



## Using PowerBI with LinkedIn data for some analysis



Chris Huntley

BLACK SHELL CONSULTING Fort Worth/San Francisco

## Table of Contents

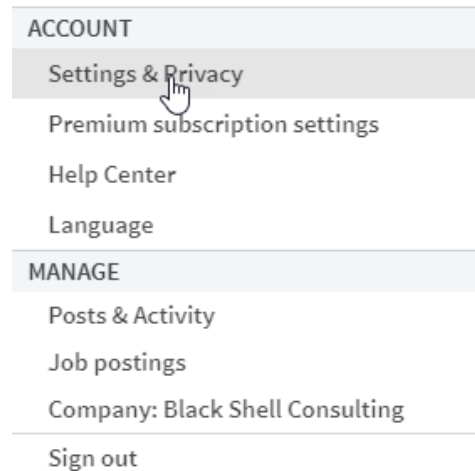
<a href="#"><u>Get The Data .....</u></a>	<a href="#"><u>2</u></a>
<a href="#"><u>Get Working Files .....</u></a>	<a href="#"><u>5</u></a>
<a href="#"><u>Transform .....</u></a>	<a href="#"><u>6</u></a>
<a href="#"><u>PowerBI .....</u></a>	<a href="#"><u>10</u></a>
<a href="#"><u>Get Custom Visuals .....</u></a>	<a href="#"><u>12</u></a>
<a href="#"><u>Extra Steps .....</u></a>	<a href="#"><u>13</u></a>
<a href="#"><u>Make a Chart .....</u></a>	<a href="#"><u>14</u></a>
<a href="#"><u>Example of Finished Dashboard .....</u></a>	<a href="#"><u>15</u></a>

## Get the Data.

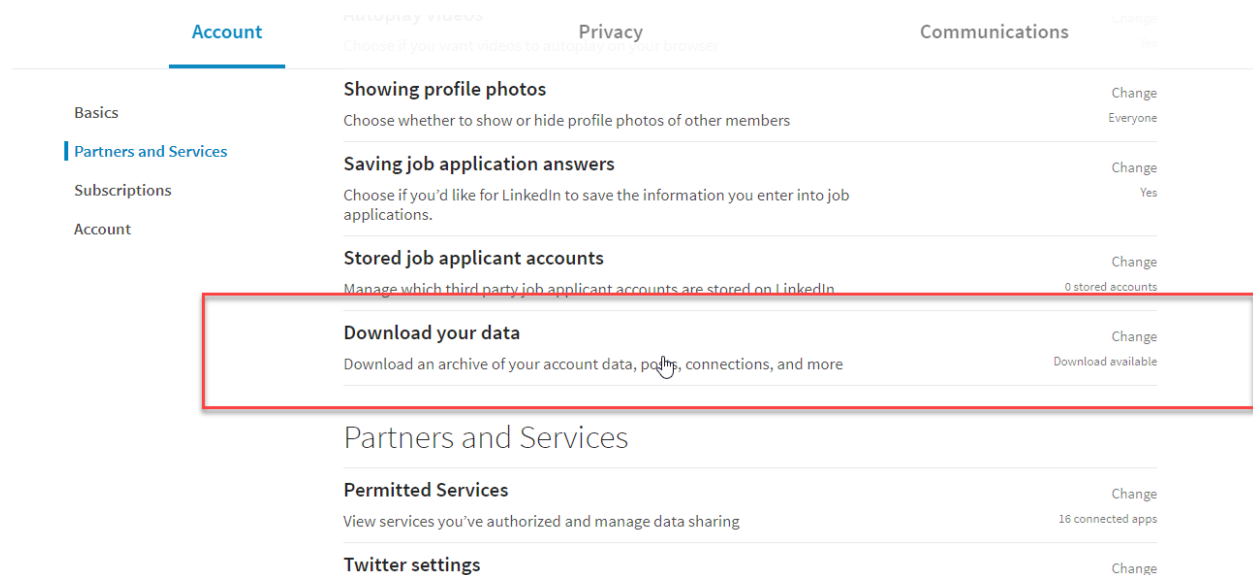
---

Download your linked in data.

