Global Cost of Living

Dataset source: https://www.kaggle.com/datasets/mvieira101/global-cost-of-living

This dataset contains information about the cost of living in almost 5000 cities across the world. The data were gathered by scraping Numbeo's website (https://www.numbeo.com).

Data Cleaning and Transformation using Excel and Power Query Editor in Power BI

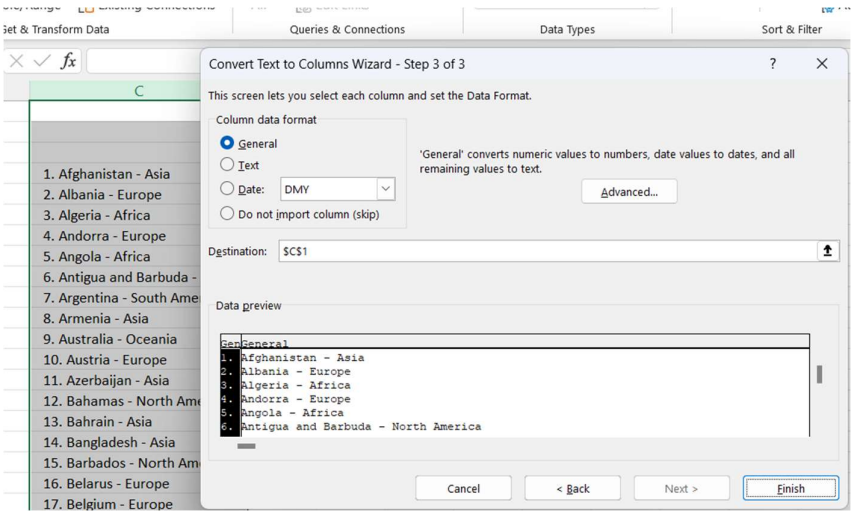
1- Using the data dictionary I replaced the columns headers with the actual description of each cost.

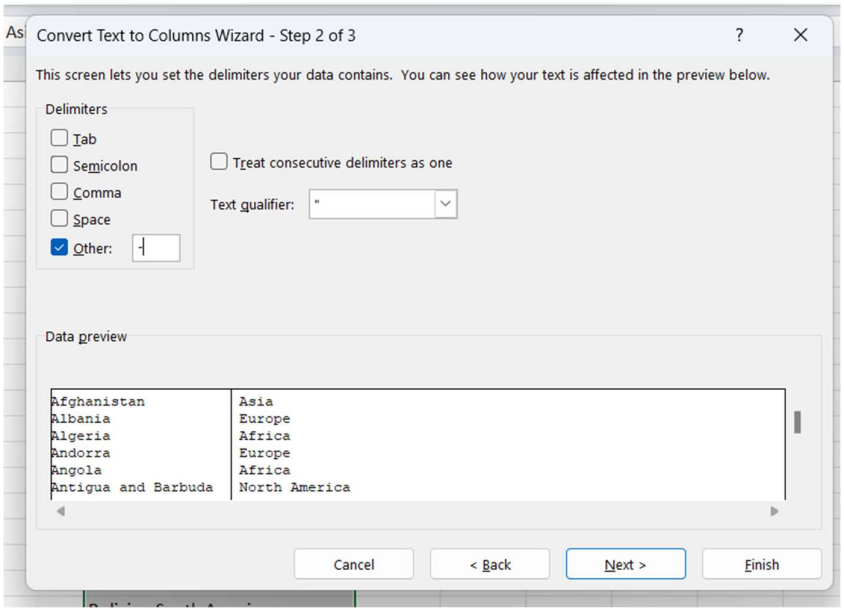
I did that copying the data dictionary into an excel sheet and copied and pasted the info in the headers using the feature transpose under the copy and paste special to transpose the columns list from rows to columns.

2- In order to add a new column to my dataset for Continents in a quick and smart way, here is what I did:

I got a list of countries and its respective continent online and copied the countries list to a new tab in my spreadsheet.

