Second Homework
Business Intelligence Short Course
Batch 2
Rakamin Academy

SQL (PostgreSQL)

Yohanes Setiawan

I am Yohanes Setiawan

Suppose I am a Business Intelligence Analyst in Superstore.

I was requested to solve problem from our department and the others, e.g. Marketing, Sales, Business Development, etc. using SQL.

Case 1 – Problem

- “SAME DAY” Ship Service is a service which provides same ship date with the order date
- However, there are several orders which is late to arrive to customers
- I was requested to show how many orders are late from “SAME DAY” Ship Service

Case 1 – Query and Result

Query

```
13  with late_same_day as
14  (
15      select  "Order ID",
16              "Order Date",
17              "Ship Date",
18              "Ship Mode"
19      from superstore_order
20      where   extract(day from "Order Date") <> extract(day from "Ship Date")
21              and "Ship Mode" = 'Same Day'
22  )
23  select count("Order ID") as "Total Late SAME DAY"
24  from late_same_day;
```

Result

Data Output | Explain | Messages | Notifications

	Total Late SAME DAY bigint
1	24

Case 1 – Brief Explanation

- Therefore, there are 24 orders which is late from “SAME DAY” Ship Service

Case 2 – Problem

- Our business team wants to analyse the profitability of the company by seeing relationship between discount and profit
- I need to show the average of profit for every discount level

Discount Level	Value
LOW	Discount < 0.2
MODERATE	$0.2 \leq$ Discount < 0.4
HIGH	Discount ≥ 0.4

Case 2 – Query and Result

Query

```
38  with discount_level as
39  (
40      select  "Order ID",
41              "Discount",
42              case when "Discount" < 0.2 then 'LOW'
43                   when "Discount" >= 0.2 and "Discount" < 0.4 then 'MODERATE'
44              else 'HIGH'
45              end as "Discount Level",
46              "Profit"
47      from superstore_order
48  )
49  select  "Discount Level",
50          round(cast(avg("Profit") as numeric), 2) as "Average Profit"
51  from discount_level
52  group by 1
53  order by 2 desc;
```

Result

Data Output | Explain | Messages | Notifications

	Discount Level text	Average Profit numeric
1	LOW	67.04
2	MODERATE	19.84
3	HIGH	-107.65

Case 2 – Brief Explanation

- The higher discount level, the lower profit average received by company
- Discount level “HIGH” has the lowest average of profit
- Discount level “LOW” has the highest average of profit

Case 3 – Problem

- Our sales team asked us to analyse the perform of category and subcategory from products
- I need to show the average of discount and the average of profit for every category and subcategory in our products

Case 3 – Query and Result (Discount)

Query

```
150  -- Order by Discount
151  with cat_subcat as
152  (
153      select  o."Product ID",
154              p."Category",
155              p."Sub-Category",
156              o."Discount",
157              o."Profit"
158      from superstore_order o
159      join superstore_product p on o."Product ID" = p."Product ID"
160  )
161  select  "Category",
162          "Sub-Category",
163          round(cast(avg("Discount") as numeric), 2) as "Average Discount",
164          round(cast(avg("Profit") as numeric), 2) as "Average Profit"
165  from  cat_subcat
166  group by 1,2
167  order by 3 desc
```

Result

Data Output | Explain | Messages

	Category character varying	Sub-Category character varying	Average Discount numeric	Average Profit numeric
1	Office Supplies	Binders	0.37	19.84
2	Technology	Machines	0.31	29.43
3	Furniture	Tables	0.26	-55.57
4	Furniture	Bookcases	0.21	-15.23
5	Office Supplies	Appliances	0.17	38.92
6	Furniture	Chairs	0.17	43.10
7	Technology	Copiers	0.16	817.91
8	Technology	Phones	0.15	50.07
9	Furniture	Furnishings	0.14	13.65
10	Technology	Accessories	0.08	54.11
11	Office Supplies	Supplies	0.08	-6.26
12	Office Supplies	Envelopes	0.08	27.42
13	Office Supplies	Fasteners	0.08	4.38
14	Office Supplies	Paper	0.07	24.86
15	Office Supplies	Storage	0.07	25.15
16	Office Supplies	Labels	0.07	15.24
17	Office Supplies	Art	0.07	8.20

Case 3 – Query and Result (Profit)

