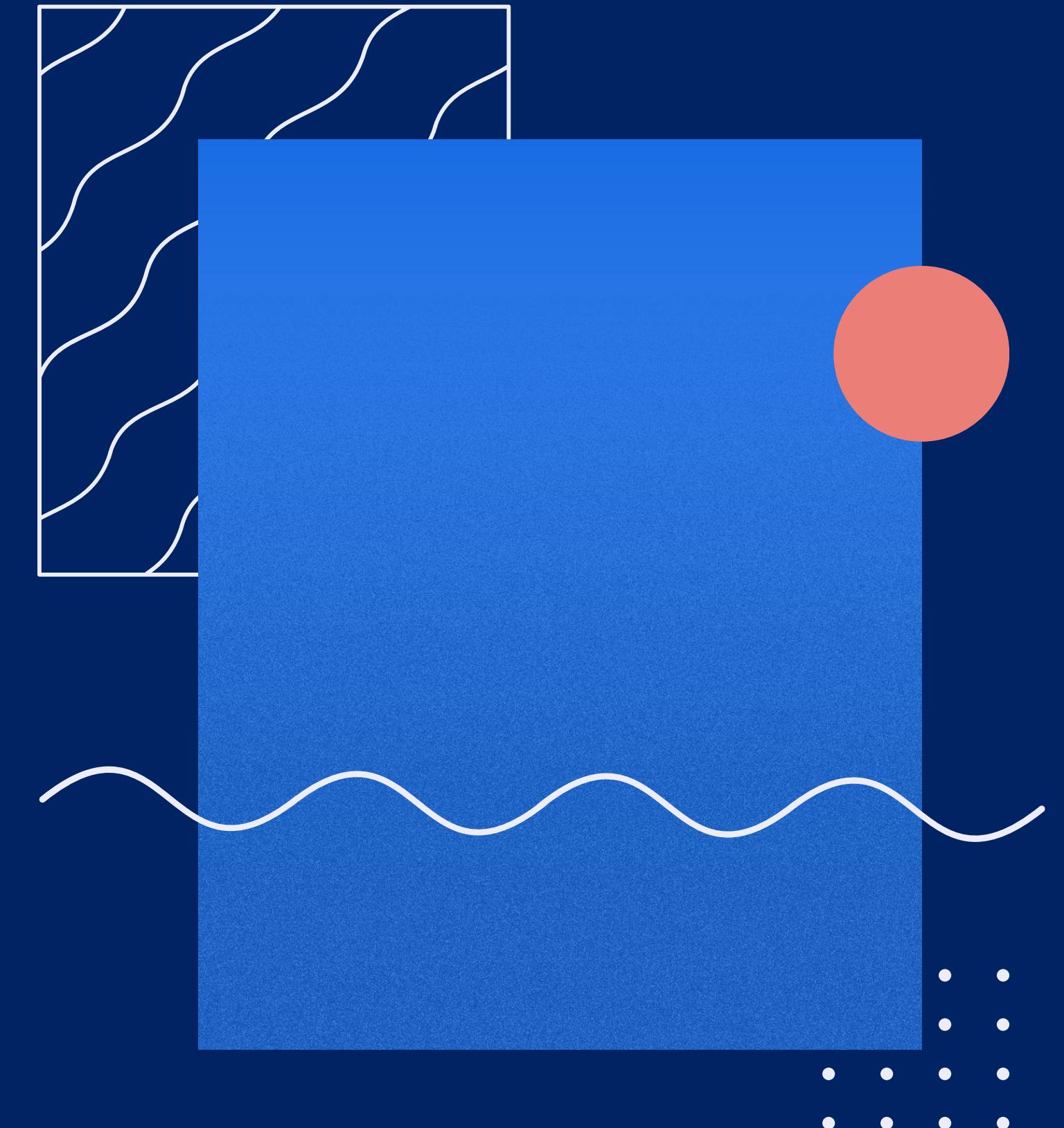


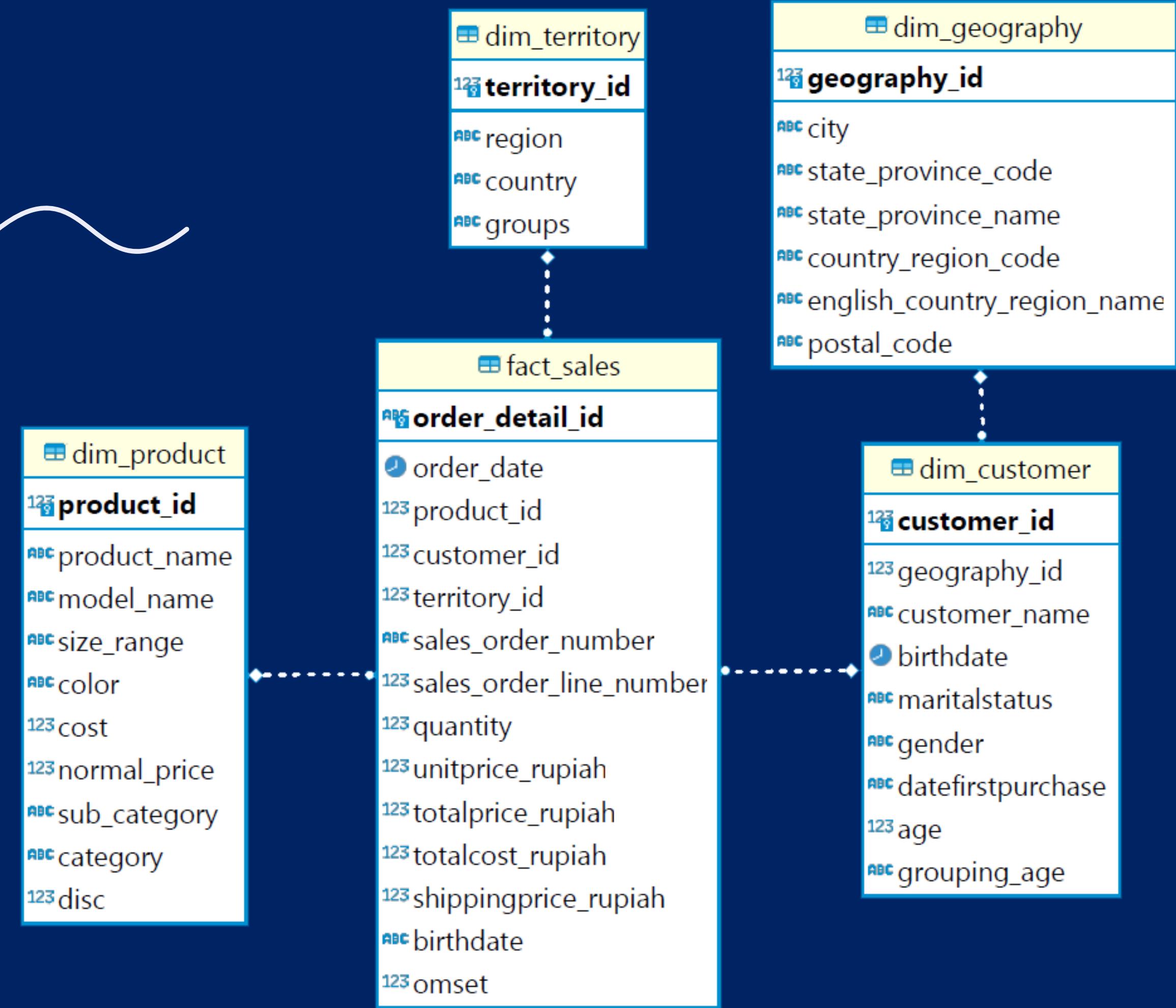
SQL Case Study

FILZAHANTI NUHA RAMADHANI



ER DIAGRAM

BEECYCLE DATASET



DATA PREVIEW

Table 1 | fact_sales

	order_detail_id	order_date	product_id	customer_id	territory_id	sales_order_number	sales_order_line_number	quantity	unitprice_rupiah	totalprice_rupiah	
0	SO43698-1	2016-07-01	346	28389.0	7	SO43698		1	1	47599860.0	47599860.0
1	SO43704-1	2016-07-02	351	11005.0	9	SO43704		1	1	47249860.0	47249860.0
2	SO43705-1	2016-07-02	344	11011.0	9	SO43705		1	1	47599860.0	47599860.0
3	SO43713-1	2016-07-05	310	27601.0	4	SO43713		1	1	50095780.0	50095780.0
4	SO43714-1	2016-07-05	311	13591.0	10	SO43714		1	1	50095780.0	50095780.0

Table 2 | dim_customer

	customer_id	geography_id	customer_name	birthdate	maritalstatus	gender	datefirstpurchase	age	grouping_age
0	11013	543	Ian Jenkins	1988-08-06		M	2018-10-15	34	Group 21 - 40
1	11014	634	Sydney Bennett	1988-05-09		S	2018-09-24	34	Group 21 - 40
2	11015	301	Chloe Young	1999-02-27		S	2018-07-22	23	Group 21 - 40
3	11016	329	Wyatt Hill	1999-04-28		M	2018-08-13	23	Group 21 - 40
