

## LESSON HANDOUT

# Introduction to Power Query

## Why

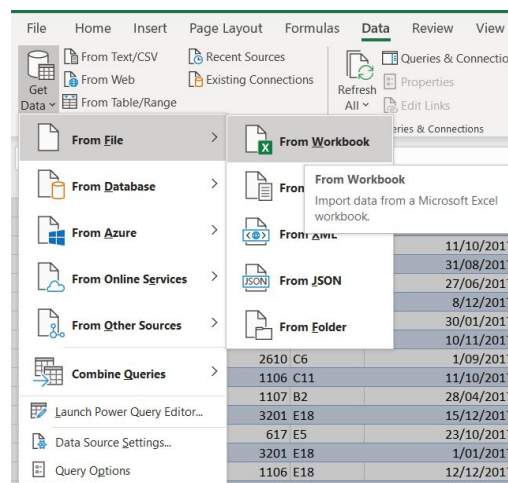
Excel Power Query (EPQ) offers the following:

- **Automation.** Excel Power Query offers the ability to automatically apply transformation steps to new data sources.
- **Integrity and reproducibility.** An EPQ query is less susceptible to errors than the use of in-cell formulas. Both understanding and amending the transformation steps of other users is far easier in EPQ (which avoids long and complex formulas) enhancing reproducibility.
- **Volume.** Data is cached within EPQ allowing the application to handle far more volume than the “front-end” Excel.
- **Auditing.** EPQ automatically records all data transformation steps and additionally allows a user to “rollback” a step.
- **Compatibility.** EPQ, like PowerBI, can connect to multiple data sources and refresh from those sources automatically.
- **Availability.** EPQ is available to anyone using Excel 2013 and onwards.

## Connecting to data sources

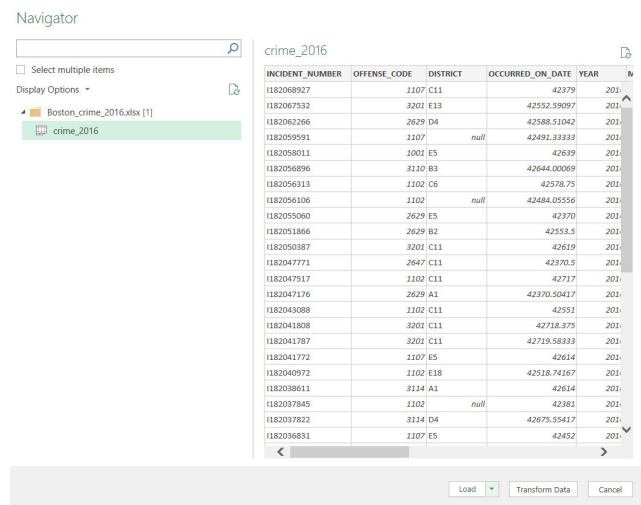
To connect to data from Power Query by clicking ‘Get Data’ in the data tab.

- Example - connecting to an excel worksheet stored locally
  - Get Data > From File > From Workbook

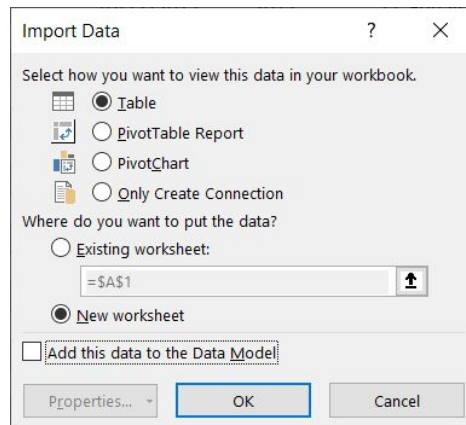


- Choose file > Import

- At the import window, select the worksheet you wish to import



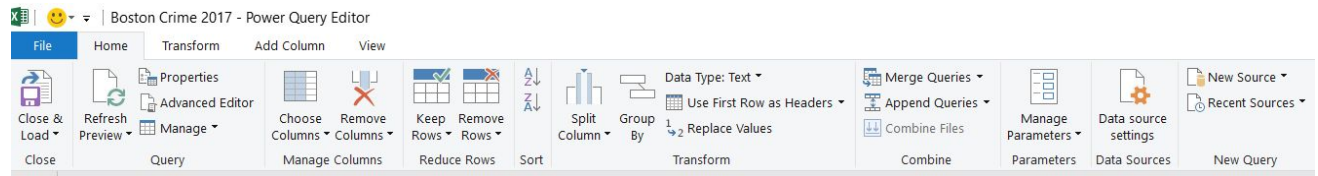
- Clicking on the chevron (drop down arrow) on the 'Load' button will present other loading options, such as 'Only Create Connection', where the data will be loaded as a query but not stored in the worksheet.



