Sales for Retail and Food Services in U.S.A.

1)Top-performing industries in terms of sales for a year 2021, and how do their sales compare month-over-month?

```
WITH monthly_sales AS (
   SELECT
      year,
      month,
      industry,
      SUM(sales) AS total_sales
   FROM
      retail_sales
   WHERE
      year = 2021
   GROUP BY
      year,
      month,
      industry
),
top_industries AS (
   SELECT
      year,
      month,
      industry,
      total_sales,
      RANK() OVER (PARTITION BY year, month ORDER BY total_sales DESC) AS
industry_rank
   FROM
      monthly_sales
)
```

```
SELECT
    year,
    month,
    industry,
    total_sales
FROM
    top_industries
WHERE
    industry_rank = 1
ORDER BY
    year,
    month;
```

Data Output Messages Notifications

	year integer	month integer	industry text	total_sales bigint
1	2021	1	Automotive	375492
2	2021	2	Automotive	369166
3	2021	3	Automotive	524244
4	2021	4	Automotive	512810
5	2021	5	Automotive	508814
6	2021	6	Automotive	486902
7	2021	7	Automotive	476584
8	2021	8	Automotive	455964
9	2021	9	Automotive	438496
10	2021	10	Automotive	443780
11	2021	11	Automotive	428848
12	2021	12	Automotive	454888

Total rows: 12 of 12 Query complete 00:00:00.079

2)Top-performing industries in terms of sales for a year 2022, and how do their sales compare month-over-month?

```
WITH monthly_sales AS (
```

```
    SELECT
        year,
        month,
        industry,
        SUM(sales) AS total_sales
    FROM
        retail_sales
    WHERE
        year = 2022
    GROUP BY
        year,
        month,
        industry
),
top_industries AS (
    SELECT
        year,
        month,
        industry,
        total_sales,
        RANK() OVER (PARTITION BY year, month ORDER BY total_sales DESC) AS
industry_rank
    FROM
        monthly_sales
)
SELECT
    year,
    month,
    industry,
    total_sales
FROM
    top_industries
WHERE
    industry_rank = 1
```

```
ORDER BY
    year,
    month;
```

Data Output Messages Notifications

	year integer	month integer	industry text	total_sales bigint
1	2022	1	Automotive	420473
2	2022	2	Automotive	431998
3	2022	3	Automotive	514582
4	2022	4	Automotive	504116
5	2022	5	Automotive	483482
6	2022	6	Automotive	484120
7	2022	7	Automotive	466716
8	2022	8	Automotive	499292
9	2022	9	Automotive	460062
10	2022	10	Automotive	467212
11	2022	11	Automotive	435198
12	2022	12	Automotive	456984

Total rows: 12 of 12 Query complete 00:00:00.098

3)Top-performing industries in terms of sales for a year 2020, and how do their sales compare month-over-month?

```
WITH monthly_sales AS (
    SELECT
        year,
        month,
        industry,
        SUM(sales) AS total_sales
    FROM
        retail_sales
    WHERE
        year = 2020
    GROUP BY
```

```
        year,
        month,
        industry
),
top_industries AS (
    SELECT
        year,
        month,
        industry,
        total_sales,
        RANK() OVER (PARTITION BY year, month ORDER BY total_sales DESC) AS
industry_rank
    FROM
        monthly_sales
)
SELECT
    year,
    month,
    industry,
    total_sales
FROM
    top_industries
WHERE
    industry_rank = 1
ORDER BY
    year,
    month;
```

Data Output Messages Notifications

	year integer	month integer	industry text	total_sales bigint
1	2020	1	Automotive	342994
2	2020	2	Automotive	358432
3	2020	3	Automotive	295174
4	2020	4	Automotive	245248
5	2020	5	Automotive	380186
6	2020	6	Automotive	402878
7	2020	7	Automotive	413128
8	2020	8	Automotive	413998
9	2020	9	Automotive	405586
10	2020	10	Automotive	411928
11	2020	11	Automotive	371776
12	2020	12	Automotive	421624

Total rows: 12 of 12 Query complete 00:00:00.104

4)Top-performing industries in terms of sales for a year 2019, and how do their sales compare month-over-month?

```
WITH monthly_sales AS (
    SELECT
        year,
        month,
        industry,
        SUM(sales) AS total_sales
    FROM
        retail_sales
    WHERE
        year = 2019
    GROUP BY
        year,
        month,
        industry
),
```

```
top_industries AS (
    SELECT
        year,
        month,
        industry,
        total_sales,
        RANK() OVER (PARTITION BY year, month ORDER BY total_sales DESC) AS
industry_rank
    FROM
        monthly_sales
)
SELECT
    year,
    month,
    industry,
    total_sales
FROM
    top_industries
WHERE
    industry_rank = 1
ORDER BY
    year,
    month;
```

