

Brazilian E-commerce store data Analysis for 2016-2018

Create the different metrics like Sales, customer acquisitions, total no. of orders for each Year across the different states they serve

Total sales for each year across the different states.

```
create view xy2 as
select
year(o.order_delivered_customer_date) as year_w,
p.payment_value,
c.customer_state
from [dbo].[olist_orders_dataset] as o
join [dbo].[olist_order_payments_dataset] as p
on o.order_id=p.order_id
join [dbo].[olist_customers_dataset] as c on o.customer_id=c.customer_id
where order_status='delivered'
and year(o.order_delivered_customer_date) in(2016,2017,2018);
select * from xy2
pivot (sum(payment_value)
for customer_state in ([SP],[SC],[MG],[PR],[RJ],[RS],[PA],[GO],[ES],[AM],[AP],[AL],[RO],[PB],[TO],[PI],[AC],[SE],[RR],[MS],[PE],[CE],[MA],[MT],[BA],[DF]))
as pvt_table;
```

year_w	SP	SC	MG	PR	RJ	RS	PA	GO	ES	AM
2016	13559.9300329685	2500.71995925903	4959.81996917725	2599.97000694275	8992.3100707531	2898.10997080803	1283.09000396729	1056.12002182007	991.80997467041	NULL
2017	2349100.1895153	245033.879865378	783320.800095383	319538.919770522	920431.0590076	378630.479903117	94816.6399567127	144871.930056095	129572.029828116	11022.8399
2018	3406421.15046232	347673.800212502	1030996.99001318	459780.660111964	1126267.08021392	480079.809838861	115927.81995225	188366.170367479	187118.809988201	16573.3400

Customer acquisition for each year across different states

```
--create view cust2 as
select order_status, customer_unique_id, customer_state,year(order_approved_at) as year_wise
from olist_customers_dataset as a left join olist_orders_dataset as b on a.customer_id = b.customer_id

select order_status, customer_unique_id, customer_state, year_wise,
row_number() over(order by year_wise)
as customer_acquired from cust2
where order_status = 'delivered' and year_wise = 2016 or year_wise = 2017 or year_wise = 2018;
```

order_status	customer_unique_id	customer_state	year_wise	customer_acquired
delivered	171d92af5bf025428dfc2fd869eb8a04	RJ	2016	1
delivered	ee1b69039bd20877692d8622dedbc3ec	SP	2016	2
delivered	99e41e69b60bd63edd83935499523063	RJ	2016	3
delivered	ad1cb06f09b27e4773870d1f76a68557	SP	2016	4
delivered	12cc890391e57d7123e004856ff1f137	PE	2016	5
delivered	d8e40a4b16990a6fd1095feae46aa1af	SP	2016	6
delivered	2e92a624a8a3f8ab5eb51d42fca60e5b	RJ	2016	7
delivered	0032c76b20340da25249092a268ce66c	SP	2016	8

Total count of orders for each year in different states.

```

select a.customer_state,year(b.order_purchase_timestamp) as years,
count(b.order_id) as total_number_of_order from
olist_orders_dataset as b
inner join olist_customers_dataset as a on b.customer_id=a.customer_id group by
a.customer_state,year(b.order_purchase_timestamp) order by count(b.order_id);

```

customer_state	years	total_number_of_order
PB	2016	1
PI	2016	1
RR	2016	2
AL	2016	2
MT	2016	3
SE	2016	3
PA	2016	4
ES	2016	4

Decreasing trends of States in different years

```

select *from
(select years,customer_state,total_sales,
dense_rank() over(partition by years order by total_sales desc) as rank from
(select year(b.order_approved_at) as years,a.customer_state,
floor(sum(c.payment_value))
as total_sales
from olist_customers_dataset
as A left join olist_orders_dataset as B
on A.customer_id=B.customer_id
left join [dbo].[olist_order_payments_dataset] as c on b.order_id=c.order_id
where year(b.order_approved_at) in ('2016','2017','2018')
group by A.customer_state,
year(b.order_approved_at))a)b
where rank in (1,2);
----bbbb trend
select *from

```

years	customer_state	total_sales	rank
2016	SP	16651	1
2016	RJ	12742	2
2017	SP	2552985	1
2017	RJ	1052810	2
2018	SP	3414707	1
2018	RJ	1074987	2

