

# Target Project

## 1.Initial exploration of dataset like checking the characteristics of data

```
SELECT * FROM `Target.customers`
```

customers							
QUERY SHARE COPY SNAPSHOT DELETE EXPORT							
CHEMA DETAILS PREVIEW LINEAGE PREVIEW							
Filter Enter property name or value							
<input type="checkbox"/>	Field name	Type	Mode	Collation	Default value	Policy tags ?	Description
<input type="checkbox"/>	customer_id	STRING	NULLABLE				
<input type="checkbox"/>	customer_unique_id	STRING	NULLABLE				
<input type="checkbox"/>	customer_zip_code_prefix	INTEGER	NULLABLE				
<input type="checkbox"/>	customer_city	STRING	NULLABLE				
<input type="checkbox"/>	customer_state	STRING	NULLABLE				

Row	customer_id	customer_unique_id	customer_zip_code_prefix	customer_city	customer_state
1	0735e7e4298a2ebbb46649346...	fc003b1bdc0df64b4d065d9b...	59650	acu	RN
2	903b3d86e3990db01619a4ebe...	46824822b15da44e983b021d...	59650	acu	RN
3	38c97666e962d4fea7fd6a83e...	b6108acc674ae5c99e29adc10...	59650	acu	RN
4	77c2f46cf580f4874c9a5751c2...	402cce5c0509000eed9e77fec...	63430	ico	CE
5	4d3ef4cfff8ad4767c199c36a...	6ba00666ab7eada5ceec279b2...	63430	ico	CE
6	3000841b86e1fbe9493b52324...	796a0b1a21f597704057184a1...	63430	ico	CE
7	3c325415ccc7e622c66dec4bc...	05d1d2d9f0161c5f397ce7fc77...	63430	ico	CE
8	04f3a7b250e3be964f01bf22bc...	c34585a0276ecc5e4fb03de75...	63430	ico	CE
9	894202b8ef01f4719a4691e79...	01a4fe5fc00bbdb0b0a4af5a53...	63430	ico	CE
10	9d715b9fb75a9d081c14126c0...	8f399f3b7ace8e6245422c9e1f...	63430	ico	CE
11	018184ac5f52a821bb00f3ef21...	54fc4ff419d5e05db5fe42906b...	63430	ico	CE

Results per page: 50 1 – 50 of 99441

```
SELECT * FROM `Target.order_items`
```

order_items							
<a href="#">QUERY</a> <a href="#">SHARE</a> <a href="#">COPY</a> <a href="#">SNAPSHOT</a> <a href="#">DELETE</a> <a href="#">EXPORT</a>							
<a href="#">SCHEMA</a> <a href="#">DETAILS</a> <a href="#">PREVIEW</a> <a href="#">LINEAGE</a> <a href="#">PREVIEW</a>							
Filter Enter property name or value							
<input type="checkbox"/>	Field name	Type	Mode	Collation	Default value	Policy tags ?	Description
<input type="checkbox"/>	order_id	STRING	NULLABLE				
<input type="checkbox"/>	order_item_id	INTEGER	NULLABLE				
<input type="checkbox"/>	product_id	STRING	NULLABLE				
<input type="checkbox"/>	seller_id	STRING	NULLABLE				
<input type="checkbox"/>	shipping_limit_date	TIMESTAMP	NULLABLE				

Row	order_id	order_item_id	product_id	seller_id	shipping_limit_date	price	freight_value
1	f09e36e258656850b926...	1	44d53f1240d6332232...	b64d51f0435e884e8d...	2018-07-09 13:31:36 UTC	3.0	12.79
2	f9ccaff7267fd0cf076e7...	1	44d53f1240d6332232...	b64d51f0435e884e8d...	2018-08-14 14:04:44 UTC	3.0	15.23
3	c79bdf061e2228860920...	1	5304ff3fa35856a156e...	cf6f6bc4df3999b9c64...	2017-05-12 19:05:20 UTC	3.5	8.72
4	37193e64eb9a46b7f319...	1	98224bfc1eaadb3a394...	ce616e1913288884e7...	2018-06-28 01:30:49 UTC	3.5	7.39
5	95d6357ffe41aa6d2998...	1	98224bfc1eaadb3a394...	ce616e1913288884e7...	2018-06-12 19:15:14 UTC	3.5	18.23
6	95d6357ffe41aa6d2998...	2	98224bfc1eaadb3a394...	ce616e1913288884e7...	2018-06-12 19:15:14 UTC	3.5	18.23
7	95d6357ffe41aa6d2998...	3	98224bfc1eaadb3a394...	ce616e1913288884e7...	2018-06-12 19:15:14 UTC	3.5	18.23
8	95d6357ffe41aa6d2998...	4	98224bfc1eaadb3a394...	ce616e1913288884e7...	2018-06-12 19:15:14 UTC	3.5	18.23
9	95d6357ffe41aa6d2998...	5	98224bfc1eaadb3a394...	ce616e1913288884e7...	2018-06-12 19:15:14 UTC	3.5	18.23
10	dde867f83e689b016778...	1	914323edd50192310d...	2c9e548be18521d1c4...	2017-10-20 14:50:12 UTC	4.5	11.85

