Q1. Import the dataset and do usual exploratory analysis steps like checking the structure & characteristics of the dataset

1. Data type of columns in a table

Customer Table

Query

```
select column_name,data_type from `target365217.target_ds.INFORMATION_SCHEMA.COLUMNS`

where table_name="Customers";
```

Output

column_name	data_type
customer_id	STRING
customer_unique_id	STRING
customer_zip_code_prefix	INT64
customer_city	STRING
customer_state	STRING

Geolocation Table

Query

```
select column_name,data_type from `target365217.target_ds.INFORMATION_SCHEMA.COLUMNS`

where table_name=" Geolocation";
```

Output

column_name	data_type
geolocation_zip_code_prefix	INT64
geolocation_lat	FLOAT64
geolocation_lng	FLOAT64
geolocation_city	STRING
geolocation_state	STRING

Order items Table

Query

select column_name,data_type from `target365217.target_ds.INFORMATION_SCHEMA.COLUMNS`

where table_name=" Order_items";

Output

column_name	data_type
order_id	STRING
order_item_id	INT64
product_id	STRING
seller_id	STRING
shipping_limit_date	TIMESTAMP
price	FLOAT64
freight_value	FLOAT64

Order reviews Table

Query

select column_name,data_type from `target365217.target_ds.INFORMATION_SCHEMA.COLUMNS`

where table_name=" Order_reviews";

Output

column_name	data_type
review_id	STRING
order_id	STRING
review_score	INT64
review_comment_title	STRING
review_creation_date	TIMESTAMP
review_answer_timestamp	TIMESTAMP

Orders Table

Query

```
select column_name,data_type from `target365217.target_ds.INFORMATION_SCHEMA.COLUMNS`
where table_name=" Orders ";
```

Output

column_name	data_type
order_id	STRING
customer_id	STRING
order_status	STRING
order_purchase_timestamp	TIMESTAMP
order_approved_at	TIMESTAMP
order_delivered_carrier_date	TIMESTAMP
order_delivered_customer_date	TIMESTAMP
order_estimated_delivery_date	TIMESTAMP

Payments Table

Query

```
select column_name,data_type from `target365217.target_ds.INFORMATION_SCHEMA.COLUMNS`
where table_name=" Payments";
```

Output

column_name	data_type
order_id	STRING
payment_sequential	INT64
payment_type	STRING
payment_installments	INT64
payment_value	FLOAT64

Products Table

Query

```
select column_name,data_type from `target365217.target_ds.INFORMATION_SCHEMA.COLUMNS`
where table_name=" Products";
```

Output

column_name	data_type
product_id	STRING
product_category	STRING
product_name_length	INT64
product_description_length	INT64
product_photos_qty	INT64
product_weight_g	INT64
product_length_cm	INT64
product_height_cm	INT64
product_width_cm	INT64

Sellers Table

Query

```
select column_name,data_type from `target365217.target_ds.INFORMATION_SCHEMA.COLUMNS`
where table_name=" Sellers";
```

Output

column_name	data_type
seller_id	STRING
seller_zip_code_prefix	INT64
seller_city	STRING
seller_state	STRING

2. Time period for which the data is given

Query

```
create view `target-365217.target_ds.order_dates_analysis` as(
  select
  order_id,
  customer_id,
  order_status,
  date(order_purchase_timestamp) as order_purchase_dt,
  date(order_approved_at) as order_approved_at_dt,
  date(order_delivered_carrier_date) as order_delivered_carrier_dt,
  date(order_delivered_customer_date) as order_delivered_customer_dt,
  date(order_estimated_delivery_date) as order_estimated_delivery_dt

  from `target-365217.target_ds.Orders`
)

select order_id,extract(year from order_purchase_dt) as order_purchase_year,
extract(year from order_delivered_carrier_dt) as order_delivered_carrier_year,
extract(year from order_delivered_customer_dt) as order_delivered_customer_year,
extract(year from order_estimated_delivery_dt) as order_estimated_delivery_year
from `target-365217.target_ds.order_dates_analysis`
where extract(year from order_purchase_dt) between 2016 and 2018 limit 10;
```

