Submitted By – Swati

Valuable Insights using SQL

Tool Used : Microsoft SQL Server

I. Database : Super Store

Q.1. What are the number of products which were bought in between the year 2014 to 2016 by various states?

```
--To select count of Products which were bought in between the year 2014 to 2016 by various states
select state, YEAR(order_date) AS order_year, COUNT(product_id) AS product_count
FROM orders
WHERE YEAR(order_date) BETWEEN 2014 AND 2016
GROUP BY state, YEAR(order_date);
```

Results | Messages

state	order_year	product_count
Rhode Island	2015	49
Nebraska	2016	93
Nevada	2016	65
Oklahoma	2016	206
New Jersey	2016	326
Iowa	2015	26
Colorado	2015	269
Arizona	2016	285
New York	2015	1722
Massachusetts	2016	158
Missouri	2014	88

Q.2. What are the number of orders coming from various regions?

```
--To fetch the number of orders coming from the various regions--
select region,count(order_id) as number_of_orders from orders group by region;
```

100 % | Results | Messages

	region	number_of_orders
1	East	18668
2	South	10604
3	West	21015
4	Central	15248

Q.3. Select the order_id,product_id and customer_name who has returned the orders from the New York state.

```
--List the Order_id , Product_id and Customer Name who has returned the orders from the New York State--
select r.order_id, o.product_id, o.customer_name
from orders o left outer join returns r
on o.order_id=r.order_id
where r.returned='Yes'
and o.state='New York';
```

Results | Messages

order_id	product_id	customer_name
CA-2015-134201	OFF-AR-10002804	Eugene Hilde...
CA-2015-134201	OFF-ST-10001476	Eugene Hilde...
CA-2015-134201	OFF-AR-10002804	Eugene Hilde...
CA-2015-134201	OFF-ST-10001476	Eugene Hilde...
CA-2015-134201	OFF-AR-10002804	Eugene Hilde...
CA-2015-134201	OFF-ST-10001476	Eugene Hilde...
CA-2015-134201	OFF-AR-10002804	Eugene Hilde...
CA-2015-140984	OFF-ST-10000585	Craig Carroll
CA-2015-140984	TEC-CO-10001571	Craig Carroll
CA-2015-140984	OFF-BI-10002012	Craig Carroll
CA-2015-140984	OFF-ST-10000585	Craig Carroll
CA-2015-140984	TEC-CO-10001571	Craig Carroll

Q.4. To fetch the list of products which were bought by the city Houston in the year 2014.

```
--To fetch the products which were bought by the city Houston in the year 2014

select product_id,product_name,city from orders where city ='Houston' and year(order_date)=2014;
```

Results | Messages

product_id	product_name	city
OFF-AR-10003183	Avery Fluorescent Highlighter Four-Color Set	Houston
OFF-SU-10000432	Acco Side-Punched Conventional Columnar Pads	Houston
TEC-AC-10003832	Imation 16GB Mini TravelDrive USB 2.0 Flash Drive	Houston
OFF-PA-10004100	Xerox 216	Houston
FUR-BO-10004467	Bestar Classic Bookcase	Houston
OFF-PA-10002005	Xerox 225	Houston
OFF-FA-10002815	Staples	Houston
FUR-CH-10001973	Office Star Flex Back Scooter Chair with White Frame	Houston
OFF-PA-10002986	Xerox 1898	Houston
OFF-PA-10001593	Xerox 1947	Houston
OFF-BI-10004040	Wilson Jones Impact Binders	Houston
FUR-FU-10003194	Eldon Expressions Desk Accessory, Wood Pencil Ho	Houston

Q.5. To fetch the top 5 sum of sales by state and categories in the year 2014 and 2015.

```
--Fetch the top 5 sum of sales by state and categories in the year 2014 and 2015
select top 5 year(order_date) as year_of_order,state,category,sum(sales) as sum_of_sales
from orders
where year(order_date) in (2014,2015)
group by year(order_date),state,category
order by sum(sales) desc;
```

esults | Messages

year_of_order	state	category	sum_of_sales
2015	New York	Technology	227898.558
2014	California	Furniture	220208.349500001
2015	California	Furniture	212438.098
2014	California	Office Supplies	201186.366
2015	California	Technology	200037.892

Q.6. Identify the duplicate rows from Orders Table

```
--Identifying duplicates from the table Orders
with cteduplicates as (
select *,
row_number() over (partition by order_id,year(order_date),customer_id, product_id order by order_id) RowNum
from orders)
select * from cteduplicates;
```