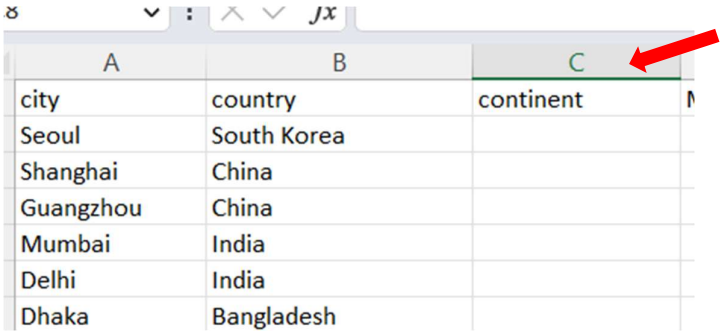
Using the 'Text to columns' feature I first split the numbers out of the list using fixed width and then split the country and continent using the dash '-' as delimiter.



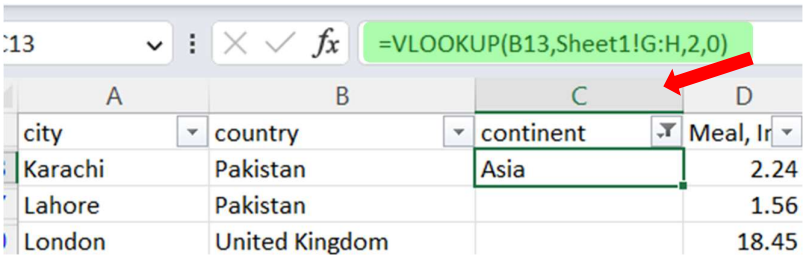


Trimming columns D and E for Country name and Continent in columns G and H

Added a column in the dataset mapping the continent for each country using a VLOOKUP based on country.



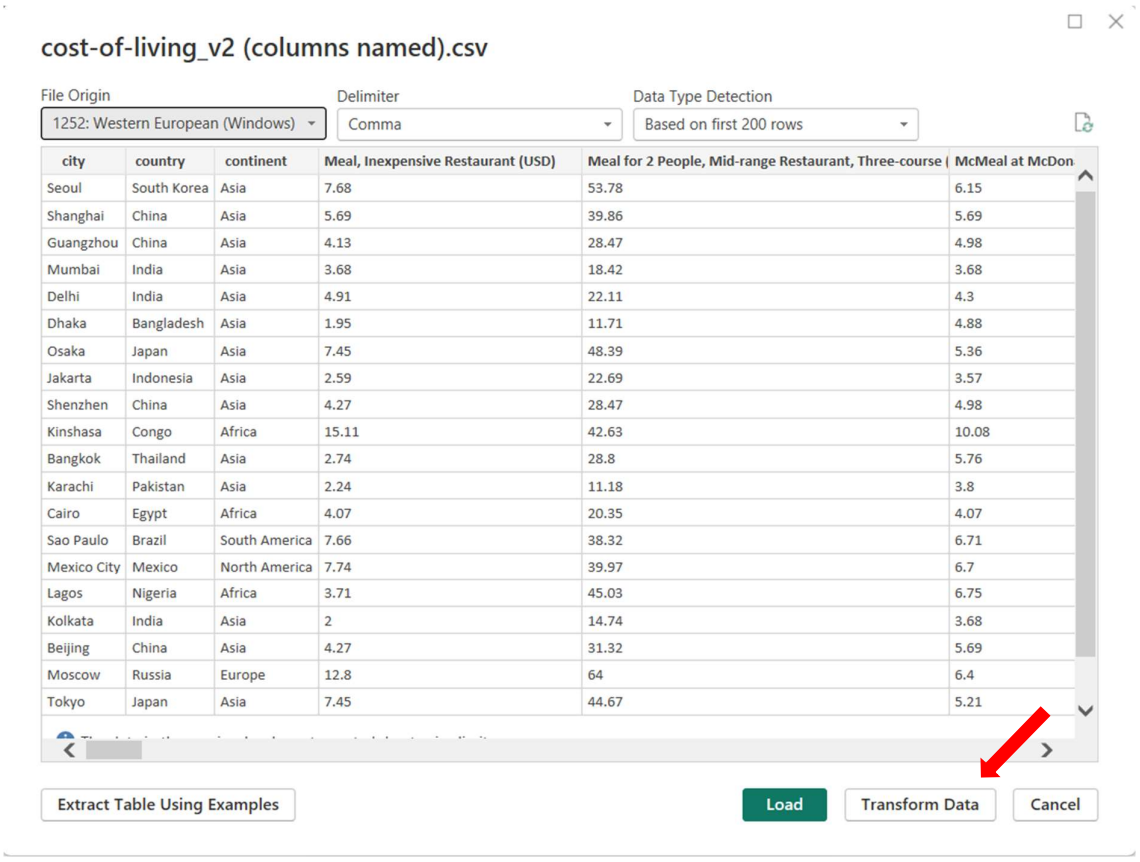
Start populating the Continent based on the list created in the other sheet on columns G (Country) and H (Continent)



