Power Query Fundamentals

Course Objectives

Basic Transformations

Import a CSV file and automate basic transformations such as Pivot and Unpivot.

Extracting Data

Learn how to extract more information fields that combine two or more values.

Consolidating Data

Learn how to group or combine data from different tables, or from files within the same folder.

Dealing with Errors

Learn how to avoid, interpret and fix errors that you experience in Power Query.

Power Query
Basic Transformations

Section Objectives

Tasks

01. Import CSV files and familiarize with the Power Query editor.

02. Transform bad data into clean data and load it to a worksheet.

Skills

Identify a CSV File

Load a CSV File

Apply Basic Filters

Delete Unwanted Data

Unpivot Values

Pivot Values

What Is a CSV File

A **comma separated values (CSV)** is a type of text file that shows each of the values in a row, separated by commas.

PQ Exercise 1a - Notepad

File Edit Format View Help

```
QueryName,MonthlySalesAndMarginSnapshot,,,,,,,,,,,,,,,,,,
SourceData,SalesView_Updated2020Jan,,,,,,,,,,,,,,,,,,
Server,CFI_Retail,,,,,,,,,,,,,,,,,,
IP,21.168.312.123,,,,,,,,,,,,,,,,,,
Config Mode,Admin,,,,,,,,,,,,,,,,,,
Query Date,January 1st 2020,,,,,,,,,,,,,,,,,,
Query Time,4:22 PM,,,,,,,,,,,,,,,,,,
Status,Complete,,,,,,,,,,,,,,,,,,
Reference,XXSYS-ENTERPRISEDATA-DOWNLOAD-20292382382,,,,,,,,,,,,,,,,,,
,,,,,,,,,,,,,,,,,,,
Metric,Store,Cat,Ref1,01/01/2017,01/02/2017,01/03/2017,,01/04/2017,01/05/2017,01/06/2017,,
Sales,1,1,11XXSYS,24924.5,46039.49,41595.55,,19403.54,21827.9,21043.39,,22136.64,26229.21,
,1,2,12XXSYS,50605.27,44682.74,47928.89,,44292.87,48397.98,43751.94,,43615.49,41892.55,474
,1,3,13XXSYS,13740.12,10887.84,11523.47,,11135.17,12275.58,10123.45,,9001.37,10366.85,1115
,1,4,14XXSYS,39954.04,35351.21,36826.95,,34660.16,38086.19,32668.67,,34118.11,33872.07,378
,2,1,21XXSYS,35034.06,60483.7,58221.52,,25962.32,27372.05,28660.87,,28446.92,32213.99,8276
,2,2,22XXSYS,74661.16,65487.46,70853.58,,64963.9,68428.64,66622.03,,64307.01,59770.18,6828
,2,3,23XXSYS,16873.2,13821.01,14607.28,,15635.95,14895.96,13061.56,,10394.28,12406.44,1340
,2,4,24XXSYS,47681.96,44197.95,46131.14,,42126.71,46937.81,42489.21,,44622.56,41455.34,479
```

More efficient storage than an Excel file.

The comma is known as a delimiter, since it separates each value.

A CSV File Is Excel

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
1	QueryNan	MonthlySalesAndMarginSnapshot													
2	SourceDat	SalesView_Updated2020Jan													
3	Server	CFI_Retail													
4	IP	21.168.312.123													
5	Config Mc	Admin													
6	Query Dat	January 1st 2020													
7	Query Tim	4:22 PM													
8	Status	Complete													
9	Reference	XXSYS-ENTERPRISEDATA-DOWNLOAD-20292382382													
10															
11	Metric	Store	Cat	Ref1	01/01/2017	01/02/2017	01/03/2017		01/04/2017	01/05/2017	01/06/2017		01/07/2017	01/08/2017	01/09/2017
12	Sales	1	1	11XXSYS	24924.5	46039.49	41595.55		19403.54	21827.9	21043.39		22136.64	26229.21	57258.43
13		1	2	12XXSYS	50605.27	44682.74	47928.89		44292.87	48397.98	43751.94		43615.49	41892.55	47450.5
14		1	3	13XXSYS	13740.12	10887.84	11523.47		11135.17	12275.58	10123.45		9001.37	10366.85	11157.08
15		1	4	14XXSYS	39954.04	35351.21	36826.95		34660.16	38086.19	32668.67		34118.11	33872.07	37809.49
16		2	1	21XXSYS	35034.06	60483.7	58221.52		25962.32	27372.05	28660.87		28446.92	32213.99	82766.07
17		2	2	22XXSYS	74661.16	65487.46	70853.58		64963.9	68428.64	66622.03		64307.01	59770.18	68286.2
18		2	3	23XXSYS	16873.2	13821.01	14607.28		15635.95	14895.96	13061.56		10394.28	12406.44	13404.23
19		2	4	24XXSYS	47681.96	44197.95	46131.14		42126.71	46937.81	42489.21		44622.56	41455.34	47951.76
20															
21															
22															

Excel is great at interpreting CSV files!

We need a way to transform this data.

Database Normalization

Data Normalization is a process of transforming data into a robust form for storage and analysis. The benefits include:

Reduced storage space

Easier maintenance

Improved query speed

Rule #1

Every attribute/value should have it's own, single column.

Database Normalization

Example Sales Table

Sales Value USD		Year	2019	2019	2019	2019
Product #	**Product Group**	**Product Color**	**01/31/2019**	**05/14/2019**	**05/21/2019**	**07/13/2019**
80725	Tools	Chrome	192	0	0	0
80726	Electric Doors	White	0	1,400	0	0
80727	Storage Boxes	Grey	0	0	1,300	0
90724	Radios	Black	0	0	0	57

There should be 4 values (not 16).

There should only be 1 column for sales.

Example Sales Table

Product #	Product Group	Product Color	Date	Sales USD
80725	Tools	Chrome	01/31/2019	192
80726	Electric Doors	White	05/14/2019	1,400
80727	Storage Boxes	Grey	05/21/2019	1300
90724	Radios	Black	07/13/2019	57

Power Query Exercise 1A: Basic CSV Unpivot

Transform a CSV file into a clean list of sales.

PQ Exercise 1a - Notepad

File Edit Format View Help

```
QueryName,MonthlySalesAndMarginSnapshot,,,,,,,,,,,,,,,,
SourceData,SalesView_Updated2020Jan,,,,,,,,,,,,,,,,
Server,CFI_Retail,,,,,,,,,,,,,,,,
IP,21.168.312.123,,,,,,,,,,,,,,,,
Config Mode,Admin,,,,,,,,,,,,,,,,
Query Date,January 1st 2020,,,,,,,,,,,,,,,,
Query Time,4:22 PM,,,,,,,,,,,,,,,,
Status,Complete,,,,,,,,,,,,,,,,
Reference,XXSYS-ENTERPRISEDATA-DOWNLOAD-20292382382,,,,,,,,,,,,,,,,
,,,,,,,,,,,,,,,,,
Metric,Store,Cat,Ref1,01/01/2017,01/02/2017,01/03/2017,,01/04/2017,01/05/2017,01/06/2017,,
Sales,1,1,11XXSYS,24924.5,46039.49,41595.55,,19403.54,21827.9,21043.39,,22136.64,26229.21,
,1,2,12XXSYS,50605.27,44682.74,47928.89,,44292.87,48397.98,43751.94,,43615.49,41892.55,474
,1,3,13XXSYS,13740.12,10887.84,11523.47,,11135.17,12275.58,10123.45,,9001.37,10366.85,1115
,1,4,14XXSYS,39954.04,35351.21,36826.95,,34660.16,38086.19,32668.67,,34118.11,33872.07,378
,2,1,21XXSYS,35034.06,60483.7,58221.52,,25962.32,27372.05,28660.87,,28446.92,32213.99,8276
,2,2,22XXSYS,74661.16,65487.46,70853.58,,64963.9,68428.64,66622.03,,64307.01,59770.18,6828
,2,3,23XXSYS,16873.2,13821.01,14607.28,,15635.95,14895.96,13061.56,,10394.28,12406.44,1340
,2,4,24XXSYS,47681.96,44197.95,46131.14,,42126.71,46937.81,42489.21,,44622.56,41455.34,479
```

	A	B	C	D
1	Store	Cat	Date	Sales
2	1	1	01/01/2017	24924.5
3	1	1	02/01/2017	46039.49
4	1	1	03/01/2017	41595.55
5	1	1	04/01/2017	19403.54
6	1	1	05/01/2017	21827.9
7	1	1	06/01/2017	21043.39
8	1	1	07/01/2017	22136.64
9	1	1	08/01/2017	26229.21
10	1	1	09/01/2017	57258.43
11	1	1	10/01/2017	42960.91
12	1	1	11/01/2017	17596.96
13	1	1	12/01/2017	16145.35
14	1	2	01/01/2017	50605.27
15	1	2	02/01/2017	44682.74
16	1	2	03/01/2017	47928.89
17	1	2	04/01/2017	44292.87
18	1	2	05/01/2017	48397.98
19	1	2	06/01/2017	43751.94
20	1	2	07/01/2017	43615.49
21	1	2	08/01/2017	41892.55
22	1	2	09/01/2017	47450.5
23	1	2	10/01/2017	46549.73
24	1	2	11/01/2017	45025.02
25	1	2	12/01/2017	44418.11
26	1	3	01/01/2017	13740.12
27	1	3	02/01/2017	10887.84
28	1	3	03/01/2017	11523.47

