**Exercise 4**

\*\* Click on following link to get all the datasets required for this exercises

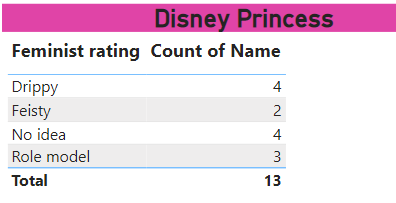
[**https://github.com/PurpleGrad/2203/blob/main/Dataset%20for%20Exercise%204.zip**](https://github.com/PurpleGrad/2203/blob/main/Dataset%20for%20Exercise%204.zip)

**Exercise on Data Source :**

**1)**Create a new Power BI file, and from the above folder load the Excel workbook containing a list of the **Disney Princesses.**

Hide a couple of columns that we're not interested in, so that the field well looks like this:

Create a table showing the number of films by feminist rating:



*You'll need to change the summarisation option to do this.*

Save this Power BI file with the name **Disney Girls**, then exit this instance of Power BI Desktop.

**2)**Import the **Events, Category and Country** tables from the **WorldEvents.xlsx**file, found in the above folder. Look in **Relationship View** to see what has happened.

*Annoyingly, no relationships have been created between the tables. There are two ways we can handle this.*

*To create a join. In this case* ***CategoryID to*** *the***E*ventCategoryID******EventCountryID****and****CountryID****pair up!*

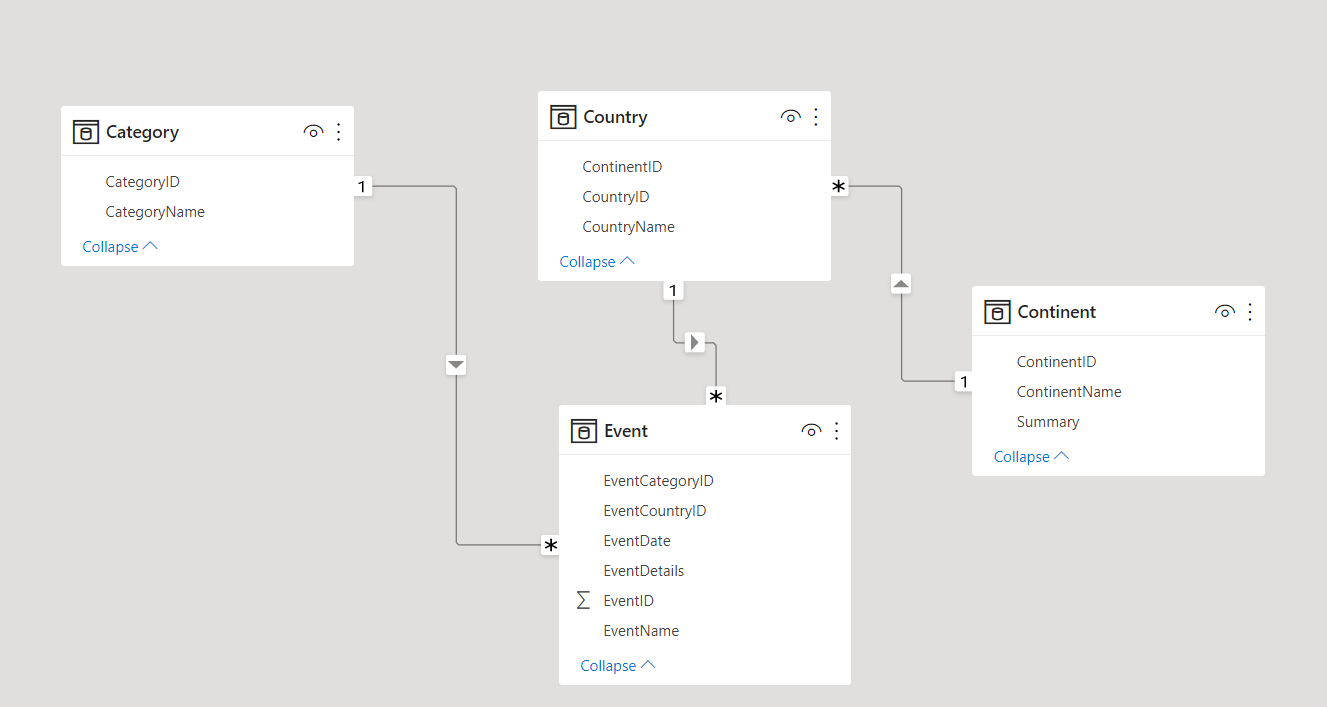
*Thankfully most of the time the names of columns make it clear what the overlap is. Modern databases have a tendency to even use the exact same column name.*

The danger of this method is if you miss. To check which columns have been joined let your mouse hover over the line you have just created. Power BI will highlight the two joined columns:

*The****1****means that****CategoryID****only has unique values while the****\*****means****EventCategoryID****has duplicates (ie more than one event can be in the same category).*

How daft, there was another table we wanted from that excel document!

