.id
group by age_group
order by 2 desc
```

### Case Study 3

Mencari tahu pengguna dari negara mana yang paling menguntungkan untuk perusahaan, sehingga perusahaan mungkin dapat menghemat waktu dan uang yang dihabiskan untuk pemasaran di negara-negara yang kurang menguntungkan dan mengalokasikan ke negara yang memberi keuntungan lebih banyak.

#### Insight

Row	country	jumlah_order	c_total_revenue	total_profit
1	China	17925	3234501.103	994596.433
2	United States	11863	2058370.771	637963.874
3	Brasil	7763	1378960.871	428647.992
4	South Korea	2743	490774.04	150151.182
5	France	2405	435141.39	135837.51
6	Germany	2313	419395.911	130285.359
7	United Kingdom	2459	415639.24	129680.018
8	Spain	2094	363816.13	116872.109
9	Japan	1304	231833.72	70328.051
10	Australia	1141	216446.79	66509.574

Dari hasil query yang dilakukan, didapatkan informasi bahwa pengguna user dari negara China adalah pengguna yang melakukan order paling banyak, dengan total pengeluaran uang pengguna sekitar 3 juta, dan didapatkan profit sebesar sekitar 1 juta.

Sehingga perusahaan dapat menyesuaikan produk yang lebih sesuai dengan keinginan customer di China, atau mungkin memperluas jaringan perusahaan seperti menambah distributor atau mitra di negara tersebut.

#### Query

```
with cte_1 as(
select
  o.order_id,
  u.country,
  o.num_of_item,
  oi.status,
  p.cost,
  p.retail_price,
  (o.num_of_item*p.retail_price) revenue,
  (p.retail_price-p.cost) profit_per_item
from bigquery-public-data.thelook_ecommerce.users u
join bigquery-public-data.thelook_ecommerce.orders o
on o.user_id = u.id
join bigquery-public-data.thelook_ecommerce.order_items oi
```

```
    on oi.user_id = u.id
  join bigquery-public-data.thelook_ecommerce.products p
    on p.id = oi.product_id)
select
  country,
  count(distinct order_id) as jumlah_order,
  round(sum(revenue),3) c_total_revenue,
  round(sum(cte_1.profit_per_item),3) total_profit
from cte_1
where status = 'Complete'
group by country
order by total_profit desc
```

## Case Study 4

Mencari tahu apakah jumlah pesanan meningkat atau menurun dalam jangka waktu tertentu, sehingga perusahaan dapat memahami tren penjualan yang sedang terjadi.

### Insight

Row	quarter_time	jumlah_order
1	2024 Quarter 1	19819
2	2023 Quarter 4	17254
3	2023 Quarter 3	13529
4	2023 Quarter 2	11037
5	2023 Quarter 1	9381
6	2022 Quarter 4	8367
7	2022 Quarter 3	7195
8	2022 Quarter 2	6203
9	2022 Quarter 1	5415

Dari hasil query yang dilakukan, didapatkan informasi bahwa setiap waktunya, jumlah order semakin bertambah.

Sehingga perusahaan dapat mempertimbangkan strategi pemasaran yang perlu dilakukan. Dilihat dari hasil query, karena order semakin bertambah maka dapat disimpulkan strategi yang saat ini dilakukan sudah berhasil.

### Query

```
with cte_1 as(
select
  o.order_id,
  date_trunc(o.created_at, quarter) AS quarter,
from bigquery-public-data.thelook_ecommerce.orders o
)
select
  concat(extract(year from quarter), ' Quarter ', extract(quarter from quarter)) quarter_time,
  count(order_id) as jumlah_order
from cte_1
group by quarter
order by quarter desc
```

## Case Study 5

Mencari tahu untuk setiap age group, produk kategori apa yang paling banyak dibeli, sehingga perusahaan dapat menargetkan produk dengan lebih tepat sesuai dengan umur user.

## Insight

Row	age_group	product_category	jumlah_order	rank
1	young adult	Jeans	3850	1
2	young adult	Intimates	3782	2
3	young adult	Tops & Tees	3741	3
4	teenager	Jeans	3228	1
5	teenager	Fashion Hoodies & Sweatshirts	3041	2
6	teenager	Intimates	2929	3
7	old	Jeans	8201	1
8	old	Tops & Tees	7872	2
9	old	Intimates	7726	3
10	middle age	Jeans	7974	1
11	middle age	Tops & Tees	7383	2
12	middle age	Intimates	7363	3

Dari hasil query yang dilakukan, didapatkan informasi bahwa kebanyakan produk kategori 'Jeans', 'Tops & Tees', dan 'Intimates', dengan pengecualian teenager (10-19) yang juga sering membeli produk kategori 'fashion hoodies & sweatshirts' sehingga masuk ke peringkat 2.

Sehingga perusahaan mungkin dapat menyesuaikan stok produk dari produk kategori yang paling banyak dibeli, dan juga penargetan pemasaran produk untuk setiap kelompok umur.

## Query

