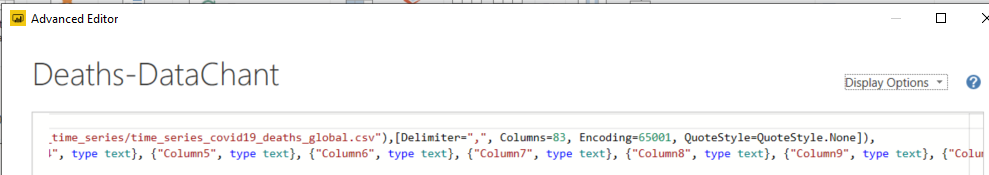
**WHAT I LEARNED TODAY**

April 10th:

DataChant

When starting out a PB project

1. Have a Date dimension table setup using CALENDARAUTO at the very minimum. The MIN and MAX date in the data set can be the parameters. The Advanced Query for this table is stored below.
2. Having the data with too many columns is s sig it should be pivoted Like in COVID data – a date with value of cases should be rows and not columns
3. Set up relationships
4. Copying steps from one table to another is easy with Advanced Queries
5. When combining tables it is important:  First, let’s ensure we will not load the separate tables into the report. Right-click on each query and uncheck **Enable Load**
6. The covid data would not refresh – this was because the number of columns in the original source was hardcoded (number of dates) So in Sources – remove this hardcoding-Luckily, if we remove **Columns=xx** from the formula, we will remove this in the advanced editor



1. To avoid failure of refresh step – Remove the Changed Type default step applied after every data source load. Read here:  
   <https://datachant.com/2017/01/11/10-common-mistakes-powerbi-powerquery-pitfall-2/>
2. Understand what the SUM displayed is. If the column is already being (AC)CUMULATED in the original data – such as in the COVID data, the confirmed each day is the sum of last day’s value AND the current day’s value. In this case displaying the final SUM which is the sum of all days is an incorrect number
3. Difference between VALUES and DISTINCT

PubNub for real time data say from Twitter

Let’s make an assumption that it takes an average of X days to recover from the virus. So, on any given day, the number of recovered cases equal to the number of confirmed cases we had X days ago minus the number of new deaths in the passing period.

Add meaning to data

Row Context – Calculated Column

Filter Context – Measure

Evaluation Context -- Iterators- FOR LOOPS

|  |
| --- |
| **ADVANCED QUERY FOR DATE DIMENSION TABLE**  Let  // The List.Dates function that returns a list of dates  Source = List.Dates,  // Invoke List.Dates function. Here you need to specify the dates you want the list to invoke.  // From Date: # Date (YYYY,MM,DD)  // To Date: Duration.Days(DateTime.Date(DateTime.FixedLocalNow())- #date(YYYY, MM, DD))+1 (It counts the number of days between today and the first date of the table and adds one day more to include today).  // To adapt this to your own calender table change the values for (YYYY; MM; DD)  // #duration (1,0,0,0) instructs List.Date to do one day increments in the list.  #"Invoke dates" = Source(#date(StartYear,StartMonth,StartDay),Duration.Days(Duration.From(#date(EndYear,EndMonth,EndDay)-#date(StartYear,StartMonth,StartDay)))+1, #duration(1, 0, 0, 0)),  // This step will convert the list to a table.  // To set the correct dates, edit the query parameters  #"List to table" = Table.FromList(#"Invoke dates", Splitter.SplitByNothing(), null, null, ExtraValues.Error),  #"Sorted Rows" = Table.Sort(#"List to table",{{"Column1", Order.Descending}}),  //Rename column to Date  Date = Table.RenameColumns(#"Sorted Rows",{{"Column1", "Date"}}),  // Extract Day from Date column and add leading zeros “00”  #"Day Added" = Table.AddColumn(Date, "Day", each Text.PadStart(Number.ToText(Date.Day([Date])),2,"0")),  // Extract Day Name from Date column using american names.  // Navigate here for other Languages not included in the parameters: https://msdn.microsoft.com/en-us/goglobal/bb896001.aspx  #"Day Name Added" = Table.AddColumn(#"Day Added", "Day Name", each Date.ToText([Date],"ddd",#"Language")),  //Extract Month No from Date column  #"Month No Added" = Table.AddColumn(#"Day Name Added", "Month No", each Date.Month([Date])),  // Extract Month Name from Date column.  // Navigate here for other Languages: https://msdn.microsoft.com/en-us/goglobal/bb896001.aspx  #"Month Name Added" = Table.AddColumn(#"Month No Added", "Month Name", each Date.ToText([Date],"MMM",#"Language")),  //Extract Quarter No from Date column  #"Quarter No Added" = Table.AddColumn(#"Month Name Added", "Quarter No", each Date.QuarterOfYear([Date])),  //Extract Week No from Date column  #"Week No Added" = Table.AddColumn(#"Quarter No Added", "Week No", each Text.PadStart(Number.ToText(Date.WeekOfYear([Date])-1),2,"0")),  // Extract Year from Date column  #"Year Added" = Table.AddColumn(#"Week No Added", "Year", each Date.Year([Date])),  // Concatenate Year and Month Name  #"Year Month Added" = Table.AddColumn(#"Year Added", "Year-Month", each Number.ToText([Year])&"-"&[Month Name]),  // Concatenate Year and Quarter No with leading zeros on quarter Nos  #"Year Quarter Added" = Table.AddColumn(#"Year Month Added", "Year-Quarter", each Number.ToText([Year]) & "Q"& Number.ToText([Quarter No],"00")),  //Change field types to TEXT  #"Change type to text" = Table.TransformColumnTypes(#"Year Quarter Added",{{"Year", type text}, {"Date", type date}, {"Month No", type text}, {"Day", type text}, {"Day Name", type text}, {"Month Name", type text}, {"Quarter No", type text}, {"Year-Quarter", type text}, {"Year-Month", type text}}),  // Extract Day No from Date column as a number to sort the previously created Day column.  //  #"Sort Day" = Table.AddColumn(#"Change type to text", "SortDay", each Date.Day([Date])),  // Extract DayName from Date column to sort the previously created Day Name column and use parameters to set if the Day starts on Monday or Sunday.  #"Sort DayName" = Table.AddColumn(#"Sort Day", "SortDayName", each Date.DayOfWeek([Date],#"StartWeekDay")),  // Extract the week number from date as a number to sort the previously created WeekNo column.  SortWeekNo = Table.AddColumn(#"Sort DayName", "SortWeekNo", each Date.WeekOfYear([Date])-1),  // Concatenate Year and Month Number to create a column to sort the previously created MonthYear column  #"Sort YearMonth" = Table.AddColumn(SortWeekNo, "SortYearMonth", each [Year]&Text.PadStart([Month No],2,"0")),  // Concatenate Year and Quarter Number to create a column to sort the previously created QuarterYear column  #"Sort YearQuarter" = Table.AddColumn(#"Sort YearMonth", "SortYearQuarter", each [Year]&Text.PadStart([Quarter No],2,"0")),  //Change the field type for the sort columns to numbers  #"Changed Type to NO" = Table.TransformColumnTypes(#"Sort YearQuarter",{{"SortYearMonth", Int64.Type}, {"SortYearQuarter", Int64.Type}, {"SortDayName", Int64.Type}, {"SortDay", Int64.Type}})  in  #"Changed Type to NO" |

CALCULATE: 3D glass filters- acts on a AGGREGATE OF A COLUMN/TABLE –WITH A FILTER OVERIDE of user or visual filter

FILTER – Acts on a table and with a expression

ALL – remove all filters

Powerbi\_desktop\_samples in Github <https://github.com/microsoft/powerbi-desktop-samples>

Power BI preview AI features for language precession and visuals

Why PowerBI

Columnar data along with encoding – helps compression Vertipaq

User ease of use – Learning curve is short

Relative low cost

Low to no code

Power platform - > Power BI suite—AI M/L

At least 2 GB 6-8 ram

64 bit ,achine and OS

Resolution

Software:

Power User

Power BI Desktop

End User

We browser

As Administrator :

Data Gateway

Power BI Report Server (On-premises)