Results per page: 50 1 – 50 of 112650

SELECT \* FROM `Target.geolocation`

geolocation							
<a href="#">QUERY</a> <a href="#">SHARE</a> <a href="#">COPY</a> <a href="#">SNAPSHOT</a> <a href="#">DELETE</a> <a href="#">EXPORT</a>							
<a href="#">SCHEMA</a> <a href="#">DETAILS</a> <a href="#">PREVIEW</a> <a href="#">LINEAGE</a> <a href="#">PREVIEW</a>							
Filter Enter property name or value							
<input type="checkbox"/>	Field name	Type	Mode	Collation	Default value	Policy tags ?	Description
<input type="checkbox"/>	geolocation_zip_code_prefix	INTEGER	NULLABLE				
<input type="checkbox"/>	geolocation_lat	FLOAT	NULLABLE				
<input type="checkbox"/>	geolocation_lng	FLOAT	NULLABLE				
<input type="checkbox"/>	geolocation_city	STRING	NULLABLE				
<input type="checkbox"/>	geolocation_state	STRING	NULLABLE				

Row	geolocation_zip_code_prefix	geolocation_lat	geolocation_lng	geolocation_city	geolocation_state	
1	49010	-10.910514518754546	-37.052400776992329	aracaju	SE	
2	49047	-10.9268145	-37.071063000000009	aracaju	SE	
3	49030	-10.970164794304576	-37.061643830745815	aracaju	SE	
4	49048	-10.940183531738979	-37.070850242714528	aracaju	SE	
5	49050	-10.927157352800547	-37.063078689600516	aracaju	SE	
6	49015	-10.923370500160772	-37.045169150380509	aracaju	SE	
7	49045	-10.930406582318476	-37.067178493623359	aracaju	SE	
8	49052	-10.922973517889163	-37.057752502914184	aracaju	SE	
9	49044	-10.992080999999946	-37.103470999999956	aracaju	SE	
10	49048	-10.940235501586523	-37.071043333891133	aracaju	SE	

Results per page: 50 1 - 50 of 1000163

SELECT \* FROM `Target.order\_reviews`

order\_reviews

QUERY

SHARE

COPY

SNAPSHOT

DELETE

EXPORT

CHEMA

DETAILS

PREVIEW

LINEAGE

PREVIEW

Filter

Enter property name or value

<input type="checkbox"/>	Field name	Type	Mode	Collation	Default value	Policy tags ?	Description
<input type="checkbox"/>	<a href="#">review_id</a>	STRING	NULLABLE				
<input type="checkbox"/>	<a href="#">order_id</a>	STRING	NULLABLE				
<input type="checkbox"/>	<a href="#">review_score</a>	INTEGER	NULLABLE				
<input type="checkbox"/>	<a href="#">review_comment_title</a>	STRING	NULLABLE				
<input type="checkbox"/>	<a href="#">review_creation_date</a>	TIMESTAMP	NULLABLE				

Row	review_id	order_id	review_score	review_comment_title	review_creation_date	review_answer_timestamp
1	be7e2989673cb2a147b87ba73...	777c67eab7c0712ccde8ffbb2...	1	null	0001-04-17 00:00:00 UTC	0001-04-17 07:40:00 UTC
2	e12151267e4594d69eda14a87...	4338a4463f7f9193d2a03a83a...	1	null	0001-04-17 00:00:00 UTC	0001-04-17 09:04:00 UTC
3	41d614b133efebcd10352001b...	b8aaeda740b17cf925d3f2a13...	1	null	0001-04-17 00:00:00 UTC	0002-04-17 03:48:00 UTC
4	c950324a42c5796d06f569f77...	b159d0ce7cd881052da94fa16...	1	null	0001-04-17 00:00:00 UTC	0001-04-17 10:24:00 UTC
5	76823ada94c8861ecebcbfc7c...	2a3007ed051b02a0e0dd0709c...	1	null	0001-04-17 00:00:00 UTC	0002-04-17 13:58:00 UTC
6	fe270df00abcb5c39fc7385143...	a39d3db795a5cf4c8b6c9dd05...	1	null	0001-04-17 00:00:00 UTC	0003-04-17 12:49:00 UTC
7	1b71e0b29ec2faa0a029ded43...	0e530f6be154c9d7e7b12f341...	1	null	0001-04-17 00:00:00 UTC	0010-04-17 12:45:00 UTC
8	efe4020a945ee6fece58606694...	264c045399fb02e9f309f715c2...	1	null	0001-04-17 00:00:00 UTC	0002-04-17 01:16:00 UTC
9	23ee6bc249280dca0e807e13...	2ba1366baecad3c3536f27546...	1	null	0001-04-17 00:00:00 UTC	0001-04-17 14:34:00 UTC
10	63c6cec8cb9a2d0990b339998...	aaae80f5b6239bd9e1b22e9aa...	1	null	0001-04-17 00:00:00 UTC	0004-04-17 12:07:00 UTC