Power Query Exercise 1A: Basic CSV Unpivot

Query Steps

	Steps	Directions
1	Open **Blank Excel File** and import data.	Data > Get Data > From File > From Text/CSV
2	Ensure the file is using a comma delimiter, and press **Transform** to load the Power Query editor.	
3	**Look around the Power Query Editor including: Data Area, List of Queries (left), Query Steps (right), Ribbon (top).**	
4	Delete 10 unwanted header rows from CSV data.	Home > Remove Rows > Remove Top Rows
5	Identify correct headers from the data.	Home > Transform > Use First Row as Headers
6	Get familiar with the query steps pane, identifying one step per transformation.	View > Query Settings
7	Delete unwanted Ref1 and Metric columns.	Click column header > Home > Remove Columns
8	Highlight Store & Cat columns, and unpivot the rest.	Transform > Unpivot Columns > Unpivot Other Columns
9	Filter Blank rows from the Value column.	Use the column header dropdown as in Excel
10	Rename Attribute and Value columns to Date and Sales, respectively.	Double click on column headers to rename
11	Change Date and Sales column types to Date and Decimal Number, respectively.	Home > Data Type: Date or Decimal Number
12	Rename query. Close and load query to worksheet.	Home > Close & Load > Close & Load To

Power Query Exercise 1B: CSV Import N Columns

We now have a query that transforms our data into a clean list of sales.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W
1	QueryNan	MonthlySalesAndMarginSnapshot																					
2	SourceDat	SalesView_Updated2020Jan																					
3	Server	CFI_Retail																					
4	IP	21.168.312.123																					
5	Config Mc	Admin																					
6	Query Dat	January 1st 2020																					
7	Query Tim	4:22 PM																					
8	Status	Complete																					
9	Reference	XXSYS-ENTERPRISEDATA-DOWNLOAD-20292382382																					
10																							
11	Metric	Store	Cat	Ref1	01/01/2017	01/02/2017	01/03/2017		01/04/2017	01/05/2017	01/06/2017		01/07/2017	01/08/2017	01/09/2017		01/10/2017	01/11/2017	01/12/2017		01/01/2018	01/02/2018	01/03
12	Sales	1	1	11XXSYS	24924.5	46039.49	41595.55		19403.54	21827.9	21043.39		22136.64	26229.21	57258.43		42960.91	17596.96	16145.35		24924.5	46039.49	415
13		1	2	12XXSYS	50605.27	44682.74	47928.89		44292.87	48397.98	43751.94		43615.49	41892.55	47450.5		46549.73	45025.02	44418.11		50605.27	44682.74	479
14		1	3	13XXSYS	13740.12	10887.84	11523.47		11135.17	12275.58	10123.45		9001.37	10366.85	11157.08		10179.29	9226.8	8868.93		13740.12	10887.84	115
15		1	4	14XXSYS	39954.04	35351.21	36826.95		34660.16	38086.19	32668.67		34118.11	33872.07	37809.49		36174.43	34740.19	33806.71		39954.04	35351.21	368
16		2	1	21XXSYS	35034.06	60483.7	58221.52		25962.32	27372.05	28660.87		28446.92	32213.99	82766.07		66560	21308.59	22163.12		35034.06	60483.7	582
17		2	2	22XXSYS	74661.16	65487.46	70853.58		64963.9	68428.64	66622.03		64307.01	59770.18	68286.2		64295.22	65172.2	66471.86		74661.16	65487.46	708
18		2	3	23XXSYS	16873.2	13821.01	14607.28		15635.95	14895.96	13061.56		10394.28	12406.44	13404.23		12873.09	11547.5	11862.26		16873.2	13821.01	146
19		2	4	24XXSYS	47681.96	44197.95	46131.14		42126.71	46937.81	42489.21		44622.56	41455.34	47951.76		45280.89	43463.55	43703.76		47681.96	44197.95	461
20																							

But what happens if the source file is updated with more data than before?

Power Query Exercise 1B: CSV Import N Columns

Query Steps

Steps		Directions
1	Duplicate the first query.	Right click on desired query > Duplicate
2	Click on the duplicate and modify the source to **File 1B.**	Use the formula bar to change the referenced source file
3	**Click on the final step to see the what the result looks like. Great, but we don't see the data for 2018 months?**	
4	Change the Source step so that it doesn't specify number of columns.	Delete **Columns=19** from the formula bar
5	**Click on the final query step. The query now returns all the data.**	
6	**But let's consider what would happen if we had less columns than we'd started with.** • Let's click on the Changed Type step. • Each step is named by date, so if dates are different, it wont work. These Changed Type steps were auto-created.	
7	Remove both Changed Type steps that were auto-created.	Delete both **Changed Type** and **Changed Type1** steps
8	Change the settings so that auto-steps are not created.	File > Options > Query Options > Data Load > Type Detection
9	Change data type of Store and Cat to whole number.	Home > Data Type: Whole Number
10	Rename the query to this slide's title.	

Power Query Exercise 1C: Filter Non-dates

	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1	QueryNan	MonthlySalesAndMarginSnapshot												
2	SourceDat	SalesView_Updated2020Jan												
3	Server	CFI_Retail												
4	IP	21.168.312.123												
5	Metric	Store	Cat	Ref1	Ref2	01/01/2017	01/02/2017	01/03/2017	Q1 2017	01/04/2017	01/05/2017	01/06/2017	Q2 2017	01/07/
6	Sales	1	1	11XXSYS	Sales11	24924.5	46039.49	41595.55	112559.5	19403.54	21827.9	21043.39	62274.83	221
7	Sales	1	2	12XXSYS	Sales12	50605.27	44682.74	47928.89	143216.9	44292.87	48397.98	43751.94	136442.8	436
8	Sales	1	3	13XXSYS	Sales13	13740.12	10887.84	11523.47	36151.43	11135.17	12275.58	10123.45	33534.2	90
9	Sales	1	4	14XXSYS	Sales14	39954.04	35351.21	36826.95	112132.2	34660.16	38086.19	32668.67	105415	341
10	Sales	2	1	21XXSYS	Sales21	35034.06	60483.7	58221.52	153739.3	25962.32	27372.05	28660.87	81995.24	284
11	Sales	2	2	22XXSYS	Sales22	74661.16	65487.46	70853.58	211002.2	64963.9	68428.64	66622.03	200014.6	643
12	Sales	2	3	23XXSYS	Sales23	16873.2	13821.01	14607.28	45301.49	15635.95	14895.96	13061.56	43593.47	103
13	Sales	2	4	24XXSYS	Sales24	47681.96	44197.95	46131.14	138011.1	42126.71	46937.81	42489.21	131553.7	446
14														

Remove unwanted data by filtering non-dates.

	A	B	C	D
1	Store	Cat	Date	Sales
2	1	1	01/01/2017	24924.5
3	1	1	02/01/2017	46039.49
4	1	1	03/01/2017	41595.55
5	1	1	04/01/2017	19403.54
6	1	1	05/01/2017	21827.9
7	1	1	06/01/2017	21043.39
8	1	1	07/01/2017	22136.64
9	1	1	08/01/2017	26229.21
10	1	1	09/01/2017	57258.43
11	1	1	10/01/2017	42960.91
12	1	1	11/01/2017	17596.96
13	1	1	12/01/2017	16145.35
14	1	1	01/01/2018	24924.5
15	1	1	02/01/2018	46039.49
16	1	1	03/01/2018	41595.55
17	1	1	04/01/2018	19403.54
18	1	1	05/01/2018	21827.9
19	1	1	06/01/2018	21043.39
20	1	1	07/01/2018	22136.64
21	1	1	08/01/2018	26229.21
22	1	1	09/01/2018	57258.43
23	1	1	10/01/2018	42960.91
24	1	1	11/01/2018	17596.96
25	1	1	12/01/2018	16145.35
26	1	2	01/01/2017	50605.27
27	1	2	02/01/2017	44682.74
28	1	2	03/01/2017	47928.89
29	1	2	04/01/2017	44292.87
30	1	2	05/01/2017	48397.98
31	1	2	06/01/2017	43751.94

Power Query Exercise 1C: Filter Non-dates

Query Steps

Steps		Directions
1	Create a new query and reference CSV **File 1C**.	Home > New Source > File > Text/CSV
2	Change the Source step so that it doesn't specify number of columns.	Delete **Columns=19,** from the formula bar
3	Delete 4 unwanted header rows from CSV data.	Home > Remove Rows > Remove Top Rows
4	Identify correct headers from the data.	Home > Transform > Use First Row as Headers
5	Delete **Ref1, Ref2** and **Metric** columns.	Click column header > Home > Remove Columns
6	Highlight Store & Cat columns, and unpivot the rest.	Transform > Unpivot Columns > Unpivot Other Columns
7	Rename Attribute & Value columns as Date and Sales, respectively.	Double click on column headers to rename them
8	**Before we start working with dates, check your date region settings.**	File Options -> Query Options -> This File -> Regional Settings -> English Canada
9	**Change Date column to Date type by clicking the ABC datatype icon next to it's name. Notice that in line 17, the 2017 TOTAL value is interpreted as a date. We don't want this. Delete this step.**	
10	Instead, let's use the Parse function to interpret dates.	Transform > Date > Parse
11	Filter out the error rows and change the data types of Store and Cat to Whole Number.	
12	Rename the query to this slide's title. **This step will not be mentioned in future exercises.**	

