Questions

1. How many total products are in the Products table?

```
SELECT COUNT(item_no)
FROM products;
```

Answer: 9977

2. Who are the top most diverse vendors (i.e. they have the highest number of distinct products)? How many different products do they have?

```
SELECT vendor_name, COUNT(DISTINCT item_no)
FROM products
GROUP BY 1
ORDER BY 2 DESC
LIMIT 5;
```

Vendor Name	No. of Products
Jim Beam Brands	925
Diageo Americas	907
Pernod Ricard USA / Austin Nichols	599
Yahara Bay Distillers Inc.	579
Heaven Hill Distilleries Inc.	388

3. Which products sell the best by total number of unit sales? (Hint: you will need to find a way of doing a cross-join between the sales and product tables.)

```
SELECT p.item_no, p.item_description, SUM(s.bottle_qty)
FROM products p
JOIN sales s
ON
    p.item_no
    =
    s.item
GROUP BY 1,2
ORDER BY 3 DESC
LIMIT 5;
```

Item No.	Item Description	Qty Sold
11788	Black Velvet	828393
36308	Hawkeye Vodka	500070
43337	Captain Morgan Spiced Rum	483621
36904	Mccormick Vodka Pet	393203
35317	Barton Vodka	288425

4. Which products sell the best by total dollar value of sales?

```
SELECT p.item_no, p.item_description, SUM(s.total)
FROM products p
JOIN sales s
ON
    p.item_no
    =
    s.item
GROUP BY 1,2
ORDER BY 3 DESC
LIMIT 5;
```

Item No.	Item Description	Total Sales
11788	Black Velvet	$ 12,863,376.81
43337	Captain Morgan Spiced Rum	$ 8,523,370.55
26827	Jack Daniels Old #7 Black Lbl	$ 6,668,863.94
11297	Crown Royal Canadian Whisky	$ 6,231,299.41
36308	Hawkeye Vodka	$ 5,380,753.20

5. What are the top 10 categories of liquor sold based on the total amount of sales revenue?

```
SELECT category_name, SUM(total)
FROM sales
GROUP BY 1
ORDER BY 2 DESC
LIMIT 10;
```

Category	Total Sales
Canadian Whiskies	$ 48,053,061.91
80 Proof Vodka	$ 48,045,532.51
Spiced Rum	$ 31,600,618.50
Imported Vodka	$ 23,879,524.63
Tequila	$ 21,411,263.64
Straight Bourbon Whiskies	$ 20,924,480.19
Whiskey Liqueur	$ 19,339,201.42
Tennessee Whiskies	$ 17,647,970.35
Puerto Rico & Virgin Islands Rum	$ 12,729,072.76
Blended Whiskies	$ 12,037,250.55

6. Which rum products have sales greater than $10,000? How about whiskey or vodka products?

Rum

```
SELECT item, description, category_name, SUM(total)
FROM sales
WHERE category_name ILIKE '%rum%'
GROUP BY 1,2,3
Having SUM(total) > 10000
ORDER BY 4 DESC
```

Item No.	Item Description	Category	Total Sales
43337	Captain Morgan Spiced Rum	SPICED RUM	$ 8,523,370.55
43338	Captain Morgan Spiced Rum	SPICED RUM	$ 4,341,131.70
1799	Captain Morgan Spiced Barrel	SPICED RUM	$ 3,774,545.06
43336	Captain Morgan Original Spiced	SPICED RUM	$ 3,497,803.08
43028	Admiral Nelson Spiced Rum	SPICED RUM	$ 2,164,286.62

Answer: 166 rum products

Whiskey

```
SELECT item, description, category_name, SUM(total)
FROM sales
WHERE category_name ILIKE '%whiskey%'
GROUP BY 1,2,3
Having SUM(total) > 10000
ORDER BY 4 DESC
```