4	11017	39	Shannon Wang	1964-06-26		S	2016-07-15	58	Group 41 - 60

DATA PREVIEW

Table 3 | dim_product

	product_id	product_name	model_name	size_range	color	cost	normal_price	sub_category	category	disc
0	210	HL Road Frame - Black, 58	HL Road Frame	54-58 CM	Black	11000.0	11000.0	Road Frames	Components	20
1	211	HL Road Frame - Red, 58	HL Road Frame	54-58 CM	Red	11000.0	11000.0	Road Frames	Components	20
2	480	Patch Kit/8 Patches	Patch kit	NA	NA	11991.0	32060.0	Tires and Tubes	Accessories	20
3	529	Road Tire Tube	Road Tire Tube	NA	NA	20892.0	55860.0	Tires and Tubes	Accessories	20
4	477	Water Bottle - 30 oz.	Water Bottle	NA	NA	26128.0	69860.0	Bottles and Cages	Accessories	20

Table 4 | dim_territory

	territory_id	region	country	groups
0	1	Northwest	United States	North America
1	2	Northeast	United States	North America
2	3	Central	United States	North America
3	4	Southwest	United States	North America
4	5	Southeast	United States	North America

DATA PREVIEW

Table 5 | dim_geography

	geography_id	city	state_province_code	state_province_name	country_region_code	english_country_region_name	postal_code