**POWER BI Interview questions:**

Merging multiple CS file reports

1)Get Data-> From Folder-> Connect to folder path (Same columns

2) Combine & Edit 🡪 Load More

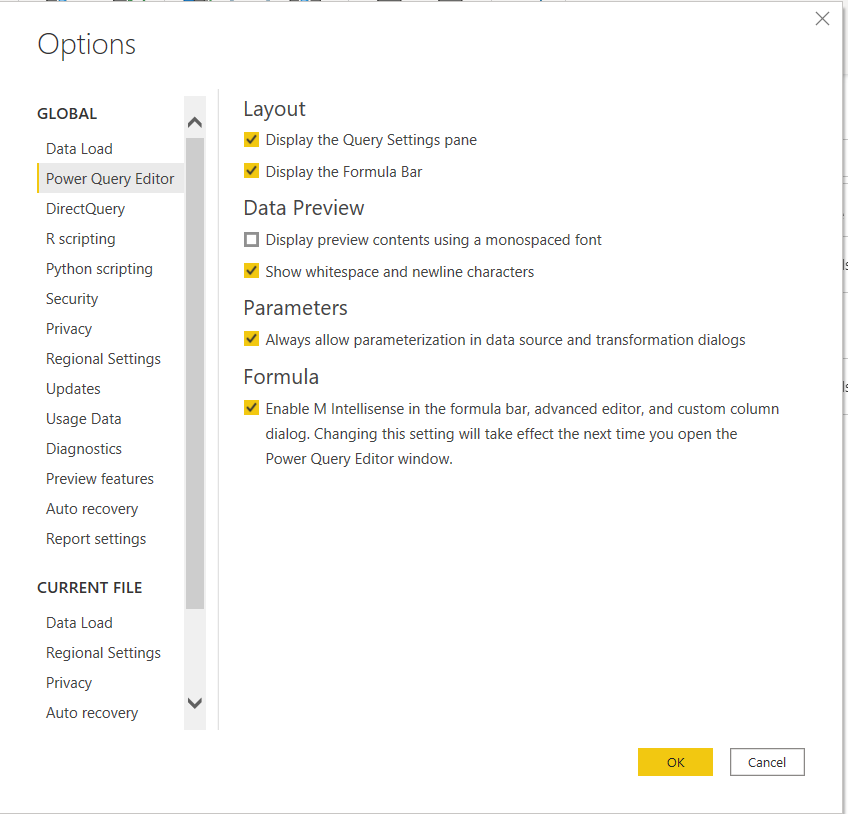
3) Automaticall will add new files dropped into the folder

**DATA SHAPING**

1. Copy and paste to transfer M Code queries from one to another (Advanced Editor)
2. Use templates – You can export a template (Only metadata) in File menu
3. Data flow

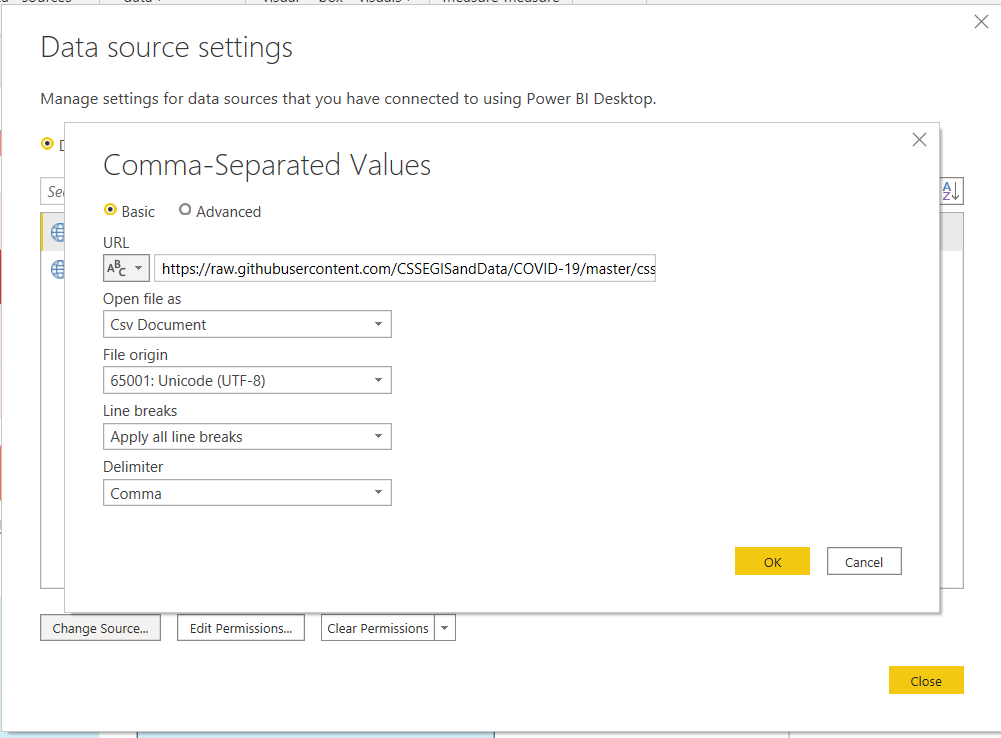
**Using parameters (Manage Parameters)**

**File-> Options and settings**



**Changing DataSource through a parameter’**

**In the menu item Transform Data – open it and you will find Data Source**



**Data Cleaning:**

**Each column must have a heading (not data)**

**No blank headings**

**No duplicate heading names**

**No merged cells : Multiple level headings X**

**No formulas in headings**

**Only one meaning per column ( Product name and code together )**

**No subtotals**

**No formatting to indicate meaning**

**Data must always grow vertically**

**Each heading in a single cell**

**New Client Deployment can use a lot of this!!!!**

**Data cleaning in Power BI:**

**Using Fuzzy merge – for spelling mistakes – User Merge Queries**

**Join the data containing misspelled State names for instance and join it to a table containing correctly spelled names. See how it is done here**

<https://www.youtube.com/watch?v=XF1ZGbeuzCE>

**You can use Oythin to clean up data such as nulls. Pythin can be used in visualizations as well**

**Power BI Dax interview questions**

<https://www.youtube.com/watch?v=VS0HJEBDmd8>

**DATA QUALITY AUDIT:**

**To check for missing data in a column- create a pivot and if the values field only allows for COUNT and not for SUM, then there is missing numeric data.**

**To check for unique values in a colum, place that column in the ROW area of pivot to highlight errors**

**In all tables, create a COUNTER COLUMN. The data in the column is a formula =”A” –a formula is automatically applied to all rows of the table (also we get a real count since all NULLS are eliminated)**

**POWER QUERY WITHIN EXCEL:**

Get Data From table: This data will also be refreshed when the source is changed

Flash Fill – to break up a column with delimiter

Pivot table Pivot chart Wizard : ALT+D, P to open the wizard

Choose this in the Wizard: Multiple consolidation ranges

I will create the Page fields

Excel – **Create excel tables.** Stop using Excel list data. This helps in instant PIVOT refresh.

Also since it has a table name, you do NOT have to select the worksheet range for the PIVOT, the entire table is the range for the PIVOT.

NAME the table so you can refer to it. No space between words. Similarly name the pivot table created.

Pivot and dragging A Field to OVER the ROWS selected- Creates a hierarchy

Dragging and rearranging fields in different orders in ROWS – If it is under a date (it disappear in the PIVOT until you open the Quarter)- again hierarchy.