Item No.	Item Description	Category	Total Sales
64867	Fireball Cinnamon Whiskey	WHISKEY LIQUEUR	$ 4,758,904.71
64866	Fireball Cinnamon Whiskey	WHISKEY LIQUEUR	$ 3,322,146.51
64868	Fireball Cinnamon Whiskey	WHISKEY LIQUEUR	$ 2,254,552.00
86887	Southern Comfort	WHISKEY LIQUEUR	$ 1,873,609.92
64858	Fireball Cinnamon Whiskey Mini Dispenser	WHISKEY LIQUEUR	$ 1,186,675.02

Answer: 35 whiskey products

Vodka

```
SELECT item, description, category_name, SUM(total)
FROM sales
WHERE category_name ILIKE '%vodka%'
GROUP BY 1,2,3
Having SUM(total) > 10000
ORDER BY 4 DESC
```

Item No.	Item Description	Category	Total Sales
36308	Hawkeye Vodka	80 PROOF VODKA	$ 5,380,753.20
34007	Absolut Swedish Vodka 80 Prf	IMPORTED VODKA	$ 3,214,878.03
34422	Grey Goose Vodka	IMPORTED VODKA	$ 3,153,668.04

37338	UV Vodka PET	80 PROOF VODKA	$ 2,308,126.71
34006	Absolut Swedish Vodka 80 Prf	IMPORTED VODKA	$ 2,171,065.18

Answer: 405 vodka products

7. Which county sold the most amount of vodka during February 2014?

```
SELECT county, SUM(bottle_qty)
FROM sales
WHERE category_name ILIKE '%vodka'
    AND DATE_PART ('year',date) = 2014
    AND DATE_PART ('month',date) = 2
GROUP BY 1
ORDER BY 2 DESC
LIMIT 5;
```

County	No. of Bottles
Polk	202012
Linn	71436
Scott	71152
Johnson	56350
Black Hawk	54476

Answer: Polk

8. Which counties were in the top 10 counties for vodka sales in any month in 2014?

```
WITH monthly_rank AS (
    SELECT
        county,
        DATE_PART ('month', date) AS month,
        SUM(total) AS monthly_sales,
        RANK () OVER (PARTITION BY DATE_PART ('month', date)
                      ORDER BY SUM(total)DESC) AS monthly_rank_number
    FROM sales
    WHERE category_name ILIKE '%vodka%' AND DATE_PART('year', date) = 2014
    GROUP BY 1,2
)
SELECT county, month, monthly_sales, monthly_rank_number
FROM monthly_rank
WHERE monthly_rank_number <= 10 AND month = 1
ORDER BY 4;
```

County	Month	Vodka Sales	Monthly Rank
Polk	1	$ 2,188,692.56	1
Linn	1	$ 975,547.22	2
Scott	1	$ 764,583.12	3

Johnson	1	$ 687,857.74	4
Black Hawk	1	$ 611,497.16	5
Pottawattamie	1	$ 425,965.76	6
Story	1	$ 358,245.36	7
Dubuque	1	$ 348,632.82	8
Woodbury	1	$ 289,273.18	9
Cerro Gordo	1	$ 222,337.54	10

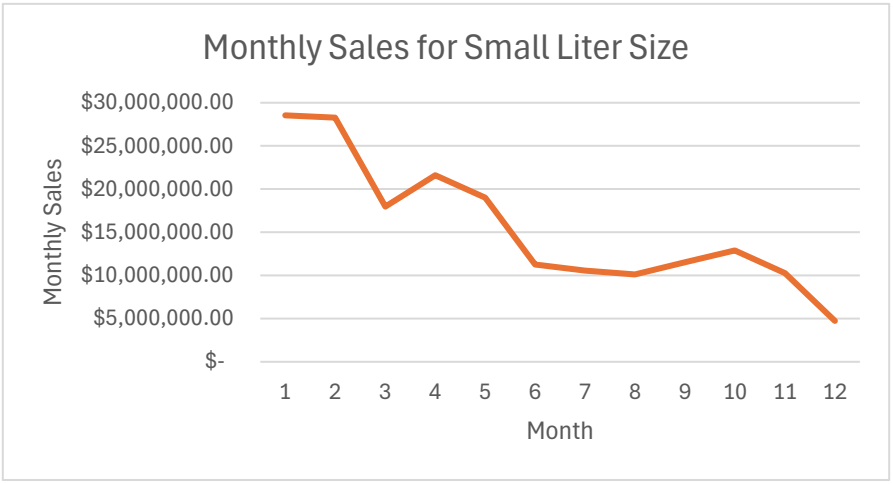
9. Create a report that shows how many times a county appeared in the "top 10 counties for vodka sales in a month" list over the course of 2014.