0	2	Coffs Harbour	NSW	New South Wales	AU	Australia	2450
1	3	Darlinghurst	NSW	New South Wales	AU	Australia	2010
2	4	Goulburn	NSW	New South Wales	AU	Australia	2580
3	5	Lane Cove	NSW	New South Wales	AU	Australia	1597
4	6	Lavender Bay	NSW	New South Wales	AU	Australia	2060

SOAL NO. 1

Tampilkan data customer dengan kriteria:

- Age harus di bawah rata-rata age seluruh customer
- Gender hanya M (male)
- Tahun kelahiran harus di bawah 1990-01-01

```
SELECT *
FROM dim_customer
WHERE age < (
    SELECT AVG(age)
    FROM dim_customer
)
AND gender = 'M'
AND birthdate < '1990-01-01'
```

PREVIEW:

	customer_id	geography_id	customer_name	birthdate	maritalstatus	gender	datefirstpurchase	age	grouping_age
0	11013	543	Ian Jenkins	1988-08-06		M	M	2018-10-15	34 Group 21 - 40
1	11094	33	Cedric Ma	1982-04-01		S	M	2018-10-09	40 Group 21 - 40
2	11095	13	Chad Kumar	1982-09-01		S	M	2016-10-02	40 Group 21 - 40
3	11096	12	Andrés Anand	1982-08-10		M	M	2016-10-12	40 Group 21 - 40
•	•	•							
•	•	•	4	11104	23	Edgar Sara	1984-03-11	38	Group 21 - 40
•	•	•							
•	•	•							

