GLOBAL ELECTRONICS RETAILER

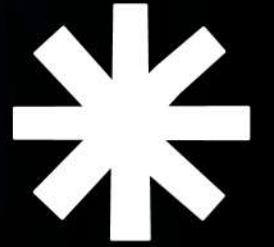# SQL QUERIES

PROJECT BACKGROUND

Global Electronic Retailer which runs a worldwide company to sell different electronics products for its customers. Retailer time-series data for Global Electronic were provided including details data from transactions, customer demographics, and store performance.

By integrating advanced SQL analysis with dynamic Power BI reporting, the project aimed to identify key growth opportunities, optimize resource allocation, and enhance customer satisfaction.

PROBLEM 1:

FIND THE TOTAL CUSTOMERS, QUANTITY, SALES, STORES AND PRODUCTS?

```
SELECT
    COUNT(DISTINCT C.CUSTOMER_ID) AS TOTAL_CUSTOMERS
    , COUNT(DISTINCT ST.STORE_ID) AS TOTAL_STORES
    , COUNT(DISTINCT P.PRODUCT_ID) AS TOTAL_PRODUCTS
    , SUM(S.QUANTITY) AS TOTAL_QUANTITY
    , ROUND(SUM(S.QUANTITY * P.UNIT_PRICE_USD * E.UNITS_PER_USD), 2) AS
TOTAL_SALES
    , ROUND(SUM(S.QUANTITY * P.UNIT_COST_USD * E.UNITS_PER_USD), 2) AS
TOTAL_MANUFACTURING_COST
    , ROUND(SUM(S.QUANTITY * P.UNIT_PRICE_USD * E.UNITS_PER_USD), 2) -
ROUND(SUM(S.QUANTITY * P.UNIT_COST_USD * E.UNITS_PER_USD), 2) AS PROFIT
FROM SALES AS S
FULL JOIN PRODUCTS AS P USING(PRODUCT_ID)
FULL JOIN EXCHANGE_RATES AS E
    ON S.ORDER_DATE = E.DATE AND S.CURRENCY = E.CURRENCY
FULL JOIN STORES AS ST USING(STORE_ID)
FULL JOIN CUSTOMERS AS C USING(CUSTOMER_ID)
```
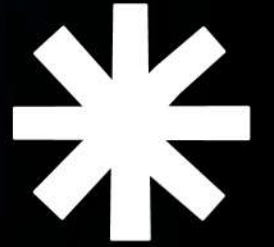
total_customers ▼	total_stores ▼	total_products ▼	total_quantity ▼
15266	67	2517	197757

total_sales ▼	total_manufacturing	profit ▼
55349466.1	22923340.47	32426125.63

PROBLEM 2:

WHAT IS THE DISTRIBUTION OF CUSTOMERS ACROSS DIFFERENT
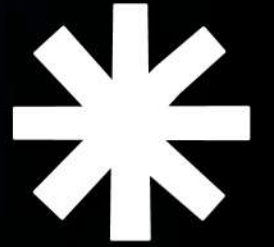CONTINENTS, AND HOW DOES IT IMPACT SALES VOLUME?

```
SELECT
    C.CONTINENT
    , COUNT(DISTINCT C.CUSTOMER_ID) AS TOTAL_CUSTOMERS
    , SUM(S.QUANTITY) AS TOTAL_QUANTITY
    , ROUND(SUM(S.QUANTITY * P.UNIT_PRICE_USD * E.UNITS_PER_USD), 2) AS
TOTAL_SALES
    , ROUND(SUM(S.QUANTITY * P.UNIT_PRICE_USD * E.UNITS_PER_USD) / 1000000, 2)
AS TOTAL_SALES_MIL
FROM SALES AS S
FULL JOIN CUSTOMERS AS C USING(CUSTOMER_ID)
FULL JOIN PRODUCTS AS P USING(PRODUCT_ID)
FULL JOIN EXCHANGE_RATES AS E
    ON S.ORDER_DATE = E.DATE AND S.CURRENCY = E.CURRENCY
GROUP BY 1
ORDER BY 4 DESC
```

Row	continent	total_customers	total_quantity	total_sales	total_sales_mil
1	North America	8381	123200	36088001.5	36.09
2	Europe	5465	65348	15481748.82	15.48
3	Australia	1420	9209	3779715.78	3.78

PROBLEM 3:

HOW DOES CUSTOMER AGE GROUP INFLUENCE PURCHASING BEHAVIOR, AND WHICH PRODUCTS ARE MOST POPULAR WITHIN EACH GROUP?
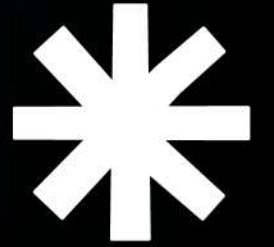
```
SELECT
    C.AGE_GROUP AS AGE_GROUP
    , P.PRODUCT_NAME AS PRODUCT_NAME
    , COUNT(S.ORDER_ID) AS TOTAL_ORDERS
    , SUM(S.QUANTITY) AS TOTAL_QUANTITY
    , ROUND(SUM(S.QUANTITY * P.UNIT_PRICE_USD * E.UNITS_PER_USD), 2) AS
TOTAL_SALES
    , ROUND(SUM(S.QUANTITY * P.UNIT_PRICE_USD * E.UNITS_PER_USD) / 1000000, 3) AS
TOTAL_SALES_MIL
FROM SALES AS S
JOIN PRODUCTS AS P USING(PRODUCT_ID)
JOIN CUSTOMERS AS C USING(CUSTOMER_ID)
JOIN EXCHANGE_RATES AS E
    ON S.ORDER_DATE = E.DATE AND S.CURRENCY = E.CURRENCY
GROUP BY 1,2
ORDER BY 5 DESC
```