**DRILLDOWN BY DOUBLE CLICKING\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\***

**When you double click on a column value in a pivot, a new sheet opens up with the drill down rows that contributed to its computation**

**SORTING BY USING CUSTOM LISTS**

**GETTING DATA FROM TXT Files- Data tab Get Data**

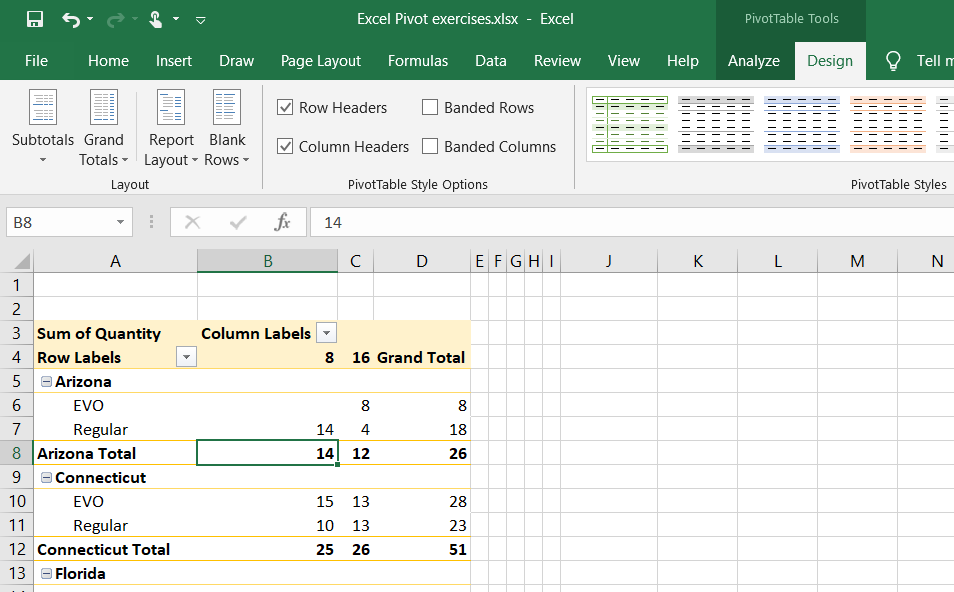
SORTING --- Cannot sort a column of the pivot, But select a cell **IN THE COLUMN WHOSE VALUES YOU WANT TO SORT BY** in the pivot and SORT the entire pivot using home tab sort

COPYING a PIVOT table:

Click on corner left and Ctrl+C and paste

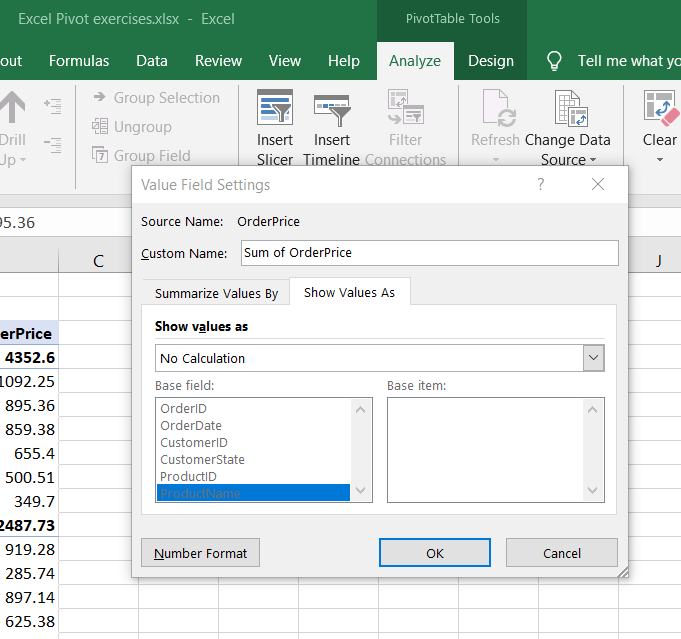
**SUBTOTALS DISPLAY**

**PivotTable Tools (If this disappear- switch tabs and come back)**

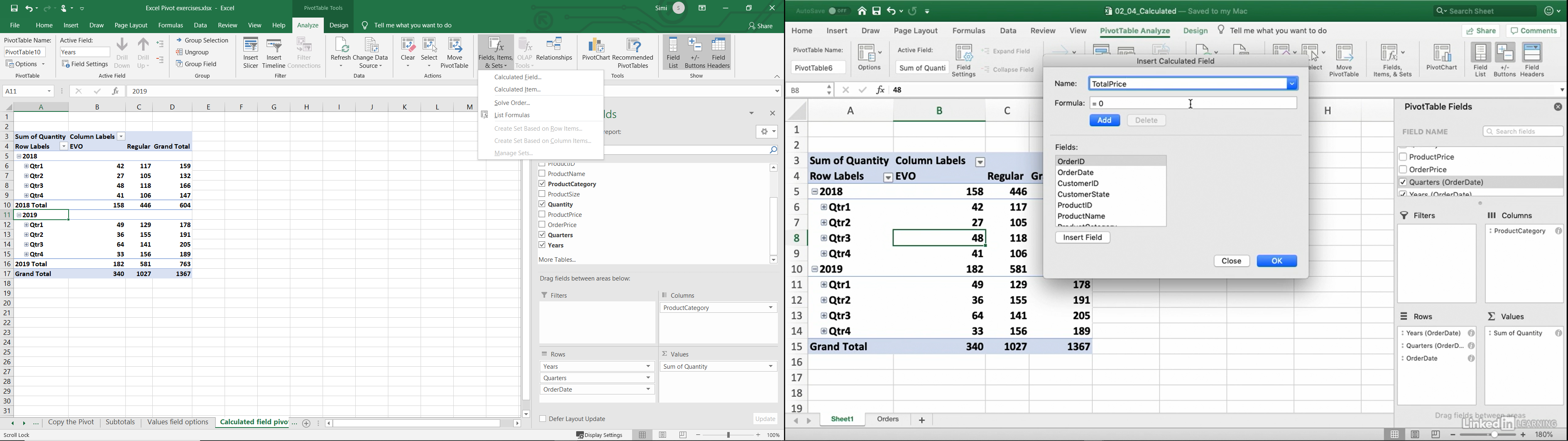


In the Sigma values, in the drop down

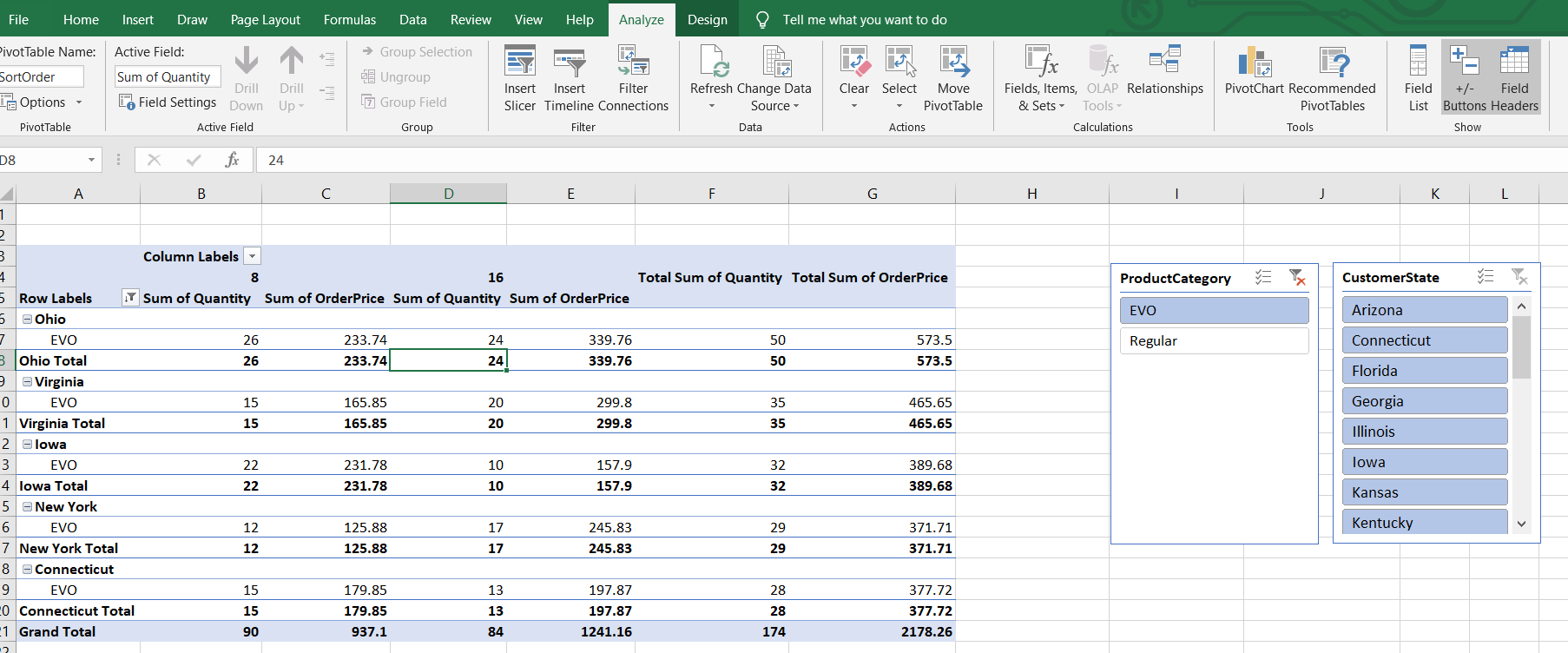
**VALUE FIELD SETTINGS**



**Calculated Field/Column Field Items/Sets**

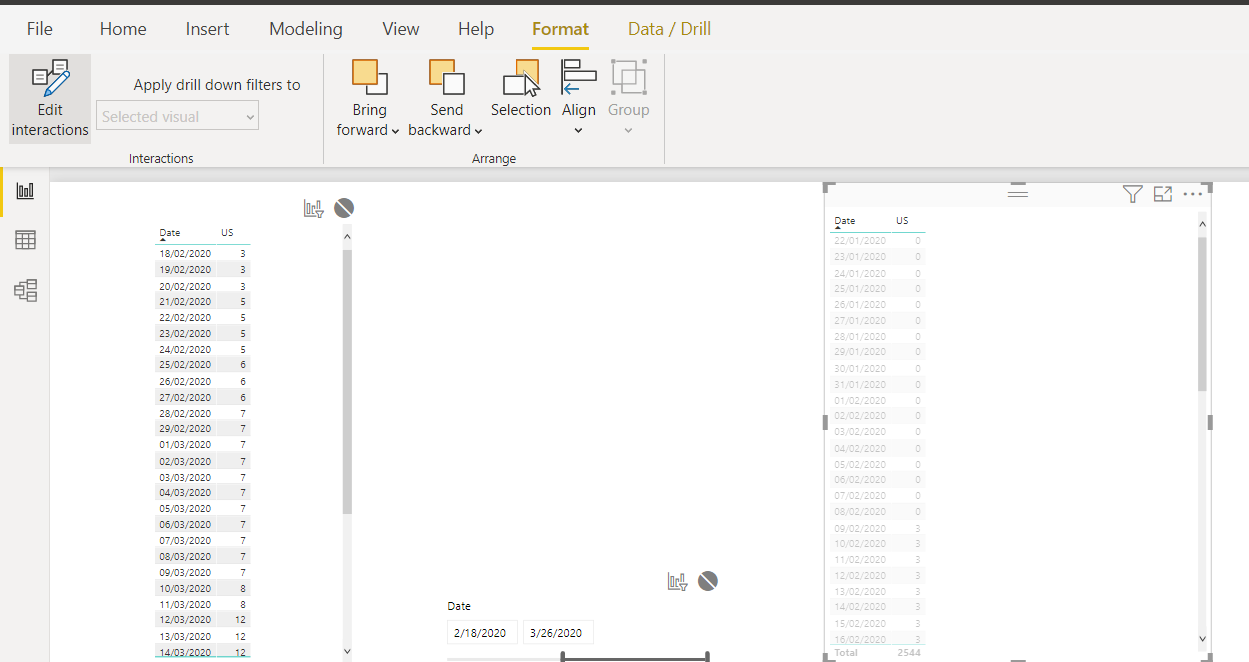


**FILTERING USING SLICER**



**FILTER OR SLICER APPLICATION TO ONLY ONE VISUAL IN POWER BI**

**Use Edit Interaction in Main Menu Format tab**



**Importing RAW CSV for COVID Coronavirus**

<https://raw.githubusercontent.com/CSSEGISandData/COVID-19/master/csse_covid_19_data/csse_covid_19_time_series/time_series_19-covid-Deaths.csv>

<https://pages.semanticscholar.org/coronavirus-research>

SHOW REPORT FILTER PAGES:

Explode the filter options into multiple worksheets- each worksheet will contain the pivot for only one filter element

**LINKING EXCEL CHART TO POWERPOINT**

**Copy a chart and then past – but CHOOSE the appropriate paste option**

**CREATING MULTIPLE PIVOT TABLES**

**When creating the pivot table you have the option of creating it in a NEW worksheet or existing worksheet**

**You can select a PIVOT table with CTRL+A and paste : Report connections**

**Slicers filter each other**

**POWER BI LEARNING:  
  
Could not convert a text date to Date:**

**Right click and split the text column with delimiter /**

**Delete the columns with date = 20**

**Create custom column with Year 2020**

**Merge the date, month and year = 2020 column with custom delimiter /**

**Convert this date in text to date – there are no more issues**

**Creating a custom List for Sort.**

**Just like FlashFill, you have to give Excel a clue. Place your list of items sorted in the way you want them to appear in a column and select that list**

**Open File-> Options->Advanced-> Edit Custom list**

**Add the selected list to this along with the Jan, Feb…**

**Cleaning Data:**

1. **Creating Group from items in a column**

**Select the items that belong to a group – right click and select Group**

**Grouping for Misspellings**

**It is better to organize all your data into a table naming each item into a group and subgroup and create this is one shot**

**TO create a table out of data – Select the table and hit Insert table**

**=VLOOKUP([@Product],Categories12,2,0)**

<https://efficiency365.wordpress.com/2014/10/19/custom-date-grouping-using-vlookup/>

**DAX VERTIPAQ learning from pluralSight Eugene Meidenger**

**Compressing data , columnar**

**Value Encoding : Have a base such as 8600000 and then store only the deltas**

**Dictionary Encoding : Hash table ( map it to a smaller key)**

**Run-length encoding—store the number of times the same value appears like all 30 rows contain Maryland**

**WHAT WOULD I CHANGE BACK IN LEASEACCELERATOR**

1. **Ask them to download and use PowerBI**
2. **Use Excel tables and NAME THE TABLES**
3. **DoubleClicking leads to contributing rows**

**EDX POWERBI**

**DEV.powerbi.com**

<https://docs.microsoft.com/en-us/rest/api/power-bi/?redirectedfrom=MSDN>

SQLSERVER PYTHON : <https://github.com/STATCowboy/snakecharmer-intro>

GIT FILES:

<https://github.com/MicrosoftLearning/Analyzing-Visualizing-Data-PowerBI/tree/master/Lab1>

WORK EMAIL FOR POWER BI

<https://www.youtube.com/watch?v=uZyy_qqRPiU>

sjtalkar@getcertified.onmicrosoft.com

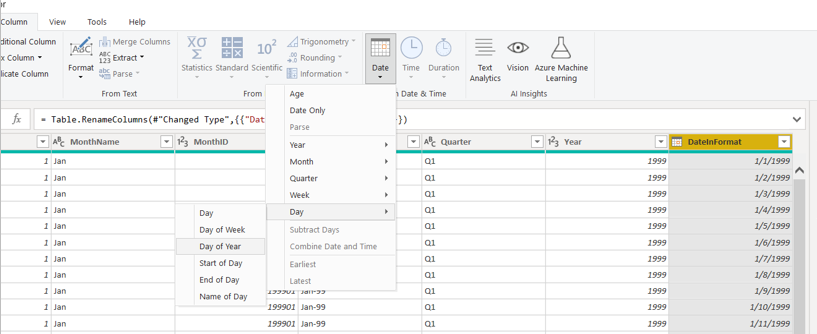
<https://gis.cdc.gov/grasp/fluview/fluportaldashboard.html>

**DataScience and bang**

**DAX FORMATTER :**

**DAXFORMATTER.COM**

1. When you have a date column formatted as a date, you can Add column --🡪 Date 🡪 This offers a variety of options to extract the compoments and with manipulations of previous as well
2. You can also do this with transform the column

