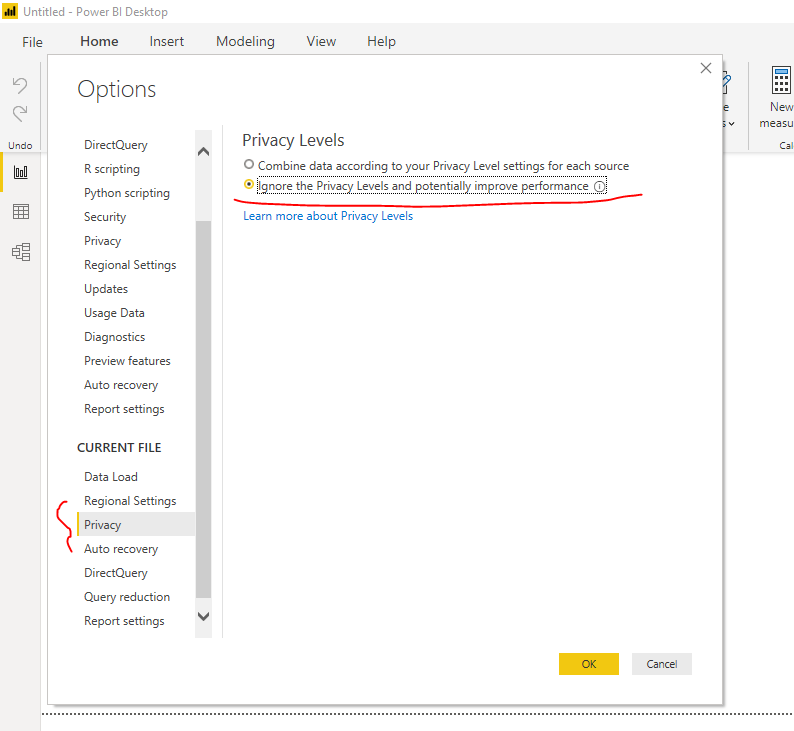
**Загрузка и трансформация данных для построения отчетов QualityOfLife**

1. В новом файле отчета включаем следующую опцию:

