

Data Analyst Portfolio Project - SQL exploration with BigQuery

1. What is the overall trend in liquor sales in Iowa over the years covered by this dataset?

```
SELECT
EXTRACT (YEAR FROM date) AS year,
ROUND(SUM (sale_dollars),2) AS sales
FROM `bigquery-public-data.iowa_liquor_sales.sales`
GROUP BY year
ORDER BY year ASC
LIMIT 1000;
```

2. Which are the most popular liquor categories sold in Iowa?

```
SELECT
category_name,
COUNT(*) AS num
FROM `bigquery-public-data.iowa_liquor_sales.sales`
GROUP BY category_name
ORDER BY num DESC
LIMIT 10;
```

For brands

```
SELECT
item_description,
COUNT(*) AS num
FROM `bigquery-public-data.iowa_liquor_sales.sales`
GROUP BY item_description
ORDER BY num DESC
LIMIT 10;
```

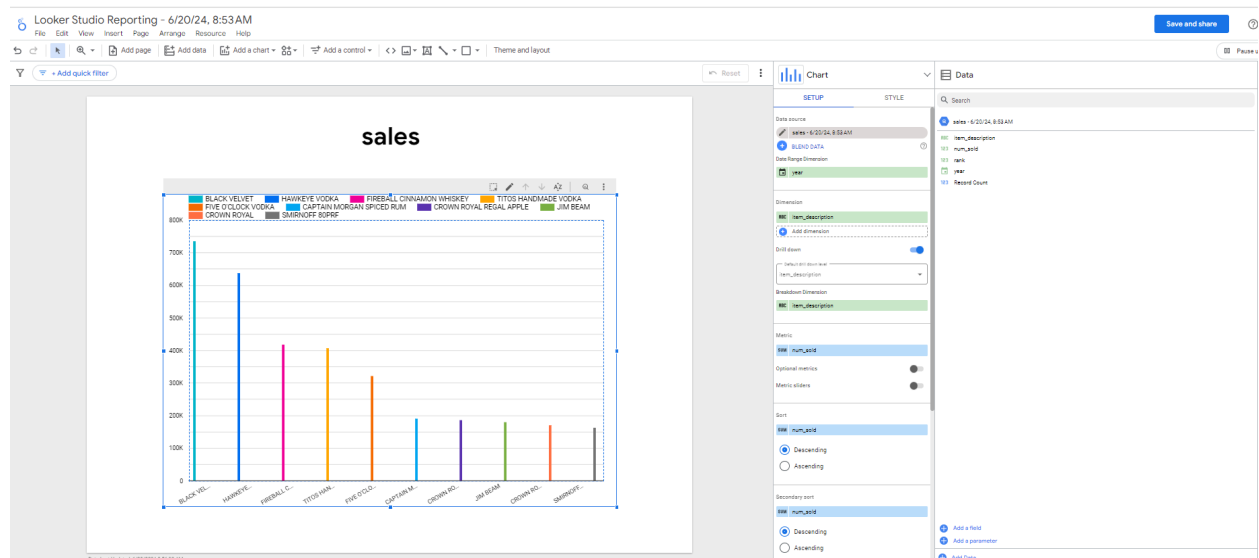
3. How has the popularity of brands changed over time?