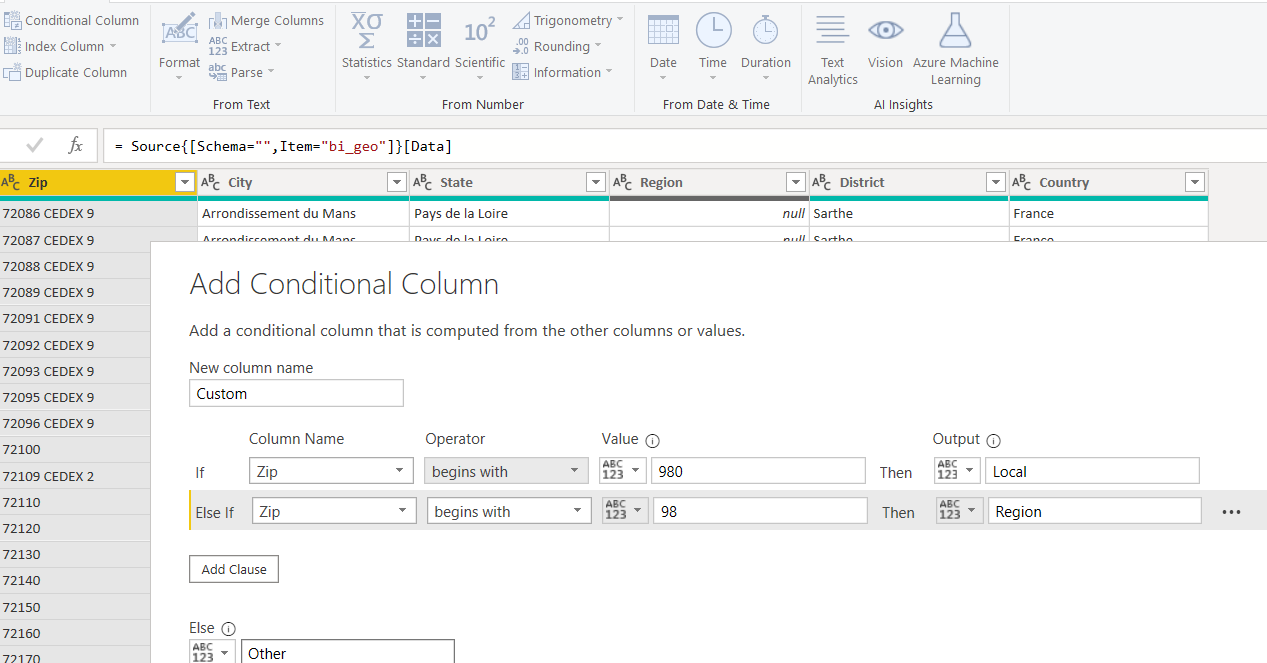


Returns a number from 1 to 7 identifying the day of the week of a date. By default the day ranges from 1 (Sunday) to 7 (Saturday).

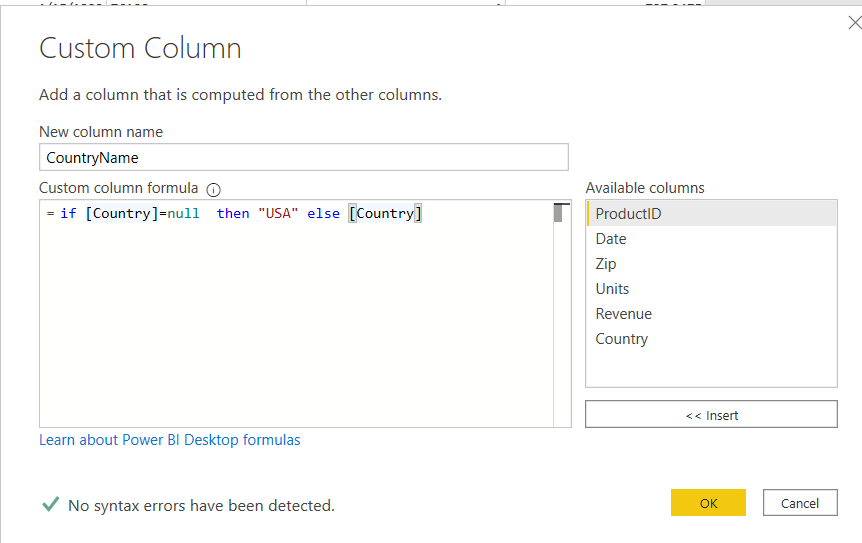
Deleting a column because it is an intermediate computation

Conditional column – It can be a value, a parameter or another column value

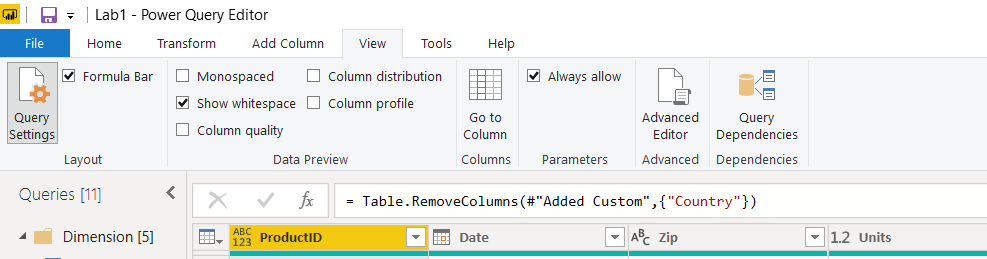
IF-THEN-ELSE with a conditional column



**Adding Custom Column**



**QUERY DEPENDENCY**



**MULTILEVEL SPREADSHEET**

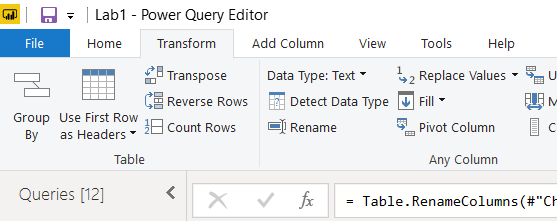
**First ensure that the multilevel headers are all marked as rows (if not use header as row transformation)**

**There should be no header row. Each column should be called Column 1, Column2….**

**Then use Transform this will change all the previous hears rows into column rows**

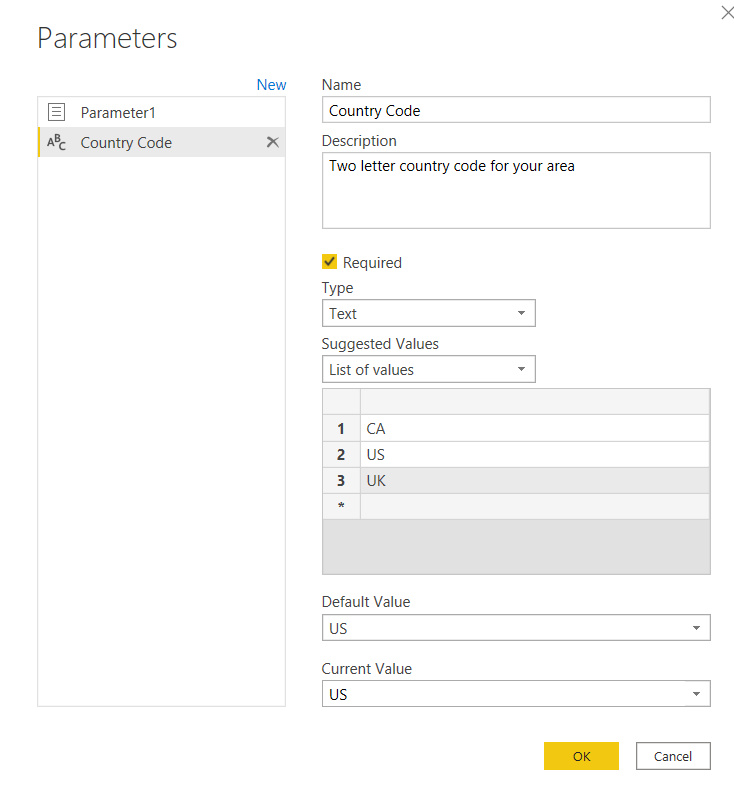
**Change Column1, Column2….using conditional column creation begins with**

**Use Fill down to fill NULL values**



**Unpivot the Other columns which are years**

**QUERY PARAMETERS AND USING THIS AS A PARAMETER FILTER AND SAVING AS TEMPLATE**



**Upgrade by April 13 and save 15% [Original price: $99, discount price:$84.15 ]**  
Use code **EDXWELCOME** at checkout! [Upgrade Now](https://ecommerce.edx.org/basket/add/?sku=89E6691)

Removed unnecessary rows – By filtering those values out in the column!