Row	age_group	product_name	total_orders	total_quantity	total_sales	total_sales_mil
1	64 - ~	WWI Desktop PC2.33 X2330 Bl...	55	198	182106.28	0.182
2	64 - ~	Adventure Works Desktop PC2....	53	177	174486.57	0.174
3	64 - ~	WWI Desktop PC2.33 X2330 Br...	57	181	166403.51	0.166
4	50 - 64	Adventure Works Desktop PC2....	40	171	162413.7	0.162
5	64 - ~	Adventure Works 52" LCD HDT...	20	54	160202.7	0.16
6	64 - ~	Adventure Works 52" LCD HDT...	16	55	153687.87	0.154
7	64 - ~	Adventure Works Desktop PC2....	48	156	149431.04	0.149
8	35 - 49	Adventure Works 52" LCD HDT...	12	52	142119.52	0.142
9	64 - ~	WWI Desktop PC2.33 X2330 W...	41	160	141333.65	0.141
10	64 - ~	Adventure Works Desktop PC2....	43	139	136631.62	0.137

PROBLEM 4:

ARE THERE SPECIFIC CITIES OR STATES WITH HIGHER SALES VOLUMES?
WHAT ARE THE CHARACTERISTICS OF THESE REGIONS?
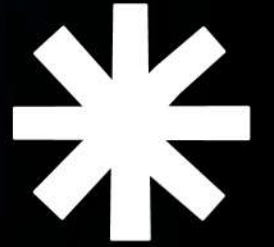
```
SELECT
    C.STATE_NAME
    , C.CITY
    , COUNT(S.ORDER_ID) AS TOTAL_ORDERS
    , ROUND(SUM(S.QUANTITY * P.UNIT_PRICE_USD * E.UNITS_PER_USD), 2) AS
TOTAL_SALES
    , ROUND(SUM(S.QUANTITY * P.UNIT_PRICE_USD * E.UNITS_PER_USD) / 1000000, 3)
AS TOTAL_SALES_MIL
FROM SALES AS S
JOIN PRODUCTS AS P USING(PRODUCT_ID)
JOIN CUSTOMERS AS C USING(CUSTOMER_ID)
JOIN EXCHANGE_RATES AS E
    ON S.ORDER_DATE = E.DATE AND S.CURRENCY = E.CURRENCY
GROUP BY 1,2
ORDER BY 4 DESC
```

Row	state_name	city	total_orders	total_sales	total_sales_mil
1	Ontario	Toronto	681	787868.19	0.788
2	New York	New York	569	492392.45	0.492
3	California	Los Angeles	579	467985.07	0.468
4	Texas	Houston	438	360714.08	0.361
5	Pennsylvania	Philadelphia	357	358600.9	0.359
6	Quebec	Montreal	361	354348.32	0.354
7	Illinois	Chicago	368	349357.12	0.349
8	Alberta	Calgary	267	284825.67	0.285
9	Texas	Dallas	282	283011.1	0.283
10	Ontario	Ottawa	184	272000.36	0.272
11	Georgia	Atlanta	248	241892.74	0.242
12	British Columbia	Vancouver	182	238417.86	0.238

PROBLEM 5:

HOW DOES THE GENDER OF CUSTOMERS INFLUENCE PRODUCT PREFERENCES AND PURCHASING FREQUENCY?
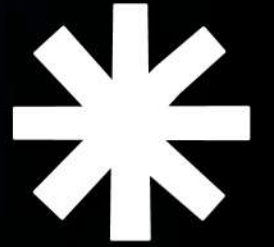
```
SELECT
    C.GENDER
    , P.PRODUCT_NAME
    , COUNT(S.ORDER_ID) AS TOTAL_ORDERS
    , SUM(S.QUANTITY) AS TOTAL_QUATITY
    , ROUND(SUM(S.QUANTITY * P.UNIT_PRICE_USD * E.UNITS_PER_USD), 2) AS
TOTAL_SALES
    , ROUND(SUM(S.QUANTITY * P.UNIT_PRICE_USD * E.UNITS_PER_USD) / 1000, 2) AS
TOTAL_SALES_K
FROM SALES AS S
JOIN CUSTOMERS AS C USING (CUSTOMER_ID)
JOIN PRODUCTS AS P USING (PRODUCT_ID)
JOIN EXCHANGE_RATES AS E
    ON S.ORDER_DATE = E.DATE AND S.CURRENCY = E.CURRENCY
GROUP BY 1,2
ORDER BY 6 DESC
```