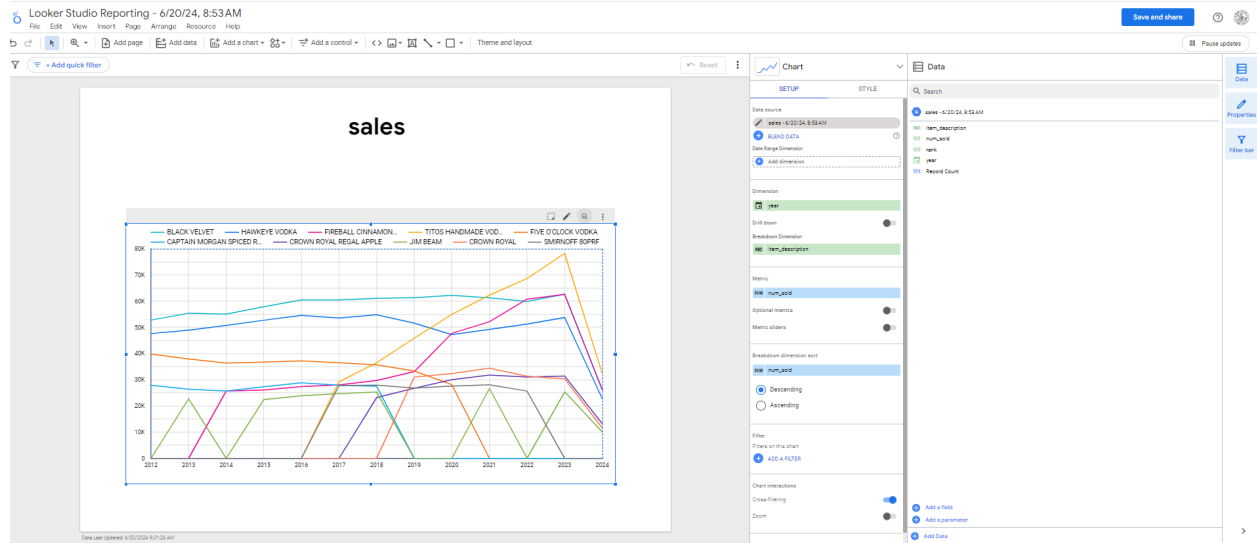
```
WITH CTE AS (
SELECT EXTRACT(YEAR FROM date) AS year,
item_description,
```

```

COUNT(*) AS num_sold
FROM `bigquery-public-data.iowa_liquor_sales.sales`
GROUP BY year, item_description
)
SELECT *
FROM (
  SELECT*,
  DENSE_RANK()OVER(PARTITION BY year ORDER BY num_sold DESC) AS rank
  FROM CTE
)ranked
WHERE rank <=10
ORDER BY year, rank;

```





4. Are there any seasonal patterns in liquor sales?

```
SELECT  
  EXTRACT(MONTH FROM date) AS month,  
  SUM(sale_dollars) AS sales  
FROM `bigquery-public-data.iowa_liquor_sales.sales`  
GROUP BY month  
ORDER BY month ASC;
```