**CREATING A RELATIONSHIP ---sometimes you need to add a calculated column (Country+Zip) to create a relationship between tables like geography and Sales so that the right data can be displayed**

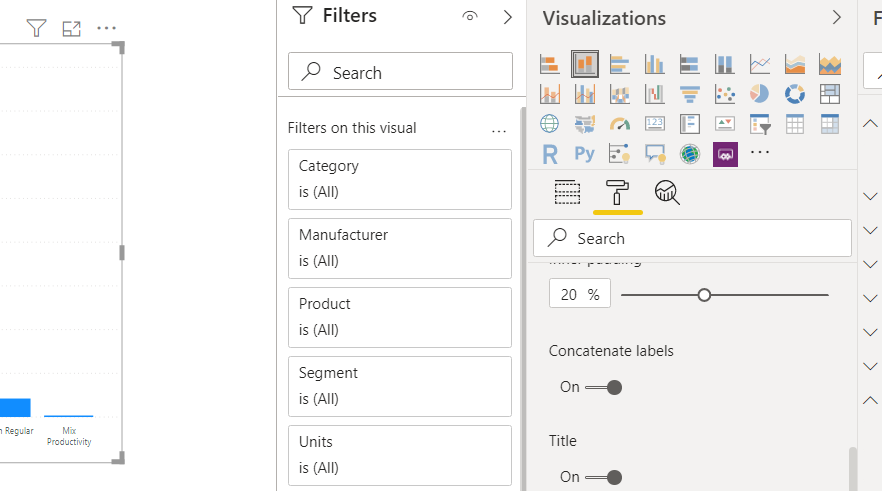
**Creating a Crosstab—The visualization is a matrix with Rows and columns that can be dragged and dropped**

**Heirarchy and usage in visualization for Right click and Drill down and Up**

**Category->Segment->Manufacturer->Product**

**Right click and it will give option to view the records contributing to that data: Show Data point as table**

**For Hierarchy labels – Concatenate or not concatemate**



**Creating a table manually**

**Can be done using Enter Data**

**To update or add to this table, In Edit Query (Power Query) click the gear next to Source This will open a window to edit the data**

**Include and Exclude (can be reincluded in Filter pane)**

**Grouping Fields**

**Import os**

**Import system**

**ConfigParser**

**ConfigSectionPath**

**cx\_Oracle.connect**

**shutil.rmtree**

**shutil.copy2**

**COMBINATION CHART: Line and Stacked column chart**

**When displaying two values with different scales – Revenue and Unit sales**

**Cluster charts and clusters help you look at many variable data and identify relations between them that might not have been obvious otherwise**

No, it is not possible to upgrade a 32bit Python installation to a 64bit one.

Still, there is something that you can do in order to speedup installation of a new 64bit version.

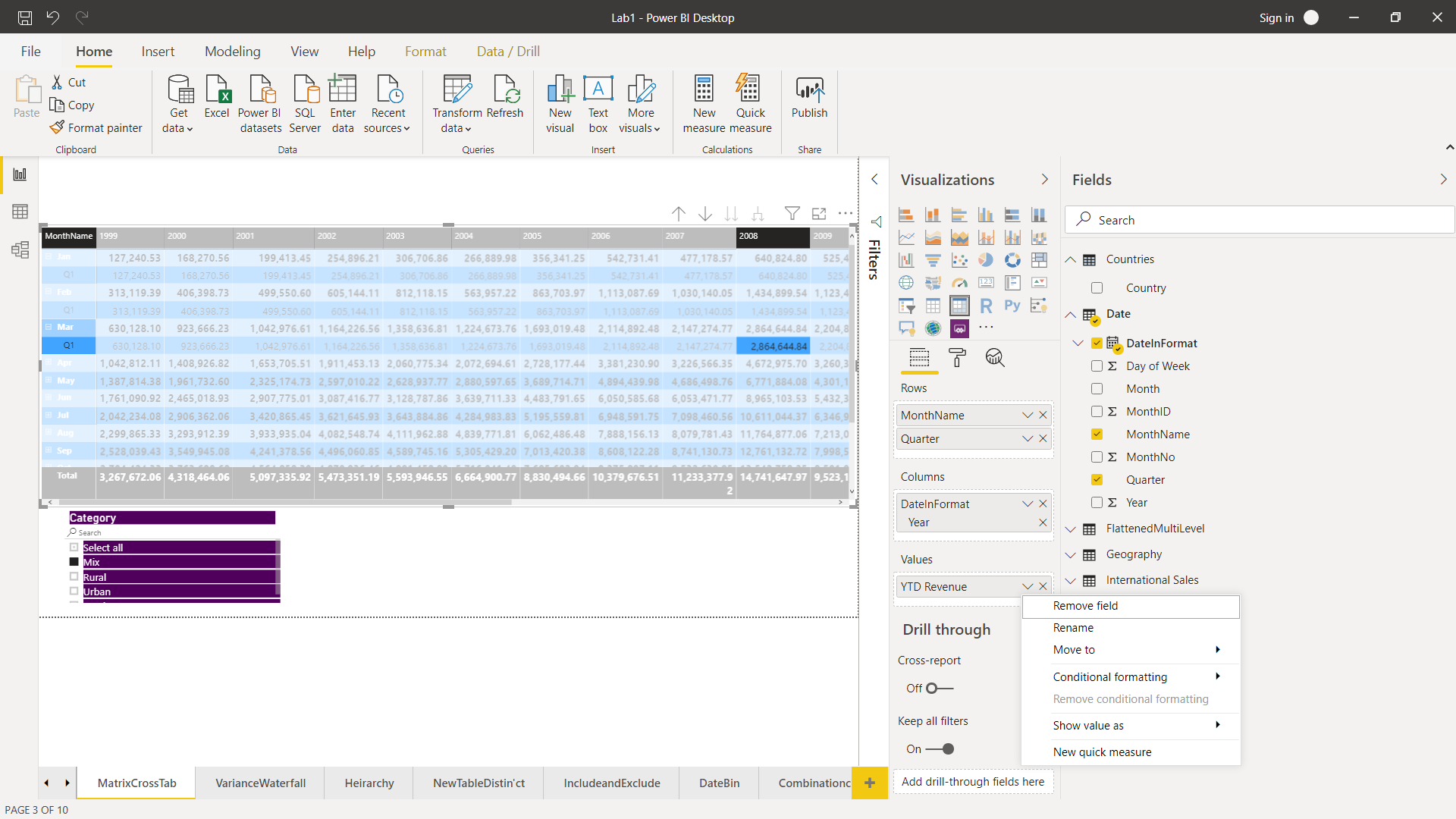
Run pip freeze > packages.txt on the old installation in order to generate a list of all installed packages and their versions.

After you install the new python version, run pip install -r packages.txt in order to install the same version of the packages that you had on the old installation.

