**Capstone Project : DATA EXPLORATION SQL**

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

**Ans.** 795

SELECT COUNT(\*) FROM "customers";

**---------------------------------------------------------------------------------**

1. **What was the city with the most profit for the company in 2015?**

**Ans.** New York City, Profit: 14753

1. SELECT
2. o.shipping\_city,
3. SUM(od.order\_profits) AS maximum\_profits
4. FROM orders AS o
5. LEFT JOIN order\_details AS od
6. ON o.order\_id = od.order\_id
7. WHERE EXTRACT(year FROM o.order\_date) = 2015
8. GROUP BY 1
9. ORDER BY 2 DESC;

**---------------------------------------------------------------------------------**

1. **In 2015, what was the most profitable city's profit?**

**Ans.** Minneapolis, Profit: 4630

1. SELECT
2. o.shipping\_city,
3. SUM(od.order\_profits) AS maximum\_profits
4. FROM orders AS o
5. LEFT JOIN order\_details AS od
6. ON o.order\_id = od.order\_id
7. WHERE EXTRACT(year FROM o.order\_date) = 2015
8. GROUP BY 1
9. ORDER BY 2 DESC;

**---------------------------------------------------------------------------------**

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

**Ans**. 531

SELECT COUNT(DISTINCT shipping\_city) FROM  orders;

**---------------------------------------------------------------------------------**

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

**Ans**. Customer\_id: 456 & 738, total\_spent:5

1. SELECT o.customer\_id, SUM(od.order\_sales) AS total\_spent
2. FROM orders AS o
3. LEFT JOIN order\_details AS od
4. ON o.order\_id = od.order\_id
5. GROUP BY 1
6. ORDER BY 2;

**---------------------------------------------------------------------------------**

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

**Ans.** Lebanon, Profit: 83

SELECT

    o.shipping\_city,

    SUM(od.order\_profits) AS maximum\_profits

 FROM orders AS o

 LEFT JOIN order\_details AS od

 ON o.order\_id = od.order\_id

 WHERE o.shipping\_state = 'Tennessee'

 GROUP BY 1

 ORDER BY 2 DESC;

**---------------------------------------------------------------------------------**

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

**Ans.** Lebanon, Avg\_annual\_Profit: 27.66

SELECT

    o.shipping\_city,

    AVG(od.order\_profits) AS AVG\_profits

FROM orders AS o

LEFT JOIN order\_details AS od

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

WHERE o.shipping\_city = ‘Lebanon’

 GROUP BY 1

 ORDER BY 2 DESC;

**---------------------------------------------------------------------------------**

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

**Ans.**

|  |  |
| --- | --- |
| **CUSTOMER SEGMENT** | **DISTRIBUTION** |
| Consumer | 410 |
| Corporate | 237 |
| Home Office | 148 |

SELECT customer\_segment,

COUNT(DISTINCT customers) AS distribution

FROM customers

GROUP BY customer\_segment;

**---------------------------------------------------------------------------------**

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

**Ans.**

|  |  |
| --- | --- |
| Product Category | Avg Profits |
| furniture | 130.25 |
| technology | 79.75 |
| office supplies | 15.72 |

 SELECT

    p.product\_category AS product\_category,

    AVG(od.order\_profits) AS AVG\_profits

 FROM orders AS o

 LEFT JOIN order\_details AS od

 ON o.order\_id = od.order\_id

 LEFT JOIN product AS p

 ON p.product\_id = od.product\_id

 WHERE o.shipping\_state = 'Iowa'

 GROUP BY 1

 ORDER BY 2 DESC;

**---------------------------------------------------------------------------------**

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

**Ans.** Produtc\_name:Global Push Button Manager's Chair, Indigo

Product\_category: Furniture

Avg\_Number\_of\_items: 22

 SELECT

    p.product\_name AS product\_name,

    --p.product\_category AS product\_category,

    SUM(od.quantity) AS num\_of\_items

 FROM orders AS o

 LEFT JOIN order\_details AS od

 ON o.order\_id = od.order\_id

 LEFT JOIN product AS p

 ON p.product\_id = od.product\_id

 WHERE EXTRACT(year FROM o.order\_date) = 2016

 AND p.product\_category = 'Furniture'