Power Query Exercise 1D: Grouped Row Headers

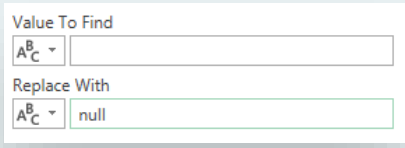
	A	B	C	D	E	F	G	H	I	J	K
1	Metric	Store	Cat	01/01/2017	01/02/2017	01/03/2017	Q1 2017	01/04/2017	01/05/2017	01/06/2017	Q2 2017
2	Sales	1	1					19403.54	21827.9	21043.39	62274.83
3		1	2	50605.27	44682.74	47928.89	143216.9	44292.87	48397.98	43751.94	136442.8
4		1	3	13740.12	10887.84	11523.47	36151.43	11135.17	12275.58	10123.45	33534.2
5		1	4	39954.04	35351.21	36826.95	112132.2	34660.16	38086.19	32668.67	105415
6		2	1	35034.06	60483.7	58221.52	153739.3	25962.32	27372.05	28660.87	81995.24
7		2	2	74661.16	65487.46	70853.58	211002.2	64963.9	68428.64	66622.03	200014.6
8		2	3	16873.2	13821.01	14607.28	45301.49	15635.95	14895.96	13061.56	43593.47
9		2	4	47681.96	44197.95	46131.14	138011.1	42126.71	46937.81	42489.21	131553.7
10	Margin	1	1					0.5432	0.5432	0.5432	0.5432
11		1	2	0.5542	0.5542	0.5542	0.5542	0.5542	0.5542	0.5542	0.5542
12		1	3	0.5212	0.5212	0.5212	0.5212	0.5212	0.5212	0.5212	0.5212
13		1	4	0.5462	0.5462	0.5462	0.5462	0.5462	0.5462	0.5462	0.5462
14		2	1	0.5432	0.5432	0.5432	0.5432	0.5432	0.5432	0.5432	0.5432
15		2	2	0.5542	0.5542	0.5542	0.5542	0.5542	0.5542	0.5542	0.5542
16		2	3	0.5212	0.5212	0.5212	0.5212	0.5212	0.5212	0.5212	0.5212
17		2	4	0.5462	0.5462	0.5462	0.5462	0.5462	0.5462	0.5462	0.5462

	A	B	C	D	E	F
1	Store	Cat	Date	Sales	Margin	Year
2	1	1	01/01/2017			2017
3	1	1	02/01/2017			2017
4	1	1	03/01/2017			2017
5	1	1	04/01/2017	19403.54	0.5432	2017
6	1	1	05/01/2017	21827.9	0.5432	2017
7	1	1	06/01/2017	21043.39	0.5432	2017
8	1	1	07/01/2017	22136.64	0.5432	2017
9	1	1	08/01/2017	26229.21	0.5432	2017
10	1	1	09/01/2017	57258.43	0.5432	2017
11	1	1	10/01/2017	42960.91	0.5432	2017
12	1	1	11/01/2017	17596.96	0.5432	2017
13	1	1	12/01/2017	16145.35	0.5432	2017
14	1	1	01/01/2018	24924.5	0.5432	2018
15	1	1	02/01/2018	46039.49	0.5432	2018
16	1	1	03/01/2018	41595.55	0.5432	2018
17	1	1	04/01/2018	19403.54	0.5432	2018
18	1	1	05/01/2018	21827.9	0.5432	2018
19	1	1	06/01/2018	21043.39	0.5432	2018
20	1	1	07/01/2018	22136.64	0.5432	2018
21	1	1	08/01/2018	26229.21	0.5432	2018
22	1	1	09/01/2018	57258.43	0.5432	2018
23	1	1	10/01/2018	42960.91	0.5432	2018
24	1	1	11/01/2018	17596.96	0.5432	2018
25	1	1	12/01/2018	16145.35	0.5432	2018

- Fix grouped row headers and pivot them to have a column each.
- Create an extra column to extract the year from the date.

Power Query Exercise 1D: Fix Grouped Row Headers

Query Steps

	Steps	Directions
1	Create a new query for **File 1D**. Remove the column number reference as in Exercise 1C.	Remove header rows > Promote headers
2	**Now that we have both sales and margin data, we need to identify each row as such by filling down the grouped row headers.** -Try to fill down. -Nothing happens. This is because blank cells in a text column are not considered empty. -We need the cells to be truly empty, which in Power Query is known as a null value. -We need to replace blanks with null values, but this needs to happen before the Filled Down step. -Ensure that we modify the data before the Filled Down step. -Replace blanks with null values . 	Click Metric Column > Transform > Fill > Down Click on the Promoted Headers Step Select Metric column > Transform > Replace Values
3	Rename the Attribute col to Date. **We can't rename the value column as it has both sales and margin.**	
4	Parse the date column and filter out the errors.	Transform > Date > Parse
5	Pivot the values from the metric column into their own column headers.	Select Metric Column > Transform > Pivot Column > Values Column: Value > Advanced: Don't Aggregate
6	Change data types for all columns to whole numbers/dates/decimals/percentages if appropriate. Notice that when we change the sales column from text to a decimal, it changes from blank to null. **Empty numbers are interpreted as null by default.**	
7	Add a custom column called **Year** that is equal to the date column.	Add Column > Custom Column > = [Date]
8	Extract the year from the new Year column.	Select Year Col > Transform > Date > Year

Power Query Exercise 1Z: Basic Transformations Review

Estimated time: 15 minutes

1. Create another CSV query to fetch data for 1Z.
2. Perform transformations to reach this layout.
3. Notice that we have only kept the quarters.
4. Use the clues or the query steps if you need them.

	A	B	C	D	E
1	Store	Cat	Sales	Margin	Quarter
2	1	1	112559.54	0.5432	Q1 2017
3	1	1	112559.54	0.5432	Q1 2018
4	1	1	62274.83	0.5432	Q2 2017
5	1	1	62274.83	0.5432	Q2 2018
6	1	1	105624.28	0.5432	Q3 2017
7	1	1	105624.28	0.5432	Q3 2018
8	1	1	76703.22	0.5432	Q4 2017
9	1	1	76703.22	0.5432	Q4 2018
10	1	2	143216.9	0.5542	Q1 2017
11	1	2	143216.9	0.5542	Q1 2018
12	1	2	136442.79	0.5542	Q2 2017
13	1	2	136442.79	0.5542	Q2 2018
14	1	2	132958.54	0.5542	Q3 2017

	A	B	C	D	E	F	G	H	I
1	QueryNan	MonthlySalesAndMarginSnapshot							
2									
3	Server	CFI_Retail							
4									
5	Config Mc	Admin							
6									
7	Query Tim	4:22 PM							
8									
9	Store	Cat	Metric	Ref1	Ref2	01/01/2017	01/02/2017	01/03/2017	Q1 2017
10	1	1	Sales	11XXSYS	Sales11	24924.5	46039.49	41595.55	112559.5
11	1	2		12XXSYS	Sales12	50605.27	44682.74	47928.89	143216.9
12	1	3		13XXSYS	Sales13	13740.12	10887.84	11523.47	36151.43
13	1	4		14XXSYS	Sales14	39954.04	35351.21	36826.95	112132.2
14	1	1	Margin	11XXSYS	Margin11	0.5432	0.5432	0.5432	0.5432
15	1	2		12XXSYS	Margin12	0.5542	0.5542	0.5542	0.5542
16	1	3		13XXSYS	Margin13	0.5212	0.5212	0.5212	0.5212
17	1	4		14XXSYS	Margin14	0.5462	0.5462	0.5462	0.5462
18	2	1	Sales	21XXSYS	Sales21	35034.06	60483.7	58221.52	153739.3
19	2	2		22XXSYS	Sales22	74661.16	65487.46	70853.58	211002.2
20	2	3		23XXSYS	Sales23	16873.2	13821.01	14607.28	45301.49
21	2	4		24XXSYS	Sales24	47681.96	44197.95	46131.14	138011.1
22	2	1	Margin	21XXSYS	Margin21	0.5432	0.5432	0.5432	0.5432
23	2	2		22XXSYS	Margin22	0.5542	0.5542	0.5542	0.5542
24	2	3		23XXSYS	Margin23	0.5212	0.5212	0.5212	0.5212
25	2	4		24XXSYS	Margin24	0.5462	0.5462	0.5462	0.5462
26	XX	Report Code 1							
27	XY	Report Code 2							
28	XZ	Report Code 3							

Power Query Exercise 1Z: Basic Transformations Review

Clues

Clue

You may need to keep errors instead of removing them.