**Conditional formatting option**

**In tables / Matrix – you can have hierarch y rows as in Excel pivot**

**For the value column open the drop dpwn to see Conditional formatting**

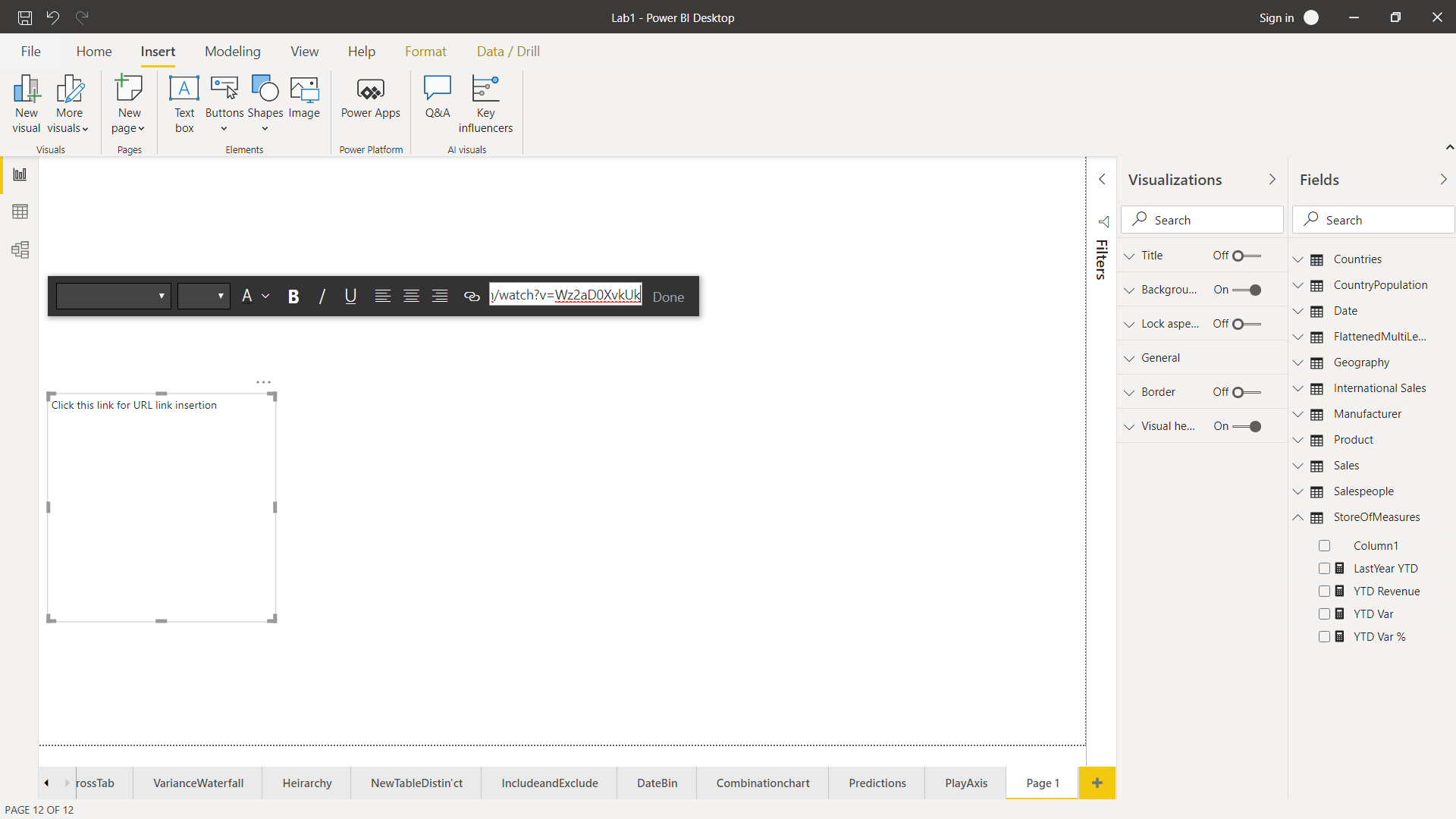


**\**

**KEY PERFORMANCE INDICATOR FOR TRENDS AND TARGETS**

**URL in visual**

**In Inset menu chose text box and write you text. Click on the intelinked venn diagram symbol to mark it as a link and place your URL there**



<https://www.youtube.com/watch?v=Wz2aD0XvkUk>

<https://www.youtube.com/watch?v=zCKIAle_i74>

POWER BI Service **AND INSIGHTS (if you create insights from the dataset, then those inssights reports can be pinned to the Dashboard)**

Use Publish to push a pbix to Power BI service

Insights.

TO pin a report, in the lower tab- go to the report you want to grab the visualization from

When you click on the visual you get an option to pin the visual on the top right of the visual itself

Similarly **within the dashboard** each visual has ellipsis in the top right and opening these ellipses will give you option to get insight for the report itself

In Dasboard - > Ellipsis🡪 Settings -> Check Turn on Tile flow – to auto rearrange dashboard tiles

Natural Language Q&A and Card display

To preserve sync and filter interactions in a dashboard, pin an entire report to the dashboard Pin a live Page

URL parameters for screen real estateS

collapseNavigation = True

chromeless= True (full screen mode

Setting Alert for threshold cross for values:

Alerts for Number Cases and

Pwer Builder personal gateway for data refreshes

Embedding the reprt into a webpage

Href (URL) or iframe generated

Sways

## Exercise 3- Schedule Data Refresh

 Bookmark this page

So far, the report you uploaded is rather static. That means, if the data in the Access database changes, the report and dashboard are not updated. You can install Power BI Personal Gateway and schedule data refresh for on-premises data sources, such as the Access database, to keep your report and dashboard on Power BI service up-to-date.

**NOTE**: Power BI Pro is required to setup scheduled refresh for on-premises data. If you do not have Power BI Pro subscription you can enroll for a free 60 day trial.

1. On the Power BI website, click **My Workspace**, click **Datasets**, click L**ab 4 – Starting**, and click the **Schedule refresh** icon.
2. Power BI Pro is required to setup scheduled refresh for on-premises data. If you do not have Power BI Pro subscription you can enrol for a 60 day trial.
3. Ensure that you are in the **Datasets** tab and that the **Lab 4 - Starting dataset** is selected.
4. Click **Gateway connection** and click **Install now** to download the Power BI Personal Gateway.
5. When the download is complete, click the **On-premises data gateway** to begin installation.
6. Click **Next**, accept the license terms, and click **Install**.
7. Sign in using your Power BI account and click **Close**.
8. In **New on-premises data gateway name**, type **Van Arsdel Sales**.
9. In **Recovery key** and **Confirm recovery key**, type **Pa$$w0rd**, and click **Configure**.
10. Click **Close**.
11. Expand **Data source credentials** and, for **Country Population by Year.xlsx**, click **Edit credentials**.
12. For **Privacy level setting** for this data source, select **Organizational**, and click **Sign in**.
13. Edit credentials for the Data sources that needs updating.
14. Now you can schedule your data source refresh.

**NOTE**: If you are using the starter file provided in the be

**POWER BI ADD IN TO EXCEL TO BE ABLE TO PIN PIVOT CHARTS AND TABLES**

**POWER BI And Oracle**

<https://www.microsoft.com/en-us/download/confirmation.aspx?id=58494>

## ercise 1- Upload Excel File with an Excel Table

 Bookmark this page

First, you will upload an Excel file containing an Excel table.

1. Download and extract the the “[Lab 5 Canada.xlsx](https://github.com/MicrosoftLearning/Analyzing-Visualizing-Data-PowerBI/raw/master/Lab5/Lab%205%20-%20Canada.zip)” file. The file contain VanArsdel's Canada sales. If you have a Microsoft Excel installed, you can open and explore the file (you don't have to).
2. Go to **https://powerbi.microsoft.com** and sign in using your account.
3. Click **Get Data**, select **Files** and click **Get**.
4. Select **Local file**, select **Lab 5 – Canada.xlsx**from the location that you extracted it to, and click **Open**.
5. Beneath **Import Excel data into Power BI** click Import.
6. When the import is complete, click **View dataset**.
7. Drag **Revenue** to the report.
8. Drag **Year** to the **Revenue** chart.
9. Drag **Units** to the report as a separate chart.
10. Drag **Year** to the **Units** chart.
11. Drag **Province** to the report as a separate chart.
12. Drag **Revenue** to the **Province** chart.
13. Now, answer the following questions by creating visualizations using the skills you learned in the previous module.

## Exercise 2- Upload an Excel File with a Data Model

 Bookmark this page

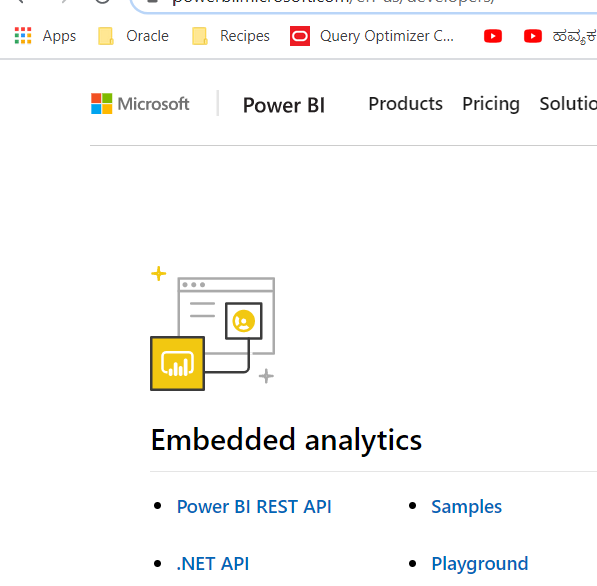
Now, let's upload an Excel file containing a Data Model and a Power View Report.