Results per page: 50 1 - 50 of 99224

SELECT \* FROM `Target.orders`

orders	QUERY	SHARE	COPY	SNAPSHOT	DELETE	EXPORT
CHEMA	DETAILS	PREVIEW	LINEAGE	PREVIEW		
Filter Enter property name or value						
<input type="checkbox"/>	Field name	Type	Mode	Collation	Default value	Policy tags ? Description
<input type="checkbox"/>	order_id	STRING	NULLABLE			
<input type="checkbox"/>	customer_id	STRING	NULLABLE			
<input type="checkbox"/>	order_status	STRING	NULLABLE			
<input type="checkbox"/>	order_purchase_timestamp	TIMESTAMP	NULLABLE			
<input type="checkbox"/>	order_approved_at	TIMESTAMP	NULLABLE			

Row	order_id	customer_id	order_status	order_purchase_timestamp	order_approved_at	order_delivered_carrier_date	order_delivered_custor	order_estimated_delivery_date
1	7a4df5d8c9ff40...	725e9c756054...	created	2017-11-25 11:10:33 UTC	null	null	null	2017-12-12 00:00:00 UTC
2	35de4050331c...	4ee64f4bfc542...	created	2017-12-05 01:07:58 UTC	null	null	null	2018-01-08 00:00:00 UTC
3	b5359909123f...	438449d4af89...	created	2017-12-05 01:07:52 UTC	null	null	null	2018-01-11 00:00:00 UTC
4	dba5062fbd3...	964a6df3d9bdf...	created	2018-02-09 17:21:04 UTC	null	null	null	2018-03-07 00:00:00 UTC
5	90ab3e7d5254...	7d61b9f4f2160...	created	2017-11-06 13:12:34 UTC	null	null	null	2017-12-01 00:00:00 UTC
6	fa65dad1b0e8...	9af2372a1e49...	shipped	2017-04-20 12:45:34 UTC	2017-04-22 09:1...	2017-04-24 11:31:17 UTC	null	2017-05-18 00:00:00 UTC
7	1df2775799ee...	1240c2e65c46...	shipped	2017-07-13 11:03:05 UTC	2017-07-13 11:1...	2017-07-18 18:17:30 UTC	null	2017-08-14 00:00:00 UTC
8	6190a94657e1...	5fc4c97dcb639...	shipped	2017-07-11 13:36:30 UTC	2017-07-11 13:4...	2017-07-13 17:55:46 UTC	null	2017-08-14 00:00:00 UTC
9	58ce513a55c7...	530d41b47b9d...	shipped	2017-07-29 18:05:07 UTC	2017-07-29 18:1...	2017-07-31 16:41:59 UTC	null	2017-08-14 00:00:00 UTC
10	088683f795a3...	58d89fd1f8638...	shipped	2017-07-13 10:02:47 UTC	2017-07-14 02:2...	2017-07-20 20:02:58 UTC	null	2017-08-14 00:00:00 UTC

Results per page: 50 1 - 50 of 99441

SELECT \* FROM `Target.payments`

payments	QUERY	SHARE	COPY	SNAPSHOT	DELETE	EXPORT
CHEMA	DETAILS	PREVIEW	LINEAGE	PREVIEW		
Filter Enter property name or value						
<input type="checkbox"/>	Field name	Type	Mode	Collation	Default value	Policy tags ? Description
<input type="checkbox"/>	order_id	STRING	NULLABLE			
<input type="checkbox"/>	payment_sequential	INTEGER	NULLABLE			
<input type="checkbox"/>	payment_type	STRING	NULLABLE			
<input type="checkbox"/>	payment_installments	INTEGER	NULLABLE			
<input type="checkbox"/>	payment_value	FLOAT	NULLABLE			