Increasing trends of States in different years

```
select *from
(select years,customer_state,total_sales,
dense_rank() over(partition by years order by total_sales ) as rank from
(select year(b.order_approved_at) as years,a.customer_state,
floor(sum(c.payment_value))
as total_sales
from olist_customers_dataset
as A left join olist_orders_dataset as B
on A.customer_id=B.customer_id
left join [dbo].[olist_order_payments_dataset] as c on b.order_id=c.order_id
where year(b.order_approved_at) in ('2016','2017','2018'))
group by A.customer_state,
year(b.order_approved_at))a)b
where rank in (1,2);
```

customer acquisition

Results			
Messages			
years	customer_state	total_sales	rank
2016	PB	74	1
2016	AL	129	2
2017	RR	2006	1
2017	AP	7331	2
2018	AC	6683	1
2018	RR	7852	2

Percentage of orders delivered earlier than the expected date

```

select order_id, customer_id, Datediff(day,order_delivered_customer_date , order_estimated_delivery_date) as pos,
Datediff(day,order_purchase_timestamp , order_estimated_delivery_date) as total_dilevery_time,
cast(Datediff(day,order_delivered_customer_date , order_estimated_delivery_date)*100 as int)/
cast(Datediff(day,order_purchase_timestamp , order_estimated_delivery_date) as int) as '% delivered_earlier'
from olist_orders_dataset
where order_status = 'delivered'
and
Datediff(day,order_delivered_customer_date, order_estimated_delivery_date) > 0;

```

Results Messages

order_id	customer_id	pos	total_dilevery_time	% delivered_earlier
e481f51cbdc54678b7cc49136f2d6af7	9ef432eb6251297304e76186b10a928d	8	16	50
53cdb2fc8bc7dce0b6741e2150273451	b0830fb4747a6c6d20dea0b8c802d7ef	6	20	30
47770eb9100c2d0c44946d9cf07ec65d	41ce2a54c0b03bf3443c3d931a367089	18	27	66
949d5b44dbf5de918fe9c16f97b45f8a	f88197465ea7920adcdbec7375364d82	13	27	48
ad21c59c0840e6cb83a9ceb5573f8159	8ab97904e6daea8866dbdbc4fb7aad2c	10	13	76
a4591c265e18cb1dcee52889e2d8acc3	503740e9ca751ccdda7ba28e9ab8f608	6	23	26
6514b8ad8028c9f2cc2374ded245783f	9bdf08b4b3b552b5526ff42d37d47222	12	22	54
76c6e866289321a7c93b82b54852dc33	f54a9f0e6b351c431402b8461ea51999	32	42	76

percentage of orders delivered later than the expected date

```

select order_id, customer_id, Datediff(day,order_delivered_customer_date , order_estimated_delivery_date) as delivertiming,
atediff(day,order_purchase_timestamp , order_estimated_delivery_date) as total_dilevery_time,
ast(Datediff(day,order_delivered_customer_date , order_estimated_delivery_date)*100 as int)/
ast(Datediff(day,order_purchase_timestamp , order_estimated_delivery_date) as int) as '% delivered_earlier'
from olist_orders_dataset
where order_status = 'delivered'
and
Datediff(day,order_delivered_customer_date, order_estimated_delivery_date) < 0;

```

Results Messages

order_id	customer_id	delivertiming	total_dilevery_time	% delivered_earlier
203096f03d82e0dffbc41ebc2e2bcfb7	d2b091571da224a1b36412c18bc3bbfe	-11	10	-110
fbf99ac61453ac646ce8ad9783d7d0af6	3a874b4d4c4b6543206ff5d892870c3	-9	20	-45
6ea2f835b4556291ffdc53fa0b3b95e8	c7340080e394356141681bd4c9b8fe31	-7	27	-25
66e4624ae69e7dc89bd50222b59f581f	684fa6da5134b9e4dab731e00011712d	-1	24	-4
a685d016c8a26f71a0bb67821070e398	911e4c37f5cafe1604fe6767034bf1ae	-7	17	-41
6a0a8bfbbe700284feb0845d95e0867f	68451b39b1314302c08c65a29f1140fc	-17	19	-89
a5474c0071dd5d1074e12d417078bbd0	ef15b3240b2083e0487762ee2978d2b8	-1	3	-33
9d531c565e28c3e0d756192f84d8731f	d4faa220408c20e53595d2950f361f3b	-32	24	-133