Go to your profile picture right click and choose settings and privacy



Scroll down to download your data



Choose what you want to download. I'd choose everything as we'll look at it in various future tutorials but at a minimum download connections for this tutorial.

## Download your data

Close

Download an archive of your account data, posts, connections, and more

Download available

Your LinkedIn data belongs to you, and you can download an archive any time. You can learn more about what data you can export by [visiting our Help Center](#).

☒ The works: All of the individual files plus more. [Learn more](#)

☐ Pick and choose: Select the data files you're most interested in. [Learn more](#)

☐ Articles

☐ Connections

☐ Imported Contacts

☐ Messages

☐ Invitations

☐ Profile

☐ Recommendations

☐ Registration
























☐ Rich Media

[Download archive](#)

[Request new archive](#)

It takes around 24 hours to get the data so once you get here set a reminder to come back and finish it.

Once you get the files extract them to somewhere you can access them.

	Articles	1/25/2018 9:58 AM	File folder	
	Media Files	1/25/2018 9:58 AM	File folder	
	Account Status History	1/25/2018 9:58 AM	Microsoft Excel C...	1 KB
	Ad Targeting	1/25/2018 9:58 AM	Microsoft Excel C...	11 KB
	Ads Clicked	1/25/2018 9:58 AM	Microsoft Excel C...	5 KB
	Causes You Care About	1/25/2018 9:58 AM	Microsoft Excel C...	1 KB
	Certifications	1/25/2018 9:58 AM	Microsoft Excel C...	2 KB
	Comments	1/25/2018 9:58 AM	Microsoft Excel C...	996 KB
	Connections	1/26/2018 7:54 AM	Microsoft Excel C...	408 KB
	Connections	1/26/2018 11:08 AM	Microsoft Excel W...	1,056 KB
	DimDate	1/25/2018 4:48 PM	Microsoft Excel W...	108 KB
	DimTime	1/25/2018 5:19 PM	Microsoft Excel W...	4,193 KB
	Education	1/25/2018 9:58 AM	Microsoft Excel C...	1 KB
	Email Addresses	1/25/2018 9:58 AM	Microsoft Excel C...	1 KB
	Endorsement Received Info	1/25/2018 9:58 AM	Microsoft Excel C...	404 KB
	Group Comments	1/25/2018 9:58 AM	Microsoft Excel C...	56 KB
	Group Likes	1/25/2018 9:58 AM	Microsoft Excel C...	19 KB
	Group Posts	1/25/2018 9:58 AM	Microsoft Excel C...	10 KB
	Imported Contacts	1/25/2018 9:58 AM	Microsoft Excel C...	26 KB
	Invitations	1/25/2018 9:58 AM	Microsoft Excel C...	248 KB
	Languages	1/25/2018 9:58 AM	Microsoft Excel C...	1 KB
	Likes	1/25/2018 9:58 AM	Microsoft Excel C...	3,801 KB
	Logins	1/25/2018 9:58 AM	Microsoft Excel C...	29 KB

## Get the working files

For the tutorial you'll need two additional files for dimensional pivoting. They're called DimDate and DimTime. You can create your own Date Dimension and Time Dimension but to make it easy I did it for you. The files are here on the github. You can download them, clone or fork the repository. It's all open for your using and licensed appropriately.

Repo: <https://github.com/Huntleychris/LinkedInConnections>

You need the ones shown in the red boxes.

Using PowerBI with the addition of a date and time dimension Edit

[Add topics](#)

5 commits 1 branch 0 releases 1 contributor MIT

Branch: master New pull request Create new file Upload files Find file Clone or download

Huntleychris adding in the powerbi custom visuals Latest commit a1633be 2 hours ago