Out of 4956 rows only 69 were not found and I will either look them up manually or adjust the name of the country in the table to fix those.

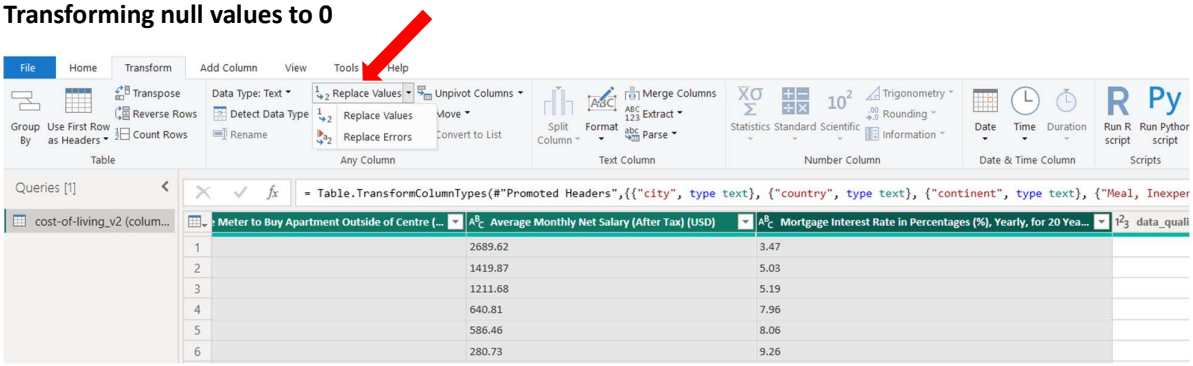
3- After the continent column is created, upload dataset into Power BI.

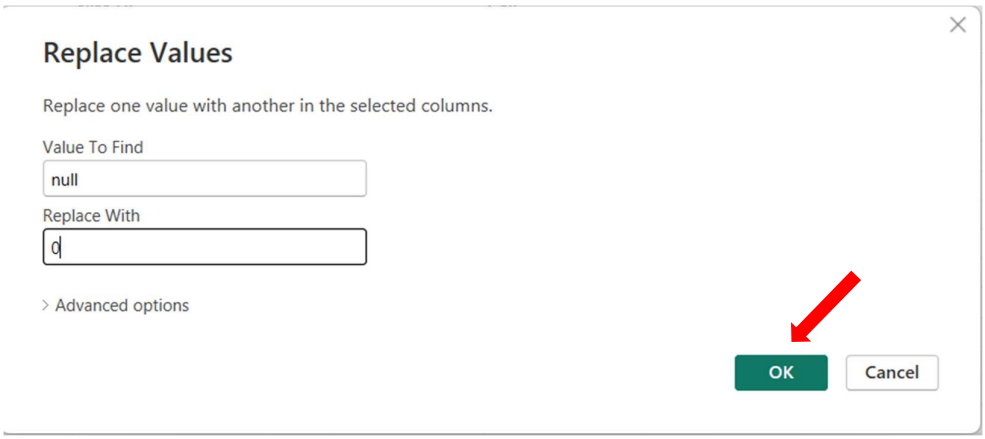
4- Transforming the data using Power BI.



