Using JOINs in SQL (BigQuery)

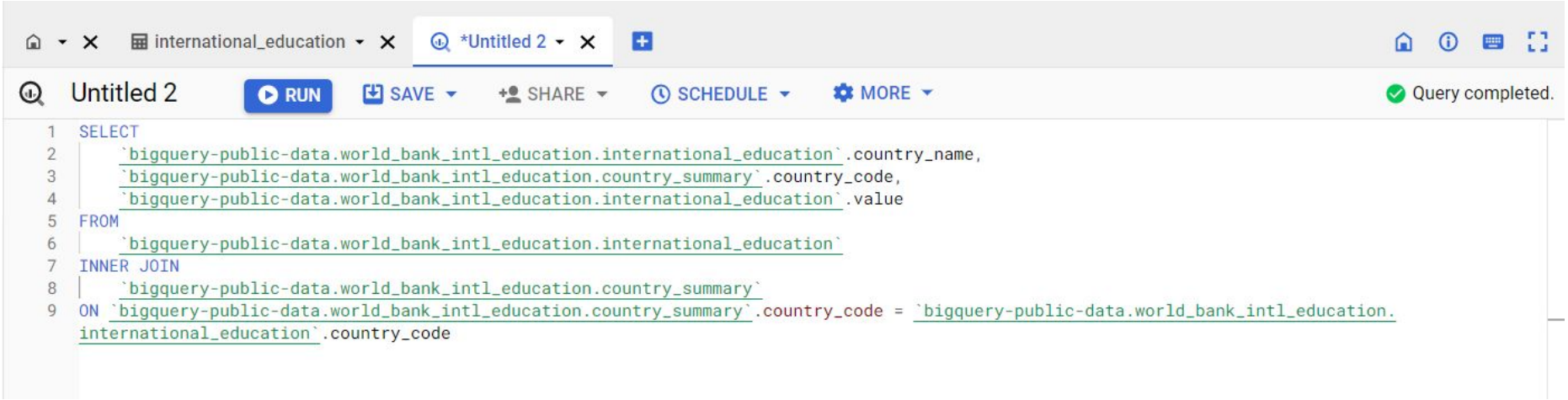
Sheba Morgan

INTRODUCTION

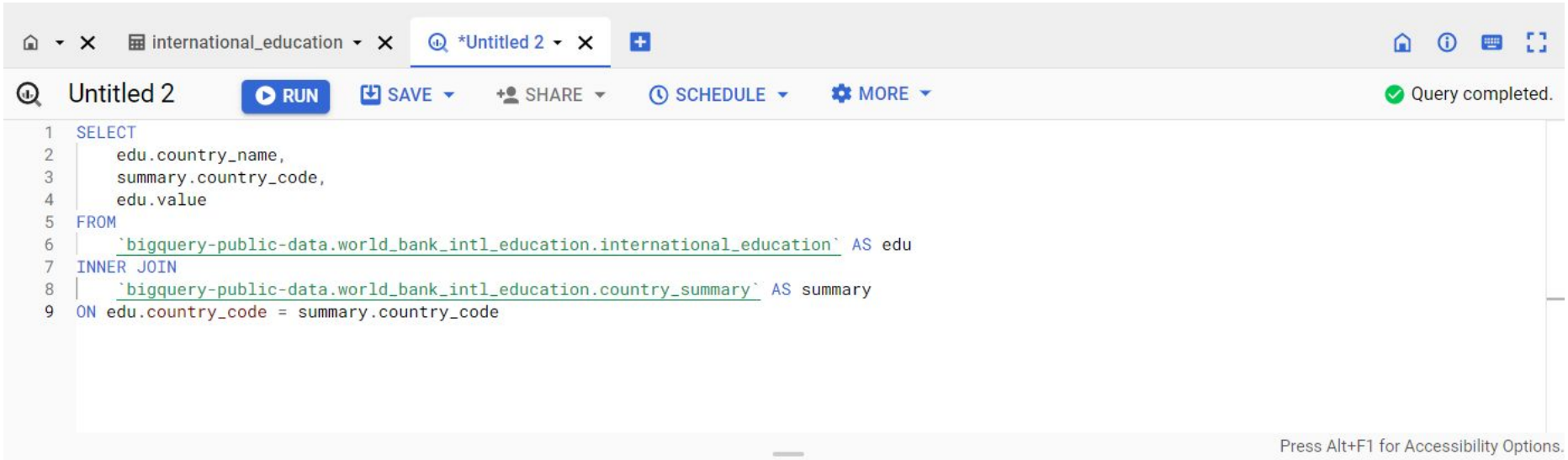
For this data analysis project, I will be working with the International Education public dataset available in BigQuery. The dataset includes multiple tables, including the international_education table and the country_summary table. Both of these tables contain the column “country_code”, according to the datasets schema.