1. Download and extract the the “[Lab 5 USA.xlsx](https://github.com/MicrosoftLearning/Analyzing-Visualizing-Data-PowerBI/raw/master/Lab5/Lab%205%20-%20USA.zip)” file. The file contain VanArsdel's USAsales. If you have a Microsoft Excel installed, you can open and explore the file (you don't have to).
2. Go to **https://powerbi.microsoft.com** and sign in using your account. Note that, for this exercise, you must use a browser that supports Silverlight such as Internet Explorer.
3. Click **Get Data**, select **Files** and click **Get**. Select **Local file**, select **Lab 5 – USA.xlsx** from the location that you extracted it to, and click **Open**.
4. Beneath **Upload your Excel file to Power BI** click **Upload**.
5. When the upload is complete click **View Workbook**, or click **Workbooks**, and then click **Lab 5 - USA**.
6. If the **MAKE SELECTION TO PIN** pop up appears, click **Got it**.
7. If you do not already have Silverlight, click **Install Silverlight** and follow the installation instructions for Silverlight. Note that the Silverlight installation must be run as an administrator.
8. Now, answer the following questions by reviewing the **Power View1** tab in the **Lab 5 USA** Report.

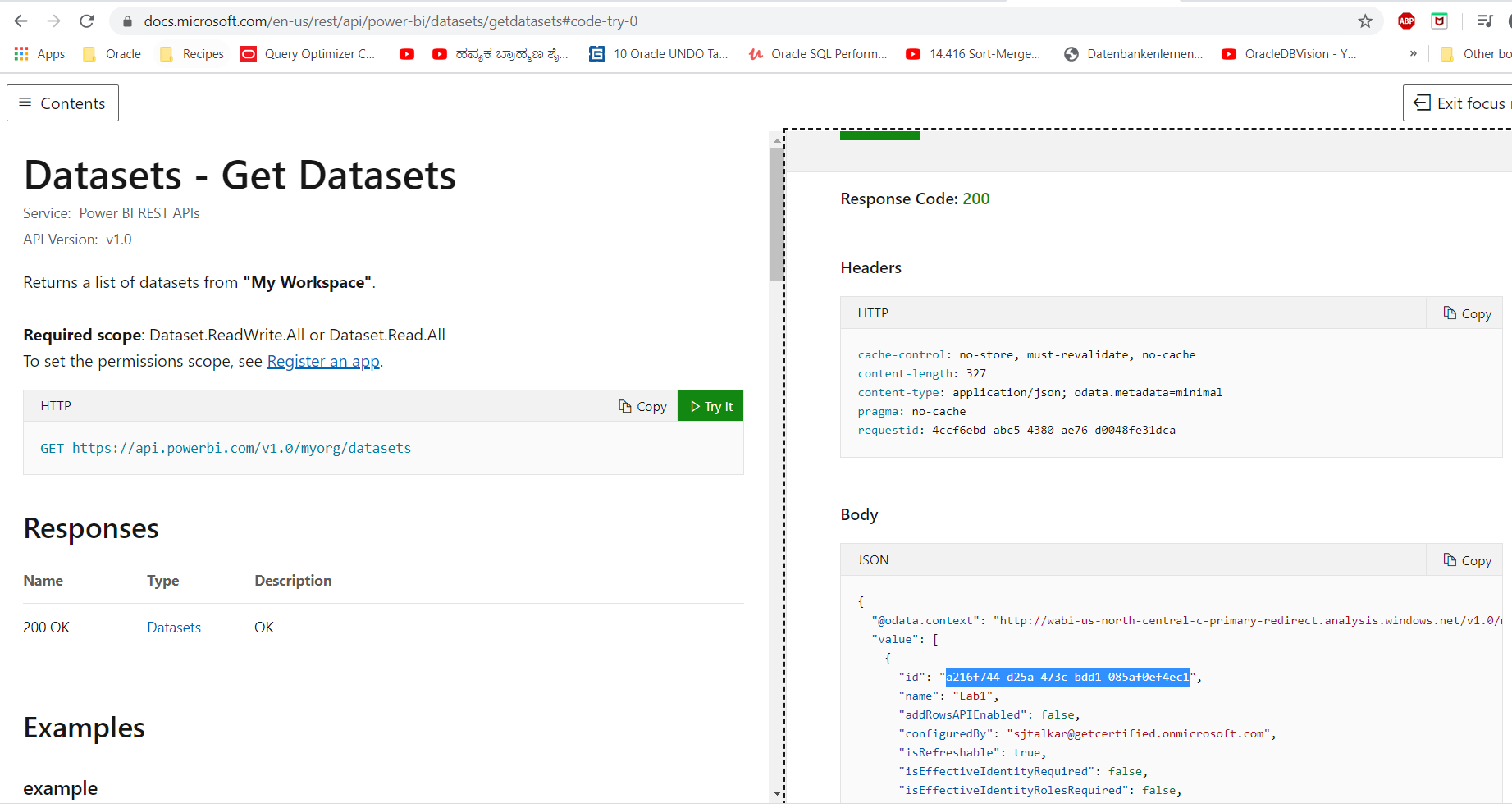
REST API Power BI

TO access the Power BI Rest API

GO to <https://powerbi.microsoft.com/en-us/developers/> and search for REST

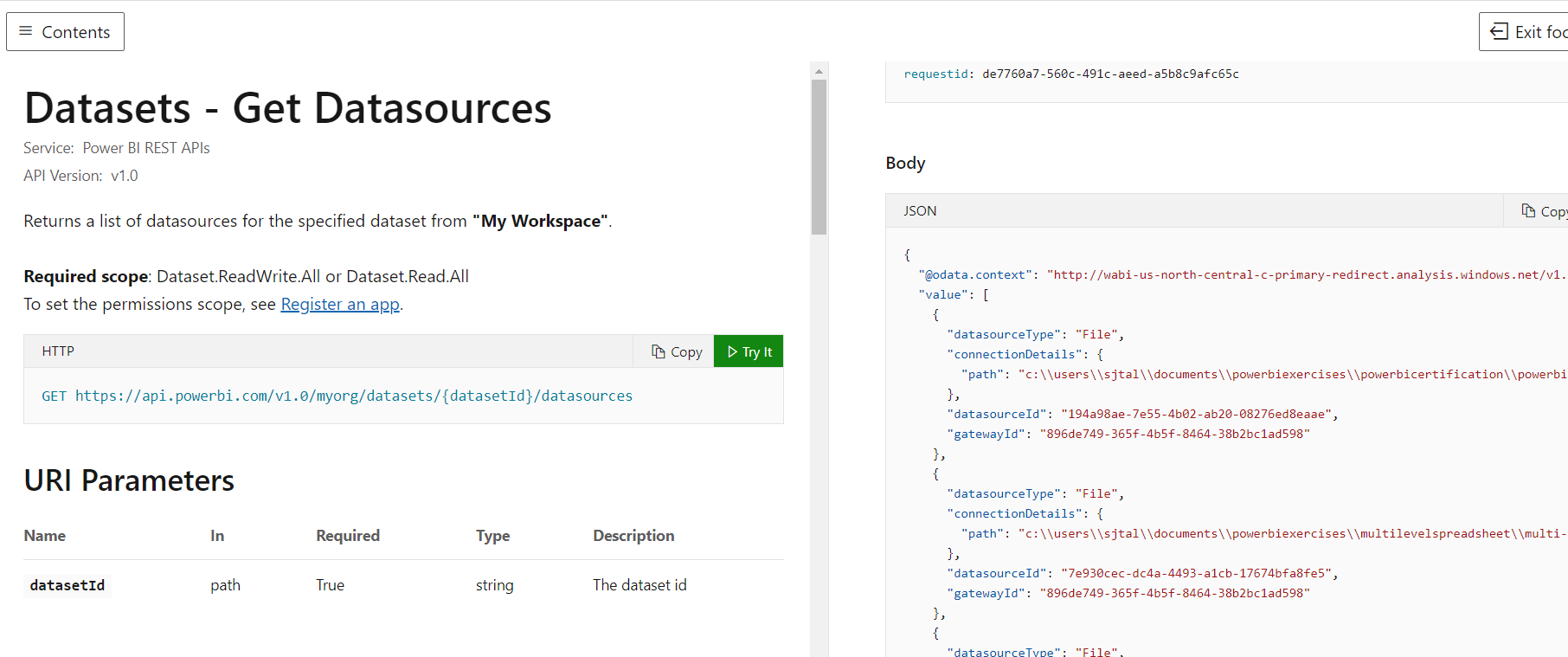


<https://docs.microsoft.com/en-us/rest/api/power-bi/datasets/getdatasources#code-try-0>



Use the ID you get above and use it in the Get Dataset API which requires an ID

You do not have to change the workspace name My Workspace



**On Github you can get source code for PowerBI D# and Power BI Java**

**Register your App**