 GROUP BY 1

 ORDER BY 2 DESC;**---------------------------------------------------------------------------**

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

**Ans.** Customer\_name: Sean Miller, Total\_discount:11769.7

Customer\_id:687, Order\_id : CA-2015-145317

 SELECT c.customer\_id AS customer\_id,

        c.customer\_name AS customer\_name,

        o.order\_id,

        SUM(od.order\_discount \* od.order\_sales) AS total\_discount

FROM customers AS c

INNER JOIN orders AS o

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

INNER JOIN order\_details AS od

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

GROUP BY 1, 2, 3

ORDER BY 4 DESC;

**---------------------------------------------------------------------------------**

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

**Ans.** Month: 02-2018, Month\_total:7137, Month\_diff:-5525

WITH month\_totals AS (

    SELECT to\_char(order\_date, 'MM-YYYY') AS month

        ,SUM(od.order\_profits) AS month\_total

FROM orders AS o

INNER JOIN order\_details AS od

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

WHERE EXTRACT(year FROM o.order\_date) = 2018

GROUP BY 1

)

SELECT month,

        month\_total,

        --LAG(month\_total,1) OVER (

        --ORDER BY month

    --) previous\_month\_sales,

    COALESCE(month\_total - LAG(month\_total,1) OVER (

        ORDER BY month), month\_total) AS moth\_diff

FROM month\_totals;

**---------------------------------------------------------------------------------**13. **Which order was the highest in 2015?**

**Ans.** Order\_id : CA-2015-145317

SELECT od.order\_id, SUM(od.order\_sales) AS highest\_sales

FROM order\_details AS od

INNER JOIN orders AS o

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

WHERE EXTRACT(year FROM o.order\_date) = 2015

GROUP BY 1

ORDER BY 2 DESC;

**---------------------------------------------------------------------------**

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

**Ans.** City: Columbus

SELECT o.shipping\_city AS city,

        SUM(od.quantity) AS amount\_of\_orders,

        rank() OVER (order by SUM(od.quantity) desc) AS rank

FROM orders AS o

LEFT JOIN order\_details AS od

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

WHERE o.shipping\_region = 'East'

AND EXTRACT(year FROM o.order\_date) = 2015

GROUP BY 1

ORDER BY 3;

**---------------------------------------------------------------------------**

**15.** **Display customer names for customers who are in the segment ‘Consumer’ or ‘Corporate.’ How many customers are there in total?**

**Ans.** 647

SELECT COUNT(\*) AS total\_customers

FROM customers c

WHERE customer\_segment IN ('Consumer','Corporate')

**---------------------------------------------------------------------------**

**16.** **Calculate the difference between the largest and smallest order quantities for product id ‘100.’**

**Ans.** 4

SELECT MAX(od.quantity)-MIN(od.quantity) AS diff

FROM order\_details AS od

WHERE od.product\_id = '100'

**---------------------------------------------------------------------------**

**17. Calculate the percent of products that have ‘Furniture’ as their category.**

**Ans.** 20.54%

SELECT (SELECT count(\*) from product WHERE product\_category = 'Furniture') AS furniture\_count,

        (SELECT count(\*) from product) AS total\_count,

        ((SELECT count(\*) from product WHERE product\_category = 'Furniture') \* 100/ (SELECT count(\*) from product) :: float ) AS percent

FROM product p

GROUP BY 1;

**---------------------------------------------------------------------------**

**18. Display the number of duplicate products based on their product manufacturer.**

**Ans.** 8

SELECT DISTINCT p.product\_manufacturer,

COUNT(p.product\_id) AS dupliate\_items

FROM product p

GROUP BY 1

ORDER BY 2 DESC;

**---------------------------------------------------------------------------**

**19. Show the product\_subcategory and the total number of products in the subcategory. Show the order from most to least products and then by product\_subcategory name ascending.**

**Ans.** Paper

SELECT product\_subcategory, COUNT(product\_subcategory) AS total\_products

FROM product

GROUP BY 1

ORDER BY 2 DESC, 1;

**---------------------------------------------------------------------------**

**20.Show the product\_id(s), the sum of quantities, where for each sale of product quantities is greater than or equal to 100.**

**Ans.** 132

SELECT p.product\_id, SUM(od.quantity) AS total\_quantity

FROM product p

INNER JOIN order\_details od

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

GROUP BY p.product\_id

HAVING SUM(od.quantity) >= 100