5.1. Calculate days between purchasing, delivering and estimated delivery

```
select order_id,
    abs(EXTRACT(day FROM (order_purchase_timestamp -
      order_delivered_customer_date))) as day_to_deliver,
    EXTRACT(day FROM (order_estimated_delivery_date -
      order_purchase_timestamp)) as estimated_day_to_deliver
from orders
order by order_id;
```

order_id	day_to_deliver	estimated_day_t...
Filter...	Filter...	Filter...
00010242fe8c5a6d1ba2dd792cb16214	7	15
00018f77f2f0320c557190d7a144bdd3	16	18
000229ec398224ef6ca0657da4fc703e	7	21
00024acbcdf0a6daa1e931b038114c75	6	11
00042b26cf59d7ce69dfabb4e55b4fd9	25	40
00048cc3ae777c65dbb7d2a0634bc1ea	6	21
00054e8431b9d7675808bcb819fb4a32	8	24
000576fe39319847cbb9d288c5617fa6	5	20
0005a1a1728c9d785b8e2b08b904576c	9	9
0005f50442cb953dcd1d21e1fb923495	2	20
00061f2a7bc09da83e415a52dc8a4af1	4	15
00063b381e2406b52ad429470734ebd5	10	10
0006ec9db01a64e59a68b2c340bf65a7	6	28
0008288aa423d2a3f00fcb17cd7d8719	12	20
0009792311464db532ff765bf7b182ae	7	13
0009c9a17f916a706d71784483a5d643	5	13
000aed2e25dbad2f9ddb70584c5a2ded	6	10
000c3e6612759851cc3cbb4b83257986	7	19
000e562887b1f2006d75e0be9558292e	18	24
000e63d38ae8c00bbcb5a30573b99628	3	12
000e906b789b55f64edcb1f84030f90d	17	15
000f25f4d72195062c040b12dce9a18a	15	34

And so on.
Here,

'day_to_deliver' represents day between purchasing & delivering Date.
'estimated_day_to_deliver' represents day between purchasing and estimated delivery.

5.2. Find time_to_delivery & diff_estimated_delivery:

```
select order_id,
    abs(EXTRACT(day FROM (order_purchase_timestamp -
      order_delivered_customer_date))) as time_to_deliver,
    EXTRACT(day FROM (order_estimated_delivery_date -
      order_delivered_customer_date)) as diff_estimated_delivery
from orders
order by order_id;
```

order_id	time_to_deliver	diff_estimated_d...
00010242fe8c5a6d1ba2dd792cb16214	7	8
00018f77f2f0320c557190d7a144bdd3	16	2
000229ec398224ef6ca0657da4fc703e	7	13
00024acbcdf0a6daa1e931b038114c75	6	5
00042b26cf59d7ce69dfabb4e55b4fd9	25	15
00048cc3ae777c65dbb7d2a0634bc1ea	6	14
00054e8431b9d7675808bcb819fb4a32	8	16
000576fe39319847cbb9d288c5617fa6	5	15
0005a1a1728c9d785b8e2b08b904576c	9	0
0005f50442cb953dcd1d21e1fb923495	2	18
00061f2a7bc09da83e415a52dc8a4af1	4	10
00063b381e2406b52ad429470734ebd5	10	0
0006ec9db01a64e59a68b2c340bf65a7	6	21
0008288aa423d2a3f00fcb17cd7d8719	12	7
0009792311464db532ff765bf7b182ae	7	5
0009c9a17f916a706d71784483a5d643	5	8
000aed2e25dbad2f9ddb70584c5a2ded	6	3
000c3e6612759851cc3cbb4b83257986	7	12
000e562887b1f2006d75e0be9558292e	18	6
000e63d38ae8c00bbcb5a30573b99628	3	8
000e906b789b55f64edcb1f84030f90d	17	-2
000f25f4d72195062c040b12dce9a18a	15	19

And so on.

Here,

'time_to_deliver' represents day between purchasing & delivering Date.

'diff_estimated_delivery' represents day between delivering and estimated delivery.

