

Pilar Kukuh Bintang Rachmadi

Assignment Chapter 1 - Foundation for Data Science

1. Kasus 1

- Studi Kasus
Pihak Marketing Ecommerce XYZ ingin mengetahui produk dengan kategori apa yang sering dilihat oleh rata-rata asal negara pengguna (terlepas dari status pembelian).
- Berikut Insight yang didapatkan:

a) Insight

country	most_frequent_category
Australia	Intimates
Belgium	Intimates
Brasil	Intimates
China	Intimates
Colombia	Pants
France	Fashion Hoodies & Sweatshirts
Germany	Intimates
Japan	Jeans
Poland	Jeans
Poland	Intimates
South Korea	Intimates
Spain	Jeans
Spain	Intimates
United Kingdom	Intimates
United States	Intimates

1) Intimates adalah Kategori Barang yang sering dilihat oleh banyak pelanggan di antara 10 dari 13 negara pengguna

2) Poland dan Spain memiliki 2 kategori di peringkat 1 nya yaitu Jeans dan Intimates

b) Query yang digunakan

```
WITH category_tot AS (  
  SELECT  
    country,  
    category1.category,  
    COUNT(*) AS total_purchases  
  FROM  
    bigquery-public-data.thelook_ecommerce.users AS users  
  INNER JOIN (  
    SELECT id, category  
    FROM bigquery-public-data.thelook_ecommerce.products AS products  
  ) AS category1  
  ON users.id = category1.id  
  GROUP BY  
    country,
```

```

        category1.category
    )
SELECT
    country,
    most_frequent_category
FROM (
    SELECT
        country,
        total_purchases,
        category AS most_frequent_category,
        RANK() OVER (PARTITION BY country ORDER BY total_purchases DESC) AS
purchase_rank
    FROM category_tot
) AS ranked_data
WHERE purchase_rank = 1;
ORDER BY country;

```

2. Kasus 2

- Studi Kasus
Stakeholder Ecommerce XYZ ingin mengetahui total pembelian di masing-masing kategori barang
- Berikut Insight yang didapatkan:
a) Insight

Row	category	total_spending
1	Outerwear & Coats	207344.26983642578
2	Jeans	195608.56028461456
3	Sweaters	130829.07000255585
4	Swim	103952.56012916565
5	Fashion Hoodies & Sweatshirts	100606.50008296967
6	Suits & Sport Coats	93524.599897384644
7	Sleep & Lounge	87166.440205812454
8	Shorts	80783.900231361389
9	Dresses	80414.130199432373
10	Intimates	79650.540111780167
11	Tops & Tees	77226.180263280869
12	Active	72489.30003619194
13	Accessories	66359.329953569919
14	Pants	62266.700231552124
15	Blazers & Jackets	51945.840113162994
16	Maternity	45617.689915180206
17	Pants & Capris	33534.500135421753
18	Underwear	29545.730034351349
19	Plus	29126.400053620338
20	Suits	21838.850017547607
21	Skirts	19203.360045909882
22	Socks	18484.560006141663
23	Leggings	15312.409976482391
24	Socks & Hosiery	11162.129980564117
25	Jumpsuits & Rompers	7358.390022277832
26	Clothing Sets	3139.2300033569336

- 1) Outwears & Coats adalah Kategori dengan pembelian terbesar
 - 2) Top 3 terbesar kategori yang ditampilkan termasuk pakaian luar dan merupakan kebutuhan primer
- b) Query yang digunakan

```
SELECT
    category,
    SUM(retail_price) AS total_spending
FROM
    bigquery-public-data.thelook_ecommerce.orders AS orders
INNER JOIN
    bigquery-public-data.thelook_ecommerce.products AS products
ON
    orders.order_id = products.id
GROUP BY
    category
ORDER BY
    total_spending DESC;
```

3. Kasus 3

- Studi Kasus
Tim developer Ecommerce XYZ ingin mengetahui rata-rata perilaku user menggunakan apa untuk mencari suatu barang dengan kategori tertentu
- Berikut Insight yang didapatkan:
 - a) Insight

category	traffic_source	total_visits
Accessories	Search	1056
Accessories	Organic	273
Accessories	Facebook	87
Accessories	Email	85
Accessories	Display	58
Active	Search	1005
Active	Organic	213
Active	Facebook	88
Active	Email	64
Active	Display	62
Blazers & Jackets	Search	391
Blazers & Jackets	Organic	88
Blazers & Jackets	Email	31
Blazers & Jackets	Facebook	26
Blazers & Jackets	Display	25
Clothing Sets	Search	20

