**SITUATION**

You are a business analyst working at Deloitte. You already know SQL, and you have both a more junior analyst and a more senior analyst working on your team. In this exercise, you will help the junior analyst make their SQL commands work correctly and help the more senior analyst finish off some of their harder queries, and write some queries of your own.

Your client, the state of Iowa, has some urgent questions - questions that need answering. They are considering cracking down on alcohol consumption, but they don’t really understand the dynamics of the alcohol market in their state.

They hire your team of SQL experts and business problem solvers to analyze the data they have (conveniently located in SQLite databases) to solve the problem. Your team divvies up the work, and assigns some specific analyses to you.

**EXERCISES**

**QUESTION 1:**

The junior analyst wants to get the item numbers and descriptions of products with a case cost of at least 100 and pack size at least 12. She writes:

SELECT item\_no item\_description

FROM products

WHERE case\_cost >= 100

WHERE pack >= 12

Why isn’t this query working?

|  |
| --- |
| SOLUTION:  Two reasons: the query needs a comma between the two columns being “selected” and an AND statement between the two WHERE clauses. |

**QUESTION 2:**

She now wants to get a list of all combinations of “category” and “vendor” in the data. She knows that she can select these two columns using:

SELECT category\_name, vendor\_name

FROM products

What else does she need to add to accomplish her goal?

|  |
| --- |
| SOLUTION:  She needs to add DISTINCT after SELECT. |

**QUESTION 3:**

Her query to get all the products where proof is greater than 85 seems right, but it’s generating an error:

SELECT \* FROM products WHERE proof > 85;

How could you fix it?

|  |
| --- |
| SOLUTION:  She needs to use CAST to tell the query to treat proof as an integer. |

**QUESTION 4:**

This analyst likes sweet drinks, and she wants to write a query that shows all the vendors that have products in a few categories she has identified:

‘PEACH BRANDIES’

‘FLAVORED VODKA’

‘FLAVORED RUM’

Her SQL has improved based on your feedback, so she correctly writes:

SELECT DISTINCT vendor\_name

FROM products

WHERE category\_name = “PEACH BRANDIES”

OR category\_name = “FLAVORED VODKA”

OR category\_name = “FLAVORED RUM”

How could you make this query shorter and more consistent with SQL best practices?

|  |
| --- |
| SOLUTION:  Change the WHERE clause to be  WHERE category\_name in (“PEACH BRANDIES”, “FLAVORED VODKA”, “FLAVORED RUM”) |

**QUESTION 5:**

Now you are given an area to analyze independently: the state of Iowa now wants you to do analysis on all “IMPORTED” products, as they want to know which foreign products are coming into their state.

First, select all products with “IMPORTED” in the category name.

|  |
| --- |
| SOLUTION:  SELECT \*  FROM products  WHERE category\_name LIKE ‘%IMPORTED%’ |

**QUESTION 6:**

*Pro Tip: The following several analyses are extremely common among analysts in the real world. Practice up!*

Now, get the top 10 vendors of imported products, ranked by sales.

|  |
| --- |
| SOLUTION:  SELECT vendor  FROM sales  WHERE category\_name LIKE ‘%IMPORTED%’  GROUP BY vendor  ORDER BY SUM(total) DESC  LIMIT 10  Answer:  "Pernod Ricard USA/Austin Nichols"  "Diageo Americas"  "Moet Hennessy USA Inc."  "Sidney Frank Importing Co."  "Bacardi U.S.A. Inc."  "REMY COINTREAU USA ."  "Sazerac Co. Inc."  "Jim Beam Brands"  "Constellation Wine Company Inc."  "Proximo" |

**QUESTION 7:**

Now, of the vendors that have greater than $100,000 in sales, which one has the highest average sales price? Order these vendors by highest to lowest average sales price.

|  |
| --- |
| SOLUTION:  SELECT SUM(total), vendor, AVG(total)  FROM sales  WHERE category\_name LIKE ‘%IMPORTED%’  GROUP BY vendor  HAVING SUM(total) > 100000  ORDER BY AVG(total) DESC  Answer:  Pernod Ricard USA/Austin Nichols |

**STRETCH QUESTION:**

You are proud of this work and you want to share it with the rest of the team. However, Question 7 results in a complicated query that would be tough for someone who wasn’t an advanced SQL user to pick up and understand.

Take the query above and add some comments explaining how it works, using the correct syntax for commenting in SQL. Aim to write a couple of full sentences at the beginning of the query explaining how it works, and add at least 3 shorter, 1 line comments throughout the code.