**Questions**

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

SELECT DISTINCT \*

FROM customers

795 customers

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

SELECT o.shipping\_city, o.order\_date, SUM(od.order\_profits)

FROM orders AS o

JOIN order\_details AS od

ON od.order\_id = o.order\_id

WHERE o.order\_date LIKE '%2015'

GROUP BY 1

ORDER BY 3 DESC

SELECT DISTINCT SUBSTR(order\_date, LENGTH(order\_date)-3, 4) AS year FROM ORDERS;

WHERE order\_date LIKE '%2015

New York city which made most profit in the year 2015 of 14753

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

SELECT DISTINCT o.shipping\_city  
FROM orders AS o

531 cities

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

SELECT c.customer\_id, c.customer\_name, SUM(od.order\_sales)

FROM customers AS c

JOIN orders AS o

ON o.customer\_id = c.customer\_id

JOIN order\_details AS od

ON od.order\_id = o.order\_id

GROUP BY 1

ORDER BY 3

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

SELECT o.shipping\_city, SUM(od.order\_profits)

FROM orders AS o

JOIN order\_details AS od

ON od.order\_id = o.order\_id

WHERE o.shipping\_state = 'Tennessee'

GROUP BY 1

ORDER BY 2 DESC

Nashville with 224 profit, Lebanon 83, Jackson 190

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

SELECT o.shipping\_city, AVG(od.order\_profits)

FROM orders AS o

JOIN order\_details AS od

ON od.order\_id = o.order\_id

WHERE o.shipping\_city = 'Lebanon'

GROUP BY 1

27.67

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

SELECT COUNT(c.customer\_segment)

FROM customers AS c

WHERE c.customer\_segment = 'Corporate'

237

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

SELECT p.product\_category, AVG(od.order\_profits)

FROM orders AS o

JOIN order\_details AS od

ON od.order\_id = o.order\_id

JOIN product AS p

ON p.product\_id = od.product\_id

WHERE o.shipping\_state = 'Iowa'

GROUP BY 1

Furniture- across all years avg profit is 130.25

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

SELECT p.product\_category, p.product\_name, SUM(od.quantity)

FROM orders AS o

JOIN order\_details AS od

ON od.order\_id = o.order\_id

JOIN product AS p

ON p.product\_id = od.product\_id

WHERE o.order\_date LIKE '%2016' AND p.product\_category = 'Furniture'

GROUP BY 1,2

ORDER BY 3 DESC

Global Push Button Manager's Chair, Indigo

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

SELECT c.customer\_id, c.customer\_name, SUM(od.order\_sales / (1 - od.order\_discount) - od.order\_sales) AS total\_of\_discount

FROM customers AS c

JOIN orders AS o

ON o.customer\_id = c.customer\_id

JOIN order\_details AS od

ON od.order\_id = o.order\_id

GROUP BY 2

ORDER BY 3 DESC

687 Sean Miller 23929.0833333333

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 num) AS month,

SUM(od.order\_profits) AS profit,

SUBSTR(o.order\_date, LENGTH(o.order\_date)-3, 4) AS year

FROM orders AS o

JOIN order\_details AS od

ON od.order\_id = o.order\_id

WHERE year LIKE '2018'

GROUP BY 1

),

t2

AS( SELECT CAST (SUBSTR(O.order\_date,INSTR(O.order\_date,'/') -2,2) AS num) AS month,

LAG(SUM(OD.order\_profits),1,0) OVER (ORDER BY substr(O.order\_date,INSTR(O.order\_date, '/') -2,2) ASC ) AS profit\_difference,

SUBSTR(O.order\_date, LENGTH(O.order\_date)-3, 4) AS year

FROM order\_details OD

JOIN orders O

USING (order\_id)

WHERE year LIKE '2018'

GROUP BY month

ORDER BY month ASC

)

SELECT t1.month, t1.profit, t2.profit\_difference, ABS(t1.profit) - ABS(t2.profit\_difference) AS change

FROM t1

JOIN t2 USING (month)

-13824

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

SELECT o.order\_id, o.order\_date, SUM(od.order\_sales)  
FROM orders AS o  
JOIN order\_details AS od  
    ON od.order\_id = o.order\_id  
WHERE o.order\_date LIKE '%2015'  
GROUP BY 1, 2  
ORDER BY 3 DESC

3/18/2015 CA-2015-111871 28105

OR CA-2015-145317 3/18/2015 23660

13.**What was the rank of each city in the East region in 2015?- quantity ordered**

SELECT o.shipping\_city, SUM(od.quantity), RANK() OVER(ORDER BY SUM(od.quantity) DESC) AS rank\_city

FROM orders AS o

JOIN order\_details AS od

ON od.order\_id = o.order\_id

WHERE o.order\_date LIKE '%2015' AND o.shipping\_region = 'East'

GROUP BY 1

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.**

CREATE TABLE unique\_table

AS (SELECT c.customer\_id, c.customer\_name, c.customer\_segment, o.order\_id, o.order\_date, o.shipping\_city, o.shipping\_state,

o.shipping\_region, o.shipping\_country, o.shipping\_postal\_code, o.shipping\_date, o.shipping\_mode, od.order\_details\_id, od.product\_id,

od.quantity, od.order\_discount, od.order\_profits, od.order\_profit\_ratio, od.order\_sales, p.product\_name, p.product\_category,

p.product\_subcategory, p.product\_manufacturer

FROM customers AS c

JOIN orders AS o

ON o.customer\_id = o.customer\_id

JOIN order\_details AS od

ON od.order\_id = o.order\_id

JOIN product AS p

ON p.product\_id = od.product\_id)

CREATE TABLE unique\_table

AS SELECT \*

FROM customers

JOIN (orders JOIN (order\_details JOIN product USING (product\_id)) USING(order\_id)) USING(customer\_id);