Clue

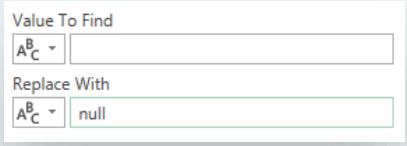
You will need to add an extra column as we did in the previous exercise.

Clue

You may need to manually filter a column, or two.

Power Query Exercise 1Z: Basic Transformations Review

Query Steps

Part 1		Part 2	
New Query for File 1Z.		Add a copy of the date column and call it DateBackup.	
Remove the column number reference as in Exercise 1C.		Parse the date column and **KEEP** the errors.	Select Date Column > Transform > Date > Parse
Remove header rows.		Delete the date column.	
Promote headers.		Filter out 2017 and 2018 from the DateBackup column.	
Remove Ref1 and Ref2 columns.		Filter the Store column to **KEEP** only values 1 and 2.	
Replace blanks with nulls.		Rename the DateBackup column to Quarter.	
Fill down nulls.			
Unpivot dates.			
Rename Attribute column.			
Pivot Metric column.	Select Metric column > Transform > Pivot Column > Values Column: Value > Advanced: Don't Aggregate	Change data types where appropriate.	

There are many ways to achieve this result.

This is possibly the quickest & easiest.

When we review this exercises we'll look at why using filters in this way may not be future proof.

Power Query
Extracting Information

Section Objectives

Tasks

01. **Import** data from Excel sheets and **learn** to futureproof them.

02. **Split** columns that contain multiple values into separate columns.

Skills

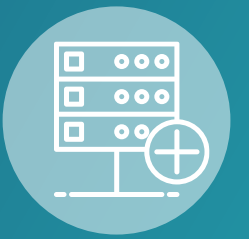
Extract Characters from Text

Remove Blank Column from Excel

Extract Text Using Delimiters

Split Columns by Position

Split Columns into Individual Rows



Add Prefixes to Columns

Dynamic Filtering

In the last exercise, we created two manual filters:

1

Keep only quarterly subtotals

([Date - Copy] <> "2017" and [Date - Copy] <> "2018")

If 2019 appeared in the dataset, the 2019-year total would not be removed from rows.

2

Keep only stores 1 and 2, and exclude rows XY, XZ numbers

([Store] = "1" or [Store] = "2")

If Store 3 appeared in the dataset, it would be incorrectly excluded from the result.

Power Query Exercise 2A: Better Filters

Query Steps

	Steps	Directions
1	Open the file named **CFI PQ Exercises – Chapter 2**.	
2	Click on the query named **PQ Exercise 2A – Better Filters**. **This is a duplicate of Exercise 1Z, except I've removed all the steps where we tried to filter our data at the end.**	
3	Identify quarters by creating a helper column to extract first character of the DateBackup column.	Select DateBackup column > Add Column > Extract > First Characters > 1
4	Filter to keep only the Qs.	
5	Delete the helper column.	
6	**We could also identify quarters using Length or Spaces, but identifying a Q seems to most reliable.**	
7	Next filter the store column as before, to include only stores 1 and 2.	
8	Amend the formula from **[Store] = "1" or [Store] = "2"** to **[Store] <> "XX" and [Store] <> "XY"**	
9	Change column types as appropriate.	
10	**Right click in a blank space in the queries area to create a folder to keep things organized into chapters.** Create one for Basic Transformations, and now one for Extracting Information.	

Power Query Exercise 2B: Import Basic Excel Files

When importing Excel files we must choose which sheets/tables to import.

Are these columns really empty?

	A	B	C	D	E	F	G	H	I	J	K
1	ID	Date	GL Account & Transaction Description	User	Journal & Ref	Amount					
2	1	01/01/2018	Share Capital Injection 1000	m pence	J10180101Ref-2369	120000 USD					
3	2	01/01/2018	Share Capital Injection 4000	m pence	J10180101Ref-5630	-120000 USD					
4	3	01/01/2018	Purchase PPE 2000	M PENCE	J10180102Ref-9778	28571 USD					
5	4	01/01/2018	Purchase PPE 3000	M PENCE	J10180102Ref-7787	-29999.55 USD					
6	5	01/01/2018	GST ITC for PPE 3300	M PENCE	J10180102Ref-3423	1428.55 USD					
7	6	01/01/2018	Purchase Inventory 1200	J HIGGINS	J10180103Ref-8268	85000 USD					
8	7	01/02/2018	Purchase Inventory 1200	J HIGGINS	J10180203Ref-7446	81775.4514600117 USD					
9	8	01/03/2018	Purchase Inventory 1200	J HIGGINS	J10180303Ref-713	99409.776775586 USD					
10	9	01/04/2018	Purchase Inventory 1200	J HIGGINS	J10180403Ref-1799	133071.383711542 USD					
11	10	01/05/2018	Purchase Inventory 1200	j higgins	J10180503Ref-3334	179531.822559164 USD					
12	11	01/06/2018	Purchase Inventory 1200	j higgins	J10180603Ref-532	207318.014707784 USD					
13	12	01/07/2018	Purchase Inventory 1200	J HIGGINS	J10180703Ref-2514	231793.872344711 USD					
14	13	01/08/2018	Purchase Inventory 1200	j higgins	J10180803Ref-4985	232808.055834635 USD					
15	14	01/09/2018	Purchase Inventory 1200	j higgins	J10180903Ref-3914	212756.514367055 USD					

Power Query Exercise 2B: Import Basic Excel File

Query Steps

	Steps	Directions
1	Create a new query and source data from **Excel File 2B**.	
2	Select the **GL transactions** sheet.	
3	Promote headers.	
4	**Sure enough we've got 35 columns of data. Only 6 are useful.**	
5	**It is tempting at this point to select and delete column 7 onwards. However, thinking again of automation and futureproofing, what if next time there are 37 columns?**	
6	A better method is to select the columns you want to keep, and to remove the others.	Select first 6 columns > Home > Remove Columns > Remove Other Columns

Power Query Exercise 2c - Splitting Columns

ID	Date	GL Account & Transaction Description	User	Journal & Ref	Amount
1	01/01/2018	Share Capital Injection 1000	m pence	J10180101Ref-2369	120000 USD
2	01/01/2018	Share Capital Injection 4000	m pence	J10180101Ref-5630	-120000 USD
3	01/01/2018	Purchase PPE 2000	M PENCE	J10180102Ref-9778	28571 USD
4	01/01/2018	Purchase PPE 3000	M PENCE	J10180102Ref-7787	-29999.55 USD
5	01/01/2018	GST ITC for PPE 3300	M PENCE	J10180102Ref-3423	1428.55 USD
6	01/01/2018	Purchase Inventory 1200	J HIGGINS	J10180103Ref-8268	85000 USD

3 have combined values that need to be split up.

Rule #1

Every attribute/value should have it's own, single column.

Power Query Exercise 2C: Splitting Columns

Query Steps

Create a new referenced query from 2B, and rename it as per the title of this slide. Notice that the source directly quotes query 2B.

Splitting the Amount Column	
Try splitting by digits and non-digits.	Select Amount Col > Transform > Split Column > Digit to Non-Digit
Delete the above step. It doesn't give us the desired result as we end up with . and – as non-digits.	
Try extracting text before delimiter.	Select Amount Col > Transform > Extract > Text Before Delimiter

Splitting the GL Account & Description Column
We could split the Description GL column using non-digits to digits, but what happens if the description contains digits?
Split the GL & Description column using a space delimiter but only at the right most occurrence.
Rename the two new columns to **Desc** and **GL Account**.

Splitting the Journal & Ref Column	
Split the Journal & Ref column after the 9th character.	Select Col > Transform > Split Column > By # of Characters > 9
Rename the Journal column and delete the Ref column.	
Change data types and transform the user column to upper case.	Transform > Format > Uppercase

Commenting Code

```
let

    // Reference Excel file data from previous query
    Source = #"PQ Exercise 2b - Import Basic Excel File",

    //Extract USD Amount from Amount Col
    #"Extracted Text Before Delimiter" = Table.TransformColumns(Source, {{"Amount", each Text.BeforeDelimiter(_, " "), type text}}),
    #"Changed Type" = Table.TransformColumnTypes(#"Extracted Text Before Delimiter",{{"Amount", type number}}),
    #"Split Column by Delimiter" = Table.SplitColumn(#"Changed Type", "GL Account & Transaction Description", Splitter.SplitTextByEachDelimit
    #"Renamed Columns" = Table.RenameColumns(#"Split Column by Delimiter",{{"GL Account & Transaction Description.1", "Desc"}, {"GL Account &

    //Remove the Ref from the Journal Col
    #"Split Column by Position" = Table.SplitColumn(#"Renamed Columns", "Journal & Ref", Splitter.SplitTextByPositions({0, 9}, false), {"Jour
    #"Renamed Columns1" = Table.RenameColumns(#"Split Column by Position",{{"Journal & Ref.1", "Journal"}, {"Journal & Ref.2", "Ref"}}),
    #"Removed Columns" = Table.RemoveColumns(#"Renamed Columns1",{"Ref"}),

    //Change data types
    #"Uppercased Text" = Table.TransformColumns(#"Removed Columns",{{"User", Text.Upper, type text}}),
    #"Changed Type3" = Table.TransformColumnTypes(#"Uppercased Text",{{"Date", type date}, {"ID", Int64.Type}})
in
    #"Changed Type3"
```

The query code can be intimidating, especially if it was created by someone else.