Bring in the **Continent** table - this should automatically create a join too:



*A relationship has been auto created since they have the same field names, data types and one is unique. Excellent!*

Optionally save this as **Creating connections.pbix** and close it down.

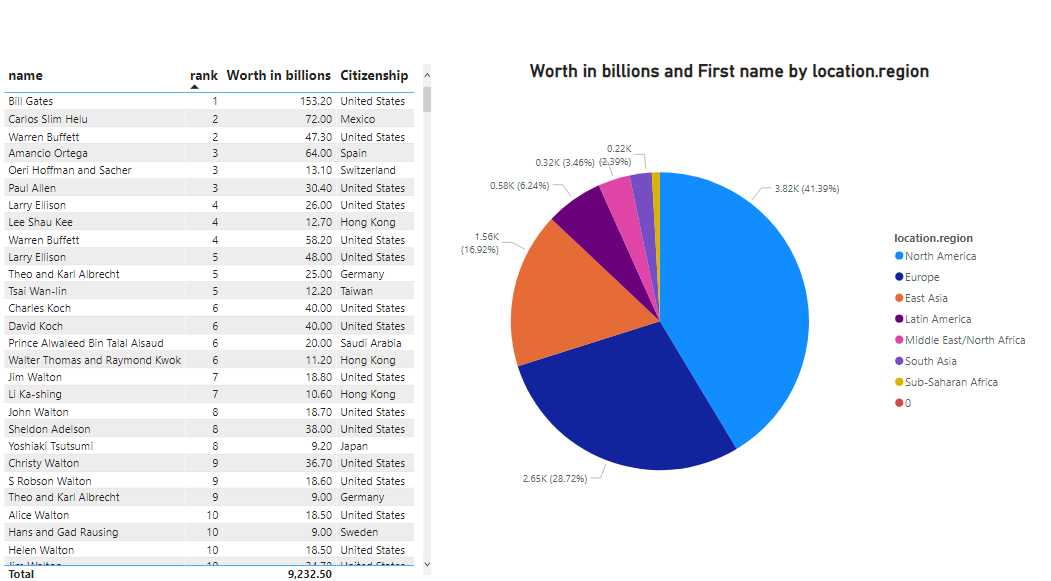
**Exercise on Query Editor**

1)Create a new Power BI report, and load in the data from **billionaires** in the above workbook:

*The raw data - your task is to massage it into something more presentable!*

Use **Query Editor** to make this data look better:

*Some columns have been renamed , Company , Name , Citizenship , Rank ,Sector and worth in billions*

Bring the data back into your Power BI report, and use it to create a simple chart and list of all richest people: 

*Create a simple chart just to prove that the billions really are being treated as numbers.*

 Save this report as **But are they happy**, then close it down.

**2)** Create a new Power BI report.

Load the data from FTSE workbook in the above folder - you should see something like this:

In Query Editor, carry out some transforms:

Rename the remaining columns as shown below.

You should now be looking at something like this:



*The remaining table of data.*

Your prices are left-aligned, which suggests that Power BI is treating them as text.  Change the data types of the last two columns



*Better - now we can do arithmetic on these columns ...*

Create a column which get difference between two price, rename it and sort by it to get:

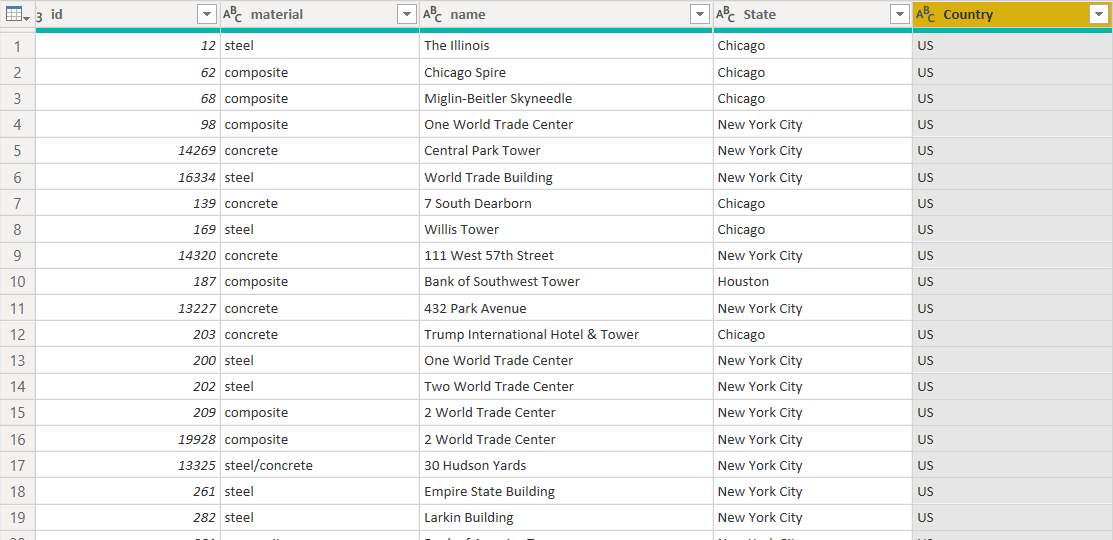


*It's been a good day today for holders of shares in****NMC Health PLC****.*

Save this report as **Equity**, then close it down.

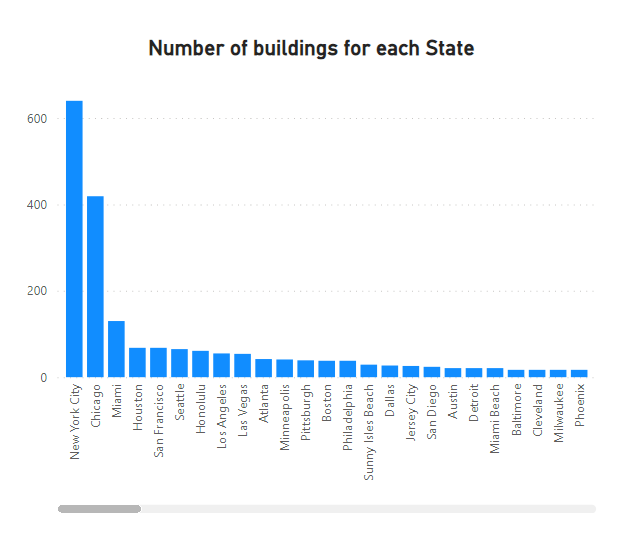
**3)** Create a new Power BI Desktop file, and load a list of the world's tallest buildings from **Skyscraper** Centre.

Add steps to your query so that it creates results like this:



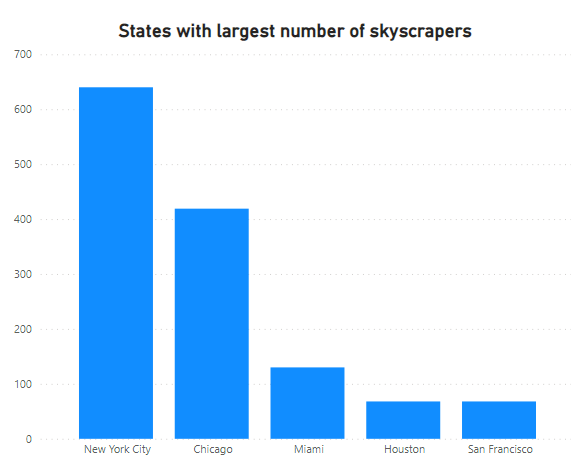
*We've split the****State****and****Country*** *, and* *renamed some other columns.*

Create a chart comparing the number of skyscrapers for each country:



*The number of buildings for each State.*

Apply a filter to the visualisation within Power BI Desktop itself to show only states having largest number of skyscrapers:



*No wonder New York is called ….****The Empire State*** *...*

 Save this file as **Skyscrapers**, and close down the Power BI instance that created it.

**4)** Load **Crouching Scyther, hidden Porygon** dataset and create Power bi report, With more *Pokémon* . it's time to clean up our data (and clear out the trash)!