Clothing Sets	Organic	9
Clothing Sets	Facebook	6
Clothing Sets	Email	2
Dresses	Search	662
Dresses	Organic	166
Dresses	Facebook	50
Dresses	Display	39
Dresses	Email	38
Fashion Hoodies & Sweatshirts	Search	1326
Fashion Hoodies & Sweatshirts	Organic	236
Fashion Hoodies & Sweatshirts	Facebook	118
Fashion Hoodies & Sweatshirts	Email	107
Fashion Hoodies & Sweatshirts	Display	79
Intimates	Search	1661
Intimates	Organic	347
Intimates	Facebook	145
Intimates	Email	125
Intimates	Display	85
Jeans	Search	1397
Jeans	Organic	282
Jeans	Facebook	119
Jeans	Email	113
Jeans	Display	88
Jumpsuits & Rompers	Search	118
Jumpsuits & Rompers	Organic	21
Jumpsuits & Rompers	Facebook	12
Jumpsuits & Rompers	Display	8
Jumpsuits & Rompers	Email	3
Leggings	Search	398
Leggings	Organic	78
Leggings	Facebook	44
Leggings	Email	23
Leggings	Display	21

Maternity	Search	615
Maternity	Organic	149
Maternity	Facebook	50
Maternity	Email	44
Maternity	Display	40
Outerwear & Coats	Search	970
Outerwear & Coats	Organic	230
Outerwear & Coats	Facebook	86
Outerwear & Coats	Email	80
Outerwear & Coats	Display	54
Pants	Search	738
Pants	Organic	144
Pants	Facebook	62
Pants	Email	56
Pants	Display	41
Pants & Capris	Search	424
Pants & Capris	Organic	97
Pants & Capris	Email	44
Pants & Capris	Facebook	31
Pants & Capris	Display	17
Plus	Search	519
Plus	Organic	117
Plus	Facebook	50
Plus	Email	41
Plus	Display	31
Shorts	Search	1224
Shorts	Organic	273
Shorts	Facebook	113
Shorts	Display	88
Shorts	Email	67
Skirts	Search	238
Skirts	Organic	58
Skirts	Facebook	35

Skirts	Email	19
Skirts	Display	17
Sleep & Lounge	Search	1221
Sleep & Lounge	Organic	274
Sleep & Lounge	Facebook	122
Sleep & Lounge	Email	87
Sleep & Lounge	Display	67
Socks	Search	640
Socks	Organic	117
Socks	Facebook	69
Socks	Email	43
Socks	Display	36
Socks & Hosiery	Search	456
Socks & Hosiery	Organic	120
Socks & Hosiery	Facebook	35
Socks & Hosiery	Email	32
Socks & Hosiery	Display	23
Suits	Search	137
Suits	Organic	26
Suits	Facebook	14
Suits	Email	9
Suits	Display	2
Suits & Sport Coats	Search	517
Suits & Sport Coats	Organic	112
Suits & Sport Coats	Facebook	37
Suits & Sport Coats	Display	37
Suits & Sport Coats	Email	36
Sweaters	Search	1242
Sweaters	Organic	247
Sweaters	Facebook	104
Sweaters	Display	75
Sweaters	Email	69
Swim	Search	1223

Swim	Organic	294
Swim	Email	112
Swim	Facebook	105
Swim	Display	64
Tops & Tees	Search	1343
Tops & Tees	Organic	260
Tops & Tees	Facebook	109
Tops & Tees	Email	89
Tops & Tees	Display	67
Underwear	Search	766
Underwear	Organic	163
Underwear	Facebook	56
Underwear	Email	55
Underwear	Display	48

- 1) Pengguna rata-rata menggunakan search dan organic untuk mencari suatu barang tersebut
 - 2) Di semua kategori metode pencarian barang, urutan ke-1 adalah search dan urutan ke-2 adalah organic
- b) Query yang digunakan

```

WITH traffic_by_source AS (
  SELECT
    category,
    traffic_source,
    COUNT(*) AS total_visits
  FROM
    bigquery-public-data.thelook_ecommerce.users AS users
  INNER JOIN
    bigquery-public-data.thelook_ecommerce.products AS products
  ON
    users.id = products.id
  GROUP BY
    category,
    traffic_source
)
SELECT
  category,
  traffic_source,
  total_visits,
FROM (
  select
    category,
    traffic_source,
    total_visits,