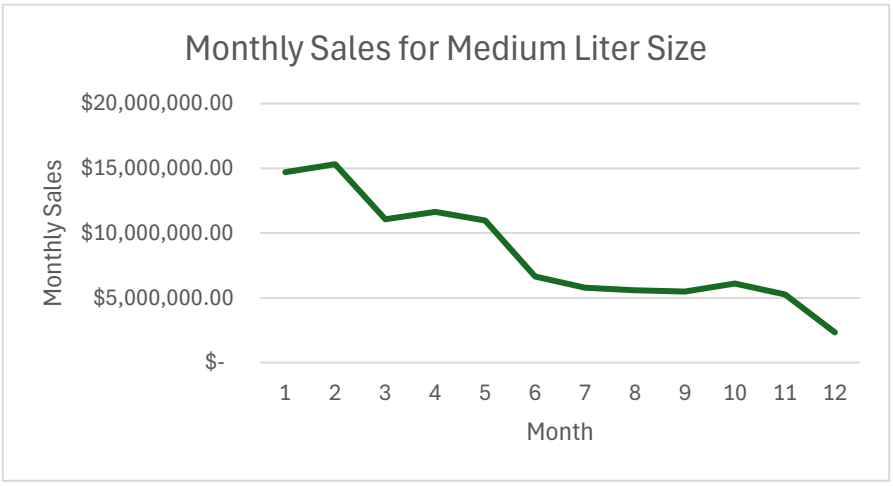
```
WITH monthly_rank AS (
    SELECT
        county,
        DATE_PART ('month', date) AS month,
        SUM(total) AS monthly_sales,
        RANK () OVER (PARTITION BY DATE_PART ('month', date)
                      ORDER BY SUM(total)DESC) AS monthly_rank_number
    FROM sales
    WHERE category_name ILIKE '%vodka%' AND DATE_PART('year', date) = 2014
    GROUP BY 1,2
)
SELECT county, COUNT(*)
FROM monthly_rank
WHERE monthly_rank_number <= 10
GROUP BY 1
ORDER BY 2 DESC;
```

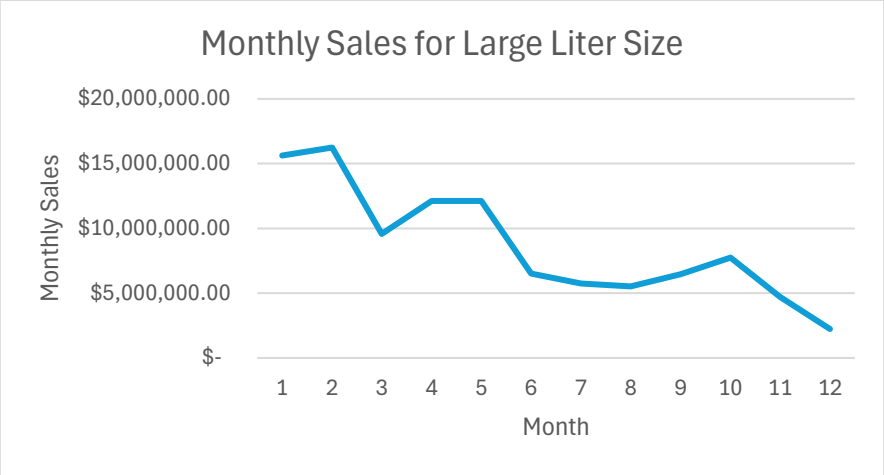
County	Number of Top 10s
Black Hawk	12
Dubuque	12
Johnson	12
Linn	12
Polk	12
Pottawattamie	12
Scott	12
Woodbury	11
Story	10
Cerro Gordo	7
Dallas	6
Des Moines	1
Dickinson	1