Output

order_id	order_purch...	order_delivered_carrier_year	order_delivered_customer_year	order_estimated_delivery_year
7a4df5d8cff4090e541401a20a...	2017	null	null	2017
35de4050331c6c644cddc86f4...	2017	null	null	2018
b5359909123fa03c50bdb0cfe...	2017	null	null	2018
dba5062fbda3af4fb6c33b1e04...	2018	null	null	2018
90ab3e7d52544ec7bc3363c82...	2017	null	null	2017
fa65dad1b0e818e3ccc5cb0e3...	2017	2017	null	2017
1df2775799eecdf9dd8502425...	2017	2017	null	2017
6190a94657e1012983a274b8...	2017	2017	null	2017
58ce513a55c740a3a81e8c8b7...	2017	2017	null	2017
088683f795a3d30bfd61152c4f...	2017	2017	null	2017

3. Cities and States covered in the dataset

Query

```
select distinct g.geolocation_city,g.geolocation_state from `target-
365217.target_ds.Geolocation` as g
right join `target-365217.target_ds.Customers` as c
on g.geolocation_zip_code_prefix = c.customer_zip_code_prefix
limit 10;
```

Output

geolocation_city	geolocation_state
acu	RN
açu	RN
ico	CE
icó	CE
ipe	RS
ipê	RS
ipu	CE
ita	SC
itá	SC
itu	SP

Query

```
select
c.customer_city,
c.customer_zip_code_prefix,
count(distinct o.order_id) as total_order,
date(o.order_purchase_timestamp) as purchase_date
from `target-365217.target_ds.Orders` as o
join `target-365217.target_ds.Customers` as c
on o.customer_id = c.customer_id
group by c.customer_city,c.customer_zip_code_prefix,purchase_date
order by total_order desc limit 10;
```

Output

customer_city	customer_state	customer_zi...	total_order	purchase_d...
curitiba	PR	80030	14	2017-01-05
curitiba	PR	82200	6	2017-01-05
rio de janeiro	RJ	22775	5	2018-01-22
ubatuba	SP	11680	4	2017-11-24
niteroi	RJ	24230	4	2017-11-24
santana de parnaiba	SP	6540	4	2017-11-24
uberlandia	MG	38408	4	2017-11-24
vila velha	ES	29101	4	2018-02-22
indaiatuba	SP	13348	3	2017-07-25
curvelo	MG	35790	3	2017-09-16

Insights

- As per the analysis I found that most of the orders is from PR,RJ,SP states in year 2017 to 2018.

Recommendation

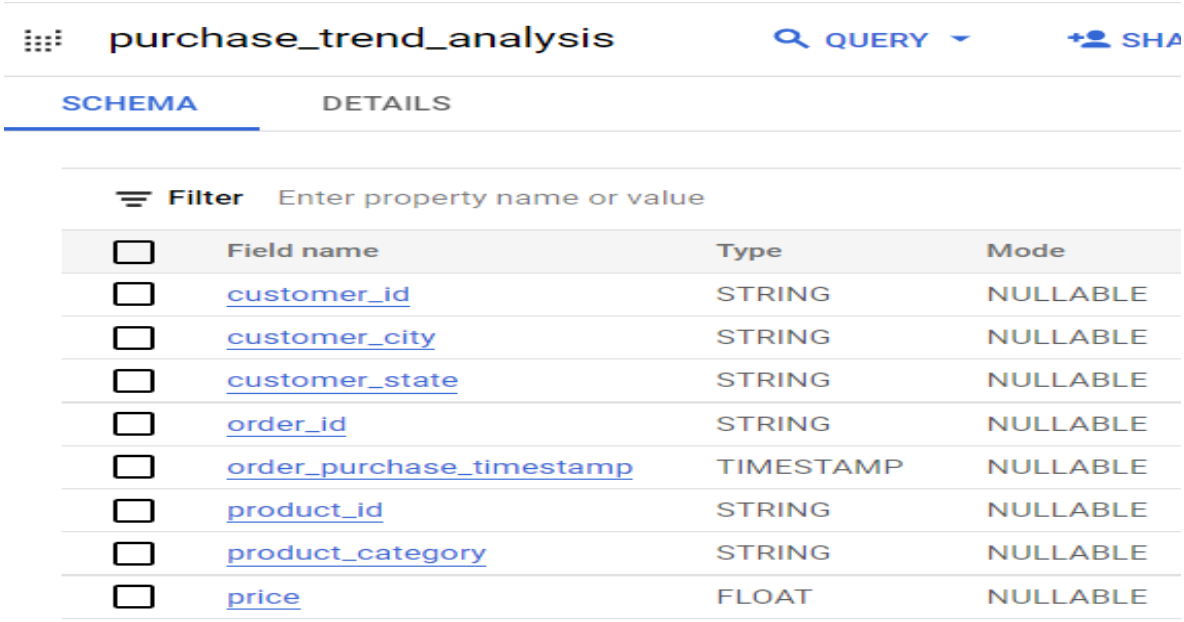
- I would like to recommend that please provide some discount in products to increase the order demand in these states.