Use comments to help the next person understand the process.

Add comments using a double //. Comments are not interpreted by the code.

Power Query Exercise 2D: Splitting Columns into Rows

Sometimes it can be more appropriate to split a column into rows.

Look at the below example of combined product orders by customer and date.

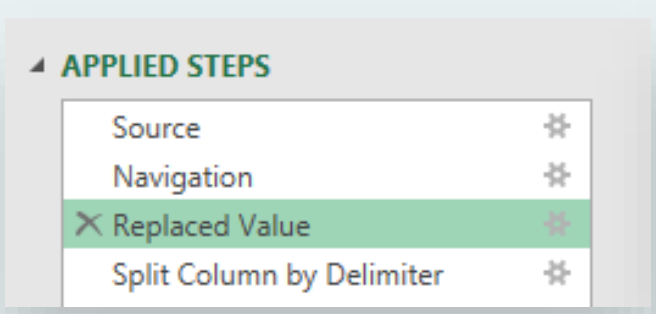
	A	B	C	D	E
1	Order ID	Date	Products	Customer	Department
2	1	01/01/2020	X6034;X9463;X7377	Robert	Kitchens
3	2	21/11/2019	X5353	James	Kitchens
4	3	08/01/2020	X9809;X8288;X6262	Mohammed	Bathroom
5	4	15/01/2020	X4271;X8273;X92930;X2626262	Pooja	Kitchens
6	5	28/01/2020	X9809;X8288;X6262	Chen	Bathroom
7	6	12/12/2019	X1123;;X8392	Flo	Bathroom

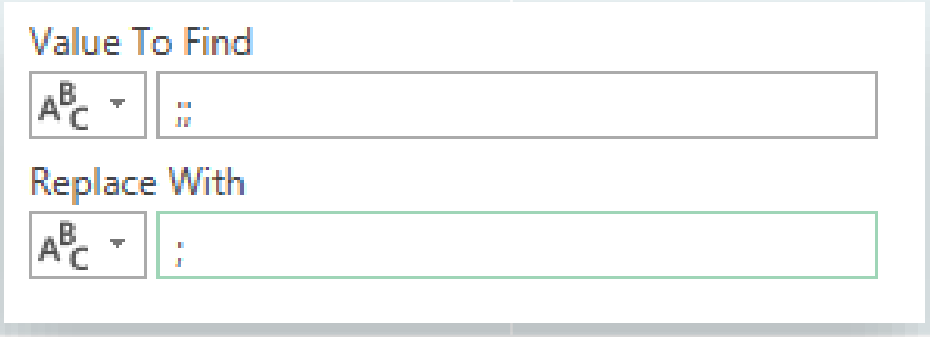
	Order ID	Date	Products	Customer	Department
1	1	01/01/2020	X6034	Robert	Kitchens
2	1	01/01/2020	X9463	Robert	Kitchens
3	1	01/01/2020	X7377	Robert	Kitchens
4	2	21/11/2019	X5353	James	Kitchens
5	3	08/01/2020	X9809	Mohammed	Bathroom
6	3	08/01/2020	X8288	Mohammed	Bathroom
7	3	08/01/2020	X6262	Mohammed	Bathroom
8	4	15/01/2020	X4271	Pooja	Kitchens
9	4	15/01/2020	X8273	Pooja	Kitchens
10	4	15/01/2020	X92930	Pooja	Kitchens
11	4	15/01/2020	X2626262	Pooja	Kitchens

Power Query Exercise 2D: Splitting Columns into Rows

Query Steps

	Steps	Directions
1	Open **File 2D** as an Excel file. **Notice that the Products Ordered tab contains a table of data called Customer Orders.**	
2	Close File 2D and return to your working Power Query file.	
3	**Create a new query to reference Excel File 2D and select the "CustomerOrders" table of data instead of a sheet.**	
4	**First, try splitting the Products column using a semi-colon delimiter. Notice that we have 4 columns of Product data.**	
5	**Delete the previous step.**	Transform > Split Column > ; > Advanced > Rows
6	Split the Products column into rows using a delimiter.	
7	**Notice that we have a blank row. Click on the previous step called Navigation to find out why.** Row 6 has two consecutive semi-colons.	
8	Insert a replace values step at this point to replace all double semi-colons with single, in the Product column.	Select Products Column > Transform > Replace Values
9	Click on the Split Column by Delimiter step to ensure any subsequent steps are placed after this one.	
10	Change data types as appropriate.	





Power Query Exercise 2Z: Extracting Information Review

Estimated time: 15 minutes

1. Transform data from 2Z DATA SET A.
2. Think about what would happen if the data changes, or grows.

Division	Department	Supplier	Brands	Quarter	Value
Division 1	ALCOHOL	Sup A	Jack Daniels Black	Q12019	60.25
Division 1	ALCOHOL	Sup A	Jack Daniels Black	Q22019	61.67
Division 1	ALCOHOL	Sup A	Jack Daniels Black	Q32019	60.53
Division 1	ALCOHOL	Sup A	Jack Daniels Black	Q42019	59.13
Division 1	ALCOHOL	Sup A	Jack Daniels	Q12019	60.25
Division 1	ALCOHOL	Sup A	Jack Daniels	Q22019	61.67
Division 1	ALCOHOL	Sup A	Jack Daniels	Q32019	60.53
Division 1	ALCOHOL	Sup A	Jack Daniels	Q42019	59.13
Division 1	ALCOHOL	Sup B	Tequilla 1	Q12019	49.07
Division 1	ALCOHOL	Sup B	Tequilla 1	Q22019	48.3
Division 1	ALCOHOL	Sup B	Tequilla 1	Q32019	50.75
Division 1	ALCOHOL	Sup B	Tequilla 1	Q42019	51.37
Division 1	ALCOHOL	Sup B	Tequilla 2	Q12019	49.07
Division 1	ALCOHOL	Sup B	Tequilla 2	Q22019	48.3
Division 1	ALCOHOL	Sup B	Tequilla 2	Q32019	50.75
Division 1	ALCOHOL	Sup B	Tequilla 2	Q42019	51.37
Division 1	ALCOHOL	Sup C	London Dry	Q12019	47.87
Division 1	ALCOHOL	Sup C	London Dry	Q22019	47.53
Division 1	ALCOHOL	Sup C	London Dry	Q32019	49.15

Duty Free Sales CAD			
2019			
Product Details	**Supplier**	**Brands**	**Q12019 | Q22019 | Q32019 | Q42019**
Division 1-Alcohol	Sup A	Jack Daniels Black; Jack Daniels	60.25 | 61.67 | 60.53 | 59.13
Division 1-Alcohol	Sup B	Tequilla 1; Tequilla 2	49.07 | 48.3 | 50.75 | 51.37
Division 1-Alcohol	Sup C	London Dry;	47.87 | 47.53 | 49.15 | 48.34
Div 1-Tobacco	Sup D	Marlboro	40.28 | 42.97 | 44.02 | 45.13
Div 1-Tobacco	Sup E	Karelia; Sherman;; Dunhill	45.8 | 43.94 | 43.48 | 42.17
Div 1-TOBACCO	Sup F	Amber Leaf	45.24 | 43.66 | 45.82 | 47.33
Div 2-Perfume	Sup G	Bvlgari; Gucci; Burberry	40.93 | 39.14 | 38.43 | 38.21
Div 2-PERfume	Sup H	Tommy; J Lo; Ariana Grande	48.63 | 47.91 | 47.14 | 45.58
Div 2-Perfume	Sup I	DIOR; Davidoff; MAC	45.4 | 43.55 | 45.86 | 46.61

Power Query Exercise 2Z: Extracting Information Review

Clues

Clue

Remove Top Rows > Split Column to Columns by Delimiter > Promote to Headers

Clue

You may need to change the order of your steps.

Clue

Remember, delimiters can have more than one character.

Clue

Transform > Format > Trim helps you remove spaces from the end of text.

Power Query Exercise 2Z: Extracting Information Review

Query Steps Part 1

Setting up for Success	
New Query for **File 2Z**– Use Data Set A.	
Remove top 4 rows.	
Notice that the headers in column 4 are also merged into one column, so we can't promote them to headers yet.	
For that same reason, we can't split the column into rows, because we need to keep the quarter attribute next to each value.	
Split column4 using a 3 character delimiter. **" | "**	
Filter out blank rows.	

Power Query Exercise 2Z: Extracting Information Review

Query Steps Part 2

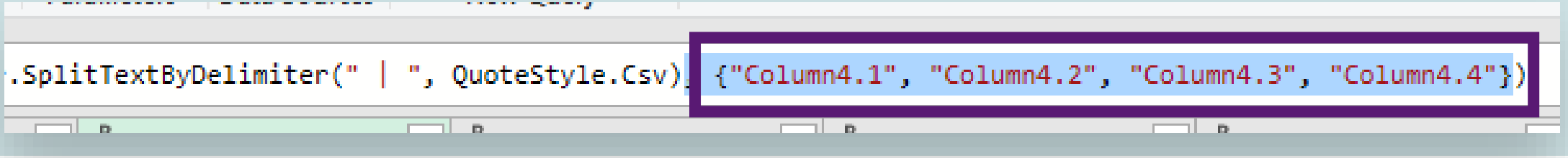
Splitting the Product Details Column	
Split the Product Details column by delimiter.	
Rename the two columns to Division and Department.	
Extract the last character from the division column to get consistent presentation of division number.	
Optional: Add "Division " as a prefix to the division column.	Transform > Format > Add Prefix
Optional: Uppercase the Department column.	

