ASSIGNMENT 2: DATA VISUALIZATION

Malaika Akram

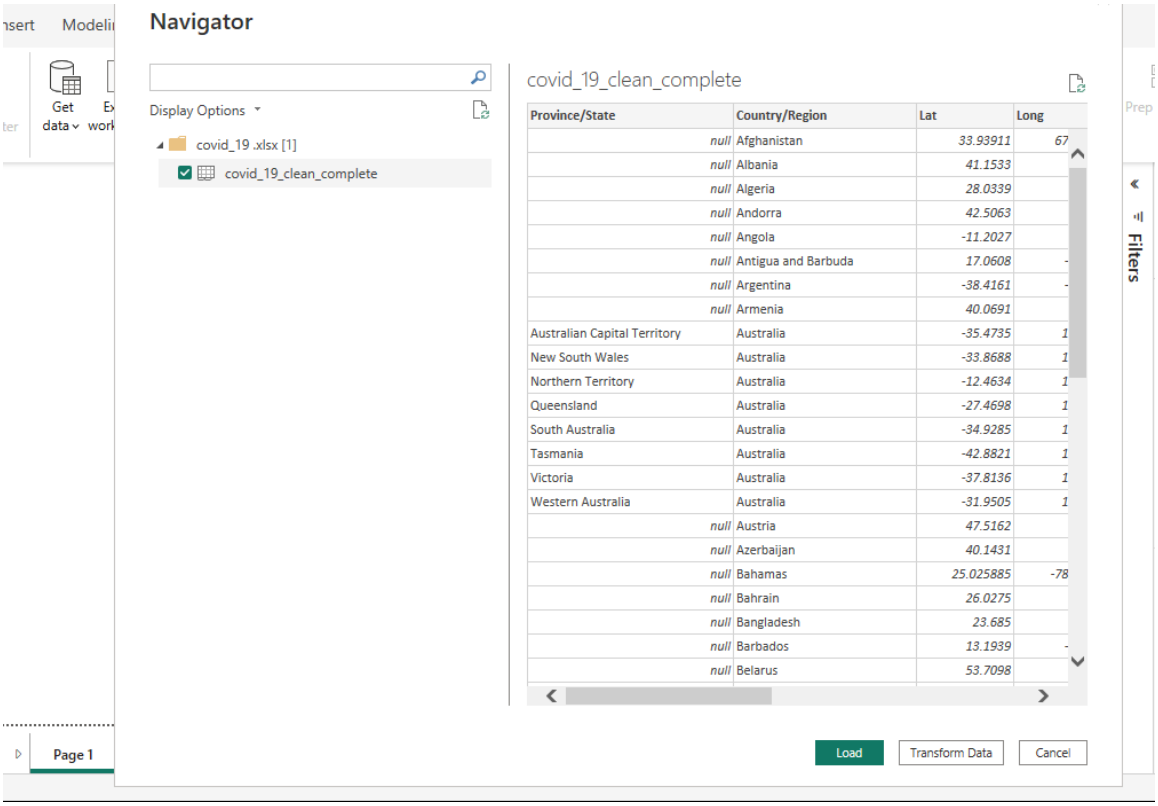
A. Load the data in the tool. Briefly explain the dataset

I used the covid_19_clean_complete.csv dataset from Kaggle. This dataset provides daily global information on COVID-19 cases from early 2020, tracking confirmed cases, deaths, recoveries, and active cases across various countries