Category wise sales

```

SELECT p.column2 AS Category,
       COUNT(DISTINCT p.order_id) AS Num_order,
       Floor(SUM(p.payment_value)) AS Revenue
FROM (
    SELECT a.product_category_name, a.product_id, a.product_description_lenght, a.product_height_cm, a.product_length_cm, a.product_name_lenght,
           b.product_photos_qty, a.product_weight_g, a.product_width_cm, b.column1, b.column2, c.order_id, e.payment_value,
           d.order_status, d.order_delivered_customer_date
    FROM olist_products_dataset AS a
    LEFT JOIN [dbo].[product_category_name_translation] AS b
    ON a.product_category_name = b.column1
    LEFT JOIN [dbo].[olist_order_items_dataset] AS c
    ON a.product_id = c.product_id
    LEFT JOIN [dbo].[olist_orders_dataset] AS d
    ON c.order_id = d.order_id
    LEFT JOIN [dbo].[olist_order_payments_dataset] AS e
    ON c.order_id = e.order_id
) as P
WHERE p.column2 IS NOT NULL
AND p.order_status <> 'canceled'
AND p.order_delivered_customer_date IS NOT NULL
GROUP BY p.column2
ORDER BY Revenue DESC;

```

Category	Num_order	Revenue
bed_bath_table	9272	1692714.28101649
health_beauty	8647	1620684.03971168
computers_accessories	6529	1549252.46649788
furniture_decor	6307	1394466.02152200

Seller wise performance of sales

```

SELECT v.seller_id,
       v.column2 AS Category,
       SUM(v.payment_value) AS Revenue
FROM (
    SELECT a.seller_id, a.seller_city, a.seller_state, b.order_id, b.price, b.order_item_id, b.shipping_limit_date, b.product_id,
           c.product_category_name, c.product_description_lenght, c.product_height_cm,
           d.column1, d.column2, e.payment_value, e.payment_type,
           f.order_status,
           f.order_estimated_delivery_date, f.order_delivered_customer_date
    FROM [dbo].[olist_sellers_dataset] AS a
    LEFT JOIN [dbo].[olist_order_items_dataset] AS b
    ON a.seller_id = b.seller_id
    LEFT JOIN [dbo].[olist_products_dataset] AS c
    ON b.product_id = c.product_id
    LEFT JOIN [dbo].[product_category_name_translation] AS d
    ON c.product_category_name = d.column2
    LEFT JOIN [dbo].[olist_order_payments_dataset] AS e
    ON b.order_id = e.order_id
    LEFT JOIN [dbo].[olist_orders_dataset] AS f
    ON b.order_id = f.order_id
) AS v
WHERE v.column2 IS NOT NULL
AND v.order_status <> 'canceled'
AND v.order_delivered_customer_date IS NOT NULL
GROUP BY v.column2, v.seller_id
ORDER BY Category, Revenue DESC;

```

seller_id	Category	Revenue
7d13fca15225358621be4086e1eb0964	audio	20610.0900192261
4869f7a5dfa277a7dca6462dcf3b52b2	audio	10423.1700353622
b33e7c55446eabf8fe1a42d037ac7d6d	audio	5894.1999206543
b2ba3715d723d245138f291a6fe42594	audio	4550.66999006271
6560211a19b47992c3666cc44a7e94c0	audio	3617.70002746582

Product wise sales and order placed

```

SELECT p.product_id AS prod,
COUNT(DISTINCT p.order_id) AS Num_order,
floor(SUM(p.payment_value)) AS Revenue
FROM (
SELECT a.product_category_name,a.product_id,a.product_description_lenght,a.product_height_cm,a.product_length_cm,a.product_name_lenght,
a .product_photos_qty,a.product_weight_g,a.product_width_cm,b.column1,b.column2,c.order_id,e.payment_value,
d.order_status,d.order_delivered_customer_date
FROM olist_products_dataset AS a
LEFT
JOIN [dbo].[product_category_name_translation] AS b
ON a.product_category_name = b.column1
LEFT
JOIN [dbo].[olist_order_items_dataset] AS c
ON a.product_id = c.product_id
LEFT
JOIN [dbo].[olist_orders_dataset] AS d
ON c.order_id = d.order_id
LEFT
JOIN [dbo].[olist_order_payments_dataset] AS e
ON c.order_id = e.order_id
)as p
WHERE p.product_id IS NOT NULL
AND p.order_status <> 'canceled'
AND p.order_delivered_customer_date IS NOT NULL
GROUP BY p.product_id
ORDER BY Revenue DESC;

```

Results			Messages	
prod	Num_order	Revenue		
5769ef0a239114ac3a854af00df129e4	1	109312		
bb50f2e236e5eea0100680137654686c	186	81539		
422879e10f46682990de24d770e783d	352	79512		
d1c427060a0f73f6b889a5c7c61f2ac4	313	68038		
d5991653e037ccb7af6ed7d94246b249	57	64143		