10. What is the trend of sales by month? Break up variables such as bottle_price or liter_size into categories (for example: cheap, medium, or expensive).

```
WITH monthly_liter_sales AS (
    SELECT CASE
        WHEN liter_size <= '750' THEN 'small'
        WHEN liter_size BETWEEN '750' AND '1000' THEN 'medium'
        ELSE 'large'
        END AS liter_category,
        DATE_PART ('month', date) AS month,
        SUM(total) AS monthly_sales
    FROM sales
    GROUP BY 1,2
)
SELECT liter_category, month, monthly_sales
FROM monthly_liter_sales
GROUP BY 1,2,3
ORDER BY 1,2;
```







Answer: There is a declining trend in sales over the months for all liter sizes.

11. Pick **one of these 4 questions**:
    a. Which stores sell one of the top five most expensive bottles of alcohol?

```
WITH btl_price_rank AS (
    SELECT s.store, str.name, s.description, s.btl_price,
    DENSE_RANK () OVER (ORDER BY btl_price DESC) AS btl_rank
    FROM sales s
    JOIN stores str
    Using (store)
)
SELECT DISTINCT store, name, description, btl_price, btl_rank
FROM btl_price_rank
WHERE btl_rank <= 5
ORDER BY 5;
```

Store	Store Name	Description	Price	Rank
2588	Hy-Vee Food and Drug #6 / Cedar Rapi	Cedar Ridge Barrel Proof Bourbon	$8,700.00	1
2590	Hy-vee Food Store #5 / Cedar Rapids	Cedar Ridge Barrel Proof Bourbon	$8,700.00	1
3354	Sam's Club 8238 / Davenport	Absolut w/ Zing Zang Bloody Mary Mix	$2,398.80	2
3385	Sam's Club 8162 / Cedar Rapids	Absolut w/ Zing Zang Bloody Mary Mix	$2,398.80	2
3420	Sam's Club 6344 / Windsor Heights	Absolut w/ Zing Zang Bloody Mary Mix	$2,398.80	2
3447	Sam's Club 6432 / Sioux City	Absolut w/ Zing Zang Bloody Mary Mix	$2,398.80	2

Answer: 23 stores sell at least one of the Top 5 most expensive bottles of alcohol.

b. How many stores have more than $2,000,000 in total sales?

```
SELECT store, SUM(total)
FROM sales
GROUP BY 1
Having SUM(total) > 2000000
ORDER BY 2 DESC
```

Answer: 24

c. How many stores have an average bottle price greater than $20?

```
SELECT store, AVG(btl_price::numeric)
FROM sales
GROUP BY 1
Having AVG(btl_price::numeric) > 20
ORDER BY 2 DESC;
```

Answer: 22

d. Which stores have the highest sales of items over 90 proof?

```
SELECT s.store, st.name, SUM(s.total)
FROM sales s
JOIN stores st
USING (store)
WHERE description ILIKE '%1__%' AND
    (description ILIKE '%proof%' OR description ILIKE '%prf%')
GROUP BY 1,2
ORDER BY 3 DESC
LIMIT 5;
```