Row	gender	product_name	total_orders	total_quatity	total_sales	total_sales_k
1	Male	WWI Desktop PC2.33 X2330 Bl...	85	318	291102.44	291.1
2	Male	Adventure Works Desktop PC2....	83	263	255683.54	255.68
3	Male	WWI Desktop PC2.33 X2330 W...	79	280	253878.99	253.88
4	Female	Adventure Works Desktop PC2...	75	265	248779.41	248.78
5	Male	WWI Desktop PC2.33 X2330 Br...	84	266	246809.95	246.81
6	Female	Adventure Works Desktop PC2...	77	254	244428.02	244.43
7	Female	Adventure Works Desktop PC2...	72	242	229226.83	229.23
8	Male	Adventure Works 52" LCD HDT...	27	77	219926.83	219.93
9	Male	Adventure Works Desktop PC2...	69	227	213472.54	213.47
10	Male	Adventure Works Desktop PC2....	69	214	210511.96	210.51

PROBLEM 6:

WHAT IS THE CORRELATION BETWEEN CUSTOMER LOCATION (CITY/STATE)
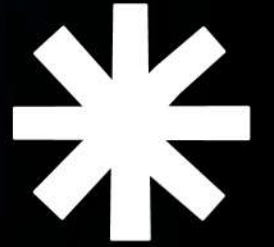AND THE AVERAGE ORDER VALUE?

```
SELECT
    C.STATE_NAME
    , C.CITY
    , ROUND(AVG(S.QUANTITY * P.UNIT_PRICE_USD * E.UNITS_PER_USD), 2) AS
AVG_TOTAL_SALES
FROM SALES AS S
JOIN CUSTOMERS AS C USING(CUSTOMER_ID)
JOIN PRODUCTS AS P USING(PRODUCT_ID)
JOIN EXCHANGE_RATES AS E
    ON S.ORDER_DATE = E.DATE AND S.CURRENCY = E.CURRENCY
GROUP BY 1,2
ORDER BY 3 DESC
```

Row	state_name	city	avg_total_sales
1	South Carolina	North Charleston	19199.94
2	Tasmania	LIFFEY	16374.82
3	Braintree	HALSTEAD	15249.87
4	Victoria	DONCASTER EAST	12766.47
5	Ontario	Blezard Valley	12215.14
6	Western Australia	BASKERVILLE	10544.4
7	British Columbia	Duncan	10453.34
8	Oklahoma	Ada	10148.5
9	New South Wales	MARRICKVILLE SOUTH	9693.71
10	Texas	Bon Wier	9667.66

PROBLEM 7:

ARE THERE ANY CORRELATIONS BETWEEN CUSTOMERS' BIRTH MONTH AND THEIR PURCHASING BEHAVIOR?

```
SELECT
    EXTRACT(MONTH FROM C.BIRTHDAY) AS MONTH
    , COUNT(S.ORDER_ID) AS TOTAL_ORDERS
    , ROUND(SUM(S.QUANTITY * P.UNIT_PRICE_USD * E.UNITS_PER_USD), 2) AS
TOTAL_SALES
    , ROUND(SUM(S.QUANTITY * P.UNIT_PRICE_USD * E.UNITS_PER_USD) / 1000000, 2)
AS TOTAL_SALES_MIL
FROM SALES AS S
JOIN CUSTOMERS AS C USING(CUSTOMER_ID)
JOIN PRODUCTS AS P USING(PRODUCT_ID)
JOIN EXCHANGE_RATES AS E
    ON S.ORDER_DATE = E.DATE AND S.CURRENCY = E.CURRENCY
GROUP BY 1
ORDER BY 1
```

Row	month	total_orders	total_sales	total_sales_mil
1	1	5255	4639146.83	4.64
2	2	5297	4690350.77	4.69
3	3	5111	4486568.82	4.49
4	4	5187	4505951.14	4.51
5	5	5706	4869352.54	4.87
6	6	5030	4471833.32	4.47
7	7	5371	4845699.61	4.85
8	8	5202	4612751.99	4.61
9	9	5074	4558621.83	4.56
10	10	5192	4579593.46	4.58

PROBLEM 8:

HOW DOES THE SALES VOLUME VARY THROUGHOUT THE YEAR ?
WHAT ARE THE PEAK MONTHS FOR DIFFERENT PRODUCT CATEGORIES?
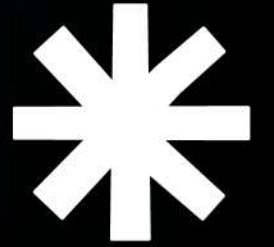
```
SELECT
    FORMAT_DATE('%Y', ORDER_DATE) AS YEAR
    , EXTRACT(MONTH FROM ORDER_DATE) AS MONTH
    , P.CATEGORY
    , SUM(QUANTITY) AS TOTAL_QUANTITY
FROM SALES AS S
JOIN PRODUCTS AS P USING(PRODUCT_ID)
GROUP BY 1,2,3
ORDER BY 1,2,4 DESC
```

Row	year	month	category	total_quantity
1	2016	1	Music, Movies and Audio Books	426
2	2016	1	Home Appliances	335
3	2016	1	Cameras and camcorders	310
4	2016	1	Computers	271
5	2016	1	TV and Video	248
6	2016	1	Games and Toys	188
7	2016	1	Audio	186
8	2016	1	Cell phones	152
9	2016	2	Music, Movies and Audio Books	464
10	2016	2	Home Appliances	428
11	2016	2	Computers	382
12	2016	2	Cameras and camcorders	341

PROBLEM 9:

