TARGET : Shopping Mart

Target is a globally renowned brand and a prominent retailer in the United States. Target makes itself a preferred shopping destination by offering outstanding value, inspiration, innovation and an exceptional guest experience that no other retailer can deliver. This particular business case focuses on the operations of Target in Brazil and provides insightful information about 100,000 orders placed between 2016 and 2018.The dataset offers a comprehensive view of various dimensions including the order status, price, payment and freight performance, customer location, product attributes, and customer reviews.

Analyse the given datasets and generate insights.

Tables present in database:

- **CUSTOMERS**

	customer_id	customer_unique_id	customer_zip_code_prefix	customer_city	customer_state
1	00012a2ce6f8dcda20d059ce98491703	248ffe10d632bebe4f7267f1f44844c9	6273	osasco	SP
2	000161a058600d5901f007fab4c27140	b0015e09bb4b6e47c52844fab5fb6638	35550	itapecerica	MG
3	0001fd6190edaaf884bcaf3d49edf079	94b11d37cd61cb2994a194d11f89682b	29830	nova venecia	ES
4	0002414f95344307404f0ace7a26f1d5	4893ad4ea28b2c5b3ddf4e82e79db9e6	39664	mendonca	MG
5	000379cdec625522490c315e70c7a9fb	0b83f73b19c2019e182fd552c048a22c	4841	sao paulo	SP
6	0004164d20a9e969af783496f3408652	104bdb7e6a6cdceaa88c3ea5fa6b2b93	13272	valinhos	SP
7	000419c5494106c306a97b5635748086	14843983d4a159080f6afe4b7f346e7c	24220	niteroi	RJ
8	00046a560d407e99b969756e0b10f282	0b5295fc9819d831f68eb0e9a3e13ab7	20540	rio de janeiro	RJ
9	00050bf6e01e69d5c0fd612f1bcfb69c	e3cf594a99e810f58af53ed4820f25e5	98700	ijui	RS
10	000598caf2ef4117407665ac33275130	7e0516b486e92ed3f3afdd6d1276cfbd	35540	oliveira	MG

- **Geolocation**

	geolocation_zip_code_prefix	geolocation_lat	geolocation_lng	geolocation_city	geolocation_state
1	1037	-23.5456212811527	-46.6392920480017	sao paulo	SP
2	1046	-23.5460811270355	-46.6448202983716	sao paulo	SP
3	1046	-23.5461289664147	-46.6429514836114	sao paulo	SP
4	1041	-23.5443921648681	-46.6394993062784	sao paulo	SP
5	1035	-23.5415779617115	-46.6416072232961	sao paulo	SP
6	1012	-23.5477623033643	-46.6353605378845	são paulo	SP
7	1047	-23.5462731124127	-46.6412251697155	sao paulo	SP
8	1013	-23.5469232084367	-46.6342636964915	sao paulo	SP
9	1029	-23.5437690557691	-46.6342778408513	sao paulo	SP
10	1011	-23.5476395503206	-46.6360316231549	sao paulo	SP

- **ORDER_ITEMS**

	order_id	order_item_id	product_id	seller_id	shipping_limit_date	price	freight_value
1	00010242fe8c5a6d1ba2dd792cb16214	1	4244733e06e7ecb4970a6e2683c13e61	48436dade18ac8b2bce089ec2a041202	2017-09-19 09:45:35.000	58.9	13.29
2	00018f77f2f0320c557190d7a144bdd3	1	e5f2d52b802189ee658865ca93d83a8f	dd7ddc04e1b6c2c614352b383efe2d36	2017-05-03 11:05:13.000	239.9	19.93
3	000229ec398224ef6ca0657da4fc703e	1	c777355d18b72b67abbeef9df44fd0fd	5b51032eddd242adc84c38acab88f23d	2018-01-18 14:48:30.000	199	17.87
4	00024acbcdf0a6daa1e931b038114c75	1	7634da152a4610f1595efa32f14722fc	9d7a1d34a5052409006425275ba1c2b4	2018-08-15 10:10:18.000	12.99	12.79
5	00042b26cf59d7ce69dfabb4e55b4fd9	1	ac6c3623068f30de03045865e4e10089	df560393f3a51e74553ab94004ba5c87	2017-02-13 13:57:51.000	199.9	18.14
6	00048cc3ae777c65dbb7d2a0634bc1ea	1	ef92defde845ab8450f9d70c526ef70f	6426d21aca402a131fc0a5d0960a3c90	2017-05-23 03:55:27.000	21.9	12.69
7	00054e8431b9d7675808bcb819fb4a32	1	8d4f2bb7e93e6710a28f34fa83ee7d28	7040e82f899a04d1b434b795a43b4617	2017-12-14 12:10:31.000	19.9	11.85
8	000576fe39319847cbb9d288c5617fa6	1	557d850972a7d6f792fd18ae1400d9b6	5996cddab893a4652a15592fb58ab8db	2018-07-10 12:30:45.000	810	70.75
9	0005a1a1728c9d785b8e2b08b904576c	1	310ae3c140ff94b03219ad0adc3c778f	a416b6a846a11724393025641d4edd5e	2018-03-26 18:31:29.000	145.95	11.65
10	0005f50442cb953dcd1d21e1fb923495	1	4535b0e1091c278dfd193e5a1d63b39f	ba143b05f0110f0dc71ad71b4466ce92	2018-07-06 14:10:56.000	53.99	11.4

- **ORDER_REVIEWS**

	review_id	order_id	review_score	review_comment_title	review_creation_date	review_answer_timestamp
1	7bc2406110b926393aa56f80a40eba40	73fc7af87114b39712e6da79b0a377eb	4	NULL	2018-01-18 00:00:00.000	2018-01-18 21:46:00.000
2	80e641a11e56f04c1ad469d5645fdfde	a548910a1c6147796b98fdf73dbeba33	5	NULL	2018-03-10 00:00:00.000	2018-03-11 03:05:00.000
3	228ce5500dc1d8e020d8d1322874b6f0	f9e4b658b201a9f2ecdecbb34bed034b	5	NULL	2018-02-17 00:00:00.000	2018-02-18 14:36:00.000
4	e64fb393e7b32834bb789ff8bb30750e	658677c97b385a9be170737859d3511b	5	NULL	2017-04-21 00:00:00.000	2017-04-21 22:02:00.000
5	f7c4243c7fe1938f181bec41a392bdeb	8e6bfb81e283fa7e4f11123a3fb894f1	5	NULL	2018-03-01 00:00:00.000	2018-03-02 10:26:00.000
6	15197aa66ff4d0650b5434f1b46cda19	b18dcdf73be66366873cd26c5724d1dc	1	NULL	2018-04-13 00:00:00.000	2018-04-16 00:39:00.000
7	07f9bee5d1b850860defd761afa7ff16	e48aa0d2dcec3a2e87348811bcfdf22b	5	NULL	2017-07-16 00:00:00.000	2017-07-18 19:30:00.000
8	7c6400515c67679fbee952a7525281ef	c31a859e34e3adac22f376954e19b39d	5	NULL	2018-08-14 00:00:00.000	2018-08-14 21:36:00.000
9	a3f6f7f6f433de0aefbb97da197c554c	9c214ac970e84273583ab523dfafd09b	5	NULL	2017-05-17 00:00:00.000	2017-05-18 12:05:00.000
10	8670d52e15e00043ae7de4c01cc2fe06	b9bf720beb4ab3728760088589c62129	4	I recommend	2018-05-22 00:00:00.000	2018-05-23 16:45:00.000

- **ORDERS**

	order_id	customer_id	order_status	order_purchase_timestamp	order_approved_at	order_delivered_carrier_date	order_delivered_customer_date	order_estimated_delivery_date
1	e481f51cbdc54678b7cc49136f2d6af7	9ef432eb6251297304e76186b10a928d	delivered	2017-10-02 10:56:33.000	2017-10-02 11:07:15.000	2017-10-04 19:55:00.000	2017-10-10 21:25:13.000	2017-10-18 00:00:00.000
2	53cdb2fc8bc7dce0b6741e2150273451	b0830fb4747a6c6d20dea0b8c802d7ef	delivered	2018-07-24 20:41:37.000	2018-07-26 03:24:27.000	2018-07-26 14:31:00.000	2018-08-07 15:27:45.000	2018-08-13 00:00:00.000
3	47770eb9100c2d0c44946d9cf07ec65d	41ce2a54c0b03bf3443c3d931a367089	delivered	2018-08-08 08:38:49.000	2018-08-08 08:55:23.000	2018-08-08 13:50:00.000	2018-08-17 18:06:29.000	2018-09-04 00:00:00.000
4	949d5b44dbf5de918fe9c16f97b45f8a	f8819765ea7920adcdbec7375364d82	delivered	2017-11-18 19:28:06.000	2017-11-18 19:45:59.000	2017-11-22 13:39:59.000	2017-12-02 00:28:42.000	2017-12-15 00:00:00.000
5	ad21c59c0840e6cb83a9ceb5573f8159	8ab97904e6daea8866dbdbc4fb7aad2c	delivered	2018-02-13 21:18:39.000	2018-02-13 22:20:29.000	2018-02-14 19:46:34.000	2018-02-16 18:17:02.000	2018-02-26 00:00:00.000
6	a4591c265e18cb1dcee52889e2d8acc3	503740e9ca751ccdda7ba28e9ab8f608	delivered	2017-07-09 21:57:05.000	2017-07-09 22:10:13.000	2017-07-11 14:58:04.000	2017-07-26 10:57:55.000	2017-08-01 00:00:00.000

- **PAYMENTS**

	order_id	payment_sequential	payment_type	payment_installments	payment_value
1	b81ef226f3fe1789b1e8b2acac839d17	1	credit_card	8	99.33
2	a9810da82917af2d9aefd1278f1dcfa0	1	credit_card	1	24.39
3	25e8ea4e93396b6fa0d3dd708e76c1bd	1	credit_card	1	65.71
4	ba78997921bbcdc1373bb41e913ab953	1	credit_card	8	107.78
5	42fdf880ba16b47b59251dd489d4441a	1	credit_card	2	128.45

- **PRODUCTS**

	product_id	product_category	product_name_length	product_description_length	product_photos_qty	product_weight_g	product_length_cm	product_height_cm	product_width_cm
1	1e9e8ef04dbcff4541ed26657ea517e5	perfumery	40	287	1	225	16	10	14
2	3aa071139cb16b67ca9e5dea641aaa2f	Art	44	276	1	1000	30	18	20
3	96bd76ec8810374ed1b65e291975717f	sport leisure	46	250	1	154	18	9	15
4	cef67bcfe19066a932b7673e239eb23d	babies	27	261	1	371	26	4	26
5	9dc1a7de274444849c219cff195d0b71	housewares	37	402	4	625	20	17	13
6	41d3672d4792049fa1779bb35283ed13	musical instruments	60	745	1	200	38	5	11

QUESTIONS

Q1. Data type of all columns in the "customers" table. Get the time range between which the orders were placed.

```
SELECT column_name,data_type
FROM shopping_mart.INFORMATION_SCHEMA.COLUMNS
WHERE table_name = 'customers'
```

1.

	column_name	data_type
1	customer_id	nvarchar
2	customer_unique_id	nvarchar
3	customer_zip_code_prefix	int
4	customer_city	varchar
5	customer_state	nvarchar

2.

	Start_Date	End_Date
1	2016-09-04	2018-10-17

Q2. Count the Cities & States of customers who ordered during the given period.

```
SELECT COUNT(DISTINCT c.customer_city) AS City, COUNT(DISTINCT
c.customer_state) AS State
FROM customers c
INNER JOIN orders o
ON c.customer_id = o.customer_id
```

	CITY	STATE
1	4119	27

Q3. Is there a growing trend in the no. of orders placed over the past years?

```
SELECT DATEPART(YEAR, order_purchase_timestamp) AS Years, COUNT(order_id)
AS Order_Placed
FROM orders
GROUP BY 1
```

	Years	Order_Placed
1	2016	329
2	2017	45101
3	2018	54011

Q4. Can we see some kind of monthly seasonality in terms of the no. of orders being placed?

```
SELECT DATEPART(MONTH, order_purchase_timestamp) AS Months,
COUNT(order_id) AS Order_Placed
FROM orders
```

```
GROUP BY 1
ORDER BY 1
```

	Months	Order_Placed
1	1	8069
2	2	8508
3	3	9893
4	4	9343
5	5	10573
6	6	9412
7	7	10318
8	8	10843
9	9	4305
10	10	4959
11	11	7544
12	12	5674

- There is indeed some type of monthly seasonality in the number of orders being placed.
- There seems to be increase in the number of orders during certain months, followed by a decrease in other. Understanding these patterns can help business of mart.
- Months 5 (May) and 8 (August) have the highest number of orders (10573 and 10843, respectively)

Q5. During what time of the day, do the Brazilian customers mostly place their orders? (Dawn, Morning, Afternoon or Night)

```
SELECT CASE
WHEN DATEPART(HOUR, order_purchase_timestamp) BETWEEN 0 AND 6 THEN 'Dawn'
WHEN DATEPART(HOUR, order_purchase_timestamp) BETWEEN 7 AND 12 THEN 'Morning'
WHEN DATEPART(HOUR, order_purchase_timestamp) BETWEEN 13 AND 18 THEN 'Evening'
WHEN DATEPART(HOUR, order_purchase_timestamp) BETWEEN 19 AND 24 THEN 'Night'
ELSE 'Unknown' END AS 'Hours', COUNT(order_id) AS Orders
FROM orders
GROUP BY 1
ORDER BY 2 DESC
```

	Hours	Orders
1	Afternoon	38135
2	Night	28331
3	Morning	27733
4	Dawn	5242

Q6. Get the month-on-month no. of orders placed in each state.

```
SELECT c.customer_state AS State, DATEPART(MONTH,
o.order_purchase_timestamp) as Months, COUNT(o.order_id) AS Orders
FROM orders o
INNER JOIN customers c
ON o.customer_id = c.customer_id
```

```
GROUP BY 1,2
ORDER BY 1,2
```

	State	Months	Orders
1	AC	1	8
2	AC	2	6
3	AC	3	4
4	AC	4	9
5	AC	5	10
6	AC	6	7
7	AC	7	9
8	AC	8	7
9	AC	9	5
10	AC	10	6
11	AC	11	5
12	AC	12	5
13	AL	1	39
14	AL	2	39

Q7. How are the customers distributed across all the states?

```
SELECT customer_state, COUNT(DISTINCT customer_unique_id) AS Customers
FROM customers
GROUP BY customer_state
ORDER BY Customers DESC
```

	customer_state	Customers
1	SP	40302
2	RJ	12384
3	MG	11259
4	RS	5277
5	PR	4882
6	SC	3534
7	BA	3277
8	DF	2075
9	ES	1964
10	GO	1952
11	PE	1609
12	CE	1313

Q8. Get the % increase in the cost of orders from year 2017 to 2018 (include months between Jan to Aug only).

```
SELECT *, ROUND(((Cost_of_Orders - Previous_Cost)/Previous_Cost)*100,2) AS
Perc_Inc
FROM
(SELECT *, LAG(Cost_of_Orders) OVER (ORDER BY Years) AS Previous_Cost
FROM
```

```
(SELECT DATEPART(YEAR, CAST(o.order_purchase_timestamp AS date)) AS Years,
ROUND(SUM(p.payment_value),2) AS Cost_of_Orders
FROM orders o
INNER JOIN payments p ON o.order_id = p.order_id
WHERE DATEPART(YEAR, CAST(ORDER_PURCHASE_TIMESTAMP AS date)) BETWEEN 2017
AND 2018
AND DATEPART(MONTH, CAST(ORDER_PURCHASE_TIMESTAMP AS date)) BETWEEN 1 AND
8
GROUP BY 1)))A)B
```