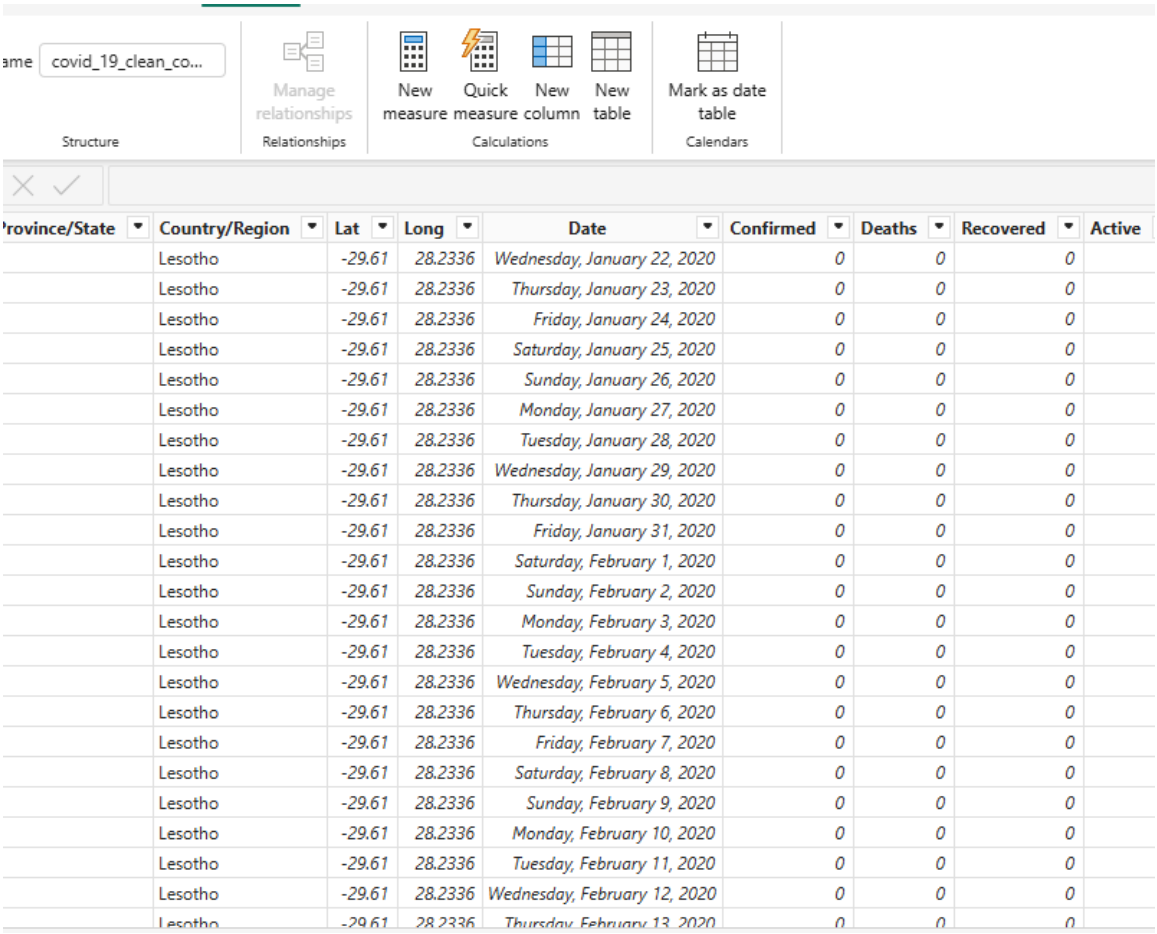
Dataset Overview

Some key pieces of information in the dataset are:
Date: The day the data was recorded
Country: The country or region being reported on
Confirmed Cases: The total number of people who tested positive for COVID-19
Deaths: The total number of people who died due to the virus
Recovered: The number of people who got better after having the virus

Active Cases: The number of people still sick (calculated as confirmed cases minus deaths and recoveries)
New Cases & New Deaths: The number of new cases or deaths reported on a specific day



Province/State	Country/Region	Lat	Long
null	Afghanistan	33.93911	67
null	Albania	41.1533	
null	Algeria	28.0339	
null	Andorra	42.5063	
null	Angola	-11.2027	
null	Antigua and Barbuda	17.0608	
null	Argentina	-38.4161	
null	Armenia	40.0691	
Australian Capital Territory	Australia	-35.4735	1
New South Wales	Australia	-33.8688	1
Northern Territory	Australia	-12.4634	1
Queensland	Australia	-27.4698	1
South Australia	Australia	-34.9285	1
Tasmania	Australia	-42.8821	1
Victoria	Australia	-37.8136	1
Western Australia	Australia	-31.9505	1
null	Austria	47.5162	
null	Azerbaijan	40.1431	
null	Bahamas	25.025885	-78
null	Bahrain	26.0275	
null	Bangladesh	23.685	
null	Barbados	13.1939	
null	Belarus	53.7098	

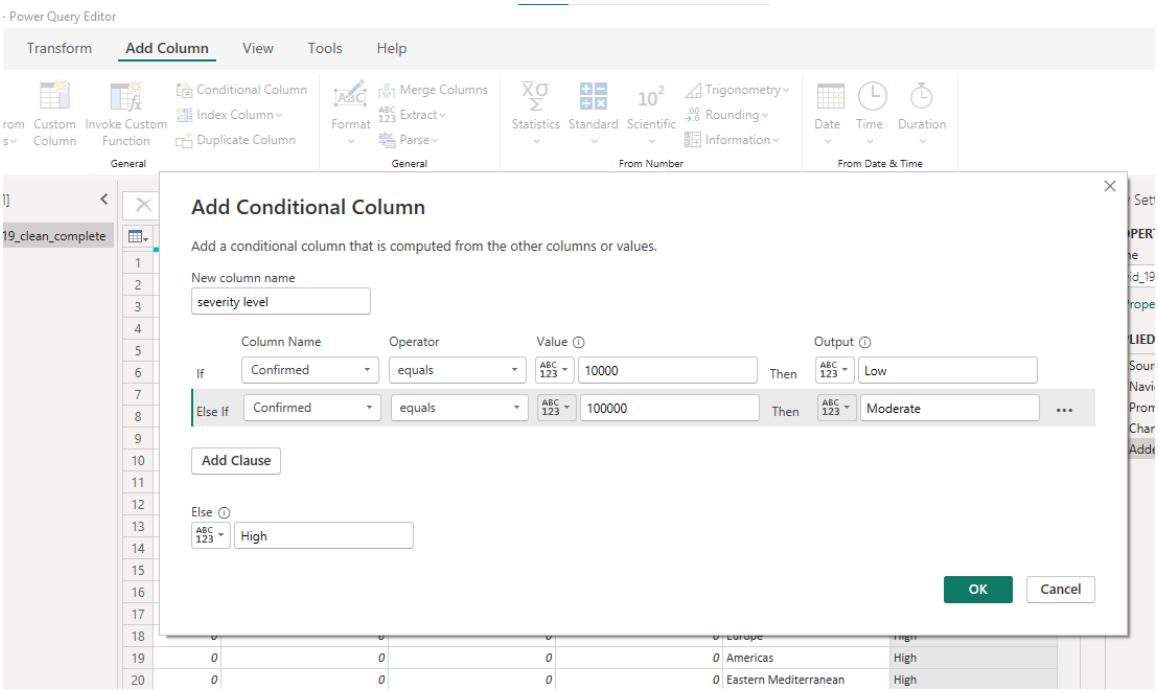


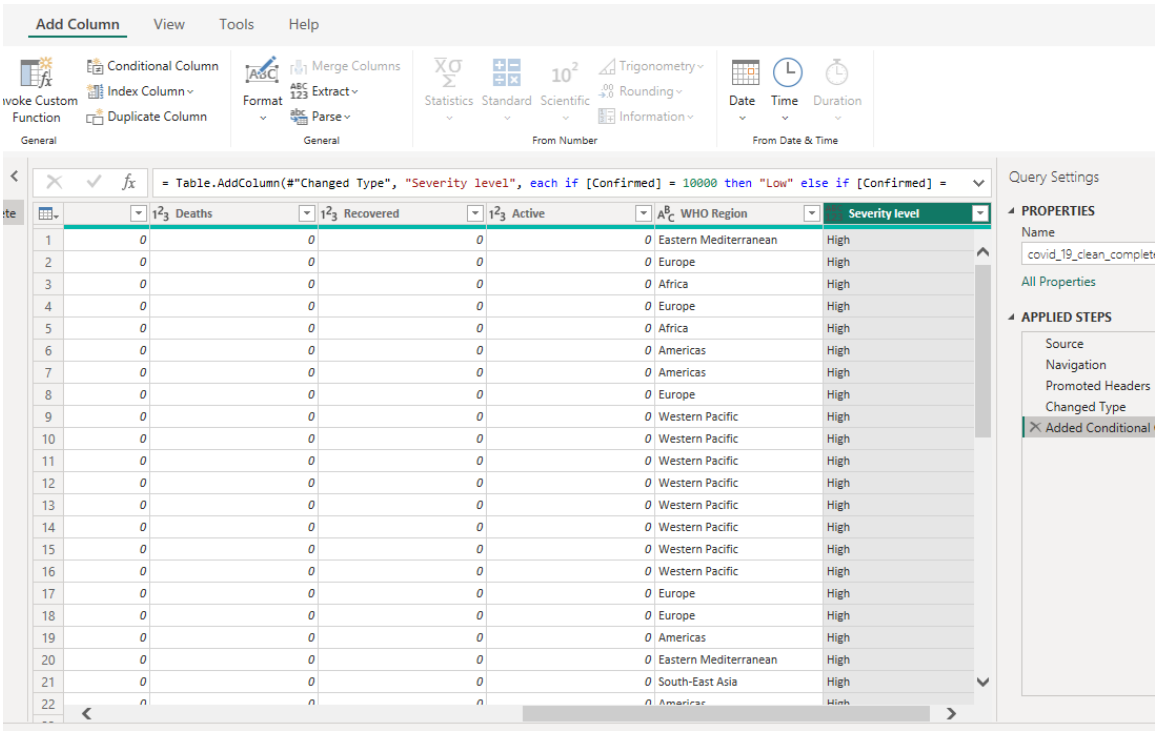
rovince/State	Country/Region	Lat	Long	Date	Confirmed	Deaths	Recovered	Active
	Lesotho	-29.61	28.2336	Wednesday, January 22, 2020	0	0	0	
	Lesotho	-29.61	28.2336	Thursday, January 23, 2020	0	0	0	
	Lesotho	-29.61	28.2336	Friday, January 24, 2020	0	0	0	
	Lesotho	-29.61	28.2336	Saturday, January 25, 2020	0	0	0	
	Lesotho	-29.61	28.2336	Sunday, January 26, 2020	0	0	0	
	Lesotho	-29.61	28.2336	Monday, January 27, 2020	0	0	0	
	Lesotho	-29.61	28.2336	Tuesday, January 28, 2020	0	0	0	
	Lesotho	-29.61	28.2336	Wednesday, January 29, 2020	0	0	0	
	Lesotho	-29.61	28.2336	Thursday, January 30, 2020	0	0	0	
	Lesotho	-29.61	28.2336	Friday, January 31, 2020	0	0	0	
	Lesotho	-29.61	28.2336	Saturday, February 1, 2020	0	0	0	
	Lesotho	-29.61	28.2336	Sunday, February 2, 2020	0	0	0	
	Lesotho	-29.61	28.2336	Monday, February 3, 2020	0	0	0	
	Lesotho	-29.61	28.2336	Tuesday, February 4, 2020	0	0	0	
	Lesotho	-29.61	28.2336	Wednesday, February 5, 2020	0	0	0	
	Lesotho	-29.61	28.2336	Thursday, February 6, 2020	0	0	0	
	Lesotho	-29.61	28.2336	Friday, February 7, 2020	0	0	0	
	Lesotho	-29.61	28.2336	Saturday, February 8, 2020	0	0	0	
	Lesotho	-29.61	28.2336	Sunday, February 9, 2020	0	0	0	
	Lesotho	-29.61	28.2336	Monday, February 10, 2020	0	0	0	
	Lesotho	-29.61	28.2336	Tuesday, February 11, 2020	0	0	0	
	Lesotho	-29.61	28.2336	Wednesday, February 12, 2020	0	0	0	
	Lesotho	-29.61	28.2336	Thursday, February 13, 2020	0	0	0	

B. Transform the data (Conditional/Custom Columns etc.)

Conditional Column

I added a column that labels the situation in each country as “Low”, “Moderate”, or “High” based on the number of confirmed cases using a conditional rule:
If Confirmed Cases < 10,000 → Low
If between 10,000 and 100,000 → Moderate
If > 100,000 → High