```

```

    RANK() OVER (PARTITION BY category ORDER BY total_visits DESC) AS
rank_visits
  from traffic_by_source
)
WHERE rank_visits = 1
ORDER BY
  category,
  total_visits DESC;

```

4. Kasus 4

- Studi Kasus
Ingin mengetahui perbedaan gender manakah di tiap kategori merupakan spender terbanyak
- Berikut Insight yang didapatkan:
a) Insight

Row	category	male_spending	female_spending
1	Outerwear & Coats	25542.65002441...	26797.84997558...
2	Jeans	22020.61000442...	23811.22005462...
3	Sweaters	16350.60006046...	17914.14998531...
4	Swim	14141.29999446...	13422.67999553...
5	Fashion Hoodies & Sweatshirts	13542.34004783...	12926.66998863...
6	Sleep & Lounge	10434.40002727...	11507.41002941...
7	Suits & Sport Coats	10684.04996585...	11248.47002601...
8	Shorts	9652.340043067...	10909.32001972...
9	Dresses	10817.97001457...	9565.730019569...
10	Tops & Tees	9300.540029048...	10929.54001712...
11	Intimates	9315.000002145...	9559.180030584...
12	Active	7066.320004701...	10083.21000146...
13	Accessories	8361.559969663...	8060.220001220...
14	Blazers & Jackets	6590.860001564...	9381.129992485...
15	Pants	7161.790033340...	7203.720000267...
16	Maternity	5870.359995841...	6496.869976043...
17	Pants & Capris	3969.970002174...	5446.380014419...
18	Plus	5595.6799877882	2700.270023107...
19	Underwear	3572.330002307...	3310.960000991...
20	Suits	2624.940006256...	3475.399999618...
21	Skirts	3469.980002403...	1552.450002670...
22	Leggings	2393.119997978...	2281.119985103...
23	Socks	2079.520000696...	1931.519999980...
24	Socks & Hosiery	1369.629996538...	1484.249999284...
25	Jumpsuits & Rompers	915.3000001907...	661.9200067520...
26	Clothing Sets	299.8399963378...	477.9599990844...

- 1) Perempuan memegang dominasi terbanyak karena merupakan spender terbesar 16 kategori dari 26 kategori
- 2) Rerata selisih spending antara male dan female paling tinggi berada di sekitaran 3000

- b) Query yang digunakan
 WITH gendered_orders AS (


```

SELECT
    o.order_id,
    o.status,
    o.user_id,
    c.gender,
    p.category,
    p.retail_price
FROM
    bigquery-public-data.thelook_ecommerce.orders AS o
INNER JOIN
    bigquery-public-data.thelook_ecommerce.users AS c
ON
    o.user_id = c.id
INNER JOIN
    bigquery-public-data.thelook_ecommerce.products AS p
ON
    o.order_id = p.id
WHERE o.status = 'Complete'
)
SELECT
    category,
    SUM(CASE WHEN gender = 'M' THEN retail_price ELSE 0 END) AS male_spending,
    SUM(CASE WHEN gender = 'F' THEN retail_price ELSE 0 END) AS
female_spending
FROM
    gendered_orders
GROUP BY
    category
ORDER BY
    (male_spending + female_spending) DESC;

```

5. Kasus 5

- Studi Kasus
Mengetahui jumlah gender yang memang mereka berbelanja hingga sampai ke tangan pelanggan
- Berikut Insight yang didapatkan:

a) Insight

gender	total_complete
F	15611
M	15760

- 1) Male merupakan pelanggan terbanyak yang berbelanja hingga sampai di tangan mereka
 - 2) Selisih jumlahnya tipis hanya berbeda 149 pesanan saja
- b) Query yang digunakan

```

SELECT
    gender,
    COUNT(*) AS total_complete
FROM
    bigquery-public-data.thelook_ecommerce.orders AS orders
WHERE

```

```

    orders.status = 'Complete'
GROUP BY
    gender;

```

6. Kasus 6

- Studi Kasus
Melakukan Komparasi keseluruhan status transaksi untuk mengetahui rata-rata status transaksi saat ini
- Berikut Insight yang didapatkan:

a) Insight

Row	status	total_count	percentage
1	Shipped	53883	29.775864546147812
2	Complete	45375	25.074324996408087
3	Processing	35971	19.877653872083641
4	Cancelled	27610	15.257346846299221
5	Returned	18123	10.01480973906124