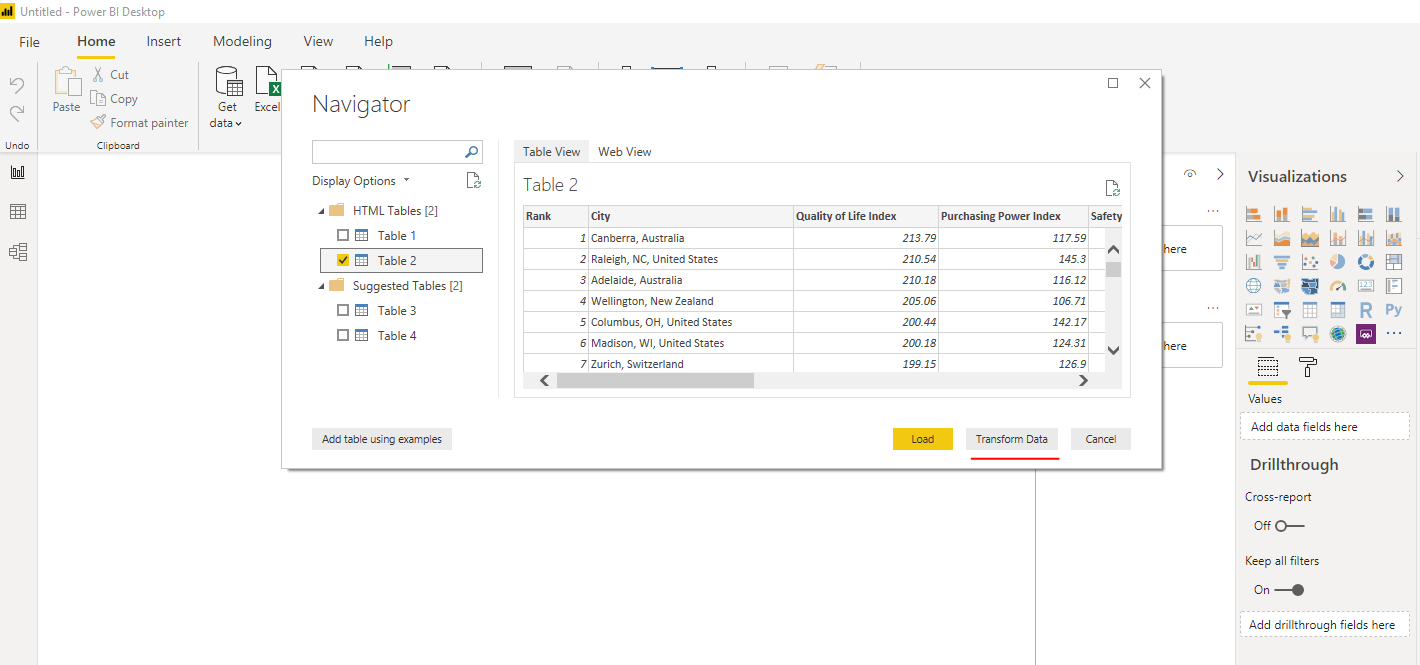


# Необходимо получить данные с сайта NUMBEO.COM, раздел Quality of Life Index by City (Изучите этот раздел, данные в таблице меняются в зависимости от выбранного года, и год отражается в адресной строке):

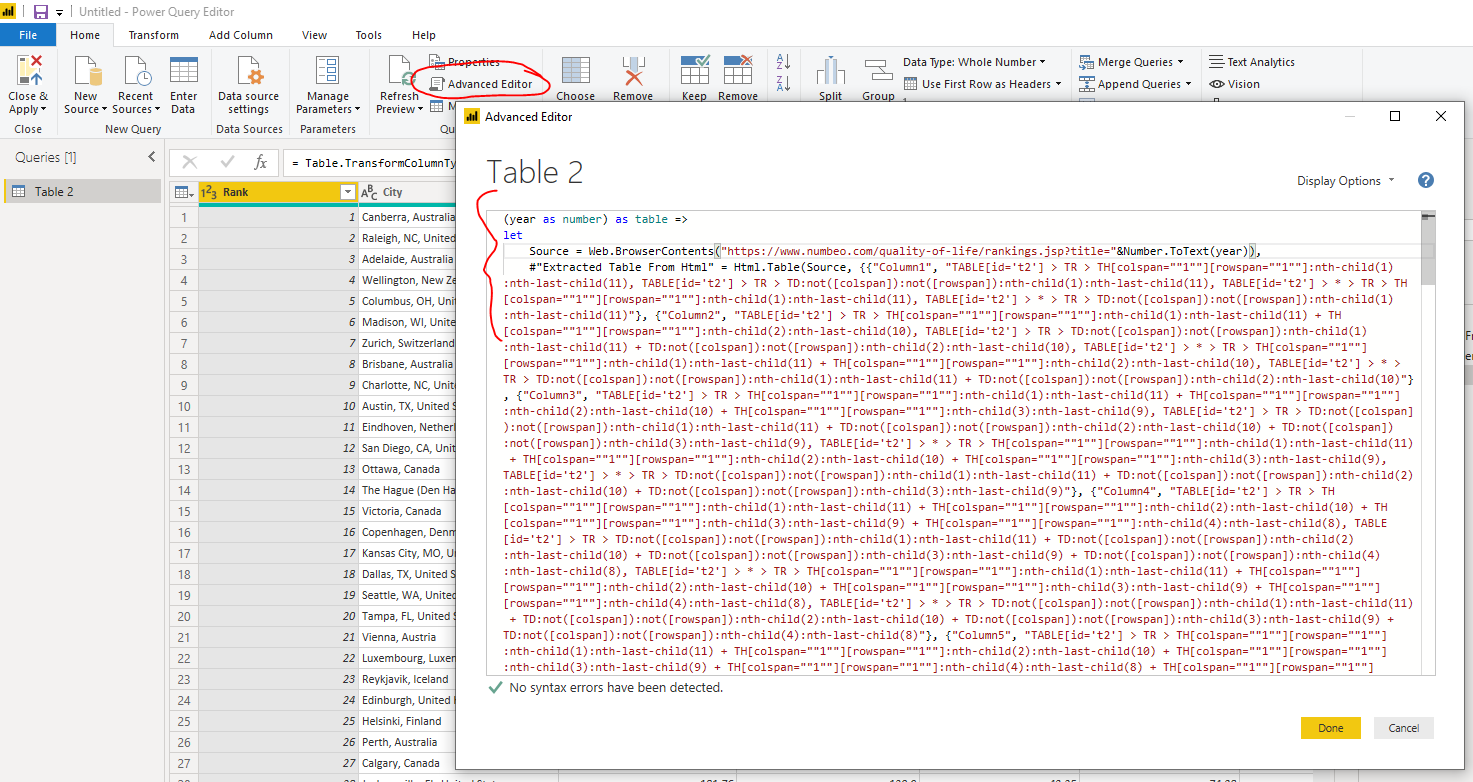
GetData -> выбираем тип источника Web -> указываем адрес <https://www.numbeo.com/quality-of-life/rankings.jsp?title=2020>

- это данные по индексам за 2020 год.