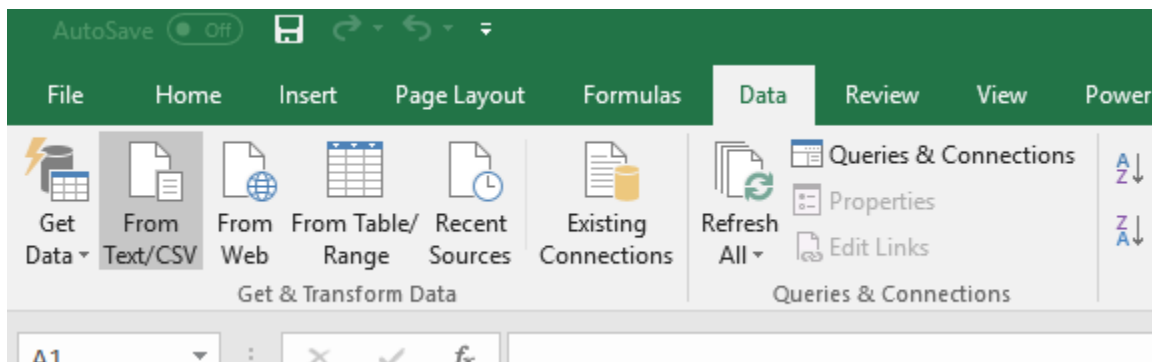
Connections.xlsx	adding in powerbi and connections file	2 days ago
DateDimension.SQL	adding in the powerbi custom visuals	2 hours ago
DimDate.xlsx	adding in the powerbi custom visuals	2 hours ago
DimTime.xlsx	Adding dimension files sql and excel	5 days ago
LICENSE	Initial commit	5 days ago
PowerbiLinkedInContacts.pbix	adding in powerbi and connections file	2 days ago
README.md	Readme update with important information	5 days ago
TimeDimension.sql	Adding dimension files sql and excel	5 days ago
WordCloud.1.4.2.0.pbiviz	adding in the powerbi custom visuals	2 hours ago
socialNetworkGraph.1.0.0.0.pbiviz	adding in the powerbi custom visuals	2 hours ago

README.md

## Transform the Data

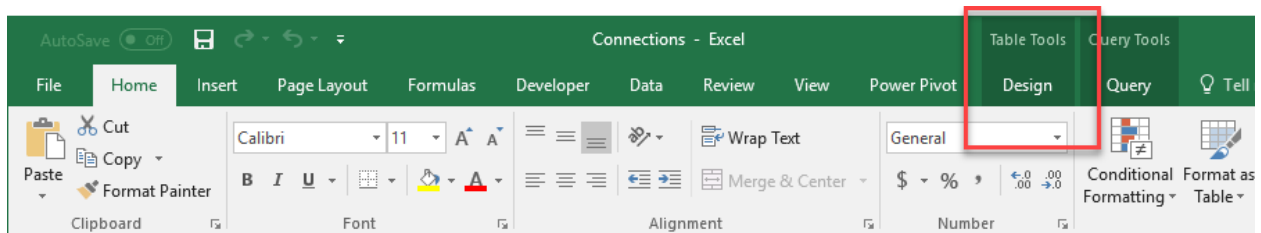
Let's open the connections file and make a couple of changes in order to get it connected properly with the dimensions.

The files come as .CSV (comma separated values) so import them into excel. Simply opening them has given me some trouble with a comma being added to the datetime but your mileage may vary.

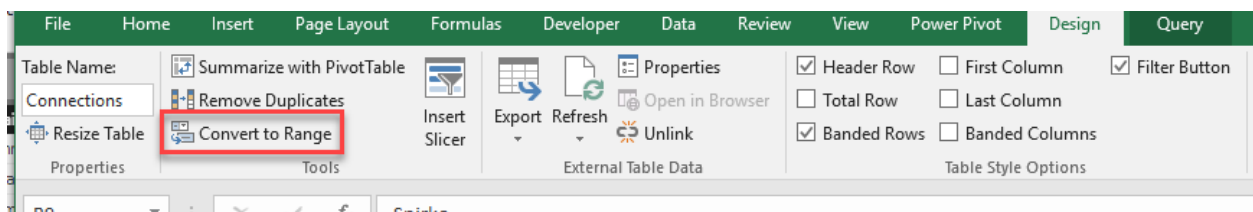


Before you begin remove the table formatting from this worksheet. It's not necessary and it's going to give you issues later.

Left click in the table area and it will activate the table tools. Click on design.



Click convert to range.



Now you can add columns without issues.

Next copy the Connected On Column twice and paste just the values ,do this in column H and I. Label one Date and One Time. Add a third label of SecondsKey in Column J (this matters for the formula later). The SecondsKey is the connection to the time dimension.