Splitting the Brands Column into Rows	
Split Brands column into rows using a delimiter.	
Trim Brands column to remove spaces.	Transform > Format > Trim
Filter blank rows from Brands column.	

Transform Quarters into a Column Attribute	
Unpivot Quarter columns.	
Rename Attribute column to Quarter.	
Change data types where appropriate.	

Power Query Exercise 2Z: Extracting Information Review

Query Steps Part 3

Splitting Into N Columns
Click on the Navigation step.
Change **DATA SET A** to **DATA SET B** in the formula bar.
Click on the final step – do we have all the data? No!
Click on the query step that splits the quarterly data in separate columns.
Notice in the formula bar that we have explicit names for the columns that we've added: 
Delete these explicit names highlighted in blue, including the curly brackets, and the comma before them.
By removing these explicit names, we allow the query to create a dynamic number of columns.
Click back on the final step. You should now see all the data.

Power Query
Consolidating Data

Section Objectives

Tasks

01. **Group and aggregate** detailed data into simple summaries.

02. **Merge data** from different tables and files into a table.

Skills

Identify Month of Each Row

Group Rows by Month

Aggregate Similar Columns

Perform Table Joins

Combine Data from Multiple Excel Files

Extract Info from File Names

Power Query Exercise 3A: Grouping Data

Sometimes the data contains too much detail, when a summary is sufficient.

Date	Time	Manager 1	Manager 2	Manager 3
01/01/2019	09:00 AM	107846.18	74921.02	30100.15
01/01/2019	10:00 AM	102922.75	70556.22	30681.21
01/01/2019	11:00 AM	108160.04	70440.04	30252.93
01/01/2019	12:00 PM	105973.11	74627.08	33187.2
01/01/2019	1:00 PM	105127.22	74170.37	30401.82
01/01/2019	2:00 PM	102505.73	71036.63	34867.47
01/01/2019	3:00 PM	108834.68	71775.75	30975.44
01/01/2019	4:00 PM	109377.75	71882.74	34421.13
01/02/2019	09:00 AM	103276.36	72440.21	32530.62
01/02/2019	10:00 AM	105555.95	71962.54	32709.55
01/02/2019	11:00 AM	104703.79	70864.9	30813.45
01/02/2019	12:00 PM	103897.64	72413.57	31862.7

Desired Table

Month	Total Sales
Jan 2019	TOTAL
Feb 2019	TOTAL

Required Transformations

Aggregate columns

Group rows

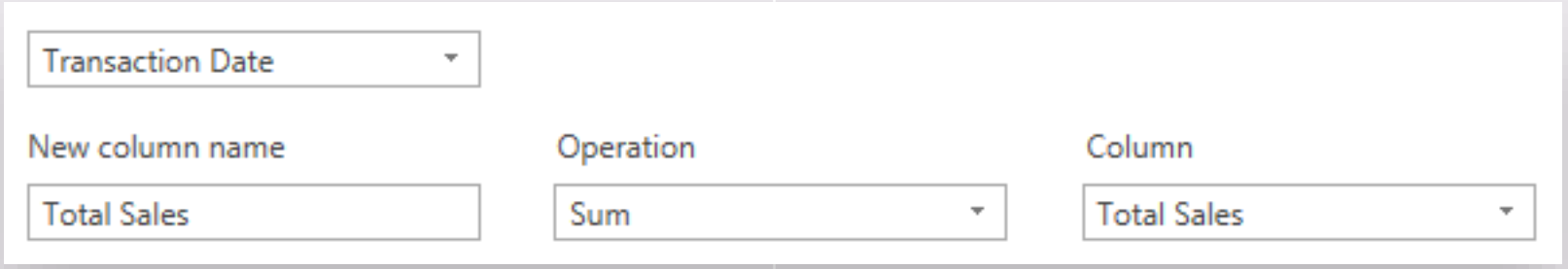
Power Query Exercise 3A: Grouping Data

Query Steps

To Start	
Create a new query to source **File 3A**, and fetch the table named **DailyManagerSales.**	
Rename the query.	

To Aggregate the Sales Columns into One	
Create a summation column of Total Sales.	Select 3 sales columns > Add Column > Standard > Add
Rename the new column to Total Sales and delete the other columns.	

To Aggregate the Sales Columns into One	
Identify the month of each row.	Select date column > Transform > Date > Month > Start of Month
Group the Date column into months.	Select date column > Group By > Transaction Date



Power Query Exercise 3B: Merging Tables

Sometimes our fact or dimension data is stored across two different tables.

Distribution Orders

Transaction ID	Order Date	Brand	Category ID	ValueAtCost
1	01/01/2018	Belvedere	Vodka	3,345.91
2	01/01/2018	Absolut	Vodka	2,358.81
3	01/01/2018	Balvenie	Whiskey	449.34
4	01/01/2018	Abelour	Whiskey	2,506.41
5	01/01/2018	Hendricks	Gin	5,498.58
6	01/01/2018	Highclere Castle	Gin	1,816.36
7	01/01/2018	Tanqueray	Gin	7,719.54

Traditional Method: Vlookup

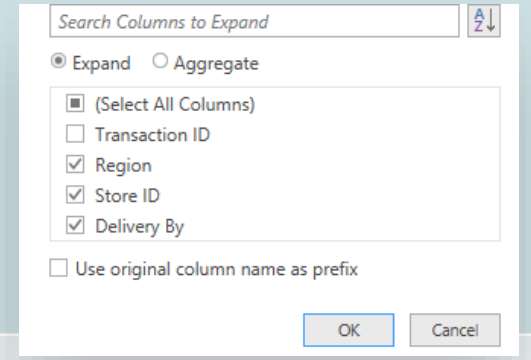
Power Query Method: Merge Tables

Order Destination Details

Transaction ID	Region	Store ID	Delivery By
1	North	1	01/29/2018
2	North	1	02/06/2018
3	North	1	02/09/2018
4	North	1	01/13/2018
5	North	1	02/01/2018
6	North	1	02/01/2018
7	North	1	01/29/2018
8	North	1	01/20/2018
9	North	1	01/11/2018
10	North	1	01/14/2018
11	North	1	01/17/2018
12	East	2	01/24/2018

Power Query Exercise 3B: Merging Tables

Query Steps

	Steps	Directions
1	Create a new query to source **File 3B,** and fetch the sheet named **DistributionOrders**.	
2	Remove top rows and promote headers.	
3	Create a new query to import **Excel File 3B, OrderDestinations** sheet.	
4	Promote headers.	
5	Ensure that Transaction ID in both tables is set to Whole Number type.	
6	**Click on the original query from Step 1.**	
7	Merge the query with additional data.	Home > Merge Queries > Merge Queries
8	**Select the Order Destinations Query to merge with. Then click on the Transaction ID column in each. This will be our lookup column.**	
9	Choose the join kind that represents a vlookup.	Join Kind > Left Outer
10	Click on the expanding arrows icon on our new column to expand the additional data from the other table.	
11	De-select Transaction ID.	
12	**Do not use original column name as prefix.**	

Power Query Exercise 3C: Combining Identical Files

North Region Data

Header 1					
Header 2					
Header 3					
Transaction	Order Date	Brand	Category ID	ValueAtCost	Region ID
1	01/01/2018	Belvedere	Vodka	5106.434805	North
2	01/01/2018	Absolut	Vodka	7316.406025	North
3	01/01/2018	Balvenie	Whiskey	8095.339585	North
4	01/01/2018	Abelour	Whiskey	7432.277286	North
5	01/01/2018	Hendricks	Gin	3543.867122	North

South Region Data

Header 1					
Header 2					
Header 3					
Transaction ID	Order Date	Brand	Category ID	ValueAtCost	Region ID
22	01/01/2018	Belvedere	Vodka	2785.383164	South
23	01/01/2018	Absolut	Vodka	4940.981403	South
24	01/01/2018	Balvenie	Whiskey	7852.254204	South
25	01/01/2018	Abelour	Whiskey	9099.641253	South
26	01/01/2018	Hendricks	Gin	4522.477689	South

East Region Data

Header 1					
Header 2					
Header 3					
Transactic	Order Date	Brand	Category ID	ValueAtCost	Region ID
12	01/01/2018	Belvedere	Vodka	6715.961603	East
13	01/01/2018	Absolut	Vodka	8085.748092	East
14	01/01/2018	Balvenie	Whiskey	6796.326432	East
15	01/01/2018	Abelour	Whiskey	6208.618074	East
16	01/01/2018	Patron	Tequila	6774.201612	East

Files can contain different data, but be formatted identically:

- **Same number of header rows**
- **Same column headers**
- **Same data types**

Power Query Exercise 3C: Joining Identical Files

Query Steps Part 1

	Steps	Directions
1	Create a new query, but instead of selecting a file, select a folder.	New Source > File > Folder
2	Browse to folder **PQ Exercise 3C.**	Browse to Folder > Press OK
3	Press Combine & Transform Data.	Combine & Transform Data
4	Select the sheet you'd like to combine from each workbook.	Press OK
5	Minimize the Helper Queries Folder. **We won't cover these supporting elements in this course. What's important is the new query called "Transform Sample File". This file allows us to transform each of the individual files, before they are combined.**	
6	Rename the Transform Sample File query to **Transform 3C Individual Files**.	
7	Remove top header rows and promote headers.	