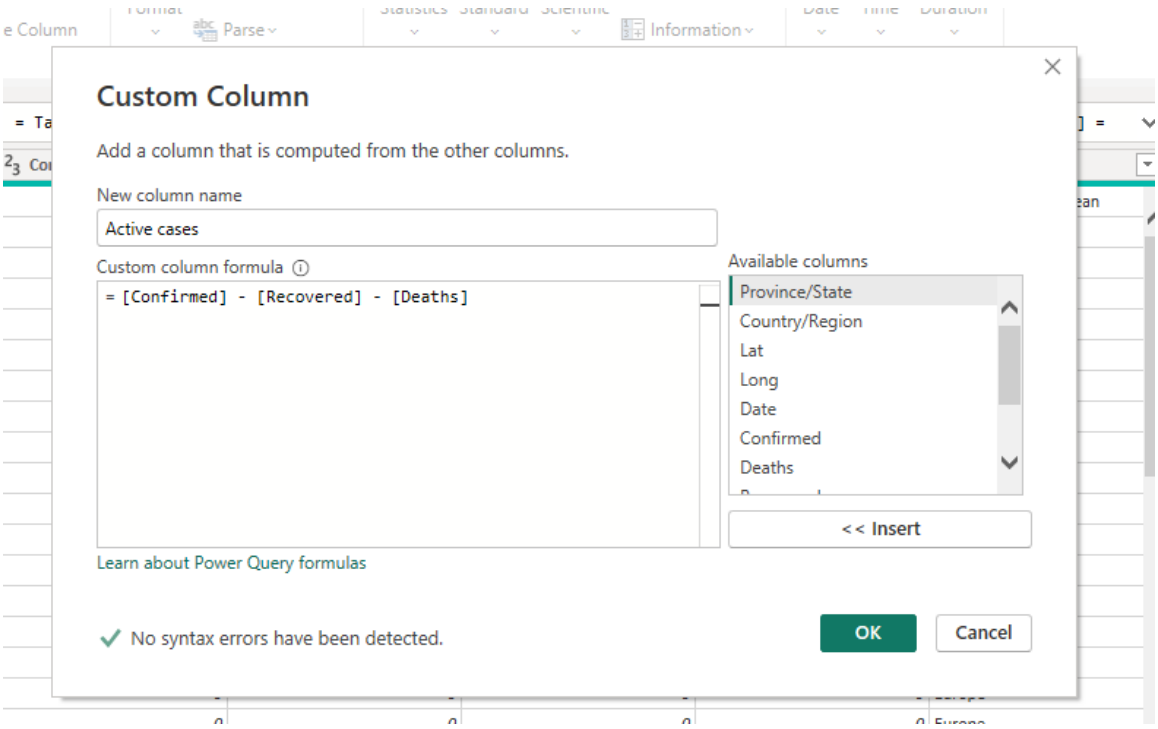


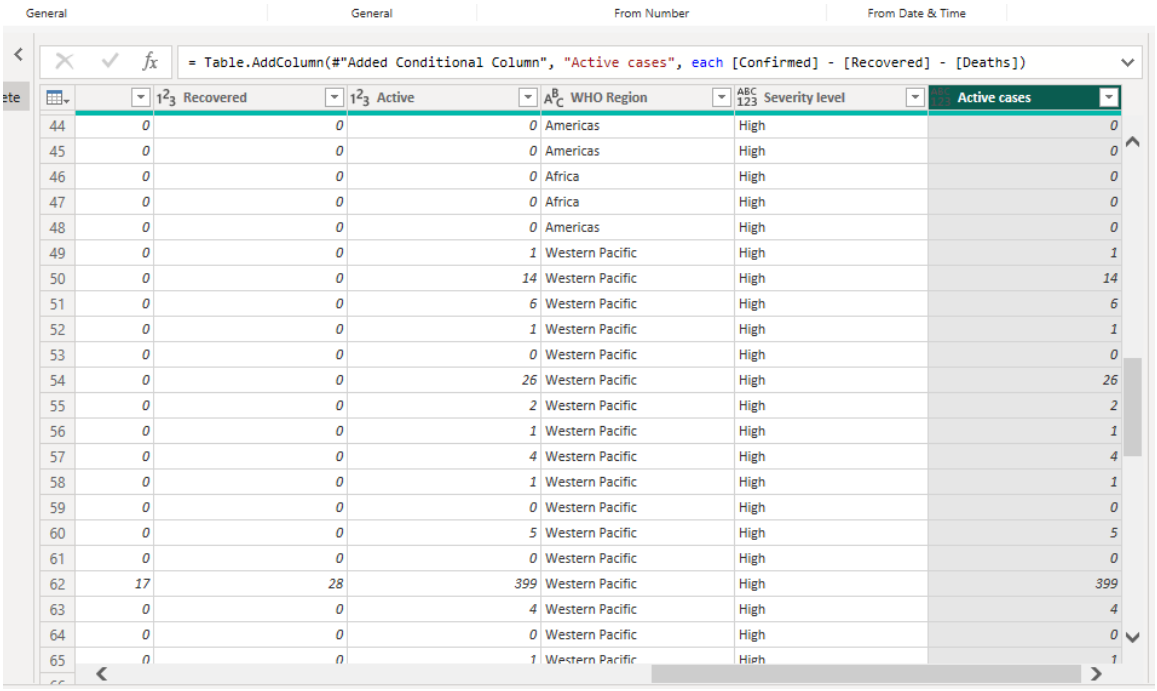


Custom column

I added a new column called Active Cases to calculate the number of currently infected people. The formula used was:

Active Cases = Confirmed Cases - Recovered - Deaths

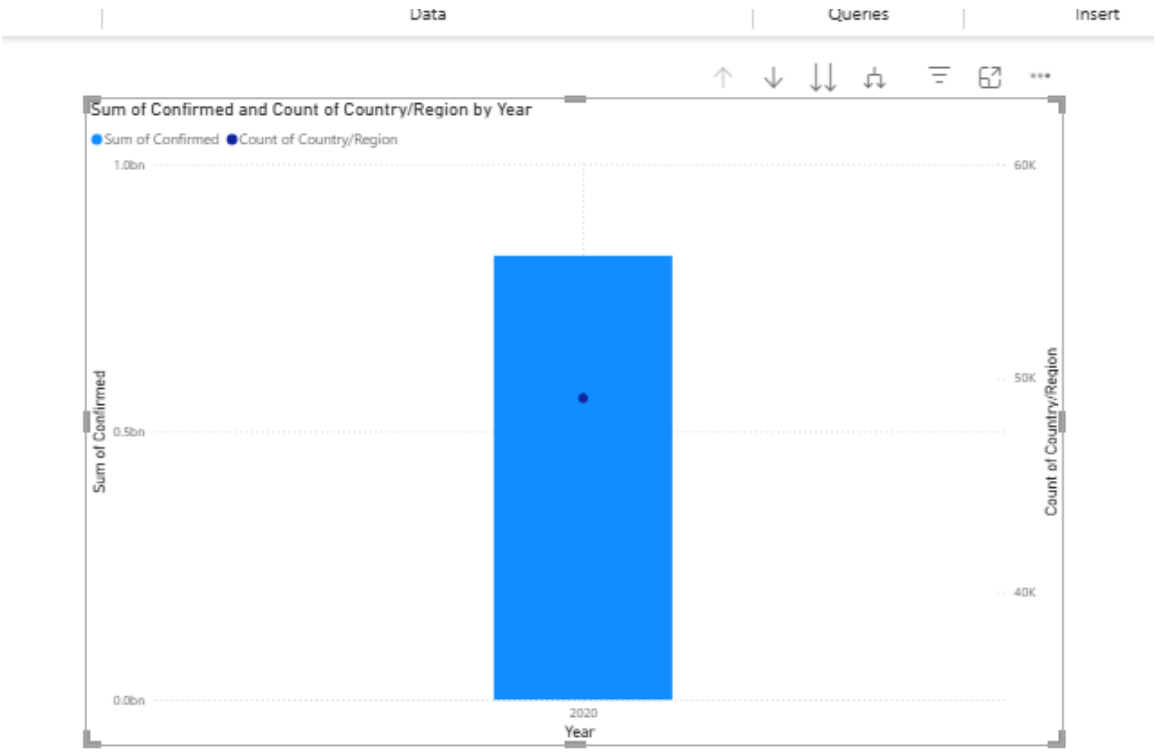




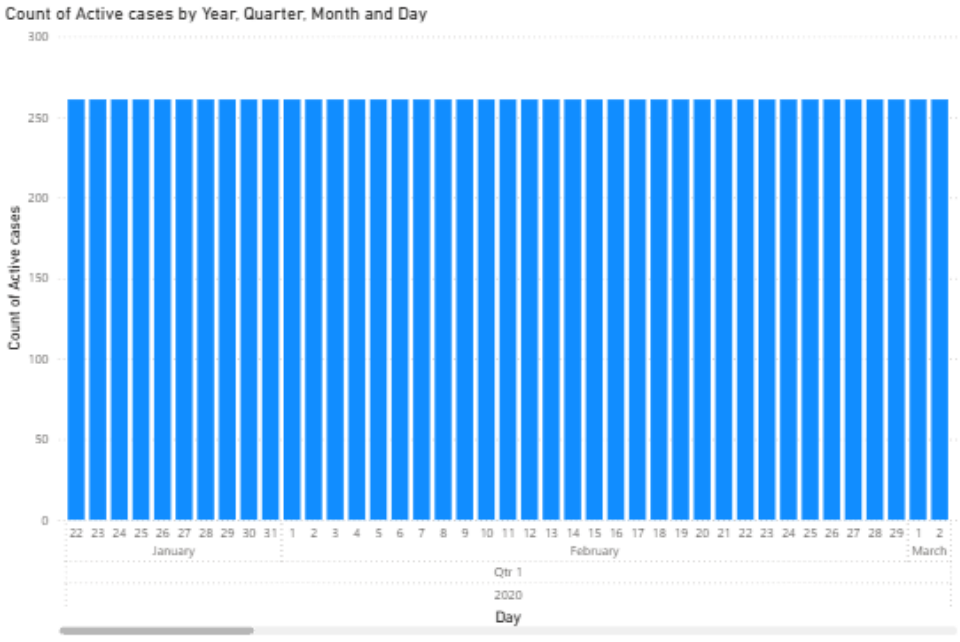
	Recovered	Active	WHO Region	Severity level	Active cases
44	0	0	0 Americas	High	0
45	0	0	0 Americas	High	0
46	0	0	0 Africa	High	0
47	0	0	0 Africa	High	0
48	0	0	0 Americas	High	0
49	0	0	1 Western Pacific	High	1
50	0	0	14 Western Pacific	High	14
51	0	0	6 Western Pacific	High	6
52	0	0	1 Western Pacific	High	1
53	0	0	0 Western Pacific	High	0
54	0	0	26 Western Pacific	High	26
55	0	0	2 Western Pacific	High	2
56	0	0	1 Western Pacific	High	1
57	0	0	4 Western Pacific	High	4
58	0	0	1 Western Pacific	High	1
59	0	0	0 Western Pacific	High	0
60	0	0	5 Western Pacific	High	5
61	0	0	0 Western Pacific	High	0
62	17	28	399 Western Pacific	High	399
63	0	0	4 Western Pacific	High	4
64	0	0	0 Western Pacific	High	0
65	0	0	1 Western Pacific	High	1