	F	G	H	I	J	K
	Connected On	Tags	Date	Time	SecondsKey	
	12/28/2017 8:53		20171228	8:53:00	31980	
	1/7/2018 16:58		20180107	16:58:00	61080	
	12/21/2017 12:25		20171221	12:25:00	18900	

Format the Date column as date

Format Cells
?
X

Number
Alignment
Font
Border
Fill
Protection

Category:

- General
- Number
- Currency
- Accounting
- Date**
- Time
- Percentage
- Fraction
- Scientific
- Text
- Special
- Custom

Sample
Date

Type:

- \*3/14/2012**
- \*Wednesday, March 14, 2012
- 3/14
- 3/14/12
- 03/14/12
- 14-Mar
- 14-Mar-12

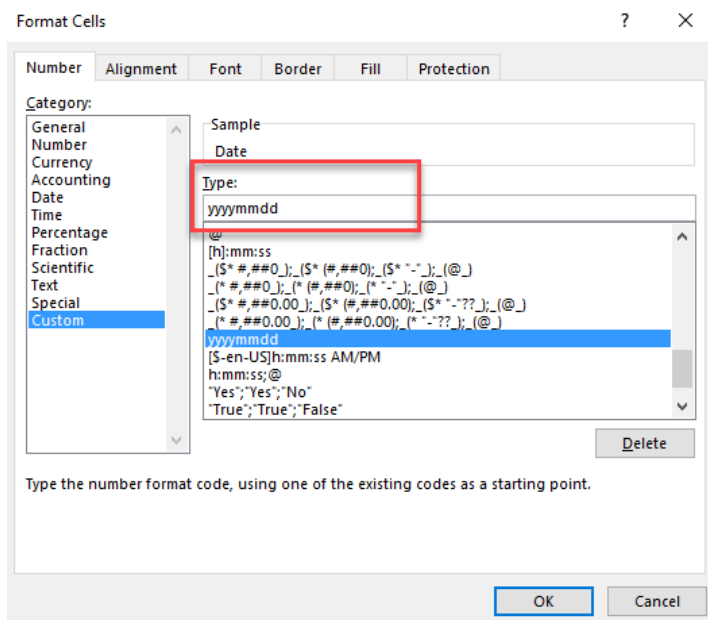
Locale (location):
English (United States)

Date formats display date and time serial numbers as date values. Date formats that begin with an asterisk (\*) respond to changes in regional date and time settings that are specified for the operating system. Formats without an asterisk are not affected by operating system settings.

OK
Cancel



Then format it as YYYYMMDD as shown



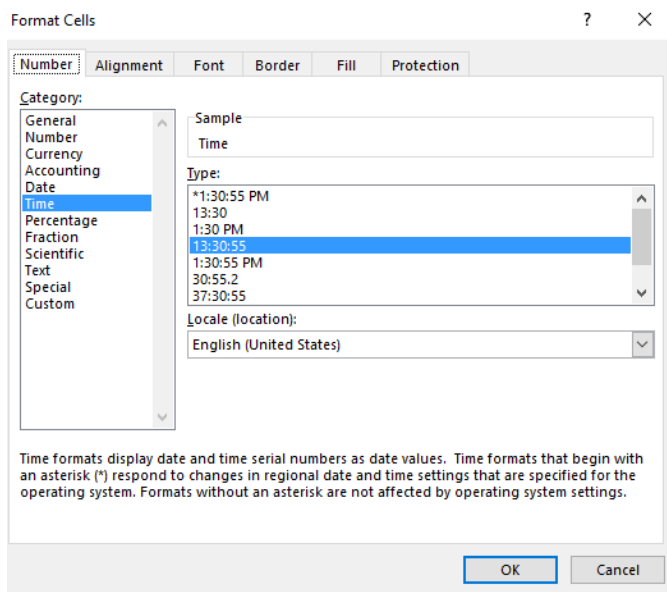
If you don't have the option yyyymmdd then just type it in the type box highlighted with the red square.

This has proven to be a headache for some people do the alternative is to use this formula

`=TEXT((H21,"YYYYMMDD"))`

That will get the date into a text column and the proper format for the joining.