Row	order_id	payment_sequential	payment_type	payment_installments	payment_value
1	1a57108394169c0b47d8f876acc9ba2d	2	credit_card	0	129.94
2	744bade1fc9ff3f1d860ace076d422	2	credit_card	0	58.69
3	8bcbe01d44d147f901cd3192671144db	4	voucher	1	0.0
4	fa65dad1b0e818e3ccc5cb0e39231352	14	voucher	1	0.0
5	6ccb433e00daae1283ccc956189c82ae	4	voucher	1	0.0
6	4637ca194b6387e2d538dc89b124b0ee	1	not_defined	1	0.0
7	00b1cb0320190ca0daa2c88b35206009	1	not_defined	1	0.0
8	45ed6e85398a87c253db47c2d9f48216	3	voucher	1	0.0
9	fa65dad1b0e818e3ccc5cb0e39231352	13	voucher	1	0.0
10	c8c528189310eaa44a745b8d9d26908b	1	not_defined	1	0.0

Results per page: 50 1 - 50 of 103886

```
SELECT * FROM `Target.products`
```

products	QUERY	SHARE	COPY	SNAPSHOT	DELETE	EXPORT
CHEMA	DETAILS	PREVIEW	LINEAGE	PREVIEW		
Filter Enter property name or value						
<input type="checkbox"/>	Field name	Type	Mode	Collation	Default value	Policy tags ? Description
<input type="checkbox"/>	product_id	STRING	NULLABLE			
<input type="checkbox"/>	product_category	STRING	NULLABLE			
<input type="checkbox"/>	product_name_length	INTEGER	NULLABLE			
<input type="checkbox"/>	product_description_length	INTEGER	NULLABLE			
<input type="checkbox"/>	product_photos_qty	INTEGER	NULLABLE			

Row	product_id	product_category	product_name_length	product_description_length	product_photos_qty	product_weight	product_length	product_height	product_width
1	a0ab96e461d745...	climatization	41	717	1	1050	18	7	8
2	4d7585daba2f8b3...	fixed telephony	53	897	2	300	15	8	9
3	20ae7c024ede61...	fixed telephony	25	455	1	330	17	11	9
4	ad7aebd205805...	Construction Tools Tools	41	2526	2	1150	22	10	9
5	16d096faa275829...	Agro Industria e Comercio	50	1153	1	4050	11	18	10
6	980ecbcc15fe174...	Agro Industria e Comercio	48	157	1	250	17	3	10
7	2b6535d32c6996...	CONSTRUCTION SECURI...	50	428	2	333	16	9	10
8	fdc45aa23bb831...	Christmas articles	39	258	1	250	16	12	11
9	33abaf0c29c1484...	fixed telephony	41	261	9	100	19	3	11
10	c023dc7b323cb5...	Christmas articles	43	262	1	600	16	15	11
Results per page: 50 1 - 50 of 32951									

```
SELECT * FROM `Target.sellers`
```

sellers	QUERY	SHARE	COPY	SNAPSHOT	DELETE	EXPORT
HEMA	DETAILS	PREVIEW	LINEAGE	PREVIEW		
Filter Enter property name or value						
	Field name	Type	Mode	Collation	Default value	Policy tags ? Description
<input type="checkbox"/>	seller_id	STRING	NULLABLE			
<input type="checkbox"/>	seller_zip_code_prefix	INTEGER	NULLABLE			
<input type="checkbox"/>	seller_city	STRING	NULLABLE			
<input type="checkbox"/>	seller_state	STRING	NULLABLE			

Row	seller_id	seller_zip_code_prefix	seller_city	seller_state
1	4be2e7f96b4fd749d52dff41f80e39dd	69900	rio branco	AC
2	327b89b872c14d1c0be7235ef4871685	69005	manaus	AM
3	4221a7df464f1fe2955934e30ff3a5a1	48602	bahia	BA
4	651530bf5c607240ccdd89a30c9c9712	44600	ipira	BA
5	2b402d5dc42554061f8ea98d1916f148	44900	irece	BA
6	d03698c2efd04a549382afa6623e27fb	45658	ilheus	BA
7	c72de06d72748d1a0dfb2125be43ba63	46430	guanambi	BA
8	fc59392d66ef99377e50356ee4f3b4e1	40243	salvador	BA
9	b00af24704019bd2e1b335e70ad11f7c	40130	salvador	BA
10	eb4a59a06b3948e851a7d7a830bcb42	41820	salvador	BA

Results per page: 50 ▼ 1 – 50 of 3095

Check the data range available:

select

min(order\_purchase\_timestamp) as min\_date,  
max(order\_purchase\_timestamp) as max\_date

from `target\_sql.orders`

Row	min_date	max_date
1	2016-09-04 21:15:19 UTC	2018-10-17 17:30:18 UTC

Customer data :