WHICH STORE LOCATIONS ARE THE MOST PROFITABLE,
AND HOW DOES THEIR PERFORMANCE COMPARE TO OTHERS?

```
SELECT
    ST.STORE_ID
    , ST.COUNTRY
    , ST.STATE_NAME
    , ROUND(SUM(S.QUANTITY * P.UNIT_PRICE_USD * E.UNITS_PER_USD), 2) AS
TOTAL_SALES
    , ROUND(SUM(S.QUANTITY * P.UNIT_PRICE_USD * E.UNITS_PER_USD) / 1000000, 2)
AS TOTAL_SALES_MIL
FROM SALES AS S
JOIN STORES AS ST USING(STORE_ID)
JOIN PRODUCTS AS P USING(PRODUCT_ID)
JOIN EXCHANGE_RATES AS E ON S.ORDER_DATE = E.DATE AND S.CURRENCY =
E.CURRENCY
GROUP BY 1,2,3
ORDER BY 4 DESC
```

Row	store_id	country	state_name	total_sales	total_sales_mil
1	0	Online	Online	11431369.95	11.43
2	9	Canada	Northwest Territories	1756645.28	1.76
3	8	Canada	Newfoundland and Labrador	1617620.45	1.62
4	55	United States	Nevada	1417885.41	1.42
5	50	United States	Kansas	1394738.06	1.39
6	54	United States	Nebraska	1384396.24	1.38
7	10	Canada	Nunavut	1377400.64	1.38
8	57	United States	New Mexico	1325611.89	1.33
9	45	United States	Connecticut	1318787.92	1.32
10	61	United States	South Carolina	1305684.43	1.31

PROBLEM 10:

HOW DOES THE SALES VOLUME OF NEW PRODUCTS COMPARE TO OLDER PRODUCTS OVER TIME?
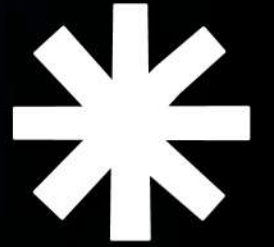
```
,  PRODUCT_LAUNCH_DATE AS (
     SELECT
        S.PRODUCT_ID
        , MIN(S.ORDER_DATE) AS LAUNCH_DATE
     FROM SALES AS S
     GROUP BY 1
     ORDER BY 1
)

SELECT
     EXTRACT(YEAR FROM S.ORDER_DATE) AS CREATED_YEAR
     , CASE
        WHEN DATE_DIFF(S.ORDER_DATE, PLD.LAUNCH_DATE, YEAR) < 1 THEN 'NEW
PRODUCT'
        ELSE 'OLD PRODUCT' END AS PRODUCT_TYPE
     , SUM(S.QUANTITY) AS TOTAL_QUANTITY
FROM PRODUCT_LAUNCH_DATE AS PLD
JOIN SALES AS S USING(PRODUCT_ID)
GROUP BY 1,2
ORDER BY 1,2,3 DESC
```

Row	created_year	product_type	total_quantity
1	2016	New Product	21761
2	2017	New Product	2307
3	2017	Old Product	22491
4	2018	New Product	989
5	2018	Old Product	43509
6	2019	New Product	181
7	2019	Old Product	68259
8	2020	New Product	58
9	2020	Old Product	34405
10	2021	New Product	2
11	2021	Old Product	3795

PROBLEM 11:

HOW DOES SALES PERFORMANCE VARY BY DAY OF
THE WEEK ACROSS DIFFERENT REGIONS?

```
SELECT
    FORMAT_DATE('%A', ORDER_DATE) AS DAY_WEEK
    , ST.COUNTRY
    , ST.STATE_NAME
    , SUM(QUANTITY) AS TOTAL_QUANTITY
    , ROUND(SUM(S.QUANTITY * P.UNIT_PRICE_USD * E.UNITS_PER_USD), 2) AS
TOTAL_SALES
    , ROUND(SUM(S.QUANTITY * P.UNIT_PRICE_USD * E.UNITS_PER_USD) / 1000, 2) AS
TOTAL_SALES_K
FROM SALES AS S
JOIN STORES AS ST USING(STORE_ID)
JOIN PRODUCTS AS P USING(PRODUCT_ID)
JOIN EXCHANGE_RATES AS E ON S.ORDER_DATE = E.DATE AND S.CURRENCY =
E.CURRENCY
GROUP BY 1,2,3
ORDER BY 5 DESC
```

Row	day_week	country	state_name	total_quantity	total_sales	total_sales_k
1	Saturday	Online	Online	9844	2758711.82	2758.71
2	Thursday	Online	Online	7971	2190491.26	2190.49
3	Wednesday	Online	Online	7194	1968564.62	1968.56
4	Friday	Online	Online	5860	1666229.44	1666.23
5	Tuesday	Online	Online	5590	1557821.22	1557.82
6	Monday	Online	Online	4164	1122328.21	1122.33
7	Saturday	Canada	Nunavut	962	394599.11	394.6
8	Thursday	Canada	Northwest Territories	958	381767.18	381.77
9	Saturday	United States	Wyoming	1170	362903.91	362.9
10	Saturday	Canada	Northwest Territories	1079	349340.92	349.34

PROBLEM 12:

ARE THERE ANY SIGNIFICANT DIFFERENCES IN SALES
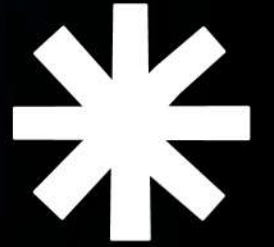PERFORMANCE FOR DIFFERENT STORE SIZES?

```
SELECT
    ST.STORE_ID
    , ST.STATE_NAME
    , ST.STORE_SIZE_M2
    , ROUND(SUM(S.QUANTITY * P.UNIT_PRICE_USD * E.UNITS_PER_USD), 2) AS
TOTAL_SALES
    , ROUND(SUM(S.QUANTITY * P.UNIT_PRICE_USD * E.UNITS_PER_USD) / 1000, 2) AS
TOTAL_SALES_K
FROM SALES AS S
JOIN STORES AS ST USING(STORE_ID)
JOIN PRODUCTS AS P USING(PRODUCT_ID)
JOIN EXCHANGE_RATES AS E ON S.ORDER_DATE = E.DATE AND S.CURRENCY =
E.CURRENCY
GROUP BY 1,2,3
ORDER BY 5 DESC
```

Row	store_id	state_name	store_size_m2	total_sales	total_sales_k
1	0	Online	0 - 500	11431369.95	11431.37
2	9	Northwest Territories	1000 - 1500	1756645.28	1756.65
3	8	Newfoundland and Labrador	1500 - ~	1617620.45	1617.62
4	55	Nevada	1500 - ~	1417885.41	1417.89
5	50	Kansas	1500 - ~	1394738.06	1394.74
6	54	Nebraska	1500 - ~	1384396.24	1384.4
7	10	Nunavut	1000 - 1500	1377400.64	1377.4
8	57	New Mexico	1500 - ~	1325611.89	1325.61
9	45	Connecticut	1500 - ~	1318787.92	1318.79
10	61	South Carolina	1500 - ~	1305684.43	1305.68

PROBLEM 13:

HOW DO SALES VOLUMES VARY BETWEEN ONLINE
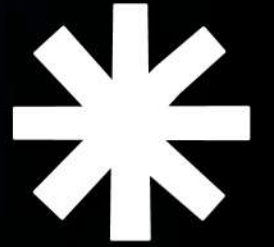AND PHYSICAL STORE PURCHASES?

```
SELECT
    S.PLATFORM
    , SUM(QUANTITY) AS TOTAL_QUANTITY
    , ROUND(SUM(S.QUANTITY * P.UNIT_PRICE_USD * E.UNITS_PER_USD), 2) AS
TOTAL_SALES
    , ROUND(SUM(S.QUANTITY * P.UNIT_PRICE_USD * E.UNITS_PER_USD) / 1000, 2) AS
TOTAL_SALES_K
FROM SALES AS S
JOIN PRODUCTS AS P USING(PRODUCT_ID)
JOIN EXCHANGE_RATES AS E ON S.ORDER_DATE = E.DATE AND S.CURRENCY =
E.CURRENCY
GROUP BY 1
ORDER BY 4 DESC
```

Row	platform	total_quantity	total_sales	total_sales_k
1	Physical_Purchase	156446	43918096.16	43918.1
2	Online_Purchase	41311	11431369.95	11431.37

15

PROBLEM 14:

WHAT ARE THE TOTAL SALES VALUES PER CUSTOMER, AND
HOW DOES EACH CUSTOMER'S SALES COMPARE TO THE HIGHEST SPENDER?
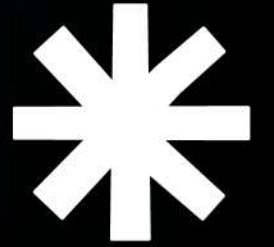
```
SELECT
    C.CUSTOMER_ID
    , ROUND(SUM(S.QUANTITY * P.UNIT_PRICE_USD * E.UNITS_PER_USD), 2) AS
TOTAL_SALES
    , ROUND(SUM(S.QUANTITY * P.UNIT_PRICE_USD * E.UNITS_PER_USD) / 1000, 2) AS
TOTAL_SALES_K
    , RANK() OVER(ORDER BY ROUND(SUM(S.QUANTITY * P.UNIT_PRICE_USD *
E.UNITS_PER_USD) / 1000, 2) DESC) AS RANK
FROM SALES AS S
JOIN CUSTOMERS AS C USING(CUSTOMER_ID)
JOIN PRODUCTS AS P USING(PRODUCT_ID)
JOIN EXCHANGE_RATES AS E ON S.ORDER_DATE = E.DATE AND S.CURRENCY =
E.CURRENCY
GROUP BY 1
ORDER BY 4
```

Row	customer_id	total_sales	total_sales_k	rank
1	1702221	61871.7	61.87	1
2	262871	49704.93	49.7	2
3	1884663	43517.8	43.52	3
4	1969704	42788.04	42.79	4
5	149955	42395.42	42.4	5
6	220519	42020.31	42.02	6
7	396944	40377.22	40.38	7
8	70912	38607.6	38.61	8
9	143033	38299.41	38.3	9
10	1928466	38191.06	38.19	10
11	261356	37367.99	37.37	11

PROBLEM 15:

WHAT IS THE SALES TREND FOR EACH MONTH OVER THE YEARS, AND HOW DOES EACH MONTH'S SALES COMPARE TO THE PREVIOUS MONTH, AND WHAT IS THE PERCENTAGE CHANGE MONTH-WISE?

```
, YEARLY_MONTHLY_SALES AS (
    SELECT
        FORMAT_DATE('%Y', ORDER_DATE) AS YEARS
        , EXTRACT(MONTH FROM ORDER_DATE) AS MONTHS
        , ROUND(SUM(S.QUANTITY * P.UNIT_PRICE_USD * E.UNITS_PER_USD), 2) AS
TOTAL_SALES
    FROM SALES AS S
    JOIN CUSTOMERS AS C USING(CUSTOMER_ID)
    JOIN PRODUCTS AS P USING(PRODUCT_ID)
    JOIN EXCHANGE_RATES AS E ON S.ORDER_DATE = E.DATE AND S.CURRENCY =
E.CURRENCY
    GROUP BY 1,2
    ORDER BY 1,2
)

, CHANGE_BY_YEAR_MONTH AS (
    SELECT
        *
        , LAG(TOTAL_SALES) OVER(ORDER BY YEARS, MONTHS) AS
PREVIOUS_MONTH_SALE
    FROM YEARLY_MONTHLY_SALES
)

SELECT
    *
    , CONCAT(ROUND(((TOTAL_SALES - PREVIOUS_MONTH_SALE) /
PREVIOUS_MONTH_SALE)*100,2), ' %') AS CHANGE_PERCENTAGE
FROM CHANGE_BY_YEAR_MONTH
ORDER BY YEARS, MONTHS
```
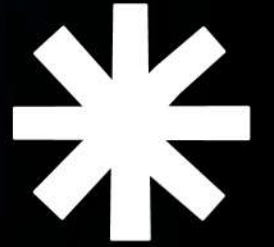
Row	years	months	total_sales	previous_month_sale	change_percentage
1	2016	1	645982.1	null	null
2	2016	2	885621.15	645982.1	37.1 %
3	2016	3	332996.94	885621.15	-62.4 %
4	2016	4	111698.41	332996.94	-66.46 %
5	2016	5	566862.2	111698.41	407.49 %
6	2016	6	502270.71	566862.2	-11.39 %
7	2016	7	437780.7	502270.71	-12.84 %
8	2016	8	534085.34	437780.7	22 %

PROBLEM 16:

WHICH PRODUCTS ARE THE TOP SELLERS BY QUANTITY, AND
HOW MUCH MORE ARE THEY SELLING COMPARED TO OTHER PRODUCTS?

```
SELECT
    P.PRODUCT_ID
    , P.PRODUCT_NAME
    , SUM(QUANTITY) AS TOTAL_QUANTITY
    , DENSE_RANK() OVER(ORDER BY SUM(QUANTITY) DESC) AS PRODUCT_RANK
FROM SALES AS S
JOIN PRODUCTS AS P USING(PRODUCT_ID)
GROUP BY 1, 2
ORDER BY 4
```

Row	product_id	product_name	total_quantity	product_rank
1	444	WWI Desktop PC2.33 X2330 Bl...	550	1
2	458	WWI Desktop PC1.80 E1800 W...	538	2
3	424	Adventure Works Desktop PC1....	521	3
4	434	Adventure Works Desktop PC2....	521	3
5	425	Adventure Works Desktop PC1....	520	4
6	423	Adventure Works Desktop PC2....	514	5
7	446	WWI Desktop PC1.60 E1600 Bl...	509	6
8	440	WWI Desktop PC1.60 E1600 Sil...	507	7
9	448	WWI Desktop PC1.80 E1801 Bl...	505	8
10	435	Adventure Works Desktop PC1....	505	8
11	452	WWI Desktop PC1.60 E1600 Red	505	8

PROBLEM 17:

WHAT IS THE TOTAL SALES PER STORE, AND HOW DOES
EACH STORE'S PERFORMANCE COMPARE TO THE AVERAGE STORE PERFORMANCE?
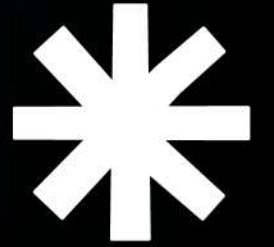
```
SELECT
    ST.STORE_ID
    , ROUND(SUM(S.QUANTITY * P.UNIT_PRICE_USD * E.UNITS_PER_USD), 2) AS
TOTAL_SALES
    , ROUND(AVG(SUM(S.QUANTITY * P.UNIT_PRICE_USD * E.UNITS_PER_USD)) OVER(),
2) AS AVG_SALES
    , ROUND(SUM(S.QUANTITY * P.UNIT_PRICE_USD * E.UNITS_PER_USD) –
AVG(SUM(S.QUANTITY * P.UNIT_PRICE_USD * E.UNITS_PER_USD)) OVER(), 2) AS
SALE_DIFFERENCE
FROM SALES AS S
JOIN STORES AS ST USING(STORE_ID)
JOIN PRODUCTS AS P USING(PRODUCT_ID)
JOIN EXCHANGE_RATES AS E ON S.ORDER_DATE = E.DATE AND S.CURRENCY =
E.CURRENCY
GROUP BY 1
ORDER BY 1
```