Now format the time as shown:



The formula to get second of the date is:

`=HOUR(I2)*3600+MINUTE(I2)*60+SECOND(I2)`

Just fill the formula and you'll have the seconds.

**Now copy the SecondsKey Column and Date Column and paste them as values only.**

This will make it much easier when it's imported

Now save it and you're ready to get into PowerBI.

# PowerBI

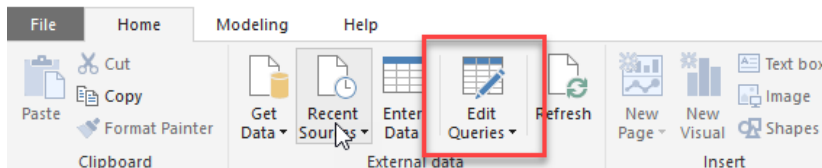
If you don't have it get the desktop here: [PowerBI Desktop](https://powerbi.microsoft.com/en-us/desktop/) (<https://powerbi.microsoft.com/en-us/desktop/>) it's free for this purpose so no worries.

Grab the premade file I loaded to the git repository or create it yourself.

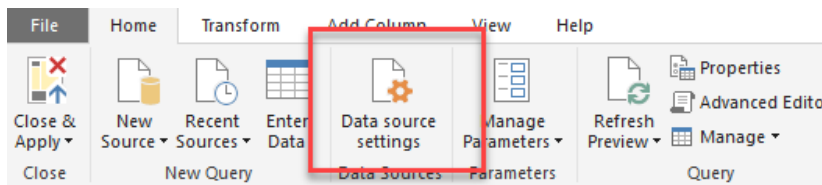
In the premade the relationships are already created and the dimensions are already existing.

If you want to update the files with the ones you pulled yourself all you have to do is:

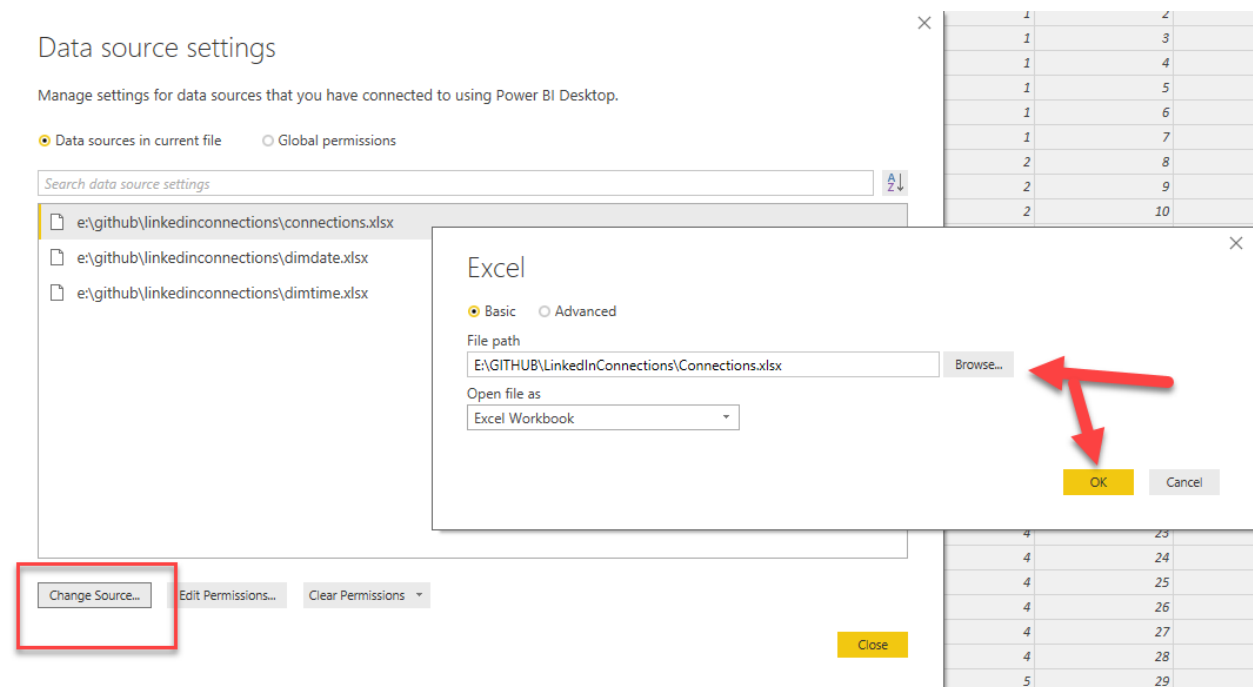
Choose Edit Queries



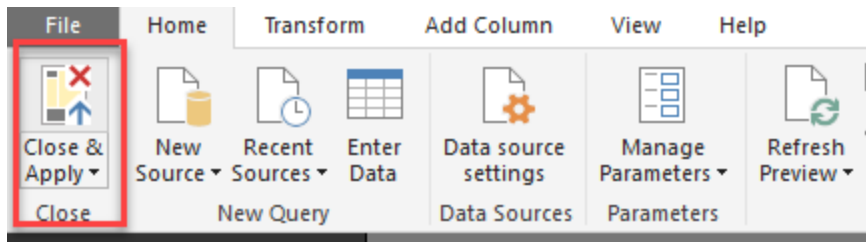
Change Data Source Settings.



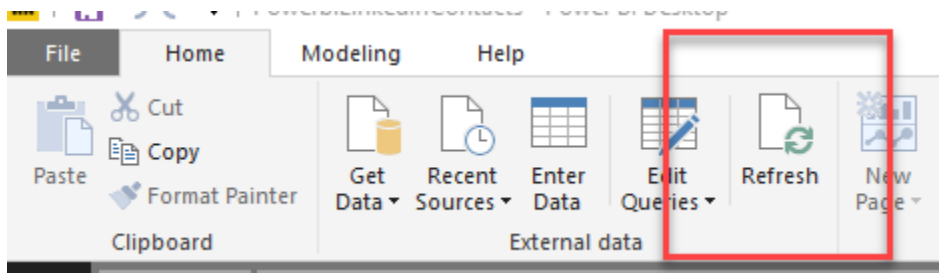
Change the paths to where you put the files



Close and apply



Refresh if it's not already up to date.



## Custom Visuals

---

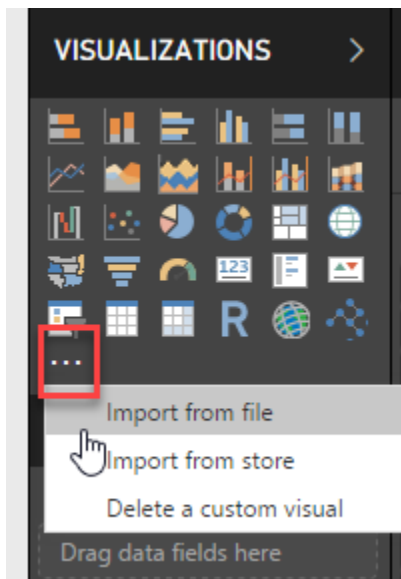
Now get custom visuals.

<https://appsource.microsoft.com/en-us/marketplace/apps?product=power-bi-visuals&page=1>

One I used was the word cloud:

I also pulled in the social graph but that's for another day. Regardless it's in the files.

Install them from files by clicking the ellipsis and import from file.



## Extra Steps

---

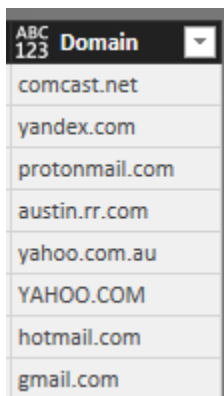
I wanted to go a little farther with the analysis so I'm going to extract the email domain to see who is using what domain in my contacts.