Power Query Exercise 3C: Joining Identical Files

Query Steps Part 2

	Steps	Directions
1	Click back on our main PQ Exercise 3C query to see the combined data. Not only do we have the combined data, but we have the Source. **Name filename so we know the region of each transaction. Let's extract the region from that column.**	
2	Extract text before delimiter.	Transform > Extract > Text Before Delimiter > .xlsx
3	**We can't extract the last X characters, because EAST and SOUTH are of different lengths.**	
4	Instead, extract text after delimiter. **Note: Make sure you include the space in the delimiter after the 3C.**	Transform > Extract > Text After Delimiter > PQ Exercise 3C
5	Rename the column to Region.	
6	**Notice that our transaction IDs start at 12. Click on the filter button to inspect all the values in that column. It only displays the first 1000, which makes it difficult to see if transaction ID #1 is present.**	
7	Sort the transaction ID column ascending to check that all the data is present.	

Power Query Exercise 3Z: Consolidating Data Review

Estimated time: 15 minutes

1. Transform data from file 3Z.
2. Summarize total Sales $ and Margin $ by month.
3. Group rows by month.

	Transaction Date	1.2 Total Sales	1.2 Total Margin
1	01/01/2019	2162720.627	762241.6291
2	01/02/2019	1969183.44	693103.8918
3	01/03/2019	2172044.93	772283.418
4	01/04/2019	2091072.847	741570.4747
5	01/05/2019	2168598.18	775265.7906
6	01/06/2019	2100409.98	742701.241
7	01/07/2019	2171143.163	770907.148
8	01/08/2019	2158635.15	759029.9007
9	01/09/2019	2095701.713	751458.7188
10	01/10/2019	2159158.243	757067.1021
11	01/11/2019	2092135.56	729373.5276
		2180690.53	763998.6856

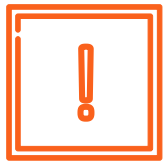
	A	B	C	D	E	F	G
1	Transaction Date	Manager 1 Sales	Manager 2 Sales	Manager 3 Sales	Manager 1 Margin	Manager 2 Margin	Manager 3 Margin
2	01/01/2019	34365.70333	24921.47	11531.04333	33%	34%	37%
3	01/02/2019	33789.92	24459.77333	10078.12667	31%	35%	35%
4	01/03/2019	34891.41667	24973.98667	10677.57	30%	39%	47%
5	01/04/2019	35016.22	24664.74333	10933.07667	33%	38%	44%
6	01/05/2019	33356.15	23614.27667	11318.41667	31%	41%	37%
7	01/06/2019	34435.27333	24441.01	11134.92333	31%	33%	44%
8	01/07/2019	35667.19333	23712.84	10917.80333	35%	37%	43%
9	01/08/2019	36463.69333	23796.64	11156.03667	33%	32%	36%
10	01/09/2019	34274.82333	23814.90333	10524.91	35%	41%	37%
11	01/10/2019	34016.6	24901.77667	10517.85	34%	40%	40%
12	01/11/2019	35850.23333	23685.75	10590.22	32%	33%	39%
13	01/12/2019	34176.88	24138.44	11143.50667	33%	33%	40%

Power Query Exercise 3Z: Extracting Information Review

Clues

Clue

You may need to use the Add Column > Multiply function



Make it harder

Do all the multiplication and addition steps in one step using the formula editor.

Power Query Exercise 2Z: Extracting Information Review

Query Steps

Steps	
1	New Query for **File 3Z** and extract the **DailyManagerSalesAndMargin** table.
2	Create a new column to multiply Manager 1 Sales by Manager 1 Margin.
3	**Instead of renaming the new column in a separate step, change it's name in the formula bar of the Insert Multiplication step.** Change the name to **M1 Margin Dollars**.
4	Repeat steps 2 & 3 for Manager 2 using M2 Margin Dollars as the new name.
5	Repeat steps 2 & 3 for Manager 3 using M3 Margin Dollars as the new name.
6	Delete the three margin percentage columns.
7	Calculate Total Sales in a new column using the sum function.
8	Calculate Total Margin in a new column using the sum function.
9	Delete all manager columns.
10	Transform the Transaction Date column by identifying the start of the month of each date.
11	Use the Group By function to group by transaction date.

Power Query
Dealing with Errors

Section Objectives

Tasks

01. **Identify and resolve** common errors in Power Query.

02. **Deal with exceptions** in the data that require a different action for each row.

Skills

Filter Errors

Create Calculated columns

Deal with Different Date Regions

Create Conditional Columns

Create Parameters to save Settings

Adapt Queries with Parameters

Power Query Exercise 4A: Basic Errors

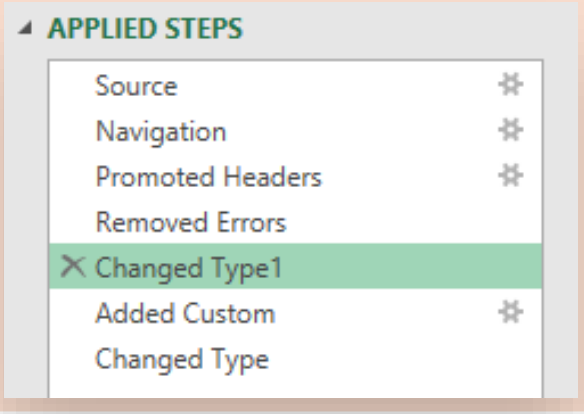
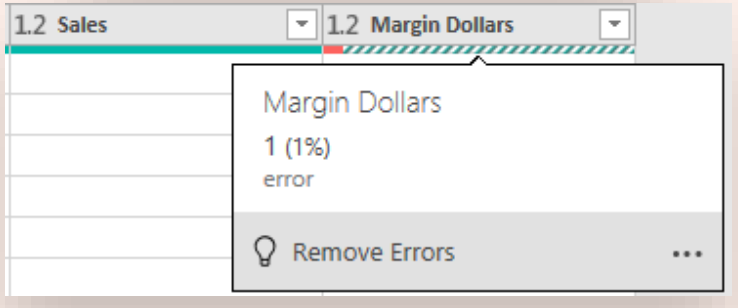
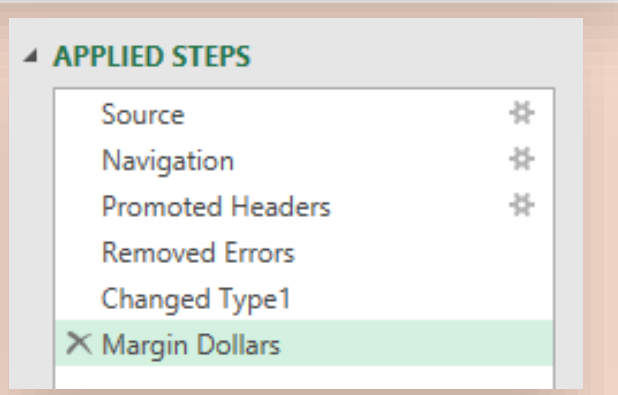
Query Steps Part 1

	Steps	Directions
1	Open file Excel **File 4A** and pull in the data from the invoices sheet. Promote headers.	
2	Remove basic errors.	Select all columns > Home > Remove Rows > Remove Errors
3	Create a new column to calculate Margin Pct.	Add Column > Custom Column > [Margin Dollars] / [Value]
4	Change the new column to decimal number.	
5	Close and load data to the worksheet if you don't have the option to change the load to parameter. **Click Queries and Connections once you're in Excel.**	
6	**Scroll down through the loaded data and you'll see that we have some blanks, even though we have data**. With a small data set, this was easy to spot, but with a larger data set this might be harder.	
7	**We need a better way to identify errors. Queries and Connections > Loaded with 7 errors. Not what we expected.**	
8	Click on the **7 errors** text. Now go to Power Query and look for the query named **Errors in ….**	
9	Click on one of the errors to see Power Query's error message.	
10	So we have an issue with numbers vs text. The clue is that the values in the sales and Margin Dollars columns are aligned differently.	