Customer State	Number of Customers
AC	81
AL	413
AM	148
AP	68
BA	3,380
CE	1,336
DF	2,140
ES	2,033
GO	2,020
MA	747
MG	11,635
MS	715
MT	907
PA	975
PB	536
PE	1,652
PI	495
PR	5,045
RJ	12,852
RN	485
RO	253
RR	46
RS	5,466
SC	3,637
SE	350
SP	41,746
TO	280
Grand Total	99,441

## In-depth Exploration: -

Cities and States of customers ordered during the year 2017

```
SELECT
  o.order_purchase_timestamp,
  o.customer_id,
  c.customer_city,
  c.customer_state
FROM
  Target.orders o
LEFT JOIN
  Target.customers c
ON
  c.customer_id=o.customer_id
WHERE
  o.order_purchase_timestamp BETWEEN '2017-01-01 00:00:00'
  AND '2017-12-31 23:59:59'
ORDER BY
  o.order_purchase_timestamp
```

Row	order_purchase_timestamp	customer_id	customer_city	customer_state
1	2017-01-05 11:56:06 UTC	c24fc5f9a446b4d8262041b9c...	curitiba	PR
2	2017-01-05 12:01:20 UTC	ff3c6d35b4927c9b965aa6a68...	curitiba	PR
3	2017-01-05 12:06:36 UTC	d2c63ad286e3ca9dd69218008...	curitiba	PR
4	2017-01-05 12:09:08 UTC	e6bee29442c761ff0dfb9a4414...	curitiba	PR
5	2017-01-05 12:11:23 UTC	42d857e22f61f10476f0cda1aa...	curitiba	PR
6	2017-01-05 12:14:58 UTC	64b56cb1662271f54a81d96a2...	curitiba	PR
7	2017-01-05 13:01:48 UTC	efdf4a7c78d7c364046efb6903...	curitiba	PR
8	2017-01-05 13:29:03 UTC	8fa33a3159dfc303b8aeccf859...	curitiba	PR
9	2017-01-05 13:31:22 UTC	3a7856f2a103364d279de7f72...	curitiba	PR
10	2017-01-05 13:33:45 UTC	63433f6eb6f0a79cbb24e940d...	curitiba	PR

Total 45101 orders made in that year.

## Cities with most customer acquisition

```
SELECT
  c.customer_city,
  (select count(customer_city) from Target.customers t where c.customer_city=t.customer_city ) as Market_share
FROM
  Target.orders o
LEFT JOIN
  Target.customers c
ON
  c.customer_id=o.customer_id
WHERE
  o.order_purchase_timestamp BETWEEN '2017-01-01 00:00:00'
  AND '2017-12-31 23:59:59'
group by customer_city
order by Market_share Desc
```

Row	customer_city	Market_share
1	sao paulo	15540
2	rio de janeiro	6882
3	belo horizonte	2773
4	brasilvia	2131
5	curitiba	1521
6	campinas	1444
7	porto alegre	1379
8	salvador	1245
9	guarulhos	1189
10	sao bernardo do campo	938
11	niteroi	849
12	santo andre	797

Total No of cities = 3290

## Cities with most Sales

```

SELECT c.customer_city, sum(oi.price) as Total_sales
FROM
  (Target.order_items oi
LEFT JOIN
  Target.orders o ON oi.order_id=o.order_id )
LEFT JOIN
  Target.customers c
ON
  c.customer_id=o.customer_id

WHERE
  o.order_purchase_timestamp BETWEEN '2017-01-01 00:00:00'
  AND '2017-12-31 23:59:59'

group by customer_city
order by Total_sales DESC

```



Row	customer_city	Total_sales
1	sao paulo	782279.340...
2	rio de janeiro	494259.250...
3	belo horizonte	150067.259...
4	brasilia	134057.439...
5	porto alegre	93697.1699...
6	curitiba	85103.7599...
7	salvador	84836.6999...
8	campinas	81850.4399...
9	guarulhos	59767.9000...
10	niteroi	48942.5500...
11	fortaleza	45993.9300...
12	sao bernardo do campo	45848.9400...

Total No of cities = 3273

Here we can see that Rio de Janeiro has market share of **6882** compared to Sao Paulo which has **15540** but total sales is **494260**. Compared to **782280**.

This mean people in Rio de Janeiro spend more than Sao Paulo per person.

\*So, we can assume that they probably buy expensive products.

**Let's confirm this theory!**

```

SELECT
  SUM(price)
FROM
  `Target.order_items`
WHERE
  price > 1000
  AND order_id IN (
    SELECT
      order_id
    FROM
      `Target.orders`
    WHERE
      order_purchase_timestamp BETWEEN '2017-01-01 00:00:00'
      AND '2017-12-31 23:59:59'
      AND customer_id IN (
        SELECT

```

```
customer_id
FROM
`Target.customers`
WHERE
customer_city = 'rio de janeiro'))
```

Row	total_sales
1	58388.2299...