Query

```
131 -- Order by Profit
132 with cat_subcat as
133 (
134     select  o."Product ID",
135             p."Category",
136             p."Sub-Category",
137             o."Discount",
138             o."Profit"
139     from superstore_order o
140     join superstore_product p on o."Product ID" = p."Product ID"
141 )
142 select  "Category",
143         "Sub-Category",
144         round(cast(avg("Discount") as numeric), 2) as "Average Discount",
145         round(cast(avg("Profit") as numeric), 2) as "Average Profit"
146 from  cat_subcat
147 group by 1,2
148 order by 4 desc;
```

Result

Data Output | Explain | Messages

	Category character varying	Sub-Category character varying	Average Discount numeric	Average Profit numeric
1	Technology	Copiers	0.16	817.91
2	Technology	Accessories	0.08	54.11
3	Technology	Phones	0.15	50.07
4	Furniture	Chairs	0.17	43.10
5	Office Supplies	Appliances	0.17	38.92
6	Technology	Machines	0.31	29.43
7	Office Supplies	Envelopes	0.08	27.42
8	Office Supplies	Storage	0.07	25.15
9	Office Supplies	Paper	0.07	24.86
10	Office Supplies	Binders	0.37	19.84
11	Office Supplies	Labels	0.07	15.24
12	Furniture	Furnishings	0.14	13.65
13	Office Supplies	Art	0.07	8.20
14	Office Supplies	Fasteners	0.08	4.38
15	Office Supplies	Supplies	0.08	-6.26
16	Furniture	Bookcases	0.21	-15.23
17	Furniture	Tables	0.26	-55.57

Case 3 – Brief Explanation

- The highest average of discount achieved by subcategory “BINDERS” from category “OFFICE SUPPLIERS”
- Category “TECHNOLOGY” has been dominating the highest average of profit by subcategory “COPIERS”, “ACCESSORIES”, and “PHONES”

Case 4 – Problem

- Our business development team wants to expand farther into California, Texas, and Georgia
- I need to show the segment of the total sales and the average of profit for every those three countries in 2016

Case 4 – Query and Result

Query

```
155 with t_cust as
156 (
157     select "Customer ID", "State", "Segment"
158     from superstore_customer
159     where "State" in ('California','Texas','Georgia')
160 ),
161 t_order as
162 (
163     select "Customer ID", "Order Date", "Sales", "Profit"
164     from superstore_order
165     where extract(year from "Order Date") = 2016
166 )
167
168 select c."Segment",
169        round(cast(sum(o."Sales") as numeric), 2) as "Sum Sales",
170        round(cast(avg(o."Profit") as numeric), 2) as "Average Profit"
171 from t_order o
172 right join t_cust c on o."Customer ID" = c."Customer ID"
173 group by 1
174 order by 3 desc;
```

Result

	Segment character varying	Sum Sales numeric	Average Profit numeric
1	Home Office	34897.95	34.66
2	Corporate	50951.91	33.57
3	Consumer	90982.32	30.33

Case 4 – Brief Explanation

- Segment “HOME OFFICE” has been the highest average of profit although it has the lowest total sales among other segments
- Segment “CONSUMER” has been the lowest average of profit although it has the highest total sales among other segments

Case 5 – Problem

- Our business team interested in analysing regions which likes discount
- I need to show how many customers which has the average of discount above 0.4 for every region

Case 5 – Query and Result

Query

```
179 with t_discount as
180 (
181     select "Customer ID",
182            round(cast(avg("Discount") as numeric), 2) as "Average Discount"
183     from superstore_order
184     group by 1
185     having round(cast(avg("Discount") as numeric), 2) > 0.4
186 ),
187 t_customer as
188 (
189     select "Customer ID",
190            "Region"
191     from superstore_customer
192 )
193 select  c."Region",
194 count(distinct d."Customer ID") as "Count Customer",
195 round(cast(avg(d."Average Discount") as numeric), 2) as "Avg Disc (Region)"
196 from t_discount d
197 left join t_customer c on d."Customer ID" = c."Customer ID"
198 group by 1
199 order by 2;
```

Result

	Region character varying	Count Customer bigint	Avg Disc (Region) numeric
1	Central	2	0.45
2	East	2	0.47
3	South	2	0.49
4	West	3	0.55

Case 5 – Brief Explanation

- “CENTRAL”, “SOUTH” and “EAST” have two customers which has got the average of discount above 0.4
- “WEST” has three customers which has got the average of discount above 0.4 and has been the highest average of discount among other regions

Thank You!

Credits: