Superstore Data Analysis Using SQL

Unlocking Business Insights through Data

-Tool Used: SQL Server

Objective:
To derive actionable business insights from Superstore sales data using SQL queries.

Dataset:
Sample Superstore Dataset with fields: order_id, order_date, sales, profit, region, segment, product_name, product_id, etc.

Deliverables:

- Top Profitable Products
- Region-wise Sales & Profit
- Customer Segment Performance
- Daily Sales Trend
- Loss-Making Orders
- Outlier Detection
- Repeat Customer Rate

	Row_ID	Order_ID	Order_Date	Ship_Date	Ship_Mode	Customer_ID	Customer_Name	Segment	Country	City	State	Postal_Code	Region	Prodı
1	1	CA-2016-152156	11/8/2016	11/11/2016	Second Class	CG-12520	Claire Gute	Consumer	United States	Henderson	Kentucky	42420	South	FUR
2	2	CA-2016-152156	11/8/2016	11/11/2016	Second Class	CG-12520	Claire Gute	Consumer	United States	Henderson	Kentucky	42420	South	FUR
3	3	CA-2016-138688	6/12/2016	6/16/2016	Second Class	DV-13045	Darrin Van Huff	Corporate	United States	Los Angeles	California	90036	West	OFF
4	4	US-2015-108966	10/11/2015	10/18/2015	Standard Class	SO-20335	Sean O'Donnell	Consumer	United States	Fort Lauderdale	Florida	33311	South	FUR
5	5	US-2015-108966	10/11/2015	10/18/2015	Standard Class	SO-20335	Sean O'Donnell	Consumer	United States	Fort Lauderdale	Florida	33311	South	OFF
6	6	CA-2014-115812	6/9/2014	6/14/2014	Standard Class	BH-11710	Brosina Hoffman	Consumer	United States	Los Angeles	California	90032	West	FUR
7	7	CA-2014-115812	6/9/2014	6/14/2014	Standard Class	BH-11710	Brosina Hoffman	Consumer	United States	Los Angeles	California	90032	West	OFF
8	8	CA-2014-115812	6/9/2014	6/14/2014	Standard Class	BH-11710	Brosina Hoffman	Consumer	United States	Los Angeles	California	90032	West	TEC
9	9	CA-2014-115812	6/9/2014	6/14/2014	Standard Class	BH-11710	Brosina Hoffman	Consumer	United States	Los Angeles	California	90032	West	OFF
10	10	CA-2014-115812	6/9/2014	6/14/2014	Standard Class	BH-11710	Brosina Hoffman	Consumer	United States	Los Angeles	California	90032	West	OFF
11	11	CA-2014-115812	6/9/2014	6/14/2014	Standard Class	BH-11710	Brosina Hoffman	Consumer	United States	Los Angeles	California	90032	West	FUR
12	12	CA-2014-115812	6/9/2014	6/14/2014	Standard Class	BH-11710	Brosina Hoffman	Consumer	United States	Los Angeles	California	90032	West	TEC
13	13	CA-2017-114412	4/15/2017	4/20/2017	Standard Class	AA-10480	Andrew Allen	Consumer	United States	Concord	North C...	28027	South	OFF
14	14	CA-2016-161389	12/5/2016	12/10/2016	Standard Class	IM-15070	Irene Maddox	Consumer	United States	Seattle	Washin...	98103	West	OFF
15	15	US-2015-118983	11/22/2015	11/26/2015	Standard Class	HP-14815	Harold Pawlan	Home O...	United States	Fort Worth	Texas	76106	Central	OFF
16	16	US-2015-118983	11/22/2015	11/26/2015	Standard Class	HP-14815	Harold Pawlan	Home O...	United States	Fort Worth	Texas	76106	Central	OFF
17	17	CA-2014-105893	11/11/2014	11/18/2014	Standard Class	PK-19075	Pete Kriz	Consumer	United States	Madison	Wiscon...	53711	Central	OFF
18	18	CA-2014-167164	5/13/2014	5/15/2014	Second Class	AG-10270	Alejandro Grove	Consumer	United States	West Jordan	Utah	84084	West	OFF
19	19	CA-2014-143336	8/27/2014	9/1/2014	Second Class	ZD-21925	Zuschuss Don...	Consumer	United States	San Francisco	California	94109	West	OFF
20	20	CA-2014-143336	8/27/2014	9/1/2014	Second Class	ZD-21925	Zuschuss Don...	Consumer	United States	San Francisco	California	94109	West	TEC
21	21	CA-2014-143336	8/27/2014	9/1/2014	Second Class	ZD-21925	Zuschuss Don...	Consumer	United States	San Francisco	California	94109	West	OFF
22	22	CA-2016-137330	12/9/2016	12/13/2016	Standard Class	KB-16585	Ken Black	Corporate	United States	Fremont	Nebras...	68025	Central	OFF
23	23	CA-2016-137330	12/9/2016	12/13/2016	Standard Class	KB-16585	Ken Black	Corporate	United States	Fremont	Nebras...	68025	Central	OFF
24	24	US-2017-156909	7/16/2017	7/18/2017	Second Class	SF-20065	Sandra Flanag...	Consumer	United States	Philadelphia	Pennsyl...	19140	East	FUR

```sql
--1.Top 10 Customers by Total Sales
select top 10 customer_id, customer_name,
round(sum(Sales), 2) as total_sales,
count(distinct order_id) as count_of_orders
from Superstore_Data
group by customer_id, customer_name
order by total_sales desc;
```

100 %

Results | Messages

	customer_id	customer_name	total_sales	count_of_orders
1	SM-20320	Sean Miller	25043.05	5
2	TC-20980	Tamara Chand	19052.22	5
3	RB-19360	Raymond Buch	15117.34	6
4	TA-21385	Tom Ashbrook	14595.62	4
5	AB-10105	Adrian Barton	14473.57	10
6	KL-16645	Ken Lonsdale	14175.23	12
7	SC-20095	Sanjit Chand	14142.33	9
8	HL-15040	Hunter Lopez	12873.3	6
9	SE-20110	Sanjit Engle	12209.44	11
10	CC-12370	Christopher Conant	12129.07	5

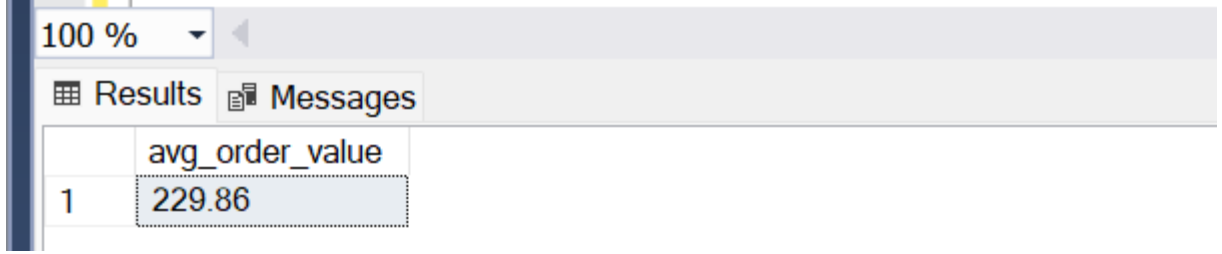
```sql
--2.Calculate total sales by year-month
select format(CAST(Order_Date AS datetime), 'yyyy-MM') as year_month,
round(sum(Sales), 2) as total_sales
from Superstore_Data
group by format(CAST(Order_Date AS datetime), 'yyyy-MM')
order by year_month
--48 rows
```

100 %

Results | Messages

	year_month	total_sales
1	2014-01	14236.89
2	2014-02	4519.89
3	2014-03	55691.01
4	2014-04	28295.34
5	2014-05	23648.29
6	2014-06	34595.13
7	2014-07	33946.39
8	2014-08	27909.47
9	2014-09	81777.35
10	2014-10	31453.39
11	2014-11	78628.72
12	2014-12	69545.62
13	2015-01	18174.08
14	2015-02	11951.41

```sql
--3.Average Order Value
select round(avg(sales), 2) as avg_order_value
from Superstore_Data
```



```sql
--4.Sales and Profit by Product Category
select category,
round(sum(sales), 2) as total_sales,
round(sum(cast(profit as float)), 2) as total_profit
from Superstore_Data
group by category
order by total_sales, total_profit;
```

100 %

Results | Messages

	category	total_sales	total_profit
1	Office Supplies	719047.03	122490.8
2	Furniture	741999.8	18451.27
3	Technology	836154.03	145454.95

```sql
--5.Orders by Shipping Mode
--what is the count & percentage distribution of orders by ship mode ?
select ship_mode,
count(ship_mode) as count_ship_mode,
round((count(ship_mode)*100.0/(select count(*) from Superstore_Data)),2) as percentage_distribution
from Superstore_Data
group by ship_mode
order by percentage_distribution desc;
```

100 %

Results | Messages

	ship_mode	count_ship_mode	percentage_distribution
1	Standard Class	5968	59.720000000000
2	Second Class	1945	19.460000000000
3	First Class	1538	15.390000000000
4	Same Day	543	5.430000000000

```sql
--6.Top 5 most profitable products
select top 5 product_name,
product_id,
round(avg(cast(profit as float)), 2) as avg_profit
from Superstore_Data
group by product_name, product_id
order by avg_profit desc;
```

100 %

Results | Messages

	product_name	product_id	avg_profit
1	Canon imageCLASS 2200 Advanced Copier	TEC-CO-10004722	5039.99
2	Canon imageCLASS MF7460 Monochrome Digital Laser Multifunction Copier	TEC-MA-10002927	1995.99
3	Ativa V4110MDD Micro-Cut Shredder	TEC-MA-10003979	1886.47
4	3D Systems Cube Printer, 2nd Generation, Magenta	TEC-MA-10001047	1858.99
5	Zebra ZM400 Thermal Label Printer	TEC-MA-10000045	1671.77

```sql
--7.Region-wise Sales & Profit Analysis
select region,
round(sum(cast(profit as float)), 2) as total_profit,
round(sum(sales), 2) as total_sales
from Superstore_Data
group by region
order by total_profit desc, total_sales desc;
```

100 %

Results | Messages

	region	total_profit	total_sales
1	West	108418.45	725457.82
2	East	91522.78	678781.24
3	South	46749.43	391721.91
4	Central	39706.36	501239.89

```sql
--8.Customer Segments Performance
select top 3 segment,
round(sum(sales), 2) as total_sales
from Superstore_Data
group by segment
order by total_sales desc
```

100 %

Results | Messages

	segment	total_sales
1	Consumer	1161401.34
2	Corporate	706146.37
3	Home Office	429653.15

```sql
--9.Daily Sales Trend
select cast(order_date as Date) as Order_day,
round(sum(sales), 2) as total_sales
from Superstore_Data
group by cast(order_date as Date)
order by total_sales desc, order_day;
--1237 rows
```

100 %

Results | Messages

	Order_day	total_sales
1	2014-03-18	28106.72
2	2016-10-02	18452.97
3	2017-10-22	15158.88
4	2017-03-23	14816.07
5	2014-09-08	14228.43
6	2017-11-17	13694.88
7	2015-11-08	12197
8	2016-12-17	12185.13
9	2014-11-17	11544.27
10	2015-09-17	11525.01
11	2017-11-04	10668.1
12	2014-09-23	10662.34
13	2016-05-23	10560.98
14	2016-12-25	10488.06
15	2017-12-02	9951.18
16	2017-08-17	9517.29
17	2017-09-02	9354.85
18	2014-09-19	9338.64

```sql
--10.Orders with Negative Profit (Loss-Making Orders)
select order_id, profit
from Superstore_Data
where TRY_CAST(profit AS FLOAT) < 0
order by order_id;
--1871 rows
```

100 %

Results | Messages

	order_id	profit
1	CA-2014-100090	-87.9354
2	CA-2014-100678	-18.1176
3	CA-2014-101147	-6.3441
4	CA-2014-101175	-1.2588
5	CA-2014-101602	-9.153
6	CA-2014-101602	-21.808
7	CA-2014-101770	-1.3083
8	CA-2014-101931	-36.294
9	CA-2014-102295	-18.1068
10	CA-2014-102869	-4.8528
11	CA-2014-102869	-31.6204
12	CA-2014-103086	-1.5936
13	CA-2014-103191	-82.884
14	CA-2014-103317	-6.867
15	CA-2014-103373	-168.9...
16	CA-2014-103492	-19.168
17	CA-2014-103492	-1.5715
18	CA-2014-103702	-46.8776

```sql
--11.Sales & Profit by top 10 States
select top 10 state,
round(sum(sales), 2) as total_sales,
round(sum(cast(profit as float)), 2) as total_profit
from Superstore_Data
group by state
order by total_sales desc, total_profit desc;
```

100 %

Results | Messages

	state	total_sales	total_profit
1	California	457687.63	76381.39
2	New York	310876.27	74038.55
3	Texas	170188.05	-25729.36
4	Washington	138641.27	33402.65
5	Pennsylvania	116511.91	-15559.96
6	Florida	89473.71	-3399.3
7	Illinois	80166.1	-12607.89
8	Ohio	78258.14	-16971.38
9	Michigan	76269.61	24463.19
10	Virginia	70636.72	18597.95

```sql
--12.Top 10 states performing worst in terms of sales & profit
select top 10 state,
round(sum(sales), 2) as total_sales,
round(sum(cast(profit as float)), 2) as total_profit
from Superstore_Data
group by state
having sum(sales) < 0 or sum(cast(profit as float)) < 0
```

100 %

Results | Messages

	state	total_sales	total_profit
1	Illinois	80166.1	-12607.89
2	Pennsylvania	116511.91	-15559.96
3	Ohio	78258.14	-16971.38
4	Arizona	35282	-3427.92
5	Oregon	17431.15	-1190.47
6	Florida	89473.71	-3399.3
7	North Carolina	55603.16	-7490.91
8	Colorado	32108.12	-6527.86
9	Texas	170188.05	-25729.36
10	Tennessee	30661.87	-5341.69

```sql
--13.Average Delivery Time by Region
select region, order_id,
avg(datediff(day, order_date, ship_date)) as avg_no_of_days,
avg(datediff(hour, order_date, ship_date)) as avg_no_of_days
from Superstore_Data
group by region, order_id
order by region, avg(datediff(day, order_date, ship_date)) desc;
```

100 %

Results | Messages

	region	order_id	avg_no_of_days	avg_no_of_days
1	Central	CA-2014-105893	7	168
2	Central	CA-2014-115336	7	168
3	Central	CA-2014-118339	7	168
4	Central	CA-2014-124723	7	168
5	Central	CA-2014-126193	7	168
6	Central	CA-2014-126802	7	168
7	Central	CA-2014-126907	7	168
8	Central	CA-2014-128888	7	168
9	Central	CA-2014-142510	7	168
10	Central	CA-2014-146283	7	168
11	Central	CA-2014-154165	7	168
12	Central	CA-2015-110863	7	168
13	Central	CA-2015-112522	7	168
14	Central	CA-2015-120551	7	168
15	Central	CA-2015-121132	7	168
16	Central	CA-2015-121041	7	168
17	Central	CA-2015-130792	7	168
18	Central	CA-2015-137064	7	168

```sql
--14.Identify Sales Outliers (High Value Orders)
select *
from Superstore_Data
where sales > (select avg(sales) + 2*STDEV(sales) from Superstore_Data)
--247 rows
```

100 %

Results | Messages

	Row_ID	Order_ID	Order_Date	Ship_Date	Ship_Mode	Customer_ID	Customer_Name	Segment	Country	City	State	Postal_Code	Region	Pro
1	11	CA-2014-115812	6/9/2014	6/14/2014	Standard Class	BH-11710	Brosina Hoffman	Consumer	United States	Los Angeles	California	90032	West	FU
2	28	US-2015-150630	9/17/2015	9/21/2015	Standard Class	TB-21520	Tracy Blumstein	Consumer	United States	Philadelphia	Pennsylvania	19140	East	FU
3	150	CA-2016-114489	12/5/2016	12/9/2016	Standard Class	JE-16165	Justin Ellison	Corporate	United States	Franklin	Wisconsin	53132	Central	FU
4	166	CA-2014-139892	9/8/2014	9/12/2014	Standard Class	BM-11140	Becky Martin	Consumer	United States	San Antonio	Texas	78207	Central	TE
5	168	CA-2014-139892	9/8/2014	9/12/2014	Standard Class	BM-11140	Becky Martin	Consumer	United States	San Antonio	Texas	78207	Central	FU
6	245	CA-2014-131926	6/1/2014	6/6/2014	Second Class	DW-13480	Dianna Wilson	Home Office	United States	Lakeville	Minnesota	55044	Central	FU
7	248	CA-2014-131926	6/1/2014	6/6/2014	Second Class	DW-13480	Dianna Wilson	Home Office	United States	Lakeville	Minnesota	55044	Central	OF
8	252	CA-2016-145625	9/11/2016	9/17/2016	Standard Class	KC-16540	Kelly Collister	Consumer	United States	San Diego	California	92037	West	TE
9	263	US-2014-106992	9/19/2014	9/21/2014	Second Class	SB-20290	Sean Braxton	Corporate	United States	Houston	Texas	77036	Central	TE
10	264	US-2014-106992	9/19/2014	9/21/2014	Second Class	SB-20290	Sean Braxton	Corporate	United States	Houston	Texas	77036	Central	TE
11	319	CA-2014-164973	11/4/2014	11/9/2014	Standard Class	NM-18445	Nathan Mautz	Home Office	United States	New York ...	New York	10024	East	TE
12	354	CA-2016-129714	9/1/2016	9/3/2016	First Class	AB-10060	Adam Bellavan...	Home Office	United States	New York ...	New York	10009	East	OF
13	378	US-2017-134481	8/27/2017	9/1/2017	Standard Class	AR-10405	Allen Rosenblatt	Corporate	United States	Franklin	Massachus...	2038	East	FU
14	393	US-2014-135972	9/21/2014	9/23/2014	Second Class	JG-15115	Jack Garza	Consumer	United States	Des Moines	Washington	98198	West	TE
15	400	CA-2016-108987	9/8/2016	9/10/2016	Second Class	AG-10675	Anna Gayman	Consumer	United States	Houston	Texas	77036	Central	FU
16	488	CA-2014-154627	10/29/2014	10/31/20...	First Class	SA-20830	Sue Ann Reed	Consumer	United States	Chicago	Illinois	60610	Central	TE
17	510	CA-2015-145352	3/16/2015	3/22/2015	Standard Class	CM-12385	Christopher Ma...	Consumer	United States	Atlanta	Georgia	30318	South	OF
18	516	CA-2017-127432	1/22/2017	1/27/2017	Standard Class	AD-10180	Alan Dominguez	Home Office	United States	Great Falls	Montana	59405	West	TE

```sql
--15.Repeat Customer Rate
select count(distinct customer_id) as total_customers,
count(distinct case when order_count > 1 then customer_id end) as repeat_customers,
round(count (distinct case when order_count > 1 then customer_id end)*100.0/count(distinct customer_id),2) as repeat_customer_rate
from (
select customer_id,
count(order_id) as order_count
from Superstore_Data
group by customer_id) A;
```

100 %

Results | Messages

	total_customers	repeat_customers	repeat_customer_rate
1	793	788	99.370000000000