SOAL NO. 2

Tampilkan omset rata-rata perbulan dengan order month descending

```
SELECT  
    DATE_PART('month', order_date)  
        AS Month,  
    AVG(omset)  
FROM fact_sales  
GROUP BY 1  
ORDER BY 1 DESC
```

	month	avg
0	12	5341222
1	11	4851475
2	10	5012181
3	9	4603845
4	8	4966882
5	7	7129073
6	6	5888909
7	5	6011665
8	4	6601117
9	3	9767824
10	2	4764898
11	1	4422646

SOAL NO. 3

Tampilkan data total omset antara laki-laki dan wanita:

- age harus lebih 30 tahun
- category product adalah Bikes

	gender	total_omset
0	M	12112080000
1	F	13760484000

```
SELECT b.gender,  
       SUM(a.omset) AS total_omset  
  FROM fact_sales a  
 JOIN dim_customer b ON  
    b.customer_id=a.customer_id  
 JOIN dim_product c ON  
    c.product_id=a.product_id  
 WHERE b.age > 30  
   AND c.category = 'Bikes'  
 GROUP BY 1
```

• • •
• • •
• • •
• • •

SOAL NO. 4

Grouping age apa dan gender apa yang memiliki transaksi paling tinggi di BeeCycle?

	grouping_age	gender	transactions
0	Group 21 - 40	F	1551
1	Group > 60	M	1430
2	Group 41 - 60	M	1320
3	Group 21 - 40	F	1280
4	Group 21 - 40	M	1270
5	Group 21 - 40	F	1260
6	Group 21 - 40	M	1250
7	Group 21 - 40	F	1240
8	Group 21 - 40	M	1230
9	Group 21 - 40	F	1220
10	Group 21 - 40	M	1210
11	Group 21 - 40	F	1200
12	Group 21 - 40	M	1190
13	Group 21 - 40	F	1180
14	Group 21 - 40	M	1170
15	Group 21 - 40	F	1160
16	Group 21 - 40	M	1150
17	Group 21 - 40	F	1140
18	Group 21 - 40	M	1130
19	Group 21 - 40	F	1120
20	Group 21 - 40	M	1110
21	Group 21 - 40	F	1100
22	Group 21 - 40	M	1090
23	Group 21 - 40	F	1080
24	Group 21 - 40	M	1070
25	Group 21 - 40	F	1060
26	Group 21 - 40	M	1050
27	Group 21 - 40	F	1040
28	Group 21 - 40	M	1030
29	Group 21 - 40	F	1020
30	Group 21 - 40	M	1010
31	Group 21 - 40	F	1000
32	Group 21 - 40	M	990
33	Group 21 - 40	F	980
34	Group 21 - 40	M	970
35	Group 21 - 40	F	960
36	Group 21 - 40	M	950
37	Group 21 - 40	F	940
38	Group 21 - 40	M	930
39	Group 21 - 40	F	920
40	Group 21 - 40	M	910
41	Group 21 - 40	F	900
42	Group 21 - 40	M	890
43	Group 21 - 40	F	880
44	Group 21 - 40	M	870
45	Group 21 - 40	F	860
46	Group 21 - 40	M	850
47	Group 21 - 40	F	840
48	Group 21 - 40	M	830
49	Group 21 - 40	F	820
50	Group 21 - 40	M	810
51	Group 21 - 40	F	800
52	Group 21 - 40	M	790
53	Group 21 - 40	F	780
54	Group 21 - 40	M	770
55	Group 21 - 40	F	760
56	Group 21 - 40	M	750
57	Group 21 - 40	F	740
58	Group 21 - 40	M	730
59	Group 21 - 40	F	720
60	Group 21 - 40	M	710
61	Group 21 - 40	F	700
62	Group 21 - 40	M	690
63	Group 21 - 40	F	680
64	Group 21 - 40	M	670
65	Group 21 - 40	F	660
66	Group 21 - 40	M	650
67	Group 21 - 40	F	640
68	Group 21 - 40	M	630
69	Group 21 - 40	F	620
70	Group 21 - 40	M	610
71	Group 21 - 40	F	600
72	Group 21 - 40	M	590