Row	store_id	total_sales	avg_sales	sale_difference
1	0	11431369.95	954301.14	10477068.81
2	1	342086.5	954301.14	-612214.64
3	2	20843.53	954301.14	-933457.61
4	4	615054.84	954301.14	-339246.3
5	5	1197946.8	954301.14	243645.67
6	6	749020.48	954301.14	-205280.66
7	8	1617620.45	954301.14	663319.31
8	9	1756645.28	954301.14	802344.14
9	10	1377400.64	954301.14	423099.5
10	12	159755.71	954301.14	-794545.43

PROBLEM 18:

WHAT IS THE LIFETIME VALUE OF EACH CUSTOMER, AND
WHO ARE THE TOP 20 CUSTOMERS BY SALES IN EACH STATE?

```
, BY_CUSTOMER AS (
    SELECT
        C.STATE_NAME
        , C.CUSTOMER_ID
        , C.CUSTOMER_NAME
        , ROUND(SUM(S.QUANTITY * P.UNIT_PRICE_USD * E.UNITS_PER_USD), 2) AS TOTAL_SALES
    FROM SALES AS S
    JOIN CUSTOMERS AS C USING(CUSTOMER_ID)
    JOIN PRODUCTS AS P USING(PRODUCT_ID)
    JOIN EXCHANGE_RATES AS E ON S.ORDER_DATE = E.DATE AND S.CURRENCY = E.CURRENCY
    GROUP BY 1, 2, 3
)

, CALCULATE_RANK_IN_STATE AS(
    SELECT
        *
        , RANK() OVER(PARTITION BY STATE_NAME ORDER BY TOTAL_SALES DESC) AS
RANK_IN_STATE
    FROM BY_CUSTOMER
    ORDER BY 1,4 DESC
)

SELECT
    *
FROM CALCULATE_RANK_IN_STATE
WHERE RANK_IN_STATE <= 20
```
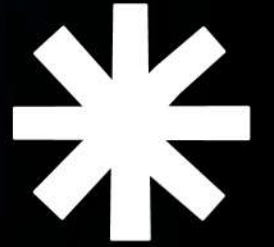
Row	state_name	customer_id	customer_name	total_sales	rank_in_state
1	Aberdeen	1031047	Louis George	6075.15	1
2	Aberdeen	1022980	Jake Patel	3546.04	2
3	Aberdeen	950923	Molly Poole	2946.79	3
4	Aberdeen	1031814	Charlie Davies	994.24	4
5	Aberdeen	975403	Spencer Wells	897.77	5
6	Aberdeen	958207	Maya Barker	611.63	6
7	Aberdeenshire	1066452	Andrew Webb	10821.19	1
8	Aberdeenshire	924256	Freddie Alexander	6793.11	2
9	Aberdeenshire	918741	Keira Grant	5470.58	3
10	Aberdeenshire	1086276	Courtney Dyer	5261.22	4
11	Aberdeenshire	1125474	Tia Hammond	4893.61	5
12	Aberdeenshire	1071326	Alexander Page	4766.03	6

PROBLEM 19:

WHAT IS THE YEAR-OVER-YEAR SALES GROWTH FOR EACH STORE?

```
, STORE_OVER_YEAR AS (
    SELECT
        ST.STORE_ID
        , FORMAT_DATE('%Y', ORDER_DATE) AS YEARS
        , ROUND(SUM(S.QUANTITY * P.UNIT_PRICE_USD * E.UNITS_PER_USD), 2) AS TOTAL_SALES
    FROM SALES AS S
    JOIN STORES AS ST USING(STORE_ID)
    JOIN PRODUCTS AS P USING(PRODUCT_ID)
    JOIN EXCHANGE_RATES AS E ON S.ORDER_DATE = E.DATE AND S.CURRENCY = E.CURRENCY
    GROUP BY 1,2
    ORDER BY 1,2
)

, CALCULATE_LAST_YEAR AS (
    SELECT
        *
        , LAG(TOTAL_SALES) OVER (PARTITION BY STORE_ID ORDER BY YEARS) AS PREVIOUS_YEAR
    FROM STORE_OVER_YEAR
)

SELECT
    *
    , CONCAT(ROUND(((TOTAL_SALES - PREVIOUS_YEAR) / PREVIOUS_YEAR)*100,2), ' %') AS
YOY_CHANGE
FROM CALCULATE_LAST_YEAR
ORDER BY STORE_ID
```
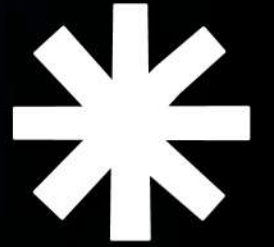
Row	store_id	years	total_sales	previous_year	yoy_change
1	0	2016	1144209.25	null	null
2	0	2017	1395704.73	1144209.25	21.98 %
3	0	2018	2537524.36	1395704.73	81.81 %
4	0	2019	3970975.87	2537524.36	56.49 %
5	0	2020	2099054.22	3970975.87	-47.14 %
6	0	2021	283901.52	2099054.22	-86.47 %
7	1	2016	22139.33	null	null
8	1	2017	32443.52	22139.33	46.54 %
9	1	2018	74407.83	32443.52	129.35 %
10	1	2019	120500.86	74407.83	61.95 %
11	1	2020	86642.03	120500.86	-28.1 %

PROBLEM 20:

WHAT IS THE YEAR-OVER-YEAR SALES GROWTH FOR EACH STORE?

```sql
, STORE_OVER_YEAR AS (
    SELECT
        ST.STORE_ID
        , FORMAT_DATE('%Y', ORDER_DATE) AS YEARS
        , ROUND(SUM(S.QUANTITY * P.UNIT_PRICE_USD * E.UNITS_PER_USD), 2) AS
TOTAL_SALES
    FROM SALES AS S
    JOIN STORES AS ST USING(STORE_ID)
    JOIN PRODUCTS AS P USING(PRODUCT_ID)
    JOIN EXCHANGE_RATES AS E ON S.ORDER_DATE = E.DATE AND S.CURRENCY =
E.CURRENCY
    GROUP BY 1,2
    ORDER BY 1,2
)

, CALCULATE_LAST_YEAR AS (
    SELECT
        *
        , LAG(TOTAL_SALES) OVER (PARTITION BY STORE_ID ORDER BY YEARS) AS
PREVIOUS_YEAR
    FROM STORE_OVER_YEAR
)

SELECT
    *
    , CONCAT(ROUND(((TOTAL_SALES - PREVIOUS_YEAR) / PREVIOUS_YEAR)*100,2), ' %')
AS YOY_CHANGE
FROM CALCULATE_LAST_YEAR
ORDER BY STORE_ID
```

ow	store_id	years	total_sales	previous_year	yoy_change
1	0	2016	1144209.25	null	null
2	0	2017	1395704.73	1144209.25	21.98 %
3	0	2018	2537524.36	1395704.73	81.81 %
4	0	2019	3970975.87	2537524.36	56.49 %
5	0	2020	2099054.22	3970975.87	-47.14 %
6	0	2021	283901.52	2099054.22	-86.47 %
7	1	2016	22139.33	null	null
8	1	2017	32443.52	22139.33	46.54 %
9	1	2018	74407.83	32443.52	129.35 %
10	1	2019	120500.86	74407.83	61.95 %
11	1	2020	86642.03	120500.86	-28.1 %

PROBLEM 21:

WHAT IS THE TOTAL SALES FOR EACH PRODUCT CATEGORY,
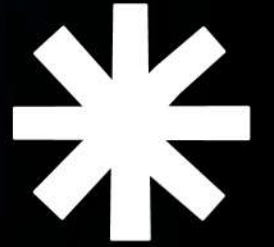AND HOW DO DIFFERENT CATEGORIES COMPARE TO EACH OTHER?

```
SELECT
    P.CATEGORY
    , ROUND(SUM(S.QUANTITY * P.UNIT_PRICE_USD * E.UNITS_PER_USD), 2) AS
TOTAL_SALES
    , DENSE_RANK() OVER(ORDER BY ROUND(SUM(S.QUANTITY * P.UNIT_PRICE_USD *
E.UNITS_PER_USD), 2) DESC) AS CATEGORY_RANK
FROM SALES AS S
JOIN CUSTOMERS AS C USING(CUSTOMER_ID)
JOIN PRODUCTS AS P USING(PRODUCT_ID)
JOIN EXCHANGE_RATES AS E ON S.ORDER_DATE = E.DATE AND S.CURRENCY =
E.CURRENCY
GROUP BY 1
ORDER BY 3
```

Row	category	total_sales	category_rank
1	Computers	19143319.65	1
2	Home Appliances	10747777.59	2
3	Cameras and camcorders	6502487.56	3
4	Cell phones	6118492	4
5	TV and Video	5857337.9	5
6	Audio	3147476.96	6
7	Music, Movies and Audio Books	3113971.62	7
8	Games and Toys	718602.82	8

PROBLEM 22:

WHAT TRENDS CAN BE OBSERVED IN SALES QUANTITY FOR WEEKDAYS VERSUS WEEKENDS OVER TIME?

```
, EXTRACT_DOW AS (
    SELECT
        *
        , FORMAT_DATE('%A', ORDER_DATE) AS DAY_OF_WEEK
    FROM SALES AS S
)

, MAKE_WEEKDAY_WEEKEND AS (
    SELECT
        *
        , CASE
            WHEN DAY_OF_WEEK IN
('MONDAY','TUESDAY','WEDNESDAY','THURSDAY','FRIDAY') THEN 'WEEKDAY'
            WHEN DAY_OF_WEEK IN ('SATURDAY','SUNDAY') THEN 'WEEKEND'
            ELSE 'UNDEFINED' END
            AS WEEKDAY_OR_WEEKEND
    FROM EXTRACT_DOW
)

SELECT
    FORMAT_DATE('%Y %B', ORDER_DATE) AS YEAR_MONTH
    , WEEKDAY_OR_WEEKEND
    , SUM(QUANTITY) AS TOTAL_QUANTITY
FROM MAKE_WEEKDAY_WEEKEND
GROUP BY 1, 2
```

Row	year_month	weekday_or_weekend	total_quantity
1	2016 Feb	Weekday	2045
2	2016 May	Weekday	1278
3	2016 May	Weekend	368
4	2016 Jun	Weekday	1482
5	2016 Jun	Weekend	382
6	2016 Aug	Weekend	304
7	2016 Aug	Weekday	1307
8	2016 Sep	Weekday	1383
9	2016 Sep	Weekend	502
10	2016 Oct	Weekday	1406