Data Output Messages Notifications

	year integer	month integer	industry text	total_sales bigint
1	2019	1	Automotive	329164
2	2019	2	Automotive	335628
3	2019	3	Automotive	411614
4	2019	4	Automotive	383028
5	2019	5	Automotive	414218
6	2019	6	Automotive	387080
7	2019	7	Automotive	404848
8	2019	8	Automotive	426178
9	2019	9	Automotive	369572
10	2019	10	Automotive	384366
11	2019	11	Automotive	375498
12	2019	12	Automotive	385392

Total rows: 12 of 12 Query complete 00:00:00.150

Business Question

5) Which specific kind of businesses contribute the most to total sales, and how does their performance vary across industries?

```
SELECT
    kind_of_business,
    industry,
    SUM(sales) AS total_sales
FROM
    retail_sales
GROUP BY
    kind_of_business,
    industry
ORDER BY
    total_sales DESC;
```

Data Output Messages Notifications

	kind_of_business text	industry text	total_sales bigint
1	Motor vehicle and parts dealers	Automotive	14531121
2	Automobile and other motor vehicle dealers	Automotive	13310756
3	Automobile dealers	Automotive	12389347
4	New car dealers	Automotive	10980583
5	Food and beverage stores	Food & Beverage	9451004
6	General merchandise stores	General Merchandise	9038491
7	Food services and drinking places	Restaurants & Bars	8635492
8	Nonstore retailers	Miscellaneous	8511856
9	Grocery stores	Food & Beverage	8463197
10	Supermarkets and other grocery (except convenience) sto...	Food & Beverage	8092220
11	Restaurants and other eating places	Restaurants & Bars	7541873

Total rows: 58 of 58 Query complete 00:00:00.261

6) Is there any seasonality in sales for specific industries, and how do they perform month-over-month?

```
SELECT
   industry,
         year,
   month,
   SUM(sales) AS total_sales
FROM
   retail_sales
GROUP BY
         year,
   industry,
   month
ORDER BY
         year,
   industry,
   month;
```

Data Output Messages Notifications

	industry text	year integer	month integer	total_sales bigint
1	Automotive	2010	1	184064
2	Automotive	2010	2	190554
3	Automotive	2010	3	247908
4	Automotive	2010	4	234756
5	Automotive	2010	5	238892
6	Automotive	2010	6	235548
7	Automotive	2010	7	245772
8	Automotive	2010	8	242440
9	Automotive	2010	9	226470
10	Automotive	2010	10	225750
11	Automotive	2010	11	216494

Total rows: 1000 of 1872 Query complete 00:00:00.213

7) How does the sales distribution vary among industries based on their North American Industry Classification System (NAICS) codes?

```
SELECT
    naics_code,
    industry,
    SUM(sales) AS total_sales
FROM
    retail_sales
GROUP BY
    naics_code,
    industry
ORDER BY
    naics_code,
    total_sales DESC;
```

Data Output Messages Notifications

	naics_code text	industry text	total_sales bigint
1	44,114,412	Automotive	13310756
2	441	Automotive	14531121
3	4411	Automotive	12389347
4	44111	Automotive	10980583
5	44112	Automotive	1408764
6	4413	Automotive	1220364
7	442	Home Goods & Electronics	1434266
8	442,443	Home Goods & Electronics	2674826
9	4421	Home Goods & Electronics	775943
10	4422	Home Goods & Electronics	514352
11	44221	Home Goods & Electronics	123189

Total rows: 58 of 58 Query complete 00:00:00.109

8) Are there any outliers or significant changes in sales for specific industries during particular months or years?

```
SELECT
    industry,
    year,
    month,
    sales
FROM
    retail_sales
WHERE
    (industry, year, month) IN (
        SELECT
            industry,
            year,
            month
        FROM (
            SELECT
                industry,
                year,
```

```
            month,
            sales,
            LAG(sales) OVER (PARTITION BY industry ORDER BY year, month) AS
prev_sales,
            LEAD(sales) OVER (PARTITION BY industry ORDER BY year, month) AS
next_sales
        FROM
            retail_sales
    ) AS sales_analysis
    WHERE
        sales > 1.5 * COALESCE(prev_sales, 0) OR sales > 1.5 * COALESCE(next_sales, 0)
  )
ORDER BY
  industry,
  year,
  month;
```

9)Which businesses all-time average sale was above 10 billion dollars?

```
SELECT
  kind_of_business,
  AVG(sales) AS average_sale
FROM
  retail_sales
GROUP BY
  kind_of_business
HAVING
  AVG(sales) > 10000;  -- 10 billion dollars in cents (1 dollar = 100 cents)
```

10)Which kind of businesses within the automotive industry had the highest sales revenue for 2022?

```
SELECT
  kind_of_business,
  SUM(sales) AS total_sales
FROM
```

```
    retail_sales
WHERE
    industry = 'Automotive' AND year = 2022
GROUP BY
    kind_of_business
ORDER BY
    total_sales DESC
;
```