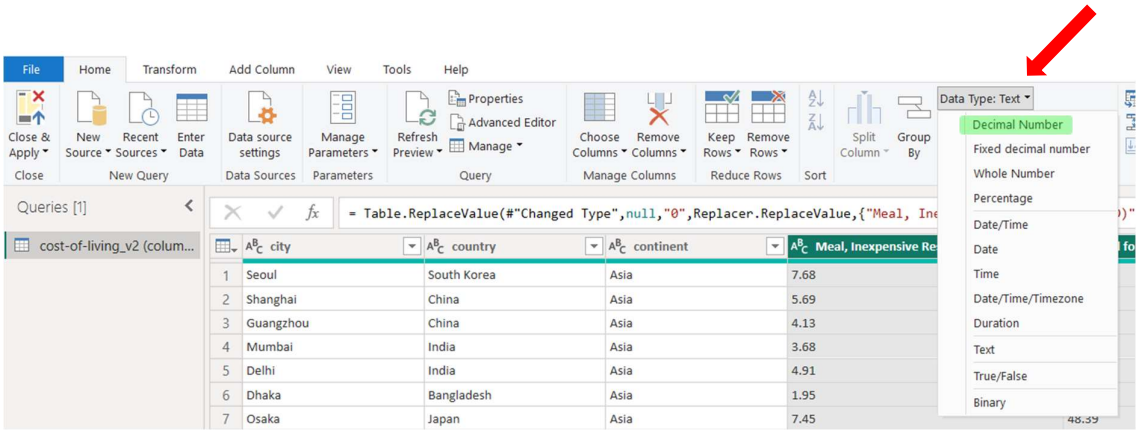
city	country	continent	Meal, Inexpensive Restaurant (USD)	Meal for 2 People, Mid-range Restaurant, Three-course (	McMeal at McDon
Seoul	South Korea	Asia	7.68	53.78	6.15
Shanghai	China	Asia	5.69	39.86	5.69
Guangzhou	China	Asia	4.13	28.47	4.98
Mumbai	India	Asia	3.68	18.42	3.68
Delhi	India	Asia	4.91	22.11	4.3
Dhaka	Bangladesh	Asia	1.95	11.71	4.88
Osaka	Japan	Asia	7.45	48.39	5.36
Jakarta	Indonesia	Asia	2.59	22.69	3.57
Shenzhen	China	Asia	4.27	28.47	4.98
Kinshasa	Congo	Africa	15.11	42.63	10.08
Bangkok	Thailand	Asia	2.74	28.8	5.76
Karachi	Pakistan	Asia	2.24	11.18	3.8
Cairo	Egypt	Africa	4.07	20.35	4.07
Sao Paulo	Brazil	South America	7.66	38.32	6.71
Mexico City	Mexico	North America	7.74	39.97	6.7
Lagos	Nigeria	Africa	3.71	45.03	6.75
Kolkata	India	Asia	2	14.74	3.68
Beijing	China	Asia	4.27	31.32	5.69
Moscow	Russia	Europe	12.8	64	6.4
Tokyo	Japan	Asia	7.45	44.67	5.21

Transforming null values to 0

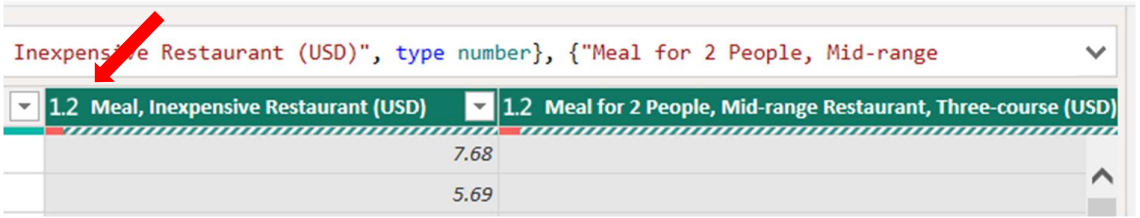




Changing these columns data type from text to decimal numbers



Columns are now decimal numbers 😊



After data is transformed it’s time to load the data to power BI and start visualizing the data.