

## PROJECT 2: SQL

NURUL NAJWA MOHD SUKRI

Questions:

1. How many distinct product names are there in the Products table? Hint: Use item\_description column to identify product names

**SQL query:**

```
SELECT      COUNT (DISTINCT item_description)
FROM        public.products
```

**Answer:** 7276

2. Which product is the best seller by total dollar value of sales? Hint: Use item column to identify product

**SQL query:**

```
SELECT      item_description, item, sum(total)
FROM        (
SELECT      products.item_description, sales.item, sales.total
FROM        products
INNER JOIN  sales
ON          products.item_no=sales.item
           ) as temp
GROUP BY    item_description, item
ORDER BY    3 DESC
LIMIT       10
```

**Answer:** Best seller by total sales is "Black Velvet", item\_no. 11788 with \$12,863,376.81 in sales.

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

**SQL query:**

```
SELECT      category, category_name, sum(total)
FROM        public.sales
GROUP BY    category, category_name
ORDER BY    3 DESC
LIMIT       10
```

**Answer:**

	category text	category_name text	sum numeric
1	1012100	CANADIAN WHISKIES	48053061.91
2	1031080	80 PROOF VODKA	48045532.51
3	1062310	SPICED RUM	31600618.50
4	1032080	IMPORTED VODKA	23879524.63
5	1022100	TEQUILA	21411263.64
6	1011200	STRAIGHT BOURBON WHISKIES	20924480.19
7	1081600	WHISKEY LIQUEUR	19339201.42
8	1011300	TENNESSEE WHISKIES	17647970.35
9	1062200	PUERTO RICO & VIRGIN ISLANDS RUM	12729072.76
10	1011100	BLENDED WHISKIES	12037250.55

4. Has anyone ever bought \$10,000 worth of rum products at a time? Hint: Use category\_name column to identify rum products.

**SQL query:**

```
SELECT      date, category_name, total
FROM        public.sales
WHERE       category_name ilike '%RUM%'
AND         total > 10000
```

**Answer: Yes**

	date timestamp without time zone	category_name text	total numeric (10,2)
1	2014-11-06 00:00:00	SPICED RUM	11635.80
2	2014-10-29 00:00:00	SPICED RUM	11826.00
3	2014-10-30 00:00:00	SPICED RUM	14904.00
4	2014-10-30 00:00:00	PUERTO RICO & VIRGIN ISLANDS RUM	15255.00
5	2014-10-30 00:00:00	PUERTO RICO & VIRGIN ISLANDS RUM	12825.00

5. Which county sold the most vodka during February 2014? Hint: Use category\_name column to identify vodka products.

**SQL query:**

```
SELECT      county, sum(bottle_qty)
FROM        (
SELECT      county, date, category_name, bottle_qty
FROM        public.sales
WHERE       category_name ilike '%VODKA%'
AND         DATE_PART('year', date) = 2014
```

```

AND                DATE_PART('month', date) = 2) as temp

GROUP BY           county

ORDER BY            2 DESC

LIMIT               1

```

**Answer:** Polk County sold the most Vodka with 202,012 bottles in February 2014.

6. We think 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. Write a query that finds records in the sales table that have a different category to the corresponding one in the products table.

**SQL query:**

```

SELECT              products.category_name, sales.category_name

FROM                products

JOIN                sales

ON                  products.item_no=sales.item

WHERE               products.category_name!=sales.category_name

```

**Answer:**

Query Editor		Query History	
1	SELECT	products.category_name, sales.category_name	
2	FROM	products	
3	JOIN	sales	
4	ON	products.item_no=sales.item	
5	Where	products.category_name!=sales.category_name	

  

Data Output		Explain	Messages	Notifications
	category_name text	category_name text		
1	MISC. IMPORTED CORDIALS & LIQUEURS	DECANTERS & SPECIALTY PACKAGES		
2	MISC. IMPORTED CORDIALS & LIQUEURS	DECANTERS & SPECIALTY PACKAGES		
3	MISC. IMPORTED CORDIALS & LIQUEURS	DECANTERS & SPECIALTY PACKAGES		
4	MISC. IMPORTED CORDIALS & LIQUEURS	DECANTERS & SPECIALTY PACKAGES		
5	MISC. IMPORTED CORDIALS & LIQUEURS	DECANTERS & SPECIALTY PACKAGES		
6	MISC. IMPORTED CORDIALS & LIQUEURS	DECANTERS & SPECIALTY PACKAGES		
7	MISC. IMPORTED CORDIALS & LIQUEURS	DECANTERS & SPECIALTY PACKAGES		
8	MISC. IMPORTED CORDIALS & LIQUEURS	DECANTERS & SPECIALTY PACKAGES		
9	MISC. IMPORTED CORDIALS & LIQUEURS	DECANTERS & SPECIALTY PACKAGES		

- Pick one of these 4 questions:

- Which stores sell one of the top five most expensive bottles of alcohol?

- How many stores have more than \$2,000,000 in total sales?

**SQL query:**

```
SELECT      COUNT (name)
FROM        (
SELECT      name, sum(total) as total_sales
FROM        (
SELECT      stores.store, stores.name, sales.store, sales.total
FROM        stores
INNER JOIN  sales
ON          stores.store=sales.store) as temp
GROUP BY    name
ORDER BY    2 DESC) as temp
WHERE       total_sales > 2000000
```

**Answer:** 24 stores have more than \$2, 000, 000 total sales

- How many stores have an average bottle price greater than \$20?
- Which stores have the highest sales of items over 90 proof (with products proof value more than 90)?

**Bonus**

- 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 northerly to the most southerly).