To combine the data from these tables based on the shared "country_code" column, I will be utilizing the INNER JOIN clause in my query. This allows me to merge the relevant information from both tables and gain comprehensive insights for analysis.

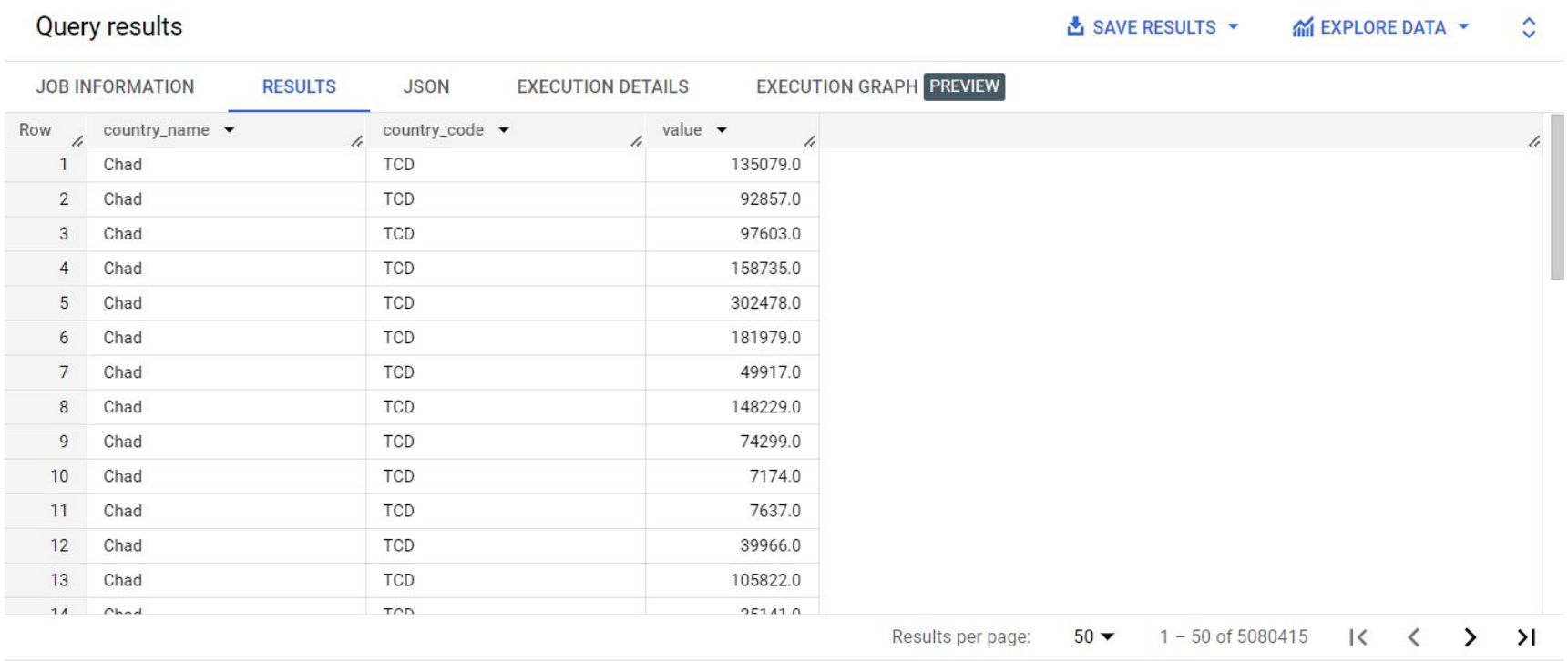
This was my initial query..