1. Выбираем нужную таблицу в открывшемся окне и нажимаем кнопку Transform Data:

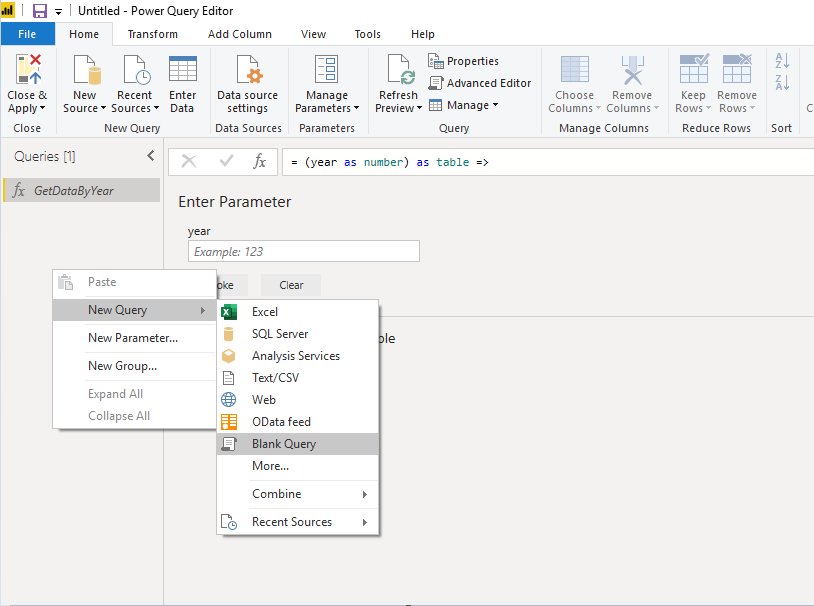


1. Создаем из запроса Table 2 функцию GetDataByYear. Открываем код запроса в Advanced Editor, добавляем параметр **year** и встраиваем его в уже существующий код, не забываем переименовать query Table 2 в GetDataByYear:



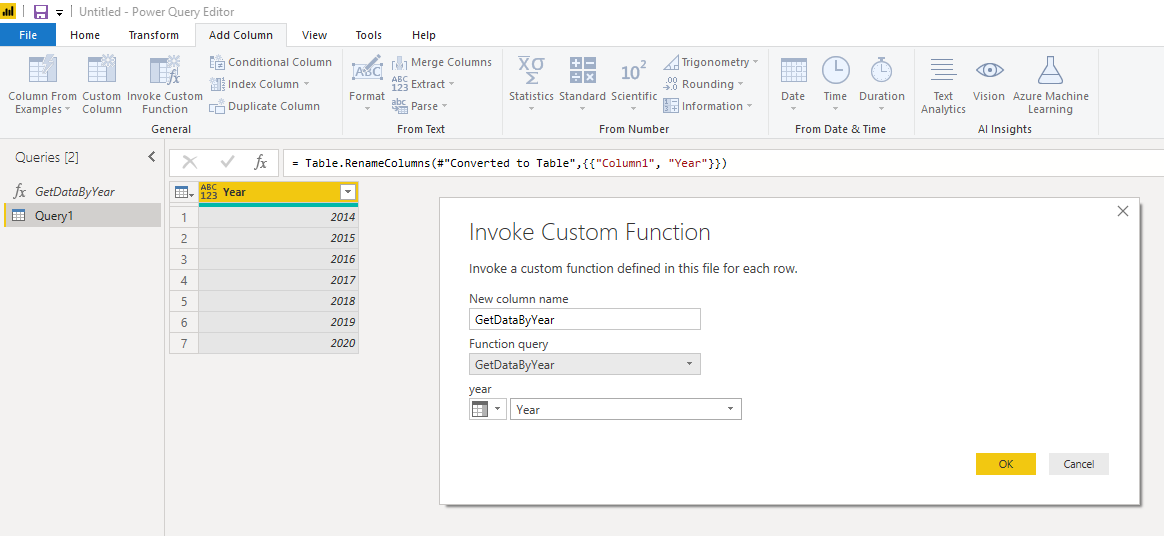
1. Создаем новый запрос, основу будущей таблицы FactQualityOfLife. Создаем пустой запрос, в котором самостоятельно вводим выражение:

***= {2014..2020}***

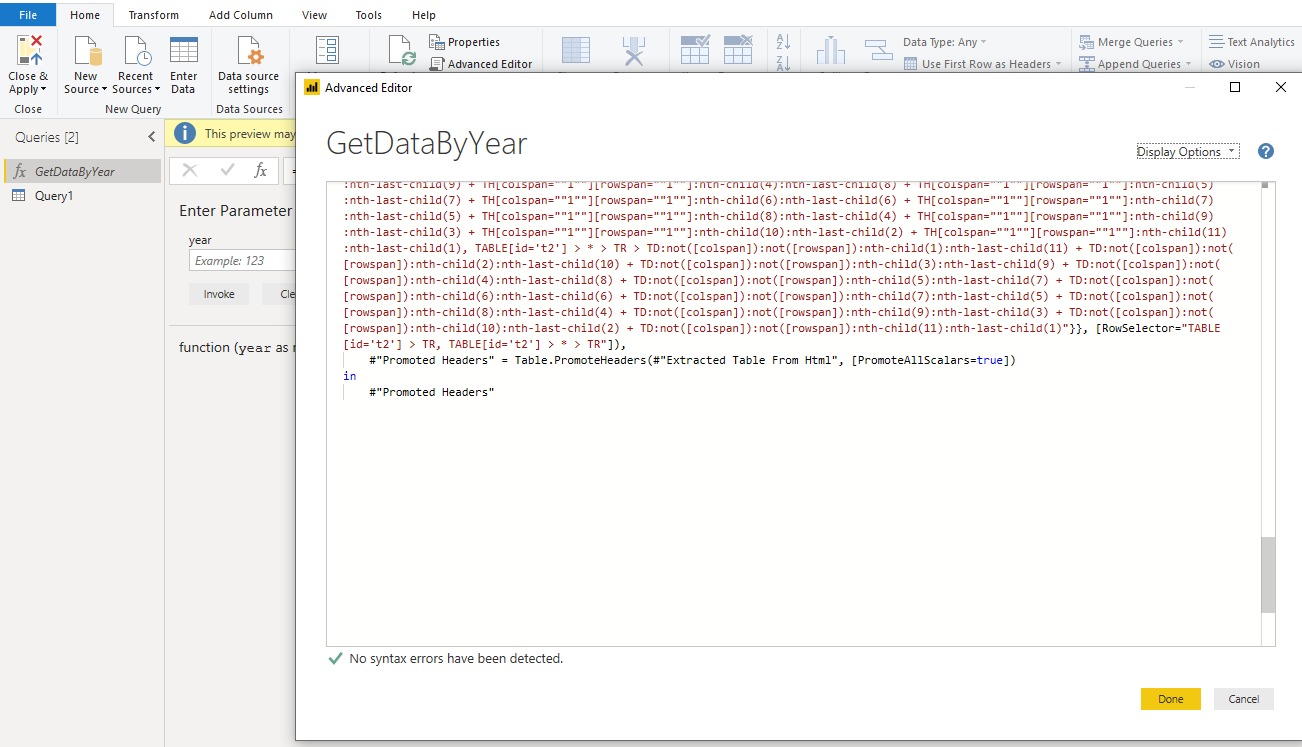


Получаем список от 2014 до 2020. Конвертируем его в таблицу, используя функцию ToTable верхней панели. Переименовываем колонку в Year.