```
with cte_1 as(
select
  case
    when u.age <10 then 'child'
    when u.age between 10 and 19 then 'teenager'
    when u.age between 20 and 29 then 'young adult'
    when u.age between 30 and 49 then 'middle age'
    else 'old'
  end as age_group,
  ii.product_category,
  count(distinct o.order_id) jumlah_order
from bigquery-public-data.thelook_ecommerce.users u
join bigquery-public-data.thelook_ecommerce.orders o
on o.user_id = u.id
join bigquery-public-data.thelook_ecommerce.order_items oi
```

```
on oi.user_id = u.id
join bigquery-public-data.thelook_ecommerce.products p
on p.id = oi.product_id
join bigquery-public-data.thelook_ecommerce.inventory_items ii
on ii.product_id = p.id
group by age_group,ii.product_category),
cte_2 as(
select
*,
rank() over (partition by age_group order by jumlah_order desc) as rank
from cte_1)
select
*
from cte_2
where rank<=3
order by age_group desc
```



## Case Study 6

Mencari tahu gender apa, laki – laki atau perempuan yang lebih banyak melakukan order di ecommerce, dan juga pengeluaran user berdasarkan gender, adakah perbedaan besar seperti di jumlah order, sehingga perusahaan dapat melakukan marketing dan mengembangkan produk yang lebih sesuai dengan kebutuhan pengguna setiap gender.

### Insight

Row	gender	jumlah_order	g_total_revenue	total_profit
1	M	26254	4981447.674	1532603.999
2	F	26499	4404071.723	1372685.719

Dari hasil query yang dilakukan, diketahui laki – laki sedikit lebih banyak dalam melakukan order dibandingkan perempuan.

Hanya saja total pengeluaran user laki – laki jauh lebih banyak dari perempuan, dengan selisih sekitar 600ribu.

Untuk profit yang diketahui, user laki – laki juga menghasilkan profit untuk penjualan lebih banyak dari pengguna perempuan.

Sehingga perusahaan mungkin dapat memperbanyak produk dan iklan yang menargetkan kepada user laki – laki.

### Query

```
with cte_1 as(
select
  o.order_id,
  u.gender,
  oi.status,
  (o.num_of_item*p.retail_price) revenue,
  (p.retail_price-p.cost) profit_per_item,
from bigquery-public-data.thelook_ecommerce.users u
join bigquery-public-data.thelook_ecommerce.orders o
on o.user_id = u.id
join bigquery-public-data.thelook_ecommerce.order_items oi
on oi.user_id = u.id
join bigquery-public-data.thelook_ecommerce.products p
on p.id = oi.product_id)
select
  gender,
  count(distinct order_id) as jumlah_order,
  round(sum(revenue),3) g_total_revenue,
  round(sum(cte_1.profit_per_item),3) total_profit
from cte_1
```

```
where status = 'Complete'  
group by gender  
order by total_profit desc
```

## Case Study 7

Mencari tahu produk kategori apa yang paling sering di cancel atau di return, sehingga dapat dianalisis lebih lanjut masalah yang ada, seperti kualitas produk yang buruk atau mungkin hal lainnya.

### Insight

Row	product_category	status	fo_
1	Intimates	Cancelled	2029
2	Jeans	Cancelled	1915
3	Tops & Tees	Cancelled	1841
4	Fashion Hoodies & Sweatshirts	Cancelled	1805
5	Sleep & Lounge	Cancelled	1732
6	Shorts	Cancelled	1690
7	Swim	Cancelled	1679
8	Sweaters	Cancelled	1582
9	Accessories	Cancelled	1516
10	Active	Cancelled	1421
11	Outerwear & Coats	Cancelled	1333
12	Intimates	Returned	1290
13	Jeans	Returned	1288
14	Tops & Tees	Returned	1187
15	Fashion Hoodies & Sweatshirts	Returned	1176
16	Swim	Returned	1137
17	Underwear	Cancelled	1113
18	Sleep & Lounge	Returned	1101

Dari hasil query yang dilakukan, didapatkan informasi lebih sering dilakukan cancel daripada return.

Selain itu produk kategori yang paling sering dicancel adalah yang termasuk ke top 3 paling banyak diorder yaitu intimates, jeans, tops & teens.

Sehingga perusahaan mungkin dapat meningkatkan kualitas customer service seperti kecepatan layanan dan pemeriksaan kualitas produk sebelum dikirim.

Perusahaan juga dapat melakukan survei feedback seperti mengapa melakukan cancel dan sebagainya.

## Query

