Queries from questions

1. **How many Customers do we have in the data?**

```
select count(1) as total_customers
from customers
```

795

1. **What was the city with the most profit for the company in 2015 and how much was it?**

```
select DISTINCT shipping_city as city, sum(order_profits)
from orders o
join order_details od
	on o.order_id = od.order_id
where order_date like '%2015%'
group by 1
order by 2 desc
```

New York City 14753

1. **How many different cities do we have in the data? CORRECT**

```
with t1 as (
	select DISTINCT(shipping_city) as cities
	from orders
)

select count(1) as total_cities
from t1
```

531

1. **Show the total spent by customers from low to high.**

```
select c.customer_id, sum(order_sales) total_spent
from customers c
join orders o
	on c.customer_id = o.customer_id
join order_details od
	on o.order_id = od.order_id
group by 1
order by 2
limit 5
```

546, 657

1. **What is the most profitable City in the State of Tennessee?**

```
select DISTINCT shipping_city as city, sum(order_profits)
from orders o
join order_details od
	on o.order_id = od.order_id
where shipping_state = 'Tennessee'
group by 1
order by 2 desc
```

Lebanon

1. **What's the average annual profit for that city across all years in that city?**

```
select DISTINCT shipping_city as city, round(avg(order_profits),2)
from orders o
join order_details od
	on o.order_id = od.order_id
where shipping_city = 'Lebanon'
group by 1
order by 2 desc
```

27.67

1. **What is the distribution of customer types in the data?**

```
with t1 as (
	select customer_segment, count(1) as total_consumers
	from customers
	where customer_segment in ('Consumer')
),

t2 as (
	select customer_segment, count(1) as total_corporate
	from customers
	where customer_segment in ('Corporate')
),

t3 as (
	select customer_segment, count(1) as total_home_office
	from customers
	where customer_segment in ('Home Office')
)

select t1.total_consumers, t2.total_corporate, t3.total_home_office
```

```
from t1, t2, t3
```

237

1. **What's the most profitable product category on average in Iowa across all years?**

```
select p.product_category, round(avg(od.order_profits), 2) as avg_profit
from orders o
	join order_details od
		on o.order_id = od.order_id
	join product p
		on od.product_id = p.product_id
where o.shipping_state like '%iowa%'
group by 1
order by 2 desc
```

Furniture

1. **What is the most popular product in that category across all states in 2016?**

```
select p.product_name, sum(od.quantity) as products_ordered
from orders o
	join order_details od
		on o.order_id = od.order_id
	join product p
		on od.product_id = p.product_id
where o.order_date like '%2016%' and p.product_category in ('Furniture')
group by 1
order by 2 DESC
limit 5
```

Global Push Button Manager's Chair, Indigo

1. **Which customer got the most discount in the data? (in total amount)**

```
with t1 as

(SELECT c.customer_id as customerid, (od.order_sales /(1-od.order_discount))
as total_price
from customers c
	join orders o
		on c.customer_id = o.customer_id
	join order_details od
		on o.order_id = od.order_id
order by 2 DESC
limit 10),
```

```
t2 as (
	select order_sales as sales
	from order_details
)

select t1.customerid, total_price - sales as totaldiscount
from t1, t2
group by 1
order by 2 desc
```

687

1. **How widely did monthly profits vary in 2018?**

```
WITH t1 AS (
SELECT
	CAST(SUBSTR(O.order_date,INSTR(O.order_date,'/') -2,2) AS INT) AS
month,
	SUM(OD.order_profits) AS profit,
	SUBSTR(O.order_date, -4, 4) as year
FROM order_details OD
JOIN orders O
	USING (order_id)
WHERE year = '2018'
GROUP BY 1
),
t2 AS
(
SELECT
	CAST(SUBSTR(O.order_date,INSTR(O.order_date,'/') -2,2) AS INT) AS
month,
	LAG(SUM(OD.order_profits), 1,0)
		OVER (ORDER BY SUBSTR(O.order_date,INSTR(O.order_date, '/')
-2,2) ) AS previous_month,
		SUBSTR(O.order_date, -4, 4) as year
FROM order_details OD
JOIN orders O
	USING (order_id)
WHERE year = '2018'
GROUP BY month
ORDER BY month ASC
)
SELECT t1.month,
	ABS(t1.profit) - ABS(t2.previous_month) AS change
FROM t1
```

JOIN t2 USING (month)

-13824

1. **Which order was the highest in 2015?**

```
select p.product_id, MAX(od.order_sales) as number_of_sales
from orders o
	join order_details od
		on o.order_id = od.order_id
	join product p
		on od.product_id = p.product_id
where o.order_date like '%2015%'
group by 1
order by 2 DESC
limit 5
```

CA-2015-145317
445

1. **What was the rank of each city in the East region in 2015?**

```
select o.shipping_city as city, SUM(od.quantity) as quantity,
	RANK () OVER (ORDER BY quantity desc) CityRank
from orders o
	join order_details od
		on o.order_id = od.order_id
where o.order_date like '%2015%' and o.shipping_region = 'East'
group by 1
order by quantity desc
```

Columbus

1. **Join all DB tables into one dataset that includes all unique columns and download it as a csv file. In the second part of the project, you're gonna work with this one table.**

```
select *
from customers c
	join orders o
		on c.customer_id = o.customer_id
	join order_details od
		on o.order_id = od.order_id
	join product p
		on p.product_id = od.product_id
```