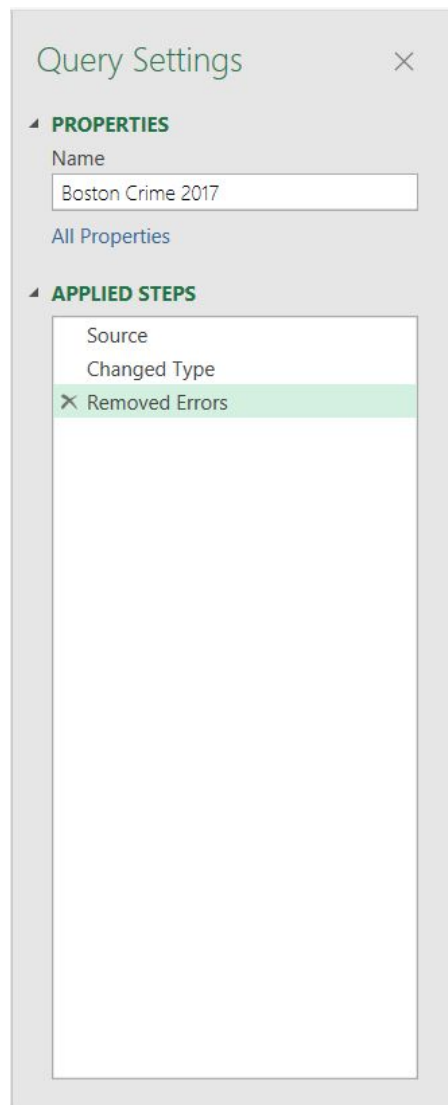
## Transforming data in power query

Power Query Editor has a number of functions to assist in manipulating data - such as removing duplicate rows, splitting by delimiter and creating calculated columns.

These can be found in the ribbon on the Home or Transform tabs.



- Transformation steps are recorded in the right-hand Query Settings bar. Using this bar you may rollback or delete steps.



## Merging data and appending data in power query

Joining data in power query is made easy by the menu interface. First, the data to be merged or appended must be created as queries.

- **Merging queries.** This is a join of one query to another based upon a common column between the two queries.
  - From the Power Query Editor Home > Merge
  - In the merge query window, highlight the common columns and select ok

**Merge**

Select a table and matching columns to create a merged table.

Boston Crime 2017

INCIDENT_NUMBER	OFFENSE_CODE	DISTRICT	OCCURRED_ON_DATE	YEAR	MONTH	DAY_OF_WEEK
I182070304	1107	D4	1/11/2017 12:00:00 AM	2017	11	Wednesday
I182070115	3114	D14	11/10/2017 5:43:00 PM	2017	10	Wednesday
I182069920	3201	B2	31/08/2017 12:08:00 PM	2017	8	Thursday
I182069657	1102	D14	27/06/2017 7:00:00 AM	2017	6	Tuesday
I182069334	3201	B3	8/12/2017 10:00:00 PM	2017	12	Friday

Boston Crime 2017 (Current)

INCIDENT_NUMBER	OFFENSE_CODE	DISTRICT	OCCURRED_ON_DATE	YEAR	MONTH	DAY_OF_WEEK
I182070304	1107	D4	1/11/2017 12:00:00 AM	2017	11	Wednesday
I182070115	3114	D14	11/10/2017 5:43:00 PM	2017	10	Wednesday
I182069920	3201	B2	31/08/2017 12:08:00 PM	2017	8	Thursday
I182069657	1102	D14	27/06/2017 7:00:00 AM	2017	6	Tuesday
I182069334	3201	B3	8/12/2017 10:00:00 PM	2017	12	Friday

Join Kind

Left Outer (all from first, matching from second)

☐ Use fuzzy matching to perform the merge

> Fuzzy matching options

✓ The selection matches 100886 of 100886 rows from the first table.

OK Cancel

- **Appending queries.** Appending is placing one query underneath another based upon common columns from both queries, such as appending multiple yearly queries.
  - From the Power Query Editor Home > Append
  - In the Append query window, add in the queries that you require to append

**Append**

Concatenate rows from three or more tables into a single table.

☐ Two tables ☒ Three or more tables

Available tables

Boston Crime 2017 (Current)

Boston Crime 2016

Add >>

Tables to append

Boston Crime 2017 (Current)

Boston Crime 2016

OK Cancel