5.3. Group data by state, take mean of freight_value, time_to_delivery, diff_estimated_delivery:

```
select
    customer_state as state
    ,round(avg(oi.freight_value), 2) as mean_freight_value
    ,round(avg(abs(EXTRACT(day FROM (o.order_purchase_timestamp -
      o.order_delivered_customer_date)))), 2) as mean_time_to_deliver
    ,round(avg(EXTRACT(day FROM (o.order_estimated_delivery_date -
      o.order_delivered_customer_date))), 2) as mean_diff_estimated_delivery
from orders o
join customers c using(customer_id)
join order_items oi using(order_id)
group by customer_state
order by customer_state;
```

state	mean_freight_va...	mean_time_to_d...	mean_diff_estim...
Filter...	Filter...	Filter...	Filter...
AC	40.07	20.33	20.01
AL	35.84	23.99	7.98
AM	33.21	25.96	18.98
AP	34.01	27.75	17.44
BA	26.36	18.77	10.12
CE	32.71	20.54	10.26
DF	21.04	12.50	11.27
ES	22.06	15.19	9.77
GO	22.77	14.95	11.37
MA	38.26	21.20	9.11
MG	20.63	11.52	12.40
MS	23.37	15.11	10.34
MT	28.17	17.51	13.64
PA	35.83	23.30	13.37
PB	42.72	20.12	12.15
PE	32.92	17.79	12.55
PI	39.15	18.93	10.68
PR	20.53	11.48	12.53
RJ	20.96	14.69	11.14
RN	35.65	18.87	13.06
RO	41.07	19.28	19.08
RR	42.98	27.83	17.43
RS	21.74	14.71	13.20
SC	21.47	14.52	10.67
SE	36.65	20.98	9.17
SP	15.15	8.26	10.27
TO	37.25	17.00	11.46

Here, **'mean_freight_value', 'mean_time_to_deliver', 'mean_diff_estimated_delivery'** represents mean of freight_value, time_to_delivery, diff_estimated_delivery respectively.

5.4. Sort the data to get the following:

```
with consolidated_sales as (select
    customer_state as state
    ,round(avg(oi.freight_value), 2) as avg_freight_value
    ,round(avg(abs(EXTRACT(day FROM (o.order_purchase_timestamp -
      o.order_delivered_customer_date)))), 2) as avg_time_to_deliver
    ,round(avg(EXTRACT(day FROM (o.order_estimated_delivery_date -
      o.order_delivered_customer_date))), 2) as avg_diff_estimated_delivery
from orders o
join customers c using(customer_id)
join order_items oi using(order_id)
group by customer_state
order by customer_state)
```

This CTEs will implement in below query.

5.5. Top 5 states with highest/lowest average freight value - sort in desc/asc limit 5

For Top 5 Cheapest:

```
select state, avg_freight_value
from consolidated_sales
order by avg_freight_value asc limit 5
```

state	avg_freight_value
abc Filter...	abc Filter...
SP	15.15
PR	20.53
MG	20.63
RJ	20.96
DF	21.04

For Top 5 not so cheap:

```
select state, avg_freight_value
from consolidated_sales
order by avg_freight_value desc limit 5
```

state	avg_freight_value
abc Filter...	abc Filter...
RR	42.98
PB	42.72
RO	41.07
AC	40.07
PI	39.15

5.6. Top 5 states with highest/lowest average time to delivery

For Top 5 fastest:

```
select state, avg_time_to_deliver
from consolidated_sales
order by avg_time_to_deliver asc limit 5
```

state	avg_time_to_deli...
Filter...	Filter...
SP	8.26
PR	11.48
MG	11.52
DF	12.50
SC	14.52

For Top 5 slowest:

```
select state, avg_time_to_deliver
from consolidated_sales
order by avg_time_to_deliver desc limit 5
```

state	avg_time_to_deli...
Filter...	Filter...
RR	27.83
AP	27.75
AM	25.96
AL	23.99
PA	23.30

5.7. Top 5 states where delivery is really fast/ not so fast compared to estimated date

For Top 5 really fast:

```
select state, avg_diff_estimated_delivery
from consolidated_sales
order by avg_diff_estimated_delivery desc limit 5
```

state	avg_diff_estimat...
Filter...	Filter...
AC	20.01
RO	19.08
AM	18.98
AP	17.44
RR	17.43

For Top 5 not so fast:

```
select state, avg_diff_estimated_delivery
from consolidated_sales
order by avg_diff_estimated_delivery asc limit 5
```

state	avg_diff_estimat...
Filter...	Filter...
AL	7.98
MA	9.11
SE	9.17
ES	9.77
BA	10.12