2. In-depth Exploration

1. Is there a growing trend on e-commerce in Brazil? How can we describe a complete scenario? Can we see some seasonality with peaks at specific months?

View Created

```
create view `target-365217.target_ds.purchase_trend_analysis` as(select
c.customer_id,c.customer_city,c.customer_state,

o.order_id,

o.order_purchase_timestamp,

oit.product_id,

p.product_category,

oit.price,

from `target-365217.target_ds.Customers` as c

left join `target-365217.target_ds.Orders` as o

on c.customer_id = o.customer_id

left join `target-365217.target_ds.Order_items` as oit

on o.order_id = oit.order_id

left join `target-365217.target_ds.Products` as p

on oit.product_id = p.product_id);
```



Few Analytical Question for more Analysis

- Which city/state has maximum orders

Query

```
-- which city/state has max order
select customer_city,customer_state ,count(order_id) as order_count from `target-365217.target_ds.purchase_trend_analysis`
group by customer_city,customer_state
order by order_count desc limit 10;
```

Output

customer_city	customer_state	order_count
sao paulo	SP	17946
rio de janeiro	RJ	7885
belo horizonte	MG	3167
brasilia	DF	2407
curitiba	PR	1762
campinas	SP	1669
porto alegre	RS	1619
salvador	BA	1419
guarulhos	SP	1340
sao bernardo do campo	SP	1070

- Which product category has max ordered by user?

Query

```
select customer_id,product_category,count(product_category) as product_ordered_count from `target-365217.target_ds.purchase_trend_analysis`
group by customer_id,product_category
order by product_ordered_count desc limit 10;
```

Output

customer_id	product_category	product_ord...
fc3d1daec319d62d49bfb5e1f8...	HEALTH BEAUTY	21
bd5d39761aa56689a265d95d...	automotive	20
be1b70680b9f9694d8c70f41fa...	computer accessories	20
10de381f8a8d23fff822753305...	Furniture Decoration	15
adb32467ecc74b53576d9d13a...	Garden tools	15
a7693fba2ff9583c78751f2b66...	telephony	14
d5f2b3f597c7ccafbb5cac0bcc...	Garden tools	14
7d321bd4e8ba1caf74c4c1aab...	telephony	13
daf15f1b940cc6a72ba558f093...	babies	12
3b54b5978e9ace64a63f90d17...	housewares	12

- Which year/month has max order count?

  Query

  ```
  select extract(year from date(order_purchase_timestamp)) as year,FORMAT_DATE('
  %B',date(order_purchase_timestamp)) as order_month,count(order_id) as total_orde
  r from `target-365217.target_ds.purchase_trend_analysis`
  group by order_month,year
  order by total_order desc
  limit 10;
  ```

  Output

year	order_month	total_order
2017	November	8758
2018	January	8257
2018	March	8240
2018	April	7980
2018	May	7945
2018	February	7706
2018	August	7308
2018	July	7111
2018	June	7085
2017	December	6357

- Which city from Brazil has maximum orders?