C. Build multiple visualization charts

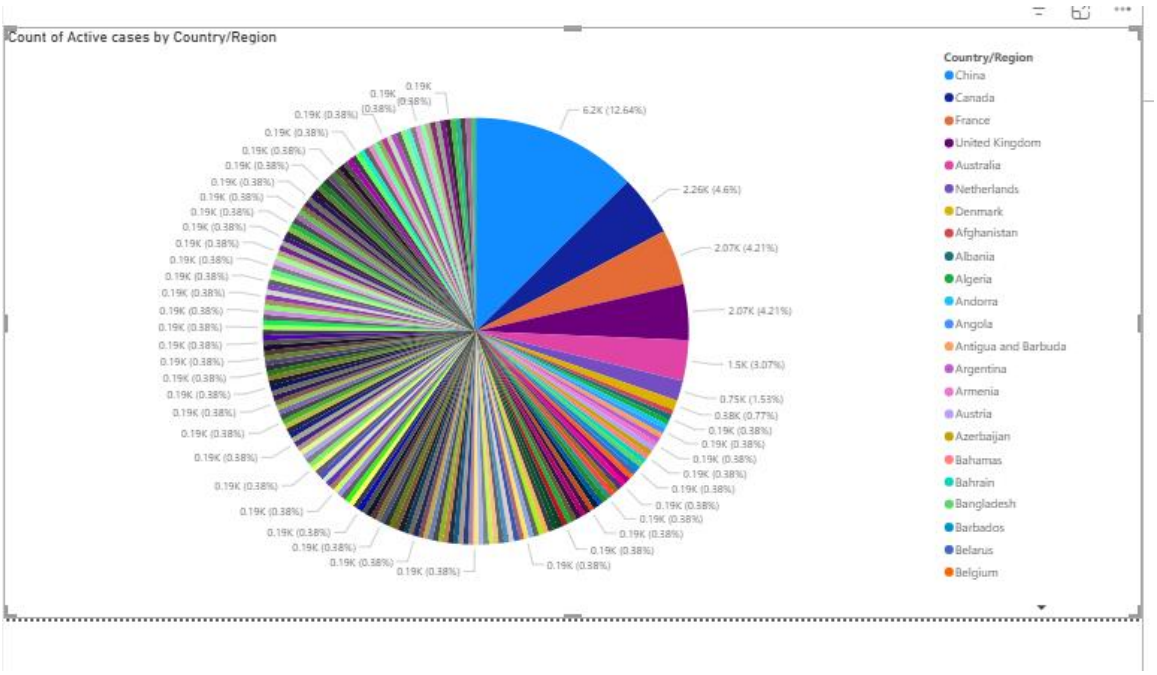
Line Chart – Confirmed Cases Over Time



Line Chart – Active Cases



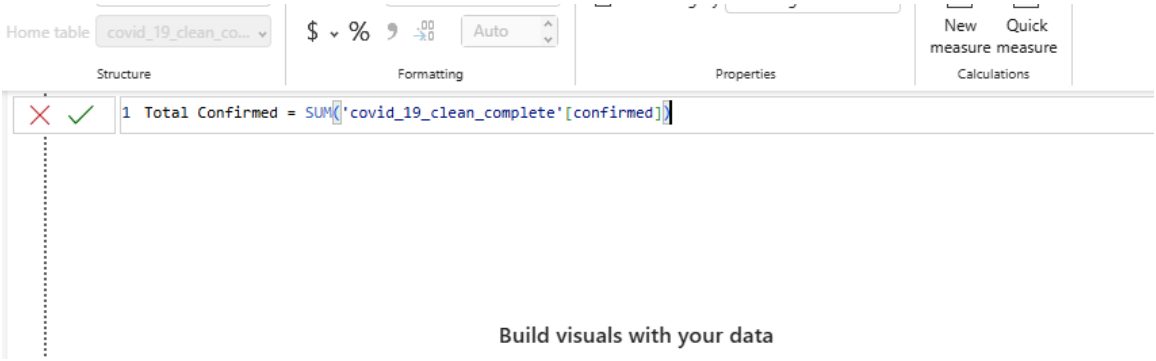
Pie Chart



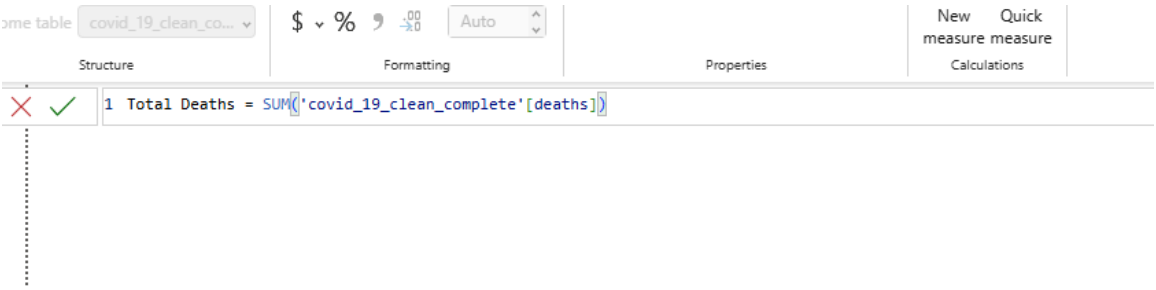
D. Build Measures and DAX functions

To perform deeper analysis on the COVID-19 dataset, I created several DAX measures in Power BI. These allow for real-time calculations based on the data shown in charts or filtered by slicers

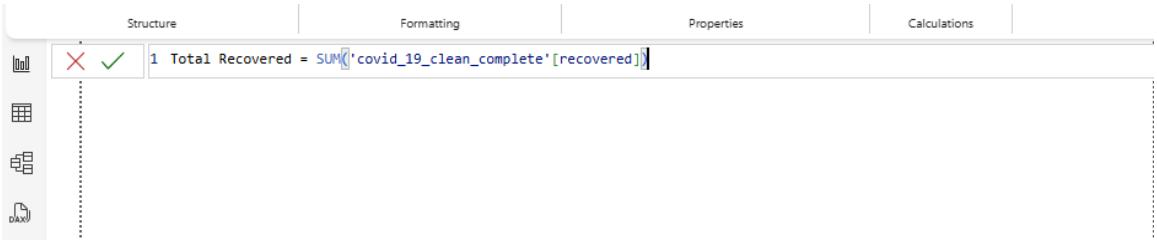
Total confirmed



Total Deaths



Total Recovered



Overall usefulness and interactive nature of the dashboard

Usefulness:

Summarizes key data like total confirmed cases, total deaths, recoveries, and active cases in one view.
Helps users identify trends through line charts (e.g., confirmed cases over time).
Allows comparison between countries, showing which were most affected.
Shows the death rate and recovery rate, which gives deeper insights into health system outcomes.
The map visual makes it easy to understand which regions were more impacted globally.

Interactivity:

I added slicers for country and date range so users can:
Focus on a specific country's data
Analyze data during a selected time period
All visuals automatically update based on the filters selected.
This allows the user to perform their own analysis without changing the report manually.