CAPSTONE PROJECT1

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

SELECT COUNT(\*) FROM customers;

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

SELECT SUM(o.order\_profits) as profit, l.shipping\_city

FROM order\_details o

JOIN orders l

ON o.order\_id = l.order\_id

WHERE l.order\_date like '%2015'

GROUP by l.shipping\_city

ORDER by profit DESC

LIMIT 1;

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

SELECT COUNT(DISTINCT shipping\_city) FROM orders;

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

SELECT c.customer\_id, c. customer\_name, SUM(l.order\_sales) as total\_spent

FROM customers c

JOIN orders o

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

JOIN order\_details l

ON o.order\_id = l.order\_id

GROUP BY c.customer\_id

ORDER BY total\_spent;

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

SELECT sum(o.order\_profits) as total\_profit, l.shipping\_city

FROM order\_details o

JOIN orders l

ON o.order\_id = l.order\_id

WHERE l.shipping\_state = 'Tennessee'

GROUP BY l.shipping\_city

ORDER BY total\_profit DESC

LIMIT 1;

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

SELECT round(Avg(o.order\_profits),2) as avg\_annual\_profit

FROM order\_details o

JOIN orders l

ON o.order\_id = l.order\_id

WHERE l.shipping\_city = 'Lebanon';

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

SELECT

SUM(CASE WHEN customer\_segment = 'Consumer' THEN 1.0 ELSE 0 END) AS consumer,

ROUND(SUM(CASE WHEN customer\_segment = 'Corporate' THEN 1.0 ELSE 0 END) AS corporate,

SUM(CASE WHEN customer\_segment = 'Home Office' THEN 1.0 ELSE 0 END) AS home\_office\_ratio

FROM customers;

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

SELECT p.product\_category, SUM(o.order\_profits) as total\_profit

FROM product p

JOIN order\_details o

ON p.product\_id = o.product\_id

JOIN orders l

ON l.order\_id = o.order\_id

WHERE l.shipping\_state = 'Iowa'

GROUP BY p.product\_category

ORDER BY total\_profit DESC

LIMIT 1;

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

SELECT SUM(o.quantity) AS total\_quantity, p.product\_name

FROM order\_details o

JOIN orders l

ON o.order\_id = l.order\_id

JOIN product p

ON o.product\_id = p.product\_id

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

GROUP BY p.product\_name

ORDER BY total\_quantity DESC

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

SELECT c.customer\_id, c.customer\_name, SUM((o.order\_sales/(1-o.order\_discount))- o.order\_sales) as total\_discount

FROM customers c

JOIN orders l

ON c.customer\_id = l.customer\_id

JOIN order\_details o

ON o. order\_id = l.order\_id

GROUP BY c. customer\_name

ORDER BY total\_discount DESC

LIMIT 1;

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

SELECT month, total\_profit, total\_profit - LAG(total\_profit ,1,1) OVER(ORDER BY month ) as profit\_difference

FROM

(SELECT CAST(SUBSTR(l.order\_date,1,INSTR(l.order\_date,'/') -1) AS int) as month, SUM(o.order\_profits) AS total\_profit

FROM order\_details o

JOIN orders l

ON o.order\_id = l.order\_id

WHERE l.order\_date LIKE '%2018'

GROUP BY month

ORDER BY month) as a

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

SELECT o.order\_id, SUM(l.order\_sales ) AS total\_sales

FROM orders o

JOIN order\_details l

ON o.order\_id = l.order\_id

WHERE order\_date LIKE '%2015'

GROUP BY o.order\_id

ORDER BY total\_sales DESC

LIMIT 1;

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

SELECT RANK() OVER( ORDER BY total DESC) AS ranking, a.shipping\_city, total

FROM

( SELECT SUM(o.quantity) AS total, l.shipping\_city

FROM order\_details o

JOIN orders l

ON o.order\_id = l.order\_id

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

GROUP BY l.shipping\_city

ORDER BY total) AS a

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 DISTINCT 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,

l.order\_details\_id,l.product\_id, l.quantity, l.order\_discount, l.order\_profits, l.order\_profit\_ratio,l.order\_sales,

p.product\_name, p.product\_category, p.product\_subcategory, p.product\_manufacturer

FROM customers c

JOIN orders o

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

JOIN order\_details l

ON o.order\_id = l.order\_id

JOIN product p

ON l.product\_id = p.product\_id