Results | Messages

		city	state	postalcode	region	product_id	category	subcategory	product_name	sales	quantity	discount	profit	RowNum
1	tes	New York City	New York	10024	East	TEC-PH-10002075	Technology	Phones	AT&T EL51110 DECT	377.97	3	0	109.6113	1
2	tes	New York City	New York	10024	East	TEC-PH-10002075	Technology	Phones	AT&T EL51110 DECT	377.97	3	0	109.6113	2
3	tes	New York City	New York	10024	East	TEC-PH-10002075	Technology	Phones	AT&T EL51110 DECT	377.97	3	0	109.6113	3
4	tes	New York City	New York	10024	East	TEC-PH-10002075	Technology	Phones	AT&T EL51110 DECT	377.97	3	0	109.6113	4
5	tes	New York City	New York	10024	East	TEC-PH-10002075	Technology	Phones	AT&T EL51110 DECT	377.97	3	0	109.6113	5
6	tes	New York City	New York	10024	East	TEC-PH-10002075	Technology	Phones	AT&T EL51110 DECT	377.97	3	0	109.6113	6
7	tes	New York City	New York	10024	East	TEC-PH-10002075	Technology	Phones	AT&T EL51110 DECT	377.97	3	0	109.6113	7
8	tes	San Francisco	California	94122	West	FUR-TA-10003715	Furniture	Tables	Hon 2111 Invitation Series Corner Table	502.488	3	0.2	-87.9354	1
9	tes	San Francisco	California	94122	West	FUR-TA-10003715	Furniture	Tables	Hon 2111 Invitation Series Corner Table	502.488	3	0.2	-87.9354	2
10	tes	San Francisco	California	94122	West	FUR-TA-10003715	Furniture	Tables	Hon 2111 Invitation Series Corner Table	502.488	3	0.2	-87.9354	3

Q.7. Delete the duplicates from the Orders Table

```
--Deleting the duplicates from the table Orders
with cteduplicates as (
select *,
row_number() over (partition by order_id,year(order_date),customer_id, product_id order by order_id) RowNum
from orders)
delete from cteduplicates where RowNum>1;
```

Messages

```
(55549 rows affected)

Completion time: 2024-06-08T13:15:19.0285379+05:30
```

Q.8. To find out the Year on Year Growth of Superstore based on Sales?

```
--To find year on year growth based on Sales

with cte_PreviousYearTotal
as
(
select year(order_date) as OrderYear, sum(sales) as TotalSales from orders
group by year(order_date)
)
select OrderYear,TotalSales, LAG(TotalSales,1) over (order by OrderYear) PrevYearTotal,
format((TotalSales-LAG(TotalSales,1) over (order by OrderYear))/
LAG(TotalSales,1) over (order by OrderYear),'P') YOYGrowth
from cte_PreviousYearTotal;
```

Results | Messages

OrderYear	TotalSales	PrevYearTotal	YOYGrowth
2014	483966.126100001	NULL	NULL
2015	470442.449000001	483966.126100001	-2.79%
2016	608532.458000001	470442.449000001	29.35%
2017	732568.539199999	608532.458000001	20.38%

II. Database: Netflix

Q.9. To display the count of type of content on Netflix.

```
--To display Number of TV shows and Movie on Netflix
select type, count(distinct(show_id)) as Count_of_Type_of_Content from netflix_titles
group by type order by type desc;
```

100 %

Results | Messages

	type	Count_of_Type_of_Content
1	TV Show	1969
2	Movie	4263

Q.10. To display the country names where movies were released in the year 2019 by the director Richard Finn.

```
--To display the country names where movies were released in the year 2019 by the director Richard Finn.

select t.type,t.title,t.release_year,d.director,c.country from netflix_titles t
join netflix_titles_countries c on t.show_id=c.show_id
join netflix_titles_directors d on t.show_id=d.show_id
where t.release_year ='2019' and d.director='Richard Finn';
```

100 %

Results | Messages

	type	title	release_year	director	country
1	Movie	Norm of the North: King Sized Adventure	2019	Richard Finn	South Korea
2	Movie	Norm of the North: King Sized Adventure	2019	Richard Finn	United States
3	Movie	Norm of the North: King Sized Adventure	2019	Richard Finn	China
4	Movie	Norm of the North: King Sized Adventure	2019	Richard Finn	India

Q.11. To display the ranks of content category wise like which released first between the year 1990 and 2020.

```
--To display the ranks of Content category wise like which released first between the year 1990 and 2020.
with cte_top_category_yearwise as
(
select t.release_year,t.date_added,t.title,c.listed_in, dense_rank() over (partition by t.release_year order by c.listed_in desc) ranks
from netflix_titles t join netflix_titles_category c
on t.show_id=c.show_id
)
select * from cte_top_category_yearwise where release_year between '1990' and '2020';
```

100 %

Results | Messages

	release_year	date_added	title	listed_in	ranks
1	1990	2017-07-01 00:00:00.000	Twin Peaks	TV Dramas	1
2	1990	2014-12-18 00:00:00.000	Pee-wee's Playhouse	TV Comedies	2
3	1990	2018-12-31 00:00:00.000	Tim Allen: Men Are Pigs	Stand-Up Comedy	3
4	1990	2019-08-01 00:00:00.000	Rocky V	Sports Movies	4
5	1990	2019-11-01 00:00:00.000	Listen Up! The Lives of Quincy Jones	Music & Musicals	5
6	1990	2014-12-18 00:00:00.000	Pee-wee's Playhouse	Kids' TV	6
7	1990	2018-08-16 00:00:00.000	Look Out, Officer	International Movies	7
8	1990	2019-12-31 00:00:00.000	Ghayal	International Movies	7
9	1990	2019-12-31 00:00:00.000	Shiva	International Movies	7
10	1990	2019-10-01 00:00:00.000	Escape from the "Liberty" Cinema	Independent Movies	8
11	1990	2020-01-01 00:00:00.000	Tremors	Horror Movies	9