Reopen the connections excel spreadsheet and enter the following formula in a new column titled domain:

```
=RIGHT(C2,LEN(C2)-FIND("@",C2))
```

Reimport the data by clicking Edit Queries

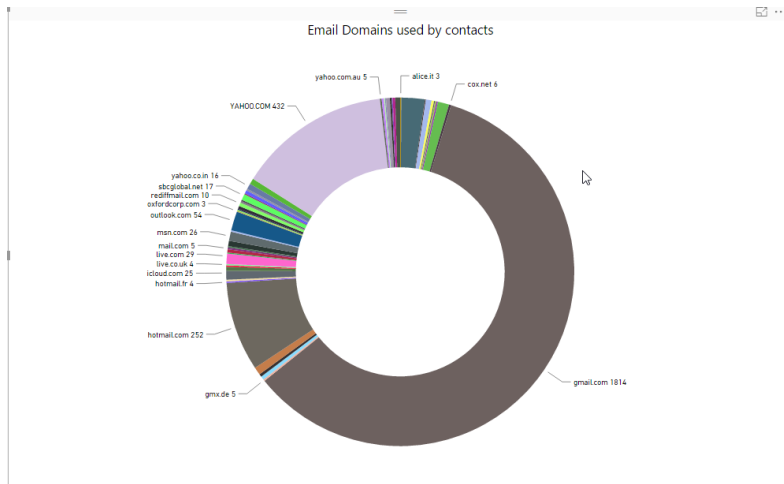
Then Refreshing the data to make sure you have a domains column



ABC 123	Domain
	comcast.net
	yandex.com
	protonmail.com
	austin.rr.com
	yahoo.com.au
	YAHOO.COM
	hotmail.com
	gmail.com

## Make a chart

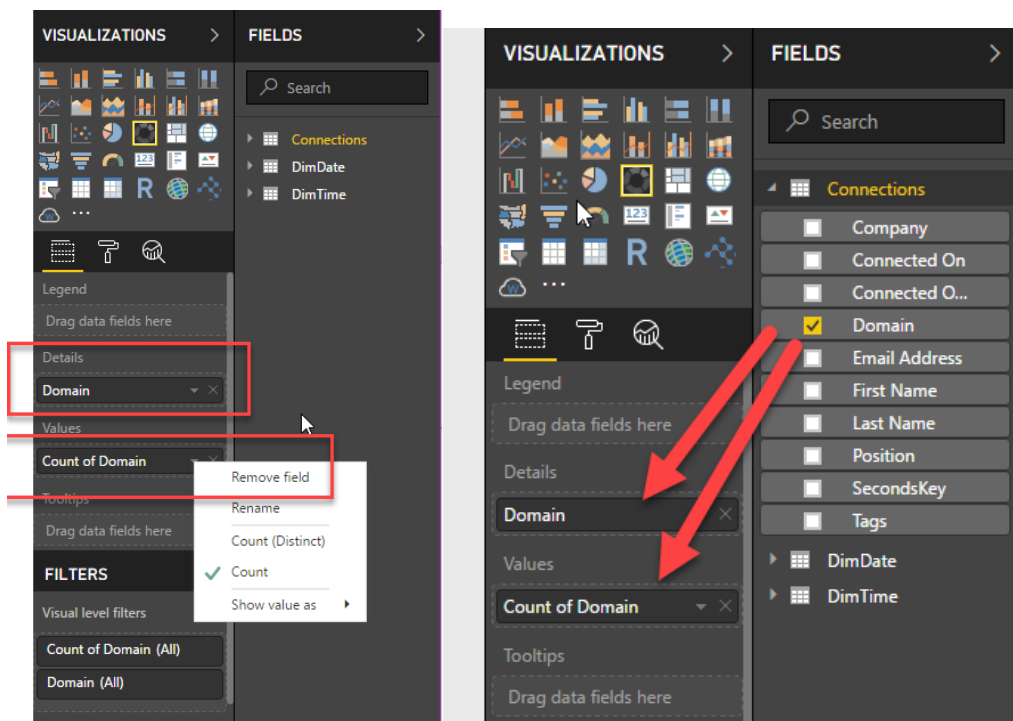
I'm going to use the Domain and make a donut chart



Drag the column Domain into the details section

Now click the donut chart above the legend text. It's got a yellow box around it below.

Last Drag the Column domain into the values section. Set it to be counted or use a percentage. Up to you.



See what you can do.

Here's what I came up with and my analysis says I really just started using LinkedIn appropriately and I probably need to get off at some point in the day.

## Example

