Embedding Simple Calculations in SQL with BigQuery

Sheba Morgan

INTRODUCTION

My goal was to find out the total number of bags of avocados sold on each date at each location using this data. There's already a column that shows me the total, but I wanted to make sure that the total column is just small, large, and extra-large bags added together. I added the values in those three columns together in my query and then compared them to the total bags column in the dataset.

I started with SELECT and then the names of the columns I wanted to use in my calculations. Then I added the calculation to the query using the names of the three columns with plus signs between them. Then I typed AS followed by the new column name to label the column with the calculated values. Finally, I ended my first query with the FROM command and the name of the table that I was pulling data from.

```
1  SELECT
2  Date,
3  Region,
4  Small_Bags,
5  Large_Bags,
6  XLarge_Bags,
7  Total_Bags,
8  Small_Bags + Large_Bags + XLarge_Bags AS Total_Bags_Calc
9  FROM `my-project-cert-382621.avocado_data.avocado_prices`
```

Untitled 2   RUN   SAVE   SHARE   SCHEDULE   MORE   Query completed.

Press Alt+F1 for Accessibility Options.

These were the results..

SAVE RESULTS ▾   EXPLORE DATA ▾   ⇕

JOB INFORMATION   RESULTS   JSON   EXECUTION DETAILS   EXECUTION GRAPH PREVIEW
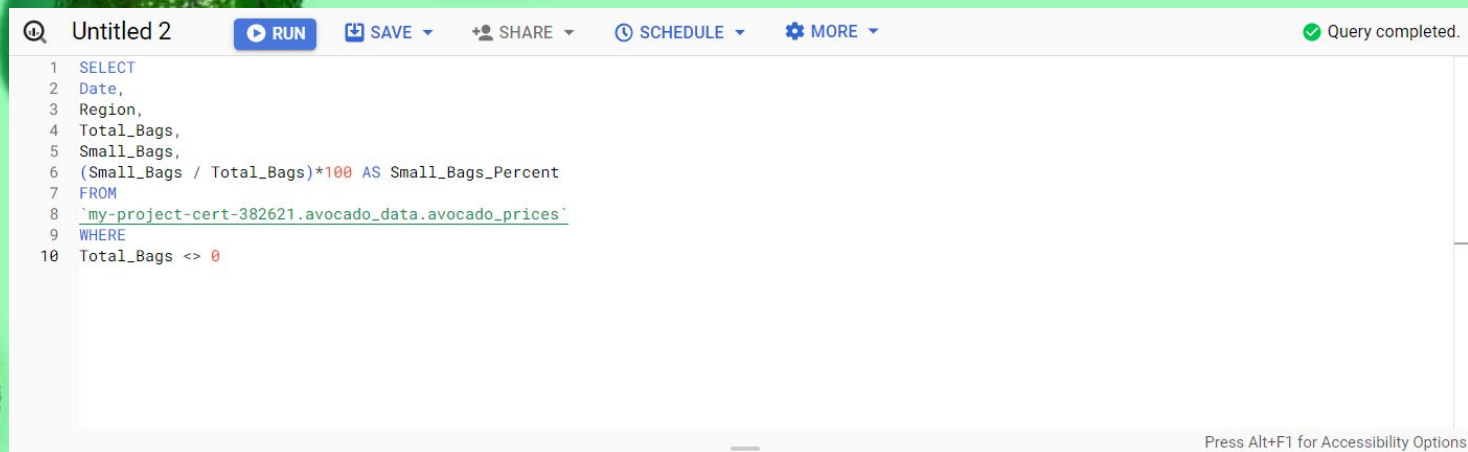
Row	Date ▾	Region ▾	Small_Bags ▾	Large_Bags ▾	XLarge_Bags ▾	Total_Bags ▾	Total_Bags_Calc ▾
1	2015-12-27	Albany	8603.62	93.25	0.0	8696.87	8696.87
2	2015-12-27	Atlanta	48605.95	17748.36	0.0	66354.31	66354.31
3	2015-12-27	BaltimoreWashington	142543.88	2367.22	0.0	144911.1	144911.1
4	2015-12-27	Boise	23520.19	5.69	35.22	23561.1	23561.1
5	2015-12-27	Boston	85913.6	99.26	0.0	86012.86	86012.86
6	2015-12-27	BuffaloRochester	55236.68	0.0	0.0	55236.68	55236.68
7	2015-12-27	California	1090140.07	110737.35	11829.59	1212707.01	1212707.010000...
8	2015-12-27	Charlotte	35130.42	2499.62	0.0	37630.04	37630.04
9	2015-12-27	Chicago	83066.75	1617.67	10056.67	94741.09	94741.09
10	2015-12-27	CincinnatiDayton	14050.64	30968.06	5350.6	50369.3	50369.29999999...
11	2015-12-27	Columbus	15032.43	3874.64	4164.71	23071.78	23071.78
12	2015-12-27	DallasFtWorth	136560.04	12277.7	15657.53	164495.27	164495.2700000...
13	2015-12-27	Denver	38620.02	227884.21	85.85	266590.08	266590.0799999...
14	2015-12-27	Detroit	48936.21	762.35	12368.92	62067.48	62067.47999999