11)What is the contribution percentage of each business in the automotive industry this year?

```
WITH automotive_sales AS (
    SELECT
        kind_of_business,
        SUM(sales) AS total_sales
    FROM
        retail_sales
    WHERE
        industry = 'Automotive' AND
        year = 2022
    GROUP BY
        kind_of_business
),
total_sales_automotive AS (
    SELECT
        SUM(sales) AS total_sales_automotive
    FROM
        retail_sales
    WHERE
        industry = 'Automotive' AND
        year = 2022
)
SELECT
    kind_of_business,
```

```
   ROUND((total_sales / total_sales_automotive.total_sales_automotive) * 100, 2) AS
contribution_percentage
FROM
   automotive_sales
CROSS JOIN
   total_sales_automotive;

----------------
--

with total_sales as(select year, industry, sum(sales) as sales_sum
from retail_sales
GROUP BY 1,2)

SELECT curr.industry, prev.year as previous_year, curr.year as current_year,
   (curr.sales_sum - prev.sales_sum) / prev.sales_sum * 100 as YoY

from total_sales as curr
join total_sales as prev
  on curr.year=prev.year+1 AND curr.industry=prev.industry
ORDER BY industry, curr.year DESC;
```

12)What are the year-over-year growth rates for each industry per year?

```
with total_sales as(select year, industry, sum(sales) as sales_sum
from retail_sales
GROUP BY 1,2)

SELECT curr.industry, prev.year as previous_year, curr.year as current_year,
   (curr.sales_sum - prev.sales_sum) / prev.sales_sum * 100 as YoY

from total_sales as curr
join total_sales as prev
  on curr.year=prev.year+1 AND curr.industry=prev.industry
ORDER BY industry, curr.year DESC;
```

```
--OR--
SELECT
   year,
   industry,
   (sales - LAG(sales) OVER (PARTITION BY industry ORDER BY year)) / LAG(sales)
OVER (PARTITION BY industry ORDER BY year) * 100 AS growth_rate
FROM
   retail_sales
ORDER BY
   industry, year;
```

13)What are the yearly total sales for women's clothing stores and men's clothing stores?

```
SELECT
   year,
   sum(CASE WHEN kind_of_business = 'Women"s clothing stores' THEN sales ELSE 0
END) as women_sales,
   sum(CASE WHEN kind_of_business = 'Men"s clothing stores' THEN sales ELSE 0 END)
as men_sales
FROM
   retail_sales
GROUP BY
   year;
```

14)What is the yearly ratio of total sales for women's clothing stores to total sales for men's clothing stores?

```
SELECT year, women_sales/men_sales as Women_to_Men_ratio
FROM (
   SELECT year,
   sum(CASE WHEN kind_of_business = 'Women"s clothing stores' THEN sales ELSE 0
END) as women_sales,
```

```
    sum(CASE WHEN kind_of_business = 'Men"s clothing stores' THEN sales ELSE 0 END)
as men_sales
    FROM retail_sales
    GROUP BY 1
) subquery;
```

15)What is the year-to-date total sale of each month for 2019, 2020, 2021, and 2022 for the women's clothing stores?

```
SELECT
    rs.month,
    rs.year,
    rs.sales,
    (
        (
            SELECT SUM(sales)
            FROM retail_sales rs2
            WHERE rs2.year = rs.year
            AND rs2.month <= rs.month
            AND rs2.kind_of_business = 'Women\'s clothing stores'
        )
    ) AS ytd_sales
FROM
    retail_sales AS rs
WHERE
    rs.kind_of_business = 'Women\'s clothing stores'
    AND rs.year IN (2019, 2020, 2021, 2022);
```

16)What is the month-over-month growth rate of women's clothing businesses in 2022?

```
-- Query 1
SELECT
    month,
    sales AS current_sales,
```

```
    -- now we want the sales from 1 previous period
    LAG(sales, 1) OVER (ORDER BY month) AS prev_sales
FROM
    retail_sales
WHERE
    kind_of_business = 'Women\'s clothing stores'
    AND year = 2022;

-- Query 2
SELECT
    month,
    sales AS current_sales,
    LAG(sales, 1) OVER (ORDER BY month) AS prev_sales,
    (sales - LAG(sales, 1) OVER (ORDER BY month)) / LAG(sales, 1) OVER (ORDER BY
month) * 100 AS growth_rate
FROM
    retail_sales
WHERE
    kind_of_business = 'Women\'s clothing stores'
    AND year = 2022;
```

Data Source: U.S. Government, Open available data
GitHub Repo: https://github.com/tushar2704/Sales-for-Retail-and-Food-Services
Author: © 2023 Tushar Aggarwal