ABC 123 Sales	ABC 123 Margin Dollars
36.65	*17.69*
454.98	*193.7*
333.73	*144.91*
86.71	*39.19*
619.41999999999996	*234.58*
704.38	*251.9*

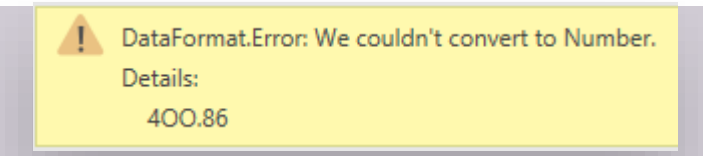
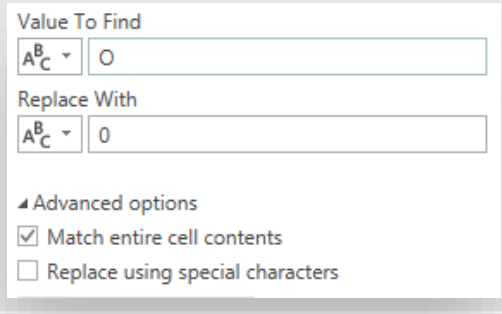
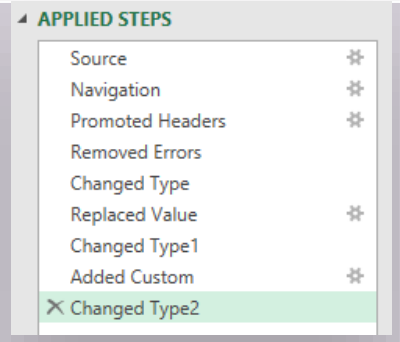
Power Query Exercise 4A: Basic Errors

Query Steps Part 2

	Steps	Screenshots
1	Ensure both columns are type decimal number before the add custom column step. **At this point, your query steps should look like this:**	
2	Ensure that you have selected the ChangedType1 step shown above. **Notice the red bar at the top of the column. This is the error indicator.** Hovering over it will tell you how many errors appear in that column.	
3	Notice that our previous steps where we'd created a new custom column have disappeared. **Be very careful when clicking on errors like this.**	

Power Query Exercise 4A: Basic Errors

Query Steps Part 3

Steps	
1	The error tells us that it couldn't convert the text to numbers. **This is because the Os are letters, not numbers.** 
2	Delete the Margin Dollars step where we investigated the error.
3	**A better method would be to right click on the error and click Add as New Query.**
4	Click on the removed errors step and insert a new step to replace **O (letter)** with **0 (number)** in the Margin Dollars column. 
5	**If it still doesn't work, we need to ensure that it is text, before we use the Replace function.**
6	Click on the removed errors step again and insert another new step to change the Margin Dollars data type to text. Click on the last step in the query steps. Re-create the deleted column to calculate Margin Pct. Change the new column to decimal number.
10	**Final step should look like this:** 

Power Query Exercise 4B: Errors with Date Locales

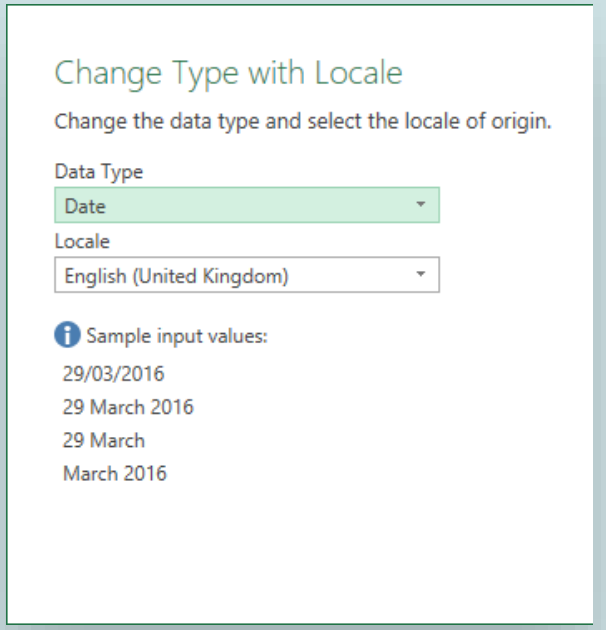
Earlier, we set the date format for our file. This is the known as the **Workbook locale**.

But what happens when the data you receive comes from a different region?

UK Format	Date	US Format
25/10/2019	25th October 2019	10/25/2019
10/02/2019	10th February 2019	02/10/2019
17/07/2019	17th July 2019	07/17/2019

Power Query Exercise 4B: Errors with Date Locales

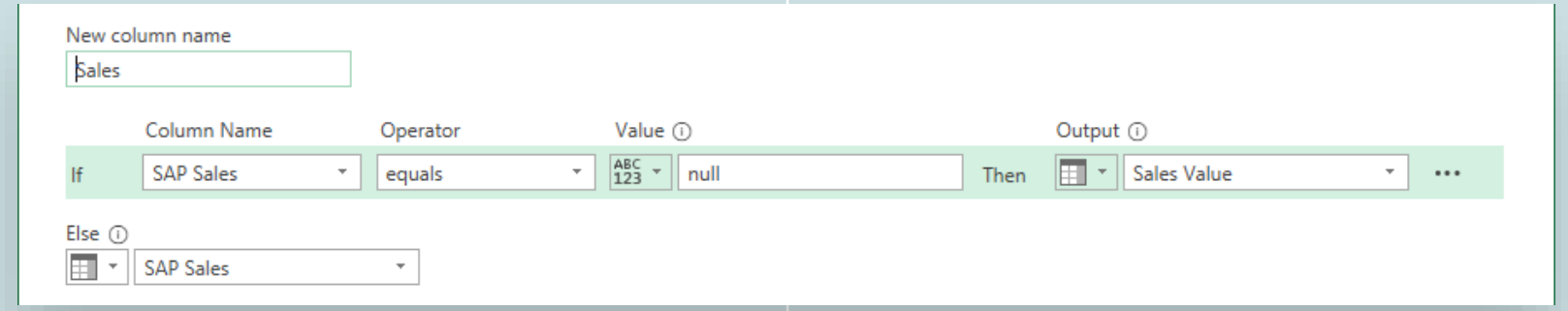
Query Steps

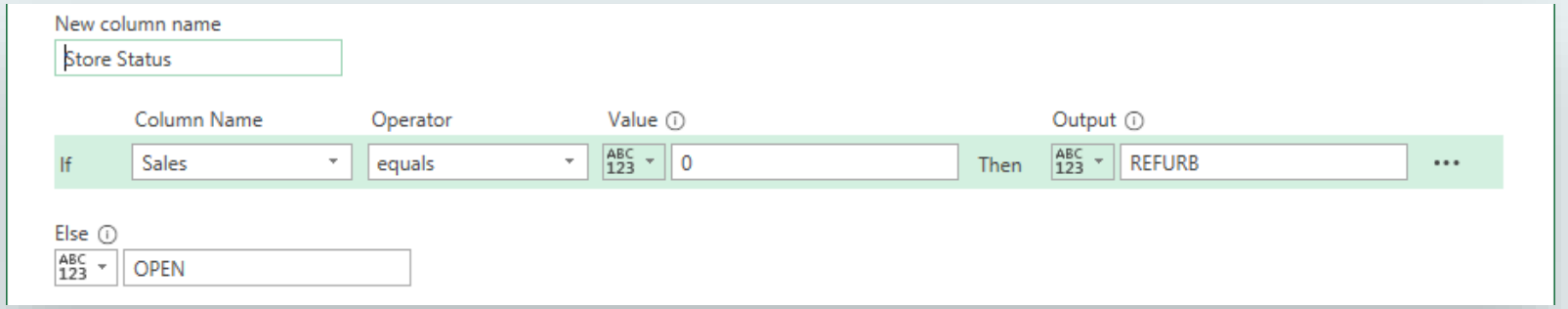
	Steps	Directions
1	Create a reference query from 4A and rename the new one with the above name.	
2	We have three date columns. **Notice that the Invoice Sent Date is in UK format, unlike the other two**.	
3	Convert all three columns to dates.	Select three date columns > Home > Date Type > Date
4	**The first column didn't work, because this file is operating in Canadian Locale.**	
5	**We need to change the Column Locale of the UK Dates – before it's interpreted as a date:** -Click on the Source step in applied steps -Right click on the Invoice Sent From UK column -Change Type > Using Locale > Data Type: Date > Locale: English (United Kingdom	
6	Now click on the final step. Dates are interpreted correctly.	

Power Query Exercise 4C: Conditional Columns for Exceptions

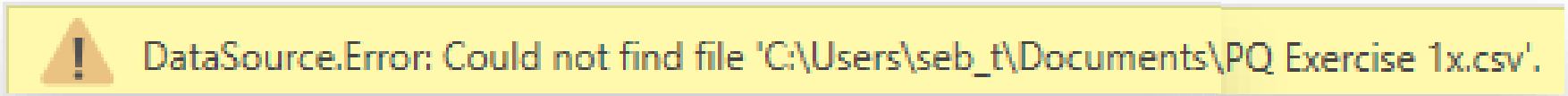
Query Steps

	Steps	Directions
1	Open the **File 4C** using a new query and load the WeeklySalesSummary table. **Change the data types of all 5 columns.**	
2	**The sales are recorded in two different columns, depending on the Store and System they came from.**	
3	Create a conditional column to consolidate the two.	Add Column > Conditional Column
4	Create a conditional column to identify when the store was open or being refurbished:	
5	Delete old sales columns and change the new sales column data type to decimal number.	





Power Query Exercise 4D: Errors with File Locations



We can use parameters to easily define the locations of our files.