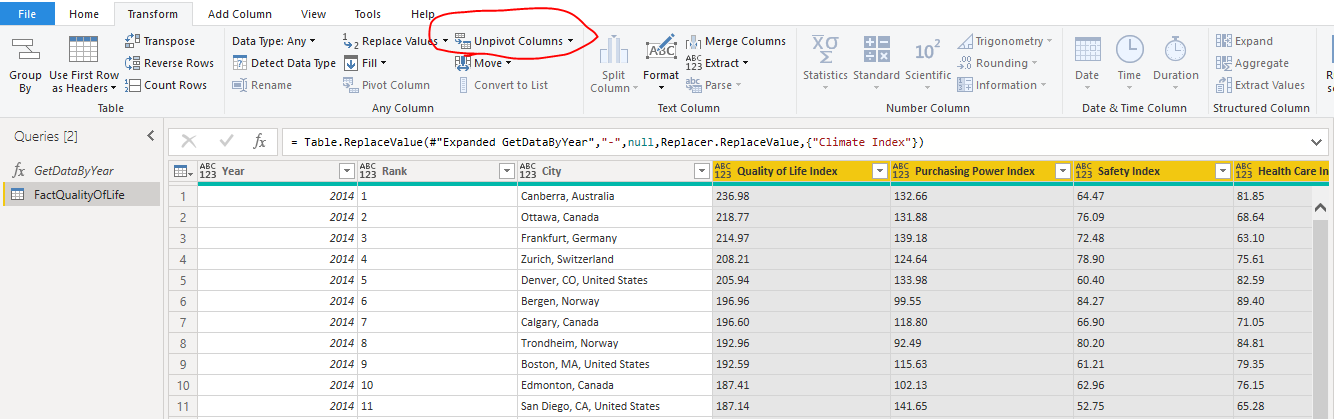
1. Загружаем данные за все годы от 2014 до 2020. В меню Add Column выбираем Invoke Custom Function, в качестве функции выбираем GetDataByYear, а в качестве параметра – значения колонки Year. Разворачиваем колонки таблицы:



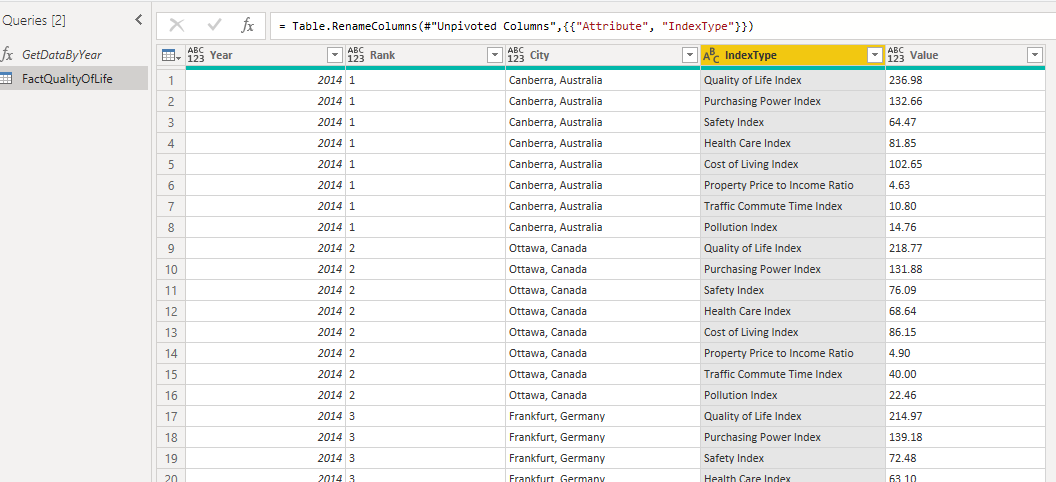
1. Устраняем ошибку в Climate Index. Открываем Advanced Editor для функции GetDataByYear и убираем трансформацию "Changed Type". Должно получиться следующее:



1. Заменяем в Query1 в колонке Climate Index знак «-» на NULL, используя функцию Replace Value. Переименовываем таблицу в FactQualityOfLife
2. Разворачиваем данные по индексам в таблице FactQualityOfLife: выделяем все столбцы с индексами и применяем трансформацию Unpivot:



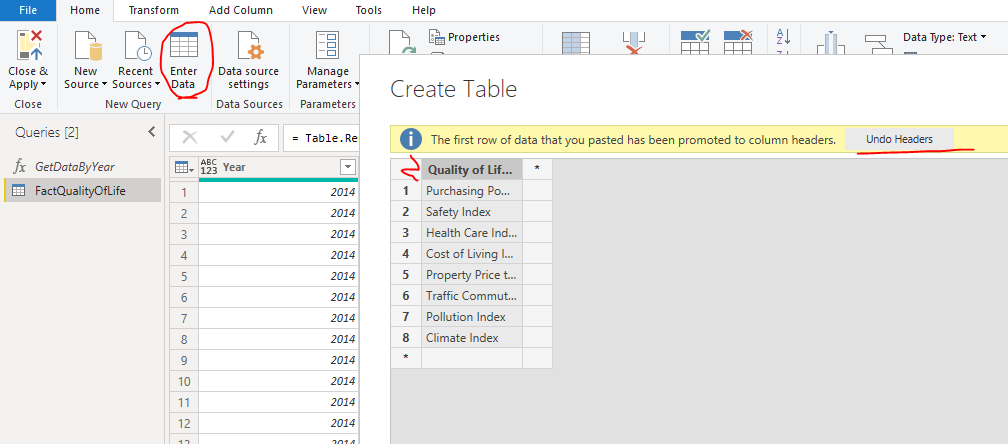
Переименовываем новый столбец в IndexType:



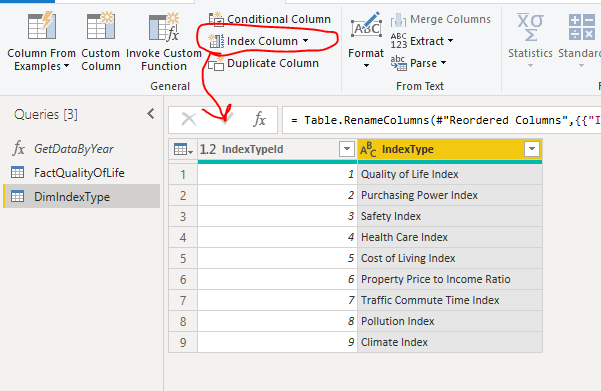
1. Создаем DimDate точно так же, как это описано в пункте 4
2. Создаем DimIndexType, используя Enter Data и следующий список индексов:

|  |
| --- |
| Quality of Life Index |
| Purchasing Power Index |
| Safety Index |
| Health Care Index |
| Cost of Living Index |
| Property Price to Income Ratio |
| Traffic Commute Time Index |
| Pollution Index |
| Climate Index |

Первую строку PowerBI попытается использовать в качестве header, это действие нужно отменить:



1. Создаем в новой таблице индексов DimIndexType столбец с идентификаторами, используя Index Column функцию (в настройках нужно указать, что начинаем от 1):



1. Создание DimLocation. Повторяем шаги 4 и 5. Удаляем из получившегося датасета ВСЕ стоблцы, кроме City.

В Advanced Editor для данного запроса добавляем операцию Distinct. Должен получиться следующий запрос (разберитесь, что в нем происходит):

***let***

***Source = {2014..2020},***

***#"Converted to Table" = Table.FromList(Source, Splitter.SplitByNothing(), null, null, ExtraValues.Error),***

***#"Invoked Custom Function" = Table.AddColumn(#"Converted to Table", "GetDataByYear", each GetDataByYear([Column1])),***

***#"Expanded GetDataByYear" = Table.ExpandTableColumn(#"Invoked Custom Function", "GetDataByYear", {"Rank", "City", "Quality of Life Index", "Purchasing Power Index", "Safety Index", "Health Care Index", "Cost of Living Index", "Property Price to Income Ratio", "Traffic Commute Time Index", "Pollution Index", "Climate Index"}, {"Rank", "City", "Quality of Life Index", "Purchasing Power Index", "Safety Index", "Health Care Index", "Cost of Living Index", "Property Price to Income Ratio", "Traffic Commute Time Index", "Pollution Index", "Climate Index"}),***