-- Rio De Janeiro Sales OF products above 1000  
 -- Total SUM(Sales) - 58388 but total sales is 494260.

## Let's check below 1000:-

```
SELECT
SUM(price)
FROM
`Target.order_items`
WHERE
price < 1000
AND order_id IN (
SELECT
order_id
FROM
`Target.orders`
WHERE
order_purchase_timestamp BETWEEN '2017-01-01 00:00:00'
AND '2017-12-31 23:59:59'
AND customer_id IN (
SELECT
customer_id
FROM
`Target.customers`
WHERE
customer_city = 'rio de janeiro'))
```

Row	total_sales
1	435871.020...

-- Rio De Janeiro Sales of products below 1000  
 -- Total Sum(Sales) – 435871 and total sales is 494260.

## After applying more ranges we got the range of products which are highly sold

-- Rio De Janeiro Sales - 2017  
 -- Total Sales - 494259

```
SELECT
SUM(price)
FROM
`Target.order_items`
WHERE
```

```

price BETWEEN 50
AND 300
AND order_id IN (
SELECT
    order_id
FROM
    `Target.orders`
WHERE
    order_purchase_timestamp BETWEEN '2017-01-01 00:00:00'
    AND '2017-12-31 23:59:59'
    AND customer_id IN (
    SELECT
        customer_id
    FROM
        `Target.customers`
    WHERE
        customer_city = 'rio de janeiro'))

```

Row	Total_sales
1	281383.219...

## Insite:

**Most sold products are in range of 50-300**

Now let's see weather any products in the above range are not available in Rio de Janeiro.

```

SELECT
    product_id
FROM
    `Target.order_items`
WHERE
    price BETWEEN 50
    AND 300
    AND order_id IN (
    SELECT
        order_id
    FROM
        `Target.orders`
    WHERE
        order_purchase_timestamp BETWEEN '2017-01-01 00:00:00'
        AND '2017-12-31 23:59:59'
        AND customer_id IN (
        SELECT
            customer_id
        FROM
            `Target.customers`
        WHERE
            customer_city = 'rio de janeiro'))

```

Row	product_id
1	12a947baa3e0ab6254ebfec78...
2	163e6400e6dadd0fe04775c5e...
3	2615fceb03b49ac6e17fa5b8f2...
4	810cb3c5af006624fa32d3620...
5	a10e0fcb1c409869c3c6da4eb...
6	a10e0fcb1c409869c3c6da4eb...
7	aea6413d9a7fadfe4a7479cafe...
8	cde89c19bf58d71bc836b1495...
9	ea44caac707f7f1325182a538...
10	f6e3aeb40c00aec24ade8333c...

Here 2324 products are there range between 50 -300.

```
select distinct product_id , price from `Target.order_items`
where product_id in
(
select product_id from `Target.products`
where product_id not in
(select product_id from `Target.order_items` where price between 50 and 300 and order_id in
(select order_id from `Target.orders` where order_purchase_timestamp BETWEEN '2017-01-01 00:00:00' AND '2017-12-
31 23:59:59' and customer_id in
(select customer_id from `Target.customers` where customer_city = 'rio de janeiro')))
) and price between 50 and 300
order by price desc
```

Row	product_id	price
1	1fd10191b0bcd5aa4020c0721...	300.0
2	50fb810523dcd3bd271eb7b6d...	300.0
3	74e901ea4caa0b0488e835043...	300.0
4	98ba4d6506268560a9b0a03b...	300.0
5	9c84b4ffd1a85bf2feb6cba0e2...	300.0
6	a1692a17fad005e22bf5a5435...	300.0
7	bf2f8d918c34551d0bc8861f37...	300.0
8	c155f6ba0dae0f11299f1e33faf...	300.0
9	c71bf92cf123eb294304073fc1...	300.0
10	cfe9e8e1f0400c8cc665d41722...	300.0

Results per page: 50 ▼ 1 – 50 of 21573

## Recommendations :-

Here 21573 products are in range of 50-300 that should also be sold in Rio de Janeiro.

---

```
select extract(year from order_purchase_timestamp) as year ,sum(p.payment_value) as yearly_sales from `Target.orders` o join `Target.payments` p on  
o.order_id = p.order_id  
group by year  
order by year
```

Row	year	yearly_sales
1	2016	59362.3400...
2	2017	7249746.72...
3	2018	8699763.04...

Here, we can see growing trend of E-Commerce in Brazil.