Q.12. Actors on Netflix with most movies in last five years till 2020.

```
--Actors on Netflix with most movies in last five years till 2020.
select top 10 t.type,c.cast,count(t.type) as count_of_movies,
dense_rank() over (partition by t.type order by count(t.type) desc) ranks
from netflix_titles t
join netflix_titles_cast c on t.show_id=c.show_id
where t.type='Movie' and t.release_year between '2015' and '2020'
group by t.type,c.cast
having count(t.type)>1;
```

100 %

Results | Messages

	type	cast	count_of_movies	ranks
1	Movie	Andrea Libman	13	1
2	Movie	Adil Hussain	11	2
3	Movie	Nassar	11	2
4	Movie	Pankaj Tripathi	10	3
5	Movie	Danny Trejo	10	3
6	Movie	Sathyaraj	9	4
7	Movie	Tamannaah Bhatia	9	4
8	Movie	Vincent Tong	8	5
9	Movie	Sardar Sohi	8	5
10	Movie	Prabhas	8	5

Q.13. Actors on Netflix with most TV shows in last five years till 2020.

```
--Actors on Netflix with most TV shows in last five years till 2020.

select top 10 t.type,c.cast,count(t.type) as count_of_movies,
dense_rank() over (partition by t.type order by count(t.type) desc) ranks
from netflix_titles t
join netflix_titles_cast c on t.show_id=c.show_id
where t.type='TV Show ' and t.release_year between '2015' and '2020'
group by t.type,c.cast
having count(t.type)>1;
```

00 %

Results | Messages

	type	cast	count_of_movies	ranks
1	TV Show	Takahiro Sakurai	13	1
2	TV Show	Ashleigh Ball	10	2
3	TV Show	Hiroshi Kamiya	9	3
4	TV Show	Vincent Tong	9	3
5	TV Show	Yuki Kaji	9	3
6	TV Show	Ai Kayano	8	4
7	TV Show	Daisuke Ono	8	4
8	TV Show	Kathleen Barr	8	4
9	TV Show	Kevin Michael Richardson	7	5
10	TV Show	Kenjiro Tsuda	7	5

III. Database: military

Dataset: World_Military_Power.xls

Q.14. To display top 30 military strength world-wide ranking.

```
--To display top 30 military stengths world wide ranking
--We observed that India comes at 25th place in terms of military strength ranking
--Minimum the number, highest will be the ranking.
select top 30 military_strength,min(military_strength_power_index)as military_strength_power_index from data$
group by military_strength
having min(military_strength_power_index)<1;
```

100 %

Results | Messages

	military_strength	military_strength_power_index
19	Ethiopia	0.8581
20	Finland	0.8498
21	France	0.1702
22	Germany	0.2186
23	Greece	0.5311
24	Hungary	0.8215
25	India	0.0953
26	Indonesia	0.2544
27	Iran	0.2191
28	Iraq	0.7911
29	Israel	0.3111
30	Italy	0.2111

We observed that India comes at 25th place in terms of military strength ranking. Minimum the number, highest will be the ranking.

Q.15. To display the top 10 countries with highest external debt value.

```
--To display the top 10 countries with highest external debt
SELECT TOP 10
    external_debt,
    MAX(CAST(REPLACE(external_debt_value, ',', '') AS BIGINT)) AS external_debt
FROM
    data$
GROUP BY external_debt_value,external_debt
ORDER BY external_debt_value desc;
```

100 %

Results | Messages

	external_debt	defense_budget
1	Vietnam	96580000000
2	Romania	95970000000
3	New Zealand	91620000000
4	Panama	91530000000
5	Albania	9505000000
6	Sweden	939900000000
7	Cameroon	9375000000
8	Zimbabwe	9357000000
9	Israel	88660000000
10	Pakistan	82190000000

Q.16. Prepare a report of having names of all countries with aircraft strength value, aircraft fleet value and helicopter fleet strength value from military data.

```
/*Prepare a report of having names of all countries with aircraft strength value,aircraft fleet value and
helicopter fleet strength value from military data*/

select aircraft_strength,aircraft_strength_value,transport_aircraft_fleet_strength_value,attack_helicopter_fleet_strength_value
from data$;
```

100 %

Results | Messages

	aircraft_strength	aircraft_strength_value	transport_aircraft_fleet_strength_value	attack_helicopter_fleet_strength_value
1	Afghanistan	260	30	0
2	Albania	19	0	0
3	Algeria	551	59	45
4	Angola	295	30	15
5	Argentina	227	9	0
6	Armenia	64	3	20
7	Australia	464	38	22
8	Austria	120	11	0
9	Azerbaijan	147	1	17
10	Bahrain	109	2	22
11	Bangladesh	177	11	0
12	Belarus	202	4	21

**********THANK YOU**********