73	Group 21 - 40	F	580
74	Group 21 - 40	M	570
75	Group 21 - 40	F	560
76	Group 21 - 40	M	550
77	Group 21 - 40	F	540
78	Group 21 - 40	M	530
79	Group 21 - 40	F	520
80	Group 21 - 40	M	510
81	Group 21 - 40	F	500
82	Group 21 - 40	M	490
83	Group 21 - 40	F	480
84	Group 21 - 40	M	470
85	Group 21 - 40	F	460
86	Group 21 - 40	M	450
87	Group 21 - 40	F	440
88	Group 21 - 40	M	430
89	Group 21 - 40	F	420
90	Group 21 - 40	M	410
91	Group 21 - 40	F	400
92	Group 21 - 40	M	390
93	Group 21 - 40	F	380
94	Group 21 - 40	M	370
95	Group 21 - 40	F	360
96	Group 21 - 40	M	350
97	Group 21 - 40	F	340
98	Group 21 - 40	M	330
99	Group 21 - 40	F	320
100	Group 21 - 40	M	310
101	Group 21 - 40	F	300
102	Group 21 - 40	M	290
103	Group 21 - 40	F	280
104	Group 21 - 40	M	270
105	Group 21 - 40	F	260
106	Group 21 - 40	M	250
107	Group 21 - 40	F	240
108	Group 21 - 40	M	230
109	Group 21 - 40	F	220
110	Group 21 - 40	M	210
111	Group 21 - 40	F	200
112	Group 21 - 40	M	190
113	Group 21 - 40	F	180
114	Group 21 - 40	M	170
115	Group 21 - 40	F	160
116	Group 21 - 40	M	150
117	Group 21 - 40	F	140
118	Group 21 - 40	M	130
119	Group 21 - 40	F	120
120	Group 21 - 40	M	110
121	Group 21 - 40	F	100
122	Group 21 - 40	M	90
123	Group 21 - 40	F	80
124	Group 21 - 40	M	70
125	Group 21 - 40	F	60
126	Group 21 - 40	M	50
127	Group 21 - 40	F	40
128	Group 21 - 40	M	30
129	Group 21 - 40	F	20
130	Group 21 - 40	M	10
131	Group 21 - 40	F	0

```
WITH age_group AS(
    SELECT *, DATE_PART('year', AGE(birthdate)) as age,
        CASE
            WHEN DATE_PART('year', AGE(birthdate)) > 60 THEN 'Group >60'
            WHEN DATE_PART('year', AGE(birthdate)) >= 41 THEN 'Group 41 - 60'
            WHEN DATE_PART('year', AGE(birthdate)) >= 21 THEN 'Group 21 - 40'
            ELSE 'Group <=20' END AS grouping_age
    FROM dim_customer
),
order_count AS(
    SELECT grouping_age, gender,
        COUNT(b.order_detail_id) AS transactions
    FROM age_group a
    JOIN fact_sales b ON b.customer_id=a.customer_id
    GROUP BY 1,2
    ORDER BY 3 DESC
)
SELECT grouping_age, gender, transactions
FROM order_count
LIMIT 1
```

SOAL NO. 5

Warna apa di tiap tahun yang menjadi warna paling populer dibeli oleh customer?

	year	color	count_color
0	2016.0	Red	216
1	2017.0	Red	223
2	2018.0	Black	503
3	2019.0	Black	281

```
WITH cte AS(
    SELECT DATE_PART('year', a.order_date) AS year,
        b.color, COUNT(b.color) as count_color,
        DENSE_RANK() OVER (PARTITION BY
            DATE_PART('year', a.order_date) ORDER BY
            COUNT(b.color) DESC) AS rank
    FROM fact_sales a
    JOIN dim_product b ON b.product_id=a.product_id
    WHERE color != 'NA'
    GROUP BY 1,2
)
SELECT year, color, count_color
FROM cte
WHERE rank = 1
ORDER BY 1
```