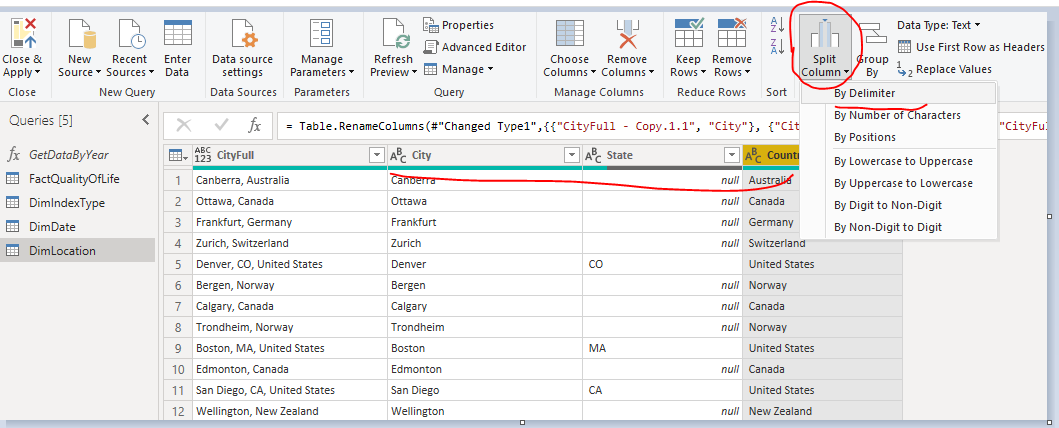
***#"Removed Columns" = Table.RemoveColumns(#"Expanded GetDataByYear",{"Column1", "Rank", "Quality of Life Index", "Purchasing Power Index", "Safety Index", "Health Care Index", "Cost of Living Index", "Property Price to Income Ratio", "Traffic Commute Time Index", "Pollution Index", "Climate Index"}),***

***#"Distinct Columns" = Table.Distinct(#"Removed Columns", "City")***

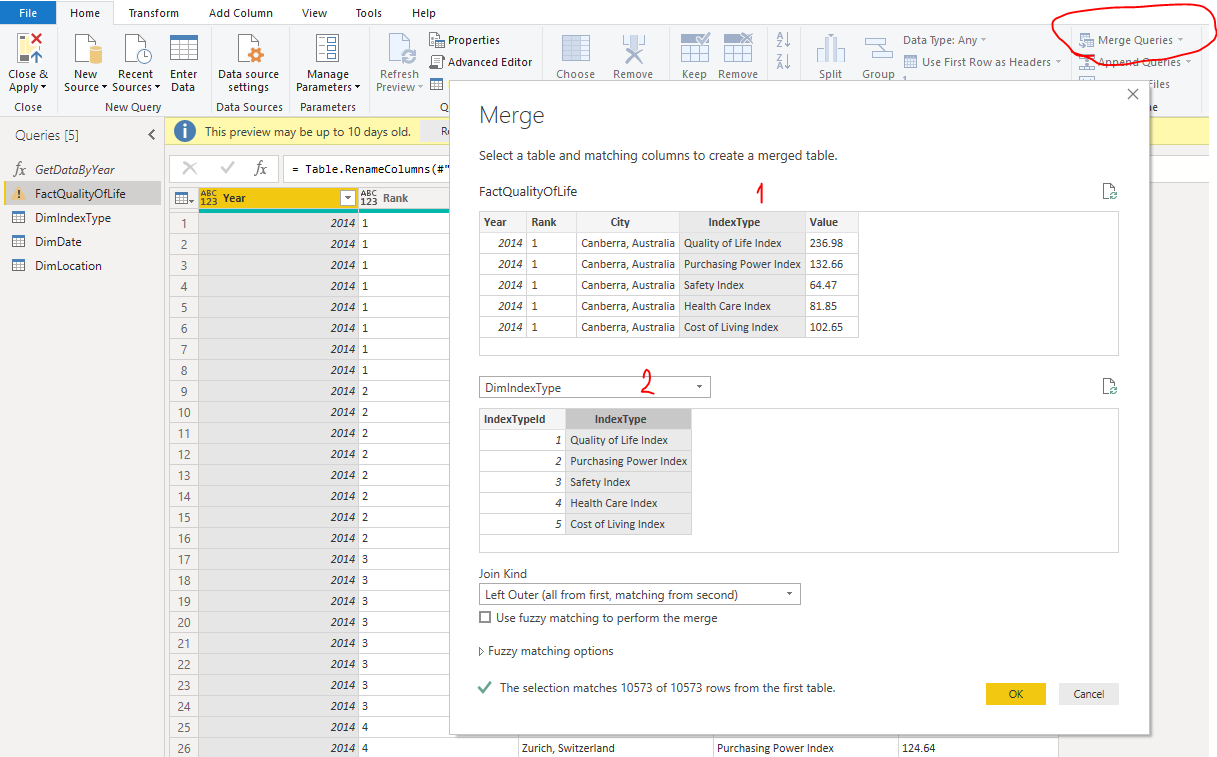
***in***

***#"Distinct Columns"***

1. Добавляем колонку LocationId с идентификаторами, используя Index Column.
2. Разделение атрибутов в DimLocation. Переименовываем City в CityFull. Дублируем колонку.Разбиваем новую колонку на Город (City) , Штат (State) и Страну (Country), используя Split Column (by Delimiter):