	Years	Cost_of_Orders	Previous_Cost	Perc_Inc
1	2017	3669022.12	NULL	NULL
2	2018	8694733.84	3669022.12	136.98

Q9. Calculate the Total & Average value of order price for each state.

```
select c.customer_state, round(sum(p.payment_value),2) as Total_Price,
round(AVG(p.payment_value),2) as Avg_Price
from customers c
inner join orders o
on c.customer_id = o.customer_id
inner join payments p
on o.order_id = p.order_id
group by 1
order by 2 desc
```

	customer_state	Total_Price	Avg_Price
1	SP	5998226.96	137.5
2	RJ	2144379.69	158.53
3	MG	1872257.26	154.71
4	RS	890898.54	157.18
5	PR	811156.38	154.15
6	SC	623086.43	165.98
7	BA	616645.82	170.82
8	DF	355141.08	161.13
9	GO	350092.31	165.76
10	ES	325967.55	154.71
11	PE	324850.44	187.99
12	CE	279464.03	199.9
13	PA	218295.85	215.92

Q10. Calculate the Total & Average value of order freight for each state.

```sql
select c.customer_state, round(SUM(oi.freight_value),2) as Total_Freight,
round(AVG(oi.freight_value),2) as Avg_Freight
from customers c
inner join orders o
on c.customer_id = o.customer_id
inner join order_items oi
on o.order_id = oi.order_id
group by 1
order by 2 desc
```

	customer_state	Total_Freight	Avg_Freight
1	SP	718723.07	15.15
2	RJ	305589.31	20.96
3	MG	270853.46	20.63
4	RS	135522.74	21.74
5	PR	117851.68	20.53
6	BA	100156.68	26.36
7	SC	89660.26	21.47
8	PE	59449.66	32.92
9	GO	53114.98	22.77
10	DF	50625.5	21.04
11	ES	49764.6	22.06
12	CE	48351.59	32.71
13	PA	38699.3	35.83

Q11. Find the no. of days taken to deliver each order from the order's purchase date as delivery time. Also, calculate the difference (in days) between the estimated & actual delivery date of an order.

```sql
select order_id, DATEDIFF(day, order_purchase_timestamp,
order_delivered_customer_date) as delivered_time_taken,
DATEDIFF(day, order_estimated_delivery_date,
order_delivered_customer_date) as diff_delivery_time
from orders
order by delivered_time_taken desc
```

	order_id	delivered_time_taken	diff_delivery_time
1	ca07593549f1816d26a572e06dc1eab6	210	181
2	1b3190b2dfa9d789e1f14c05b647a14a	208	188
3	440d0d17af552815d15a9e41abe49359	196	165
4	2fb597c2f772eca01b1f5c561bf6cc7b	195	155
5	285ab9426d6982034523a855f55a885e	195	166
6	0f4519c5f1c541ddec9f21b3bddd533a	194	161
7	47b40429ed8cce3aee9199792275433f	191	175
8	2fe324febf907e3ea3f2aa9650869fa5	190	167
9	2d7561026d542c8dbd8f0daeadf67a43	188	159
10	c27815f7e3dd0b926b58552628481575	188	162

Q12. Find out the top 5 states with the highest & lowest average freight value.

```
select customer_state, freight_val from
(select c.customer_state, round(AVG(oi.freight_value),2) as freight_val,
ROW_NUMBER() over (order by AVG(oi.freight_value) desc) as h_rnk,
ROW_NUMBER() over (order by AVG(oi.freight_value) asc) as l_rnk
from customers c
inner join orders o
on c.customer_id = o.customer_id
inner join order_items oi
on o.order_id = oi.order_id
group by c.customer_state)a
where h_rnk <= 5 or l_rnk <= 5
```

	customer_state	freight_val
1	SP	15.15
2	PR	20.53
3	MG	20.63
4	RJ	20.96
5	DF	21.04
6	PI	39.15
7	AC	40.07
8	RO	41.07
9	PB	42.72
10	RR	42.98