  Query

  ```
  select extract(year from date(order_purchase_timestamp)) as year,FORMAT_DATE('
  %B',date(order_purchase_timestamp)) as order_month,count(order_id) as total_orde
  r,customer_city from `target-365217.target_ds.purchase_trend_analysis`
  group by order_month,year,customer_city
  order by total_order desc
  limit 10;
  ```

  Output

year	order_month	total_order	customer_city
2018	August	1493	sao paulo
2018	May	1434	sao paulo
2018	April	1354	sao paulo
2018	March	1321	sao paulo
2017	November	1306	sao paulo
2018	January	1245	sao paulo
2018	July	1241	sao paulo
2018	June	1237	sao paulo
2018	February	1210	sao paulo
2017	December	961	sao paulo

2. What time do Brazilian customers tend to buy (Dawn, Morning, Afternoon or Night)?

Query

```
with purchase_time as (select distinct time(order_purchase_timestamp) as purchase_time,
case
  when time(order_purchase_timestamp) between "00:00:00" and "05:00:00" then "Dawn"
  when time(order_purchase_timestamp) between "05:00:00" and "11:58:00" then "Morning"
  when time(order_purchase_timestamp) between "12:00:00" and "16:58:00" then "Afternoon
"
  else "night"
end as result
from `target-365217.target_ds.purchase_trend_analysis`)
select result,count(purchase_time)as total_purchase from purchase_time group by result
order by total_purchase ;
```

Output

result	total_purcha...
Dawn	3650
Morning	12095
Afternoon	14906
night	20167

Insights

- User are more looking for health beauty products
- User had ordered more products in year 2018 as compare to 2017.
- User from sao paulo have high order rate as compare to another city.
- Brazilian User are more active in night for purchase as compare to other time.

Recommendation

- Show more products that is related to beauty health followed by computer accessories.
- Show less discount on products at night to users so that we can manage traffic at night and it will help to divert the traffic to another timestamp too.

3. Evolution of E-commerce orders in the Brazil region

1. Get month on month orders by region, states?

Query

```
select customer_id,customer_city,customer_state,FORMAT_DATE('%B',date(order_
purchase_timestamp)) as order_month,count(order_id) as total_order from `target-
365217.target_ds.purchase_trend_analysis`
group by order_month,customer_state,customer_city,customer_id
order by total_order desc
limit 12;
```

Output

customer_id	customer_city	customer_state	order_month	total_order
fc3d1daec319d62d49bfb5e1f8...	sao paulo	SP	July	21
bd5d39761aa56689a265d95d...	goiania	GO	August	20
be1b70680b9f9694d8c70f41fa...	sao paulo	SP	February	20
10de381f8a8d23fff822753305...	uniao da vitoria	PR	November	15
adb32467ecc74b53576d9d13a...	goiania	GO	January	15
d5f2b3f597c7ccafbb5cac0bcc...	santos	SP	December	14
a7693fba2ff9583c78751f2b66...	indaiatuba	SP	February	14
7d321bd4e8ba1caf74c4c1aab...	sao paulo	SP	April	13
3b54b5978e9ace64a63f90d17...	sao paulo	SP	May	12
9eb3d566e87289dcb0acf28e1...	belo horizonte	MG	June	12
daf15f1b940cc6a72ba558f093...	celso ramos	SC	April	12
0d93f21f3e8543a9d0d8ece01...	juiz de fora	MG	October	12

2. How are customers distributed in Brazil?

- User distribution on the basis of product they bought or interested?

  Query

  ```
  select count(customer_id) user_count ,product_category from `target-
  365217.target_ds.purchase_trend_analysis`
  group by product_category
  order by user_count desc limit 10;
  ```

  Output

user_count	product_category
11115	bed table bath
9670	HEALTH BEAUTY
8641	sport leisure
8334	Furniture Decoration
7827	computer accessories
6964	housewares
5991	Watches present
4545	telephony
4347	Garden tools
4235	automotive

- User Distribution on the basis of purchase time?

Query

```
with purchase_time as (select distinct time(order_purchase_timestamp) as purchase
_time,
case
  when time(order_purchase_timestamp) between "00:00:00" and "05:00:00" then "Da
wn"
  when time(order_purchase_timestamp) between "05:00:00" and "11:58:00" then "M
orning"
  when time(order_purchase_timestamp) between "12:00:00" and "16:58:00" then "Aft
ernoon"
  else "night"
end as result
from `target-365217.target_ds.purchase_trend_analysis`)
select result,count(purchase_time)as total_purchase from purchase_time group by re
sult
order by total_purchase ;
```

Output

result	total_purcha...
Dawn	3650
Morning	12095
Afternoon	14906
night	20167

Insights

- Users interest is mainly towards health and home accessories products.
- User are highly active in night time as compare to dawn or morning time.