1) Kebanyakan transaksi saat ini menunjukkan masih dalam tahap pengiriman

b) Query yang digunakan

```

SELECT
    status,
    COUNT(status) AS total_count,
    (COUNT(status) / (SELECT COUNT(*) FROM bigquery-public-
data.thelook_ecommerce.order_items)) * 100 AS percentage
FROM
    bigquery-public-data.thelook_ecommerce.order_items
GROUP BY
    status
ORDER BY
    total_count DESC;

```

7. Kasus 7

- Studi Kasus
Mengetahui pembelian terbanyak pada macam-macam kategori barang dengan status transaksi sudah selesai
- Berikut Insight yang didapatkan:

a) Insight

category	total_purchases
Intimates	3274
Jeans	3231
Tops & Tees	2992
Fashion Hoodies & Sweatshirts	2972
Swim	2892
Sweaters	2846

Shorts	2765
Sleep & Lounge	2744
Accessories	2368
Outerwear & Coats	2359
Active	2322
Underwear	1877
Pants	1839
Socks	1555
Dresses	1301
Suits & Sport Coats	1257
Maternity	1256
Plus	1017
Socks & Hosiery	878
Pants & Capris	868
Blazers & Jackets	838
Leggings	828
Skirts	542
Suits	272
Jumpsuits & Rompers	226
Clothing Sets	56

- 1) Intimates merupakan pembelian terbanyak dengan total 3274 transaksi
- 2) Pembelian Intimates dan Jeans merupakan top 2 pembelian kategori terbanyak sekaligus memiliki selisih yang tipis

b) Query yang digunakan

```
SELECT
    p.category,
    COUNT(oi.product_id) AS total_purchases
FROM
    bigquery-public-data.thelook_ecommerce.orders AS o
INNER JOIN
    bigquery-public-data.thelook_ecommerce.order_items AS oi
ON
    o.order_id = oi.order_id
INNER JOIN
    bigquery-public-data.thelook_ecommerce.products AS p
ON
    oi.product_id = p.id
WHERE
    o.status = 'Complete'
GROUP BY
    p.category
```

ORDER BY
total_purchases DESC;

8. Kasus 8

- Studi Kasus
Mengetahui status transaksi yang sudah selesai di semua kategori barang, terbanyak dikirim ke negara tujuan mana
- Berikut Insight yang didapatkan:
 - a) Insight

category	ship_country	total_shipments
Intimates	China	1166
Fashion Hoodies & Sweatshirts	China	1036
Jeans	China	1023
Sleep & Lounge	China	1014
Tops & Tees	China	1008
Swim	China	941
Shorts	China	938
Sweaters	China	937
Accessories	China	870
Outerwear & Coats	China	799
Active	China	799
Underwear	China	708
Pants	China	559
Socks	China	526
Dresses	China	486
Suits & Sport Coats	China	449
Maternity	China	447
Plus	China	357
Socks & Hosiery	China	341
Pants & Capris	China	333
Leggings	China	292
Blazers & Jackets	China	277
Skirts	China	201
Suits	China	77
Jumpsuits & Rompers	China	69
Clothing Sets	China	22

- 1) Negara China mendominasi seluruh tujuan pengiriman kategori barang pada ecommerce ini
- 2) Jika dikomparasi dengan kasus 7, Tujuan negara China hampir memegang sepertiga Riwayat transaksi pembelian barang pada ecommerce ini

b) Query yang digunakan

```
WITH joined_data AS (
  SELECT
    o.order_id,
    o.status,
    q.category,
    u.country AS ship_country
  FROM `bigquery-public-data.thelook_ecommerce.order_items` o
  INNER JOIN `bigquery-public-data.thelook_ecommerce.orders` p
    ON o.order_id = p.order_id
  INNER JOIN `bigquery-public-data.thelook_ecommerce.products` q
    ON o.product_id = q.id
  INNER JOIN `bigquery-public-data.thelook_ecommerce.users` u
    ON o.user_id = u.id
)
SELECT
  category,
  ship_country,
  total_shipments,
FROM
  (
    SELECT
      category,
      ship_country,
      total_shipments,
      RANK() OVER (PARTITION BY category ORDER BY total_shipments DESC) AS
rank_shipments
    FROM
      (
        SELECT
          category,
          status,
          ship_country,
          COUNT(*) AS total_shipments
        FROM joined_data
        WHERE status = 'Complete'
        GROUP BY
          category,
          status,
          ship_country
      )
  )
WHERE rank_shipments = 1
GROUP BY
  category,
  ship_country,
  total_shipments
ORDER BY
  total_shipments DESC
```