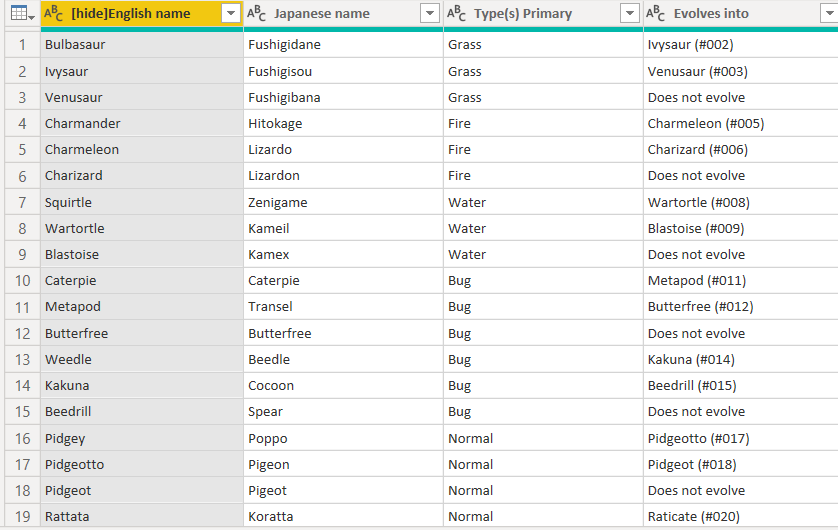
Load this into the query editor and get rid of a couple of columns:

*The****Notes****column is gumpf and the idea of****Type(s) Secondary****is offensive. We've also moved the****National Pokedex number****column.*

Next get rid of any *Pokémon* who didn't first appear in **Red and Green**.

Finally to make our list completely free of any non-originals do a find and replace to remove any reference to **Mega Evolution**:

Since we no longer need the **First appeared column,** delete it. The column was only there to remove non-essential *Pokémon*.



*We millennials might not have houses, but we do have avocado toast and PokÃ©mon!*

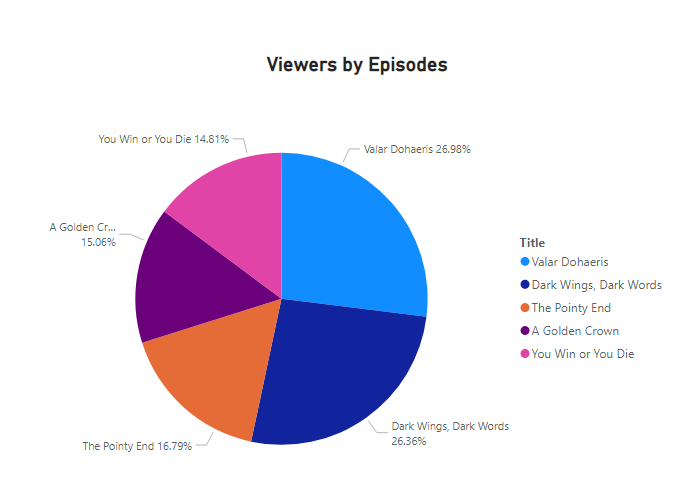
Save this as ***PokÃ©mon!*,** then close it down.

**5)** Create a new Power BI report, and load data from **Episodes** workbook.  You should see two tables (**Episodes** and **Series**).

Use Query Editor to tidy up the **Episodes** table so that it looks something like this:



Create a pie chart based on this table showing viewing figures by episode:



*Your pie chart should show how many people watched each episode directed by Daniel Minahan.*

 Save this report as **Danny Boy**, then close it down.

**6)** Create a new Power BI report.  Load data from the Excel workbook

Make the first row the header row and rename and remove columns to get:

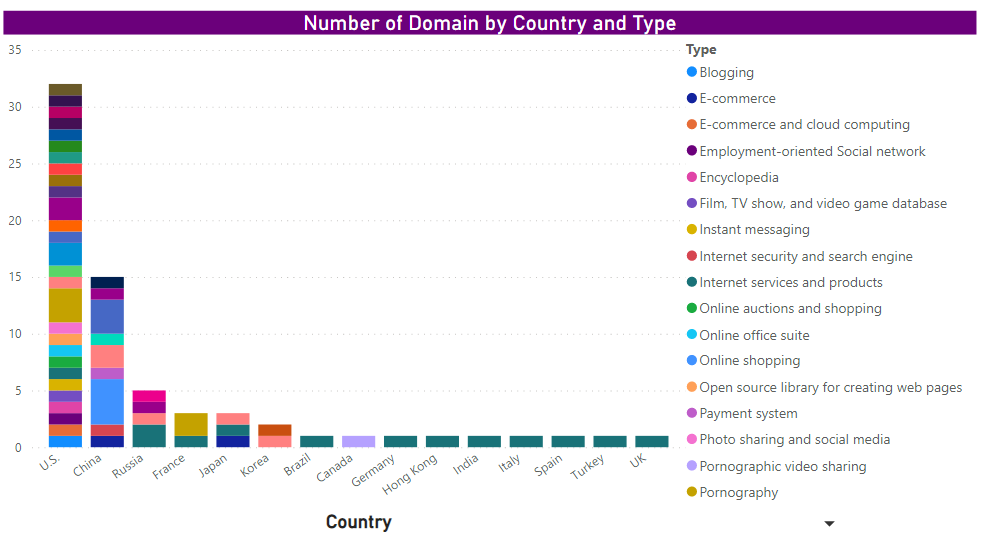


*Better - now the fun begins!*

Add in a column to number each row, and rename it to **Ranking**:

Use splitting columns to derive the current Alexa rank, keeping the number but losing the bit in brackets after it:

Load this data back into Power BI, and use it to create a chart something like this:

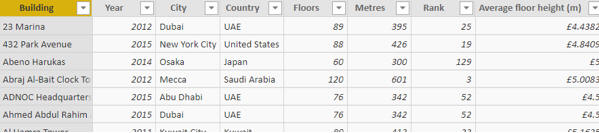


*This chart shows how many domains there are for each country and type.*

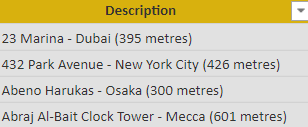
Save this report as **Google rules**, then close it down.

**7)** Create a new Power BI report, and load data from **Wiki buildings** file

Tidy the data up in **Query Editor**to get these columns:



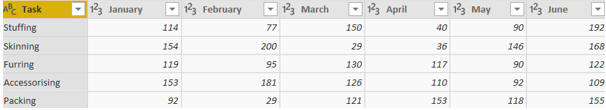
Back in Query Editor, use Column from Examples to add another column giving a description of each building:



Save this report as **Reach for the sky**, then close it down.

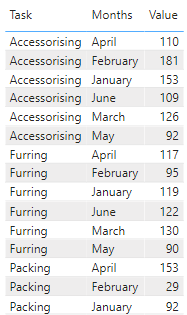
**8)** This exercise covers unpivoting data, changing data sources and more besides!  To begin with, create a new report and load the data in the **First Half.xlsx**workbook.

User Query Editor to tidy up this data so that you have something like this:

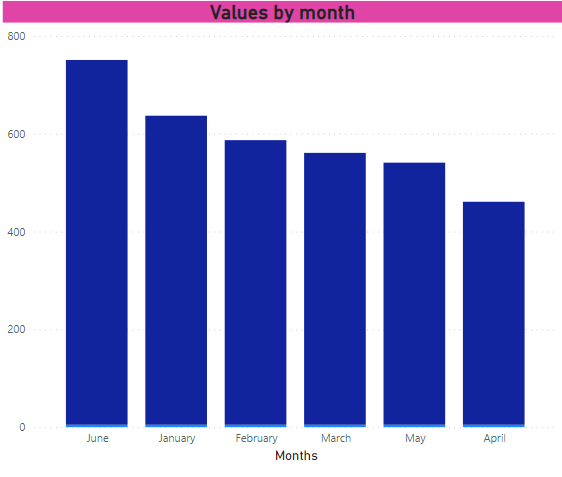


Unpivot the monthly data:

After a bit of renaming, you should have this table in Power BI:



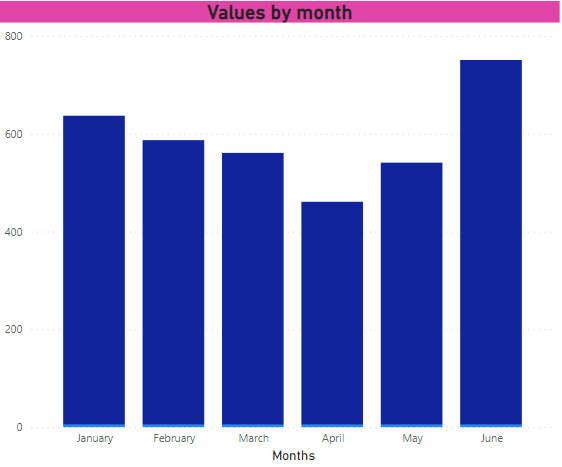
Use this to create a chart showing value by month:



*The option to add a conditional column, returning the correct month number for each row.*

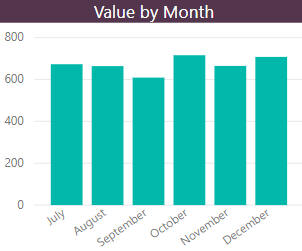
You can now choose to sort your month names by number:

After checking your chart is sorting by the correct column, you should now have this!

****

*The same chart, but with the correct sort order.*

Now,  Change your M to refer to the file called **Second half.xlsx** ,then check your chart still works:



Save your report as **The knees of a bee**, then close it down.