Recommendation

- Showcase health and home accessories products in night time with good discount.

4. **Impact on Economy: Analyze the money movemented by e-commerce by looking at order prices, freight and others.**

    1. Get % increase in cost of orders from 2017 to 2018 (include months between Jan to Aug only)?

    Query

```
with order_cost_analysis_2017 as (select FORMAT_DATE('%B',date(order_purchas
e_timestamp)) as order_month,extract(year from date(order_purchase_timestamp) )
as year,
count(order_id) as order_count from `target-
365217.target_ds.purchase_trend_analysis`
where FORMAT_DATE('%B',date(order_purchase_timestamp)) in ("January","Febru
ary","March","April","May","June","July","August")
and extract(year from date(order_purchase_timestamp) )=2017
group by order_month,year
order by order_month desc
),
order_cost_analysis_2018 as (
  select FORMAT_DATE('%B',date(order_purchase_timestamp)) as order_month,extr
act(year from date(order_purchase_timestamp) ) as year,
count(order_id) as order_count from `target-
365217.target_ds.purchase_trend_analysis`
where FORMAT_DATE('%B',date(order_purchase_timestamp)) in ("January","Febru
ary","March","April","May","June","July","August")
and extract(year from date(order_purchase_timestamp) )=2018
group by order_month,year
order by order_month desc
)select ocas.order_month ,ocas.order_count as `order_count_2017` ,
ocae.order_count as `order_count_2018`,((ocae.order_count - ocas.order_count)/100
) as percentage_increase
from order_cost_analysis_2017 as ocas,
order_cost_analysis_2018 as ocae
where ocas.order_month = ocae.order_month
order by ocas.order_month ;
```

    Output

order_month	order_count_2017	order_count_2018	percentage_increase
April	2697	7980	52.83
August	4948	7308	23.6
February	1998	7706	57.08
January	966	8257	72.91
July	4576	7111	25.35
June	3611	7085	34.74
March	3041	8240	51.99
May	4176	7945	37.69

2. Mean & Sum of price and freight value by customer state

Query

```
select customer_state,round(avg(freight_value),2) as avg_freight_value_per_stat
e,
round(avg(price),2) as avg_price_per_state,
round(sum(freight_value),2) as total_freight_value_per_state,
round(sum(price),2) as total_price_state_per_state,
from target-365217.target_ds.purchase_trend_analysis
group by customer_state
order by avg_freight_value_per_state
limit 10;
```

Output

customer_state	avg_freight_...	avg_price_p...	total_freight...	total_price_...
SP	15.15	109.65	718723.07	5202955.05
PR	20.53	119.0	117851.68	683083.76
MG	20.63	120.75	270853.46	1585308.03
RJ	20.96	125.12	305589.31	1824092.67
DF	21.04	125.77	50625.5	302603.94
SC	21.47	124.65	89660.26	520553.34
RS	21.74	120.34	135522.74	750304.02
ES	22.06	121.91	49764.6	275037.31
GO	22.77	126.27	53114.98	294591.95
MS	23.37	142.63	19144.03	116812.64

Insights

- Increase in orders around 30% to 70% in between 2017 and 2018 from January to august.
- As per analysis I found that SP state is the on which has less freight_price which in it self says the delivery cost will be less.

Recommendation

- Give more offers in between month of January and August since it shows that users order count had increased in these months allot.