Q13. Find out the top 5 states with the highest & lowest average delivery time.

```
select customer_state, avg_delivery
from
(select c.customer_state, avg(DATEDIFF(DAY, o.order_purchase_timestamp,
o.order_delivered_customer_date)) as avg_delivery,
ROW_NUMBER() over (order by avg(DATEDIFF(DAY, o.order_purchase_timestamp,
o.order_delivered_customer_date)) desc) as h_rnk,
ROW_NUMBER() over (order by avg(DATEDIFF(DAY, o.order_purchase_timestamp,
o.order_delivered_customer_date)) asc) as l_rnk
from customers c
inner join orders o
on c.customer_id = o.customer_id
group by c.customer_state)a
where h_rnk <=5 or l_rnk <= 5
```

	customer_state	avg_delivery
1	SP	8
2	PR	11
3	MG	11
4	DF	12
5	SC	14
6	PA	23
7	AL	24
8	AM	26
9	AP	27
10	RR	29

Q14. Find out the top 5 states where the order delivery is really fast as compared to the estimated date of delivery.

```
select c.customer_state,
avg(DATEDIFF(day,
o.order_estimated_delivery_date,o.order_delivered_customer_date)) as
fastest_delivery
from customers c
inner join orders o
on c.customer_id = o.customer_id
group by c.customer_state
order by fastest_delivery
```

	customer_state	fastest_delivery
1	AC	-20
2	RO	-20
3	AP	-19
4	AM	-19
5	RR	-17

Q15. Find the month-on-month no. of orders placed using different payment types.

```
select p.payment_type, datepart(month, CAST(o.order_purchase_timestamp as
date)) as Month, COUNT(o.order_id) as Orders
from orders o
inner join payments p on
o.order_id = p.order_id
group by 1, 2
order by 1, 2
```

	payment_type	Month	Orders
1	credit_card	1	6103
2	credit_card	2	6609
3	credit_card	3	7707
4	credit_card	4	7301
5	credit_card	5	8350
6	credit_card	6	7276
7	credit_card	7	7841
8	credit_card	8	8269
9	credit_card	9	3286
10	credit_card	10	3778
11	credit_card	11	5897
12	credit_card	12	4378
13	debit_card	1	118
14	debit_card	2	82
15	debit_card	3	109
16	debit_card	4	124

Q16. Find the no. of orders placed on the basis of the payment installments that have been paid.

```
select p.payment_installments, COUNT(o.order_id) as Orders
from orders o
inner join payments p on
o.order_id = p.order_id
where p.payment_installments >=1
group by p.payment_installments
order by p.payment_installments
```

	payment_installments	Orders
1	1	52546
2	2	12413
3	3	10461
4	4	7098
5	5	5239
6	6	3920
7	7	1626
8	8	4268
9	9	644
10	10	5328
11	11	23
12	12	133
13	13	16
14	14	15
15	15	74

INSIGHTS

- It is clear that orders placing was started from “September 4 ,2016” and ended on “August 17,2018”.
- There are 4119 cities and 27 states of different customers who have ordered the products.
- There is indeed some type of monthly seasonality in the number of orders being placed.
- Months 5 (May) and 8 (August) have the highest number of orders (10573 and 10843,respectively).
- During ‘Afternoon’, the Brazilian customers mostly place their orders.
- States SP, RJ, MG have the greatest number of customers.
- States RR, AP, AC have the lowest number of customers, a negative from the company point of view.
- The cost of orders in 2018 increased by approximately 137.0% compared to 2017.
- AC, AP, RR have the lowest total freight values which implies that shipping to and from these states is less expensive.
- SP, RJ, MG have the highest total freight values, which means that shipping to and from these states is more expensive.
- The data shows trends in the number of orders for different payment types, including "credit_card," "UPI," "voucher," and "debit_card.