Store	Store Name	Total Sales
2512	Hy-vee Wine and Spirits / Iowa City	$ 18,431.40
4829	Central City 2	$ 17,525.45
2633	Hy-vee #3 / Bdi / Des Moines	$ 17,094.09
2613	Hy-vee Food Store #1 / Council Bluff	$ 15,986.22
3618	Wal-Mart 2716 / Cedar Rapids	$ 15,715.08

- The data might have been corrupted in some way. The category listed in the sales table doesn't always match up with the category in the products table. How many times has this happened, and can you find any patterns to it?

```
WITH sales_combined AS (
    SELECT *,
        s.category_name AS store_category_name,
        p.category_name AS product_category_name,
        CASE
            WHEN s.category_name = p.category_name THEN 'match'
            ELSE 'not match'
            END AS match_check
    FROM sales s
    JOIN products p
    ON
        s.item
        =
        p.item_no
)
SELECT *
FROM sales_combined
WHERE match_check = 'not match' AND
      (store_category_name IS NOT NULL OR product_category_name IS NOT NULL);
```

Answer: 2285 times with no apparent pattern.

- Store 2238 (Adventureland Inn at 3200 Adventureland Dr) sold $883.24 in April and $27,526.38 in May, for a 3017% growth rate. That was the highest percentage month-on-month growth rate. Create a query that shows this and the next 9 highest after that.

```
WITH store_monthly_sales AS (
    SELECT
        s.store,
        str.name,
        DATE_PART ('month', date) AS month,
        SUM(total) AS monthly_sales
    FROM sales s
    JOIN stores str
    Using (store)
    GROUP BY 1,2,3
)
SELECT
    store,
    name,
    month,
    monthly_sales,
    LAG(monthly_sales, 1) OVER (PARTITION BY store ORDER BY month) AS prev_month_sales,
    ROUND(((monthly_sales - LAG(monthly_sales, 1) OVER (PARTITION BY store ORDER BY month))
          / LAG(monthly_sales, 1) OVER (PARTITION BY store ORDER BY month))*100, 2)
          AS month_on_month_perc
FROM store_monthly_sales
ORDER BY 6 DESC NULLS LAST
LIMIT 10;
```

Store	Store Name	Month	Monthly Sales	Prev Month Sales	Month-on-Month Growth
2238	Adventureland Inn	5	$ 27,526.38	$ 883.24	3016.52%
4275	Fareway Stores #980 / Knoxville	9	$ 41,674.59	$ 2,925.30	1324.63%
4169	Super Quick 2 / Hubbell	11	$ 12,794.70	$ 985.92	1197.74%
4669	Vom Fass / Des Moines	10	$ 7,231.01	$ 649.24	1013.77%
4717	Walgreens #05512 / Bettendorf	8	$ 14,064.95	$ 1,412.70	895.61%
2667	Hy-Vee Mainstreet	3	$ 77,712.78	$ 7,878.87	886.34%
5058	Kum & Go # 542 / Urbandale	11	$ 11,686.97	$ 1,372.99	751.21%
5009	New Pioneer Food Co-op	8	$ 11,661.63	$ 1,423.98	718.95%
4899	Graettinger Market	7	$ 9,321.60	$ 1,172.52	695.01%
4598	Quik Trip #501 / E Euclid DM	5	$ 30,868.46	$ 4,166.58	640.86%

- The store_address field in the stores table actually contains three rows of text. Quite often the latitude and longitude are in the last line of text. Create a query that shows the geo-locatable stores in latitude order (i.e. show the stores from the most northernly to the most southernly).

```
SELECT store, name,
    split_part(split_part(store_address, '(', 2), ',', 1) AS latitude
FROM stores
ORDER BY 3 DESC;
```

Store No.	Store Name	Latitude
3596	Laddy's Bar And Grill	43.49114240400007
4030	Buy Rite Foods	43.45480634900008
3914	Market Street Market	43.45480634900008
4006	Larchwood Offsale	43.45384534900006
4904	Larchwood Quick Stop	43.45355826500003