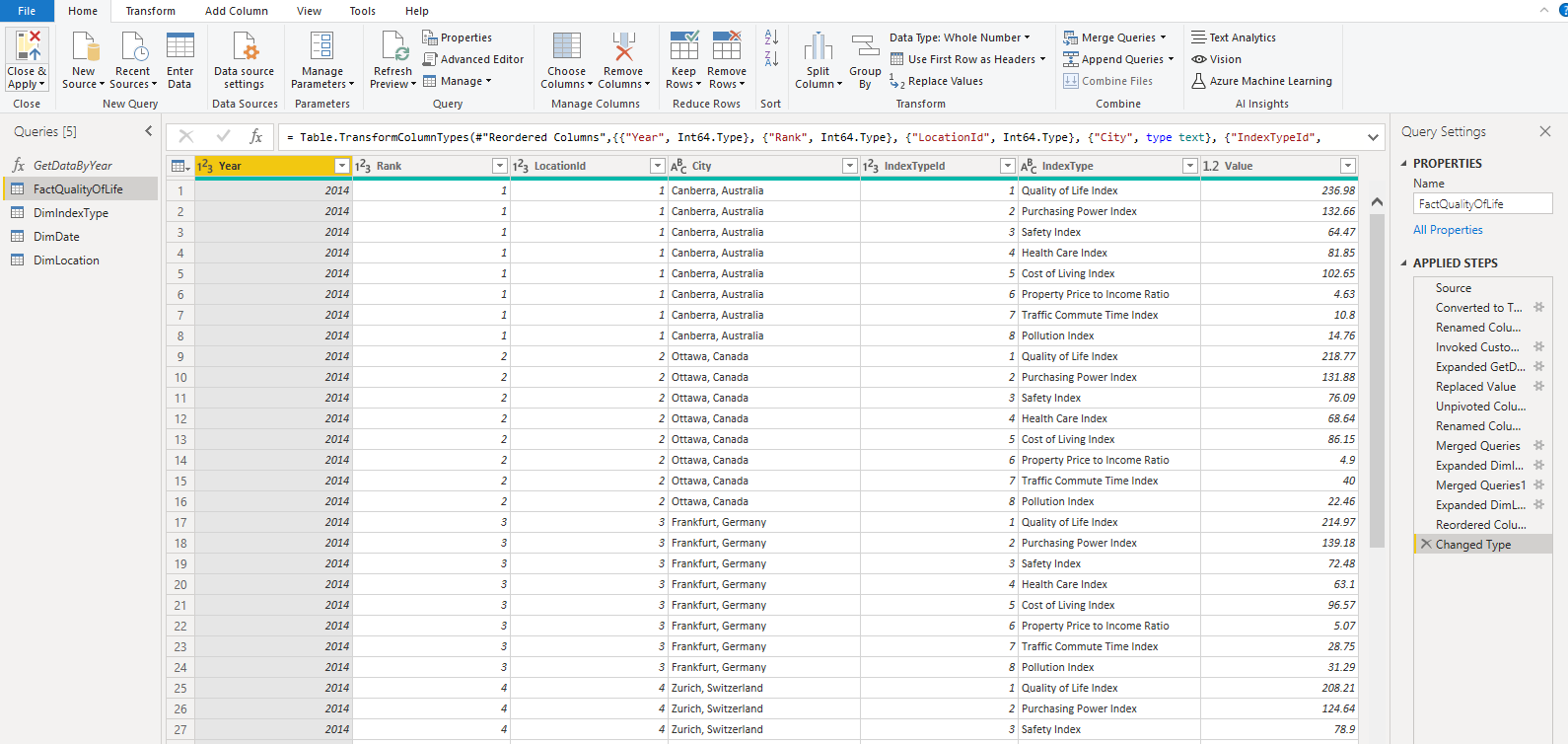


1. Добавляем в FactQualityOfLife идентификаторы из DimIndexType, используя Merge Queries и оставляя в присоединенной таблице только колонку с идентификаторами:



1. Аналогично добавляем идентификаторы из DimLocation.
2. По всем query проверяем тип данных, и если он не соответствует данным – меняем на нужный.

Должен получитьcя следующий набор queries:



Жмем Close & Apply, загружаем данные и сохраняем заготовку отчета.