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

```
select count(*) as No_of_Purchases, Preferred_Time  
from  
(select  
case  
when extract(HOUR from order_purchase_timestamp) < 6  
then 'Dawn'  
when extract(HOUR from order_purchase_timestamp) < 12  
then 'Morning'  
when extract(HOUR from order_purchase_timestamp) < 18  
then 'Afternoon'  
else 'Night'  
end as Preferred_Time  
from `Target.orders`)
```

group by Preferred\_Time

Row	No_of_Purchase	Preferred_Time
1	22240	Morning
2	4740	Dawn
3	38361	Afternoon
4	34100	Night

Here Brazilians prefer to order at **Afternoon** then at **Night**, then at **Morning** and least during **Dawn**.

Time\_to\_delivery = order\_purchase\_timestamp-order\_delivered\_customer\_date

```
select o.customer_id,c.customer_state, (o.order_purchase_timestamp-order_delivered_customer_date) as time_to_delivery
from `Target.orders` o
join `Target.customers` c on c.customer_id =o.customer_id
order by time_to_delivery desc
```

Row	customer_id	customer_state	time_to_delivery
1	b19da0df0271e8a3553e3670f...	RJ	0-0 0 -12:48:7
2	922a46283625e9c096bfd9989...	SP	0-0 0 -18:45:10
3	6aef84c09844a371d82a49152...	SP	0-0 0 -20:31:39
4	ff58662c328f81d3ee549c9caa...	BA	0-0 0 -20:43:20
5	118295a853acb536efff13740f...	SP	0-0 0 -21:22:41
6	198f511b5a75bf936a96f1d476...	SP	0-0 0 -21:25:11
7	225aed9e773953084b09cf496...	SP	0-0 0 -22:27:49
8	344423c2e26d47d2b6d3dd36...	SP	0-0 0 -22:30:59
9	c5e200d485ae35a7036cc2e7c...	SP	0-0 0 -23:37:30
10	18c934f4cdc994cd04eb13bce...	SP	0-0 0 -23:39:18

Diff\_estimated\_delivery=order\_estimated\_delivery\_date\_order\_delivered\_customer\_date

```
select o.customer_id,c.customer_state, (o.order_estimated_delivery_date-
o.order_delivered_customer_date) as diff_estimated_delivery
from `Target.orders` o
join `Target.customers` c on c.customer_id = o.customer_id
order by diff_estimated_delivery desc
```

Row	customer_id	customer_state	diff_estimated_delivery
1	a5fbb6579eacbeb02752a143b...	SP	0-0 0 3504:23:13
2	964253ff0e4e08180064764a4...	MA	0-0 0 3345:32:15
3	32cef4bdd6bfa50612d81dc77...	RS	0-0 0 3223:24:17
4	6357fffb5704244d552615bbfc...	SP	0-0 0 2962:24:6
5	6210a37f9d6a265a4f3fbe2c21...	RJ	0-0 0 2602:10:53
6	b882cbae40f34e60a3ec3efefa...	SP	0-0 0 1993:41:37
7	816642e9995c2461f2172469e...	SP	0-0 0 1990:7:26
8	5fd5cb96f515996e88d84c610...	ES	0-0 0 1855:47:10
9	0f9043c635f86f7eb1e8fc0770...	MG	0-0 0 1852:38:21
10	6e5b6ba2e8de70d9e920b541a...	MG	0-0 0 1851:37:50

-- Average Freight Value of 5 states

```
select c.customer_state,avg(freight_value) as avg_freight from
(select oi.freight_value, o.order_id, o.customer_id
from `Target.orders` o
join `Target.order_items` oi on o.order_id = oi.order_id) as a
join `Target.customers` c on c.customer_id = a.customer_id
group by c.customer_state
order by avg_freight desc limit 5
```

Row	customer_state	avg_freight
1	RR	42.9844230...
2	PB	42.7238039...
3	RO	41.0697122...
4	AC	40.0733695...
5	PI	39.1479704...

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

```
select o.customer_id,c.customer_state, (o.order_estimated_delivery_date-
o.order_delivered_customer_date) as diff_estimated_delivery
from `Target.orders` o
join `Target.customers` c on c.customer_id=o.customer_id
order by diff_estimated_delivery desc limit 5
```

Row	customer_id	customer_state	diff_estimated_delivery
1	a5fbb6579eacbeb02752a143b...	SP	0-0 0 3504:23:13
2	964253ff0e4e08180064764a...	MA	0-0 0 3345:32:15
3	32cef4bdd6bfa50612d81dc77...	RS	0-0 0 3223:24:17
4	6357fffb5704244d552615bbfc...	SP	0-0 0 2962:24:6
5	6210a37f9d6a265a4f3fbe2c21...	RJ	0-0 0 2602:10:53

