1. I’ve retrieved daily stock data from all global indices from 1920s (where available) all the way to 2023. The data is already very clean – and does not have any apparent null / error values. This is a promising start, but some transformation is required to ensure that the data can be well interfaced.A screenshot of a graph

   Description automatically generated
2. To help make the database easier to query, and increase organisation, I’ve decided to make a new table that contains the indices ticker as the primary key, as well as the indices full name, date of creation, and other meta details.

In order to do this, I need a way to source all the ticker information for all the indices. There does not appear to be a publicly available dataset of ticker names corresponding to indices (there is one for each stock), therefore, I chose to use Python to begin web scraping on Yahoo Finance. This allows me to gather information about each index, such as the name, ticker name, and other information.

Please refer to the code ‘scraper.py’ alongside the output ‘world\_indices.csv’ to see the information scraped.

1. I then assigned the relationship between the merged table and this new table by creating a one-many relationship using the ticker attribute as the foreign key.

A screenshot of a computer

Description automatically generated

We no longer need the original tables. We cannot delete them as the form the basis of our new tables, but we can at least hide them from the report view and relocate them into a new group in the Power Query Editor.

A screenshot of a search box

Description automatically generated A screenshot of a computer

Description automatically generated

A white rectangular sign with black text

Description automatically generated