SOAL NO. 6

TOP 10 product name apa saja yang paling populer dari masing-masing territory

```
WITH cte AS(
    SELECT product_name, territory_id,
        COUNT(product_name) as count_product,
        ROW_NUMBER() OVER (PARTITION BY territory_id
        ORDER BY COUNT(product_name) DESC) AS rank
    FROM fact_sales a
    JOIN dim_product b ON b.product_id=a.product_id
    GROUP BY 1,2
)
SELECT territory_id, product_name, count_product
FROM cte
WHERE rank<=10
```

• • •
• • •
• • •
• • •

SOAL NO. 6

Jawaban

territory_id	product_name	count_product
0	HL Mountain Tire	39
1	Patch Kit/8 Patches	34
2	Mountain Tire Tube	28
3	Road-150 Red, 62	26
4	Road-150 Red, 48	21
...
70	Road Bottle Cage	36
71	Sport-100 Helmet, Black	23
72	Mountain-200 Black, 42	22
73	Mountain-200 Silver, 42	21
74	Touring Tire	20

SOAL NO. 7

Apa saja produk-produk yang paling banyak muncul pada transaksi tahun 2018 dan 2019 di setiap negara?

```
WITH cte AS(
    SELECT model_name, sub_category, cost,
           normal_price-cost AS profit,
           country AS branch_country,
           COUNT(model_name) as count_product,
           DENSE_RANK() OVER (PARTITION BY country
                               ORDER BY COUNT(model_name) DESC) AS rank
    FROM fact_sales a
    JOIN dim_product b ON b.product_id=a.product_id
    JOIN dim_territory c ON a.territory_id=c.territory_id
    WHERE DATE_PART('year', a.order_date) = '2018.0'
      OR DATE_PART('year', a.order_date) = '2019.0'
    GROUP BY 1,2,3,4,5
)
SELECT branch_country, model_name, cost,
       profit, sub_category, count_product
FROM cte
WHERE rank<=3
```

SOAL NO. 7

Jawaban

	branch_country	model_name	cost	profit	sub_category	count_product
0	Australia	Touring-1000	20747132.0	12629848.0	Touring Bikes	203
1	Australia	Sport-100	183208.0	306652.0	Helmets	181
2	Australia	Water Bottle	26128.0	43732.0	Bottles and Cages	124
3	Canada	Sport-100	183208.0	306652.0	Helmets	91
4	Canada	Road Tire Tube	20892.0	34968.0	Tires and Tubes	83
5	Canada	Patch kit	11991.0	20069.0	Tires and Tubes	79
6	France	Sport-100	183208.0	306652.0	Helmets	82
7	France	Water Bottle	26128.0	43732.0	Bottles and Cages	58
8	France	Mountain-200	17527738.0	14602122.0	Mountain Bikes	43
9	Germany	Sport-100	183208.0	306652.0	Helmets	75
10	Germany	Mountain-200	17718672.0	14761188.0	Mountain Bikes	54
11	Germany	Water Bottle	26128.0	43732.0	Bottles and Cages	52
12	United Kingdom	Sport-100	183208.0	306652.0	Helmets	106
13	United Kingdom	Water Bottle	26128.0	43732.0	Bottles and Cages	64
14	United Kingdom	Mountain-200	17527738.0	14602122.0	Mountain Bikes	50
15	United States	HL Mountain Tire	183260.0	306740.0	Tires and Tubes	87
16	United States	Patch kit	11991.0	20069.0	Tires and Tubes	77
17	United States	Mountain-200	17527738.0	14602122.0	Mountain Bikes	67

SOAL NO. 8

Buatlah data distribusi total order per customer, setelah didapatkan kemudian buat summary dengan tampilan:

- min_total_order
- avg_total_order
- p25_total_order
- median_total_order
- p75_total_order
- max_total_order

```
WITH cte AS(
    SELECT a.customer_id,
           COUNT(order_detail_id) AS total_order
      FROM fact_sales a
     JOIN dim_customer b ON a.customer_id=b.customer_id
     GROUP BY 1
)
SELECT MIN(total_order) AS min_total_order,
       AVG(total_order) AS avg_total_order,
       PERCENTILE_CONT(0.25) WITHIN GROUP
          (ORDER BY total_order) AS p25_total_order,
       PERCENTILE_CONT(0.5) WITHIN GROUP
          (ORDER BY total_order) AS median_total_order,
       PERCENTILE_CONT(0.75) WITHIN GROUP
          (ORDER BY total_order) AS p75_total_order,
       MAX(total_order) AS max_total_order
     FROM cte
```

SOAL NO. 8

Jawaban

PREVIEW:

	customer_id	avg_diff_order_date
0	11000.0	6.000000
1	11001.0	2.500000
2	11002.0	9.666667
3	11003.0	0.666667
4	11004.0	8.000000
5	11005.0	16.333333
6	11006.0	12.666667
7	11007.0	4.500000
8	11008.0	6.250000
9	11009.0	5.000000
..		
..		
..		
..		
..		

SOAL NO. 9

Berapa jumlah rata-rata omset dari:

- category = Accessories
- country_region_code = AU

Kemudian tampilkan data fact sales dengan kolom:

order_detail_id, order_date, customer_id, sales_order_number, omset, avg_omset, status.

Status adalah jika omset < avg_omset maka omset lebih kecil dari rata-rata dan omset lebih besar dari rata-rata untuk yang lainnya.

```
WITH cte AS (
    SELECT order_detail_id, order_date, customer_id,
           sales_order_number, omset,
           (SELECT AVG(omset)
            FROM fact_sales a
            LEFT JOIN dim_product b ON a.product_id=b.product_id
            LEFT JOIN dim_customer c ON a.customer_id=c.customer_id
            LEFT JOIN dim_geography d ON
                  c.geography_id=d.geography_id
            WHERE category='Accessories'
            AND country_region_code='AU') AS avg_omset
     FROM fact_sales
)
SELECT *, CASE WHEN omset<avg_omset
                THEN 'omset lebih kecil dari rata-rata'
                ELSE 'omset lebih besar dari rata-rata'
           END AS status
   FROM cte
```

SOAL NO. 9

Jawaban

PREVIEW:

	order_detail_id	order_date	customer_id	sales_order_number	omset	avg_omset	case
0	SO43698-1	2016-07-01	28389.0	SO43698	20829698.0	188433.258065	omset lebih besar dari rata-rata
1	SO43704-1	2016-07-02	11005.0	SO43704	20676538.0	188433.258065	omset lebih besar dari rata-rata
2	SO43705-1	2016-07-02	11011.0	SO43705	20829698.0	188433.258065	omset lebih besar dari rata-rata
3	SO43713-1	2016-07-05	27601.0	SO43713	19697660.0	188433.258065	omset lebih besar dari rata-rata
4	SO43714-1	2016-07-05	13591.0	SO43714	19697660.0	188433.258065	omset lebih besar dari rata-rata
5	SO43719-1	2016-07-06	27612.0	SO43719	19697660.0	188433.258065	omset lebih besar dari rata-rata
6	SO43720-1	2016-07-06	13264.0	SO43720	19697660.0	188433.258065	omset lebih besar dari rata-rata
7	SO43721-1	2016-07-06	13590.0	SO43721	19697660.0	188433.258065	omset lebih besar dari rata-rata
8	SO43725-1	2016-07-08	13258.0	SO43725	19697660.0	188433.258065	omset lebih besar dari rata-rata
9	SO43728-1	2016-07-09	27666.0	SO43728	19697660.0	188433.258065	omset lebih besar dari rata-rata

THANK YOU