5. Analysis on sales, freight and delivery time

1. Calculate days between purchasing, delivering and estimated delivery

Query

```
select order_id,
date(order_purchase_timestamp) as order_date,
date(order_delivered_customer_date) as delivery_date,
date(order_estimated_delivery_date) as est_delivery_date,
date_diff(date(order_estimated_delivery_date),date(order_delivered_customer_date)
,day) as order_estimate_delivery_date_diff,

date_diff(date(order_delivered_customer_date),date(order_purchase_timestamp),da
y) as order_purchase_delivery_date_diff,

date_diff(date(order_estimated_delivery_date),date(order_purchase_timestamp),day
) as order_purchase_estimate_date_diff
from target-365217.target_ds.Orders
order by order_id;
```

Output

order_id	order_date	delivery_date	est_delivery...	order_estim...	order_purch...	order_purch...
00010242fe8c5a6d1ba2dd792...	2017-09-13	2017-09-20	2017-09-29	9	7	16
00018f77f2f0320c557190d7a1...	2017-04-26	2017-05-12	2017-05-15	3	16	19
000229ec398224ef6ca0657da...	2018-01-14	2018-01-22	2018-02-05	14	8	22
00024acbcdf0a6daa1e931b03...	2018-08-08	2018-08-14	2018-08-20	6	6	12
00042b26cf59d7ce69dfabb4e...	2017-02-04	2017-03-01	2017-03-17	16	25	41
00048cc3ae777c65dbb7d2a06...	2017-05-15	2017-05-22	2017-06-06	15	7	22
00054e8431b9d7675808bcb8...	2017-12-10	2017-12-18	2018-01-04	17	8	25
000576fe39319847cbb9d288c...	2018-07-04	2018-07-09	2018-07-25	16	5	21
0005a1a1728c9d785b8e2b08...	2018-03-19	2018-03-29	2018-03-29	0	10	10
0005f50442cb953dcd1d21e1f...	2018-07-02	2018-07-04	2018-07-23	19	2	21

2. Create columns

- time_to_delivery = order_purchase_timestamp-order_delivered_customer_date
- diff_estimated_delivery = order_estimated_delivery_date-order_delivered_customer_date

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

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

CTE Prepared for below Questions

```
with sales_freight_analysis as (select
c.customer_state,
o.order_purchase_timestamp,
o.order_delivered_customer_date,
o.order_estimated_delivery_date,
date_diff(order_delivered_customer_date,o.order_purchase_timestamp,day) as time_to_deli
very,
abs(date_diff(o.order_delivered_customer_date,order_estimated_delivery_date,day)) as diff
_estimated_delivery,
avg(oi.freight_value) as avg_freight_value,
avg(date_diff(order_delivered_customer_date,o.order_purchase_timestamp,day)) as avg_ti
me_to_delivery,
avg(date_diff(o.order_delivered_customer_date,order_estimated_delivery_date,day)) as avg
_diff_estimated_delivery
from target-365217.target_ds.Customers as c
left join target-365217.target_ds.Orders as o
on c.customer_id = o.customer_id
left join target-365217.target_ds.Order_items as oi
on o.order_id = oi.order_id
group by c.customer_state,
o.order_purchase_timestamp,
o.order_delivered_customer_date,
o.order_estimated_delivery_date)

select * from sales_freight_analysis;
```