Results per page: 50 ▾   1 – 50 of 18249 |< < > >|

In my second query, I set up a new column starting with my calculation. To find the percentage of small bags, I needed to first divide the number of small bags by the number of total bags using a slash as the operator. I put that part of the calculation in parentheses to let the server know that this calculation would need to be performed first. Then I multiplied the total by 100 using an asterisk as our operator. Multiplying by 100 gave me a value that's a percentage instead of a decimal. I used the AS command to name this new column, "Small Bags Percent." Then I added FROM and the name of the set I was pulling from and WHERE with the Total_Bags column followed by a less than sign, a greater than sign, and a zero.


```
       Untitled 2        ▶ RUN      💾 SAVE ▾    +● SHARE ▾    🕐 SCHEDULE ▾    ⚙ MORE ▾                              ✅ Query completed.
  1    SELECT
  2    Date,
  3    Region,
  4    Total_Bags,
  5    Small_Bags,
  6    (Small_Bags / Total_Bags)*100 AS Small_Bags_Percent
  7    FROM
  8    `my-project-cert-382621.avocado_data.avocado_prices`
  9    WHERE
 10    Total_Bags <> 0
```

And here are the results for the second query..

Query results

JOB INFORMATION   **RESULTS**   JSON   EXECUTION DETAILS   EXECUTION GRAPH PREVIEW

Row	Date ▾	Region ▾	Total_Bags ▾	Small_Bags ▾	Small_Bags_Percent
1	2015-12-27	Albany	8696.87	8603.62	98.92777516508...
2	2015-12-27	Atlanta	66354.31	48605.95	73.25213689962...
3	2015-12-27	BaltimoreWashington	144911.1	142543.88	98.36643293715...
4	2015-12-27	Boise	23561.1	23520.19	99.82636634113...
5	2015-12-27	Boston	86012.86	85913.6	99.884598651876
6	2015-12-27	BuffaloRochester	55236.68	55236.68	100.0
7	2015-12-27	California	1212707.01	1090140.07	89.89311193970...
8	2015-12-27	Charlotte	37630.04	35130.42	93.35738149627...
9	2015-12-27	Chicago	94741.09	83066.75	87.67763807657...
10	2015-12-27	CincinnatiDayton	50369.3	14050.64	27.89524571514...
11	2015-12-27	Columbus	23071.78	15032.43	65.15505088900...
12	2015-12-27	DallasFtWorth	164495.27	136560.04	83.01760895617...

Load more

Results per page: 50 ▾   1 – 50 of 18234 |<   <   >   >|

CONCLUSION

The queries performed successfully enabled the analysis of the total number of bags of avocados sold on each date at each location, specifically focusing on the small, large, and extra-large bags. By summing up the values from these three columns and comparing them to the "Total Bags" column in the dataset, the queries ensured the accuracy and integrity of the "Total Bags" column. Additionally, the introduction of a new column for the percentage of small bags allowed for a comprehensive examination of the data. These queries provided a systematic and efficient means of evaluating the avocados sales data, enabling informed decision-making based on reliable information.