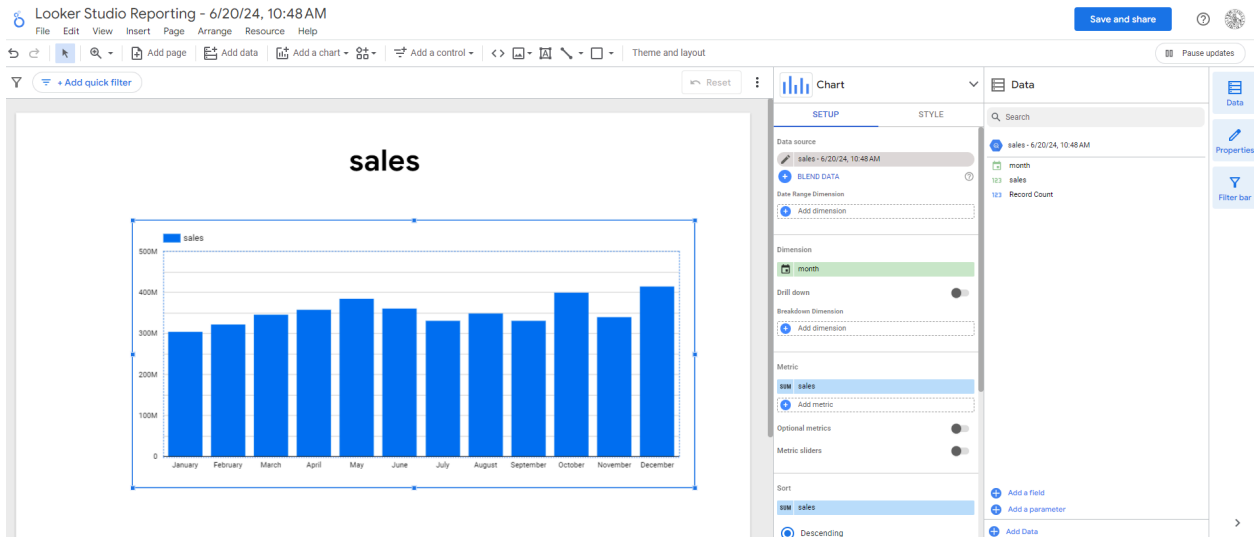
Query results

SAVE RESULTS EXPLORE DATA

JOB INFORMATION		RESULTS	CHART	JSON	EXECUTION DETAILS	EXECUTION GRAPH
Row	month	sales				
1	1	304037967.1700...				
2	2	322202080.7900...				
3	3	346193184.01				
4	4	358284223.0499...				
5	5	385723600.7499...				
6	6	361236349.2300...				
7	7	331128191.4900...				
8	8	348808679.12				
9	9	330866319.9800...				
10	10	399984873.2500...				
11	11	340521996.3500...				
12	12	414271179.9300...				

Results per page: 50 1 - 12 of 12

Job history REFRESH



5. Which counties have the biggest sales?

SELECT

county,

SUM(sale_dollars) AS sales

FROM `bigquery-public-data.iowa_liquor_sales.sales`

GROUP BY county

ORDER BY sales DESC

LIMIT 10;

Google Cloud My First Project

Search (/) for resources, docs, products, and more

Search

Explorer

Viewing resources.

SHOW STARRED ONLY

- epa_historical_air_qu...
- ethereum_blockchain...
- etsi_technical_standa...
- fas
- fcc_political_ads
- fdic_drug
- fdic_food
- fdic_banks
- iowa_liquor_sales
- sales

SUMMARY

Nothing currently selected

Untitled query

RUN SAVE DOWNLOAD SHARE SCHEDULE MORE

Processing location: US

Query results

SAVE RESULTS EXPLORE DATA

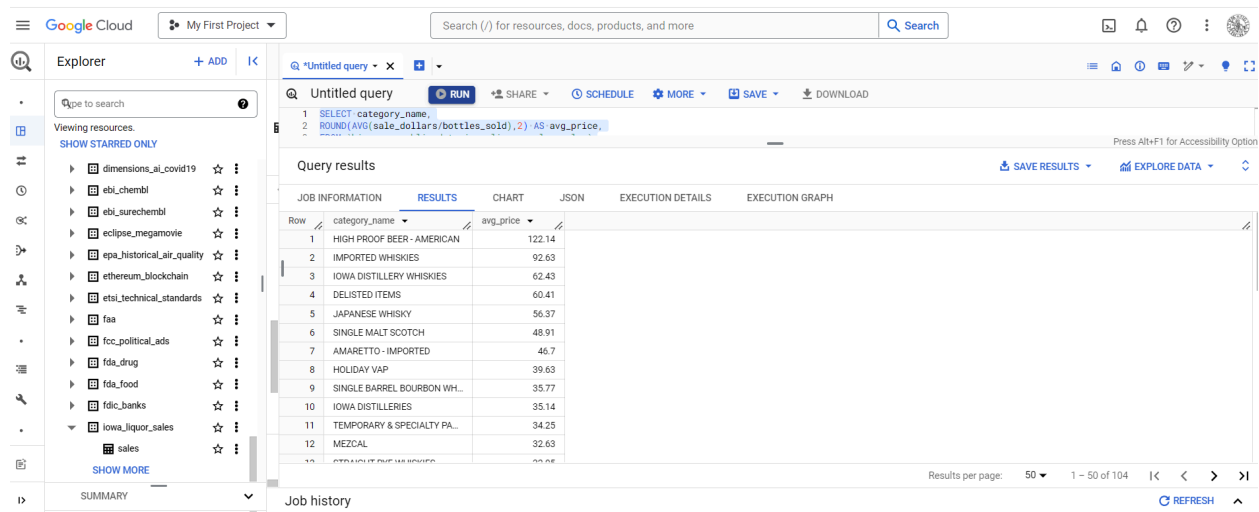
Row	county	sales
1	POLK	967714481.4599...
2	LINN	365659936.7399...
3	SCOTT	295351741.4700...
4	JOHNSON	249207780.5100...
5	BLACK HAWK	233783462.5500...
6	POTTAWATTAMIE	151618557.8600...
7	WOODBURY	145321818.4900...
8	STORY	126467188.1800...
9	DUBUQUE	124596529.6100...
10	DALLAS	100870589.6599...