9. Kasus 9

- Studi Kasus
Mengetahui Negara dengan pengeluaran terbanyak untuk berbelanja di ecommerce ini
- Berikut Insight yang didapatkan:

a) Insight

Row	country	total_cost
1	China	1761655.0660778631
2	United States	1157282.6506376066
3	Brasil	750515.6956184177
4	South Korea	265931.60421582137
5	France	247092.97661426882
6	United Kingdom	237355.03067381
7	Germany	216701.62171718941
8	Spain	211719.4532770256
9	Japan	119687.73287128261
10	Australia	110989.88975390955
11	Belgium	65013.481361241968
12	Poland	13299.215402080183
13	Colombia	707.60100597204837
14	España	46.999999787658446
15	Deutschland	45.687699746421728
16	Austria	31.178000012412667

- 1) Negara China merupakan negara dengan pengeluaran terbanyak untuk berbelanja di ecommerce ini
- 2) Jika kita kaitkan antara kasus 7, kasus 8, dan kasus 9 ini, China merupakan negara yang mendominasi dalam penggunaan ecommerce ini

b) Query yang digunakan

```
WITH gabungan AS (  
  SELECT  
    u.country,  
    q.cost  
  FROM `bigquery-public-data.thelook_ecommerce.order_items` o  
  INNER JOIN `bigquery-public-data.thelook_ecommerce.products` q  
    ON o.product_id = q.id  
  INNER JOIN `bigquery-public-data.thelook_ecommerce.users` u  
    ON o.user_id = u.id  
)  
SELECT  
  country,  
  SUM(cost) AS total_cost  
FROM gabungan  
GROUP BY  
  country  
ORDER BY  
  total_cost DESC;
```

10. Kasus 10

- Studi Kasus
Mengetahui jumlah Gender terbanyak dalam pembelian masing-masing kategori barang yang ada di ecommerce ini
- Berikut Insight yang didapatkan:

a) Insight

category	gender	total_purchases
Intimates	F	3331
Underwear	M	1923
Tops & Tees	M	1895
Jeans	M	1890
Pants	M	1742
Sweaters	M	1690
Fashion Hoodies & Sweatshirts	M	1685
Sleep & Lounge	M	1683
Shorts	M	1593
Swim	M	1544
Socks	M	1511
Outerwear & Coats	M	1500
Accessories	M	1415
Dresses	F	1375
Active	M	1276
Suits & Sport Coats	M	1271
Maternity	F	1266
Plus	F	1077
Socks & Hosiery	F	986
Pants & Capris	F	913
Leggings	F	801
Blazers & Jackets	F	800
Skirts	F	526
Suits	F	268
Jumpsuits & Rompers	F	204
Clothing Sets	F	61

- 1) Mayoritas gender pelanggan membeli barang sesuai dengan kategori barangnya

2) Mayoritas pelanggan di ecommerce ini membeli barang yang sering digunakan untuk melindungi kemaluan mereka

b) Query yang digunakan

```
WITH gabungan AS (  
  SELECT  
    o.order_id,  
    o.status,  
    q.category,  
    u.gender  
  FROM `bigquery-public-data.thelook_ecommerce.order_items` o  
  INNER JOIN `bigquery-public-data.thelook_ecommerce.orders` p  
    ON o.order_id = p.order_id  
  INNER JOIN `bigquery-public-data.thelook_ecommerce.products` q  
    ON o.product_id = q.id  
  INNER JOIN `bigquery-public-data.thelook_ecommerce.users` u  
    ON o.user_id = u.id  
)  
SELECT  
  category,  
  gender,  
  total_purchases  
FROM (  
  SELECT  
    category,  
    gender,  
    total_purchases,  
    RANK() OVER (PARTITION BY category ORDER BY total_purchases DESC) AS  
ranking  
  FROM  
  (  
    SELECT  
      category,  
      gender,  
      COUNT(*) AS total_purchases,  
      FROM gabungan  
      WHERE status = 'Complete'  
      GROUP BY  
        category,  
        gender  
    )  
  ) AS ranked_data  
WHERE ranking = 1  
GROUP BY  
  category,  
  gender,  
  total_purchases  
ORDER BY  
  total_purchases DESC;
```