```
select
  ii.product_category,
  o.status,
  count(o.order_id)
from bigquery-public-data.thelook_ecommerce.orders o
  join bigquery-public-data.thelook_ecommerce.order_items oi
    on oi.order_id = o.order_id
  join bigquery-public-data.thelook_ecommerce.inventory_items ii
    on oi.inventory_item_id = ii.id
where oi.status = 'Cancelled' or oi.status = 'Returned'
group by 1,2
order by 3 desc
```

## Case Study 8

Mencari tahu user di negara mana yang paling sering melakukan cancel atau return, sehingga dapat dilakukan analisa lebih lanjut seperti jika ada negara tertentu di mana user sering melakukan pengembalian produk, hal ini mungkin menandakan masalah dengan kualitas atau kecocokan produk di pasar negara tersebut.

### Insight

Row	country	status	fo_
1	China	Cancelled	6366
2	United States	Cancelled	4331
3	China	Returned	4138
4	United States	Returned	2778
5	Brasil	Cancelled	2718
6	Brasil	Returned	1837
7	South Korea	Cancelled	1043
8	United Kingdom	Cancelled	888

Dari hasil query yang dilakukan, didapatkan informasi bahwa user di China dan United States adalah user yang paling sering melakukan cancel dan return.

Sehingga perusahaan mungkin dapat meningkatkan kualitas customer service seperti kecepatan layanan dan pemeriksaan kualitas produk sebelum dikirim.

Perusahaan juga dapat melakukan analisis lebih lanjut seperti kenapa user di China dan United States banyak melakukan cancel daripada negara lain.

Bisa juga melakukan alternatif lain seperti menawarkan produk lain atau variasi produk yang lebih sesuai dengan preferensi user di negara tersebut

Perusahaan juga dapat melakukan survei feedback seperti mengapa melakukan cancel dan sebagainya.

### Query

```
select
  u.country,
  o.status,
  count(o.order_id)
from bigquery-public-data.thelook_ecommerce.orders o
  join bigquery-public-data.thelook_ecommerce.users u
  on o.user_id = u.id
where o.status = 'Cancelled' or o.status = 'Returned'
group by 1,2
order by 3 desc
```

## Case Study 9

Mencari tahu traffic source user paling banyak yang melakukan order, sehingga dapat dilakukan dan mengalokasikan iklan dengan lebih efisien.

### Insight

Row	traffic_source	jumlah_order
1	Search	87283
2	Organic	18813
3	Facebook	7397
4	Email	6237
5	Display	5049

Dari hasil query yang dilakukan, didapatkan informasi bahwa user paling banyak melakukan order dengan traffic source yaitu search, dengan perbedaan yang sangat jauh dengan traffic source lainnya.

Sehingga perusahaan dapat menambah dana iklan untuk traffic source yang paling banyak order tersebut.

### Query

```
select
  u.traffic_source,
  count(o.order_id) jumlah_order
from bigquery-public-data.thelook_ecommerce.users u
  join bigquery-public-data.thelook_ecommerce.orders o
    on u.id = o.user_id
group by 1
order by 2 desc
```

## Case Study 10

Mencari tahu kapan waktu user kebanyakan melakukan order, sehingga dapat dilakukan strategi marketing seperti flash sale pada waktu tertentu.

### Insight

Row	waktu	jumlah_order
1	Subuh (00:00-04:59)	32834
2	Pagi (05:00-09:59)	32449
3	Siang (10:00-13:59)	26443
4	Sore (14:00-17:59)	26328
5	Malam (18:00-23:59)	6725

Dari hasil query yang dilakukan, didapatkan informasi kebanyakan user melakukan order ketika subuh dan pagi (00.00 – 09.59 satuan waktu UTC) dengan total jumlah order mencapai 60 ribu lebih.

Sehingga perusahaan dapat menambah penawaran seperti flash sale pada jam – jam tersebut.

Perusahaan juga perlu memastikan platform dapat menerima traffic yang banyak dalam jangka waktu tersebut sehingga tidak terjadi crash atau lag.

Perusahaan juga dapat memastikan ketersediaan stok produk supaya menghindari kekurangan stok yang dapat menyebabkan hilangnya profit.

### Query

```
select
  case
    when extract(hour from timestamp (o.created_at)) between 0 and 4 then "Subuh (00:00-04:59)"
    when extract(hour from timestamp (o.created_at)) between 5 and 9 then "Pagi (05:00-09:59)"
    when extract(hour from timestamp (o.created_at)) between 10 and 13 then "Siang (10:00-13:59)"
    when extract(hour from timestamp (o.created_at)) between 14 and 17 then "Sore (14:00-17:59)"
    when extract(hour from timestamp (o.created_at)) between 18 and 23 then "Malam (18:00-23:59)"
  end as waktu,
  count(o.order_id) jumlah_order
from bigquery-public-data.thelook_ecommerce.users u
join bigquery-public-data.thelook_ecommerce.orders o
on u.id = o.user_id
group by 1
order by 2 desc
```