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

```
select o.customer_id,c.customer_state, avg(o.order_purchase_timestamp-
order_delivered_customer_date) as avg_time_to_delivery
from `Target.orders` o
join `Target.customers` c on c.customer_id=o.customer_id
group by customer_state,customer_id
order by avg_time_to_delivery desc limit 5
```

Row	customer_id	customer_state	avg_time_to_delivery
1	b19da0df0271e8a3553e3670f...	RJ	0-0 0 -12:48:7
2	922a46283625e9c096bfd9989...	SP	0-0 0 -18:45:10
3	6aef84c09844a371d82a49152...	SP	0-0 0 -20:31:39
4	ff58662c328f81d3ee549c9caa...	BA	0-0 0 -20:43:20
5	118295a853acb536efff13740f...	SP	0-0 0 -21:22:41

```
select o.customer_id,c.customer_state, avg(o.order_purchase_timestamp-
order_delivered_customer_date) as avg_time_to_delivery
from `Target.orders` o
join `Target.customers` c on c.customer_id=o.customer_id
group by customer_state,customer_id
order by avg_time_to_delivery limit 5
```

Row	customer_id	customer_state	avg_time_to_delivery
1	725e9c75605414b21fd8c8d5a...	RJ	null
2	4ee64f4bfc542546f422da0aeb...	RS	null
3	438449d4af8980d107bf04571...	SP	null
4	964a6df3d9bdf60fe3e7b8bb69...	DF	null
5	7461b0f4f216052ba664f22e0a...	RR	null

## Payment type analysis:

### Count of orders based on the no. of payment installments

```
select payment_installments,max(Total_orders) as Total_orders from
(select payment_installments,
count(*) over(partition by payment_installments order by payment_installments )as Total_orders
from `Target.payments`
)
group by payment_installments
```

Row	payment_installments	Total_orders
1	0	2
2	1	52546
3	2	12413
4	3	10461
5	4	7098
6	5	5239
7	6	3920
8	7	1626
9	8	4268
10	9	644
11	10	5328
12	11	23



## Month over Month count of orders for different payment types-

### Credit Card: -

```
SELECT extract (month from o.order_purchase_timestamp) as Months ,  
COUNT(p.order_id) as monthly_order  
FROM `Target.orders` o join `Target.payments` p on p.order_id=o.order_id  
where p.payment_type = 'credit_card'  
GROUP BY Months  
order by Months
```

Row	Months	monthly_order
1	1	477
2	2	424
3	3	591
4	4	572
5	5	613
6	6	563
7	7	645
8	8	589
9	9	302
10	10	318
11	11	387
12	12	294

### Voucher: -

```
SELECT extract (month from o.order_purchase_timestamp) as Months ,  
COUNT(p.order_id) as monthly_order  
FROM `Target.orders` o join `Target.payments` p on p.order_id=o.order_id  
where p.payment_type = 'voucher'  
GROUP BY Months  
order by Months
```

Row	Months	monthly_order
1	1	477
2	2	424
3	3	591
4	4	572
5	5	613
6	6	563
7	7	645
8	8	589
9	9	302
10	10	318
11	11	387
12	12	294

#### Not\_defined: -

```
SELECT extract (month from o.order_purchase_timestamp) as Months ,
COUNT(p.order_id) as monthly_order
FROM `Target.orders` o join `Target.payments` p on p.order_id=o.order_id
where p.payment_type = 'not defined'
GROUP BY Months
order by Months
```

Row	Months	monthly_order
1	8	2
2	9	1

#### debit\_card:-

```
SELECT extract (month from o.order_purchase_timestamp) as Months ,
COUNT(p.order_id) as monthly_order
FROM `Target.orders` o join `Target.payments` p on p.order_id=o.order_id
where p.payment_type = 'debit_card'
GROUP BY Months
order by Months
```

Row	Months	monthly_order
1	1	118
2	2	82
3	3	109
4	4	124
5	5	81
6	6	209
7	7	264
8	8	311
9	9	43
10	10	54
11	11	70
12	12	64

#### UPI :-

```

SELECT extract (month from o.order_purchase_timestamp) as Months ,
COUNT(p.order_id) as monthly_order
FROM `Target.orders` o join `Target.payments` p on p.order_id=o.order_id
where p.payment_type = 'UPI'
GROUP BY Months
order by Months

```

Row //	Months //	monthly_order //
1	1	1715
2	2	1723
3	3	1942
4	4	1783
5	5	2035
6	6	1807
7	7	2074
8	8	2077
9	9	903
10	10	1056
11	11	1509
12	12	1160