Output

customer_state	order_purchase_timestamp	order_delivered_customer_date	order_estimated_delivery_date	time_to_deli...	diff_estimat...	avg_freight_...	avg_time_to...	avg_diff_est...
CE	2018-02-16 18:31:12 UTC	2018-04-03 23:50:57 UTC	2018-03-20 00:00:00 UTC	46	14	27.59	46.0	14.0
CE	2018-03-03 16:51:20 UTC	2018-04-20 20:58:44 UTC	2018-04-04 00:00:00 UTC	48	16	27.1	48.0	16.0
CE	2018-05-12 14:03:02 UTC	2018-06-18 17:19:58 UTC	2018-06-12 00:00:00 UTC	37	6	38.14	37.0	6.0
CE	2018-03-19 08:48:09 UTC	2018-04-17 17:31:50 UTC	2018-05-07 00:00:00 UTC	29	19	68.99	29.0	-19.0
SP	2017-11-18 16:10:29 UTC	2018-01-04 18:04:30 UTC	2017-12-07 00:00:00 UTC	47	28	12.89	47.0	28.0
SP	2018-01-06 08:13:47 UTC	2018-02-14 19:41:34 UTC	2018-03-07 00:00:00 UTC	39	20	17.29	39.0	-20.0
SP	2017-11-30 18:53:51 UTC	2018-01-03 22:41:26 UTC	2018-01-03 00:00:00 UTC	34	0	8.31	34.0	0.0
MG	2017-10-25 15:07:24 UTC	2017-11-25 12:57:07 UTC	2017-11-16 00:00:00 UTC	30	9	25.11	30.0	9.0
MG	2018-01-18 12:49:48 UTC	2018-02-16 16:23:00 UTC	2018-02-16 00:00:00 UTC	29	0	32.08	29.0	0.0
MG	2017-11-28 13:50:19 UTC	2017-12-27 21:43:23 UTC	2017-12-21 00:00:00 UTC	29	6	14.43	29.0	6.0
MG	2017-09-27 16:03:02 UTC	2017-11-10 19:03:45 UTC	2017-10-24 00:00:00 UTC	44	17	17.24	44.0	17.0
MG	2018-01-08 17:51:23 UTC	2018-02-07 23:22:33 UTC	2018-02-05 00:00:00 UTC	30	2	17.61	30.0	2.0
MG	2018-03-13 13:18:23 UTC	2018-04-19 18:48:29 UTC	2018-04-04 00:00:00 UTC	37	15	15.19	37.0	15.0
BA	2018-05-04 08:41:08 UTC	2018-06-12 12:32:08 UTC	2018-05-30 00:00:00 UTC	39	13	22.82	39.0	13.0
RS	2018-02-28 23:30:36 UTC	2018-04-04 17:12:47 UTC	2018-03-27 00:00:00 UTC	34	8	16.29	34.0	8.0
RS	2018-02-25 18:53:53 UTC	2018-03-27 22:06:48 UTC	2018-03-27 00:00:00 UTC	30	0	15.1	30.0	0.0

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

Query

select customer_state from sales_freight_analysis order by avg_freight_value desc limit 5;

Output

customer_state
PI
SC
PR
SP
MT

select customer_state from sales_freight_analysis order by avg_freight_value limit 5;

Output

customer_state
SP
SP
RJ
RS
RS

Top 5 states with highest/lowest average time to delivery?

Query

select customer_state from sales_freight_analysis order by avg_time_to_delivery desc limit 5
;

Output

customer_state
ES
RJ
PA
PI
PI

Query

select customer_state from sales_freight_analysis order by avg_time_to_delivery limit 5;

Output

customer_state
RN
SP
SP
SP
SP

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

Query

select customer_state from sales_freight_analysis order by avg_diff_estimated_delivery desc limit 5;

Output

customer_state
RJ
ES
SP
SP
SE

select customer_state from sales_freight_analysis order by avg_diff_estimated_delivery limit 5;

Output

customer_state
RN
SP
SP
SP
SP

Insights

- As per analysis it is found that there is a huge gap between estimated date and delivery date of product.

Recommendation

- Please try to reduce the gap between estimated date and delivery date of product.

6. Payment type analysis

1. Month over Month count of orders for different payment types

Query

```
select p.payment_type,FORMAT_DATE('%B',date(o.order_purchase_timestamp)) as ord
er_month,
count(o.order_id) as order_count, from target-365217.target_ds.Payments as p
left join target-365217.target_ds.Orders as o
on p.order_id = o.order_id
group by p.payment_type,order_month
order by order_count desc
limit 10;
```

Output

payment_type	order_month	order_count
credit_card	May	8350
credit_card	August	8269
credit_card	July	7841
credit_card	March	7707
credit_card	April	7301
credit_card	June	7276
credit_card	February	6609
credit_card	January	6103
credit_card	November	5897
credit_card	December	4378

2. Distribution of payment installments and count of orders?

Query

```
select payment_installments,count(order_id) as order_count from target-
365217.target_ds.Payments
group by payment_installments
limit 10;
```

Output

payment_in...	order_count
0	2
1	52546
2	12413
3	10461
4	7098
5	5239
6	3920
7	1626
8	4268
9	644

Insights

- Most of the users are relying on credit card payment
- Most of the User refer to do payment by installments.

Recommendation

- Provide more credit card option if not then provide UPI payment as secondary payment option.