Because that query was too long and a bit hard to work with, I decided to use an alias for each table instead..



Here are the results of my query..

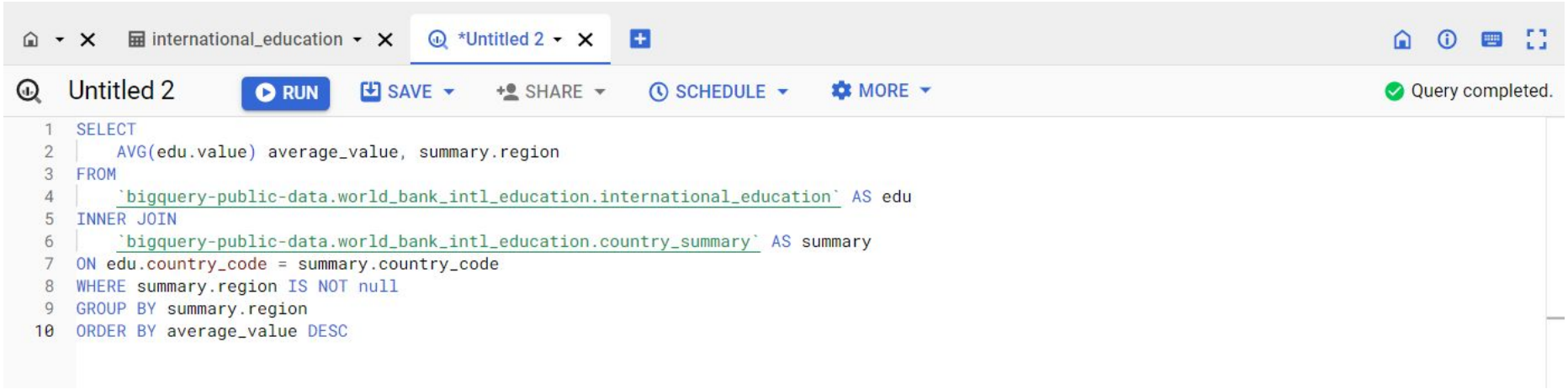


Row	country_name	country_code	value
1	Chad	TCD	135079.0
2	Chad	TCD	92857.0
3	Chad	TCD	97603.0
4	Chad	TCD	158735.0
5	Chad	TCD	302478.0
6	Chad	TCD	181979.0
7	Chad	TCD	49917.0
8	Chad	TCD	148229.0
9	Chad	TCD	74299.0
10	Chad	TCD	7174.0
11	Chad	TCD	7637.0
12	Chad	TCD	39966.0
13	Chad	TCD	105822.0
14	Chad	TCD	25141.0

DATA QUESTION

What is the average amount of money spent per region on education?

I was able to answer the data question using this query..



These were the results..

Query results

SAVE RESULTS | EXPLORE DATA

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

Row	average_value	region
1	41653437607.78...	North America
2	3882406275.688...	East Asia & Pacific
3	2696534813.883...	South Asia
4	2374149029.045...	Europe & Central Asia
5	973477665.3314...	Middle East & North Africa
6	965999068.5475...	Latin America & Caribbean
7	211692028.7065...	Sub-Saharan Africa

An alias was also set to give the AVG(edu.value) a more descriptive name for the temporary table the query returned.

The WHERE statement also excluded any rows with null information.

CONCLUSION

In conclusion, the project aimed to determine the average amount of money spent per region on education. By employing aliases and excluding null values, valuable insights into regional spending patterns were obtained. The results contribute to a better understanding of education expenditures and showcase the effectiveness of the analysis techniques employed.