Job history

REFRESH

6. What is the average price of liquor by category?

```
SELECT category_name,  
ROUND(AVG(sale_dollars/bottles_sold),2) AS avg_price,  
FROM `bigquery-public-data.iowa_liquor_sales.sales`  
WHERE bottles_sold !=0  
GROUP BY category_name  
ORDER BY avg_price DESC;
```



The screenshot shows the Google Cloud BigQuery interface. On the left is the Explorer pane with a search bar and a list of datasets. The main area displays a query titled 'Untitled query' with the following SQL code:

```
1 SELECT category_name,  
2 ROUND(AVG(sale_dollars/bottles_sold),2) AS avg_price,
```

Below the query, the 'Query results' section shows a table with two columns: 'category_name' and 'avg_price'. The table contains 12 rows of data. At the bottom right, it indicates 'Results per page: 50' and '1 - 50 of 104'.

Row	category_name	avg_price
1	HIGH PROOF BEER - AMERICAN	122.14
2	IMPORTED WHISKIES	92.63
3	IOWA DISTILLERY WHISKIES	62.43
4	DELISTED ITEMS	60.41
5	JAPANESE WHISKY	56.37
6	SINGLE MALT SCOTCH	48.91
7	AMARETTO - IMPORTED	46.7
8	HOLIDAY VAP	39.63
9	SINGLE BARREL BOURBON WH...	35.77
10	IOWA DISTILLERIES	35.14
11	TEMPORARY & SPECIALTY PA...	34.25
12	MEZCAL	32.63

7. Which liquor brands have the highest profit margins from retailers?

```
SELECT  
item_description AS brand,  
ROUND(AVG(state_bottle_cost),2) AS avg_cost,  
ROUND(AVG(state_bottle_retail),2) AS avg_revenue,  
ROUND(AVG(state_bottle_retail) - AVG (state_bottle_cost),2) AS profit,  
  
ROUND((((AVG(state_bottle_retail)-AVG(state_bottle_cost))/AVG(state_bottle_retail  
)*)*100,2) AS margin  
FROM `bigquery-public-data.iowa_liquor_sales.sales`  
GROUP BY item_description
```

ORDER BY margin DESC
LIMIT 10;

Google Cloud

My First Project

Search (/) for resources, docs, products, and more

Search

Explorer

+ ADD

⌕

Type to search

Viewing resources.

SHOW STARRED ONLY

augmented-vim-418614

bigquery-public-data

SUMMARY

Nothing currently selected

sales

Untitled query

RUN

SAVE

DOWNLOAD

SHARE

SCHEDULE

MORE

1 SELECT

2 item_description AS brand,

3 avg(cost) AS avg_cost,

Processing location: US

Press Alt+F1 for Accessibility Options

Query results

SAVE RESULTS

EXPLORE DATA

↕

JOB INFORMATION

RESULTS

CHART

JSON

EXECUTION DETAILS

EXECUTION GRAPH

Row	brand	avg_cost	avg_revenue	profit	margin
1	CONNEMARA 12 YEAR	14.66	56.25	41.59	73.94
2	EL DORADO SPECIAL RESERVE...	18.27	44.93	26.66	59.34
3	BIRD DOG PEACH W/2-50MLS	10.0	21.61	11.61	53.72
4	EL DORADO SPECIAL RESERVE...	18.27	34.52	16.25	47.08
5	OLE SMOKY APPLE PIE W/FLA...	8.98	15.57	6.59	42.32
6	STRAIGHT EDGE BOURBON W...	25.84	43.25	17.41	40.25
7	PLANTATION RHUM GRAND R...	13.6	22.05	8.45	38.32
8	WISCONSIN CLUB BRANDY	6.44	10.4	3.96	36.05
9	SLAUGHTER HOUSE AMERICA...	21.5	34.5	13.0	37.68
10	HOT ROSE CINNAMON CREAM...	8.47	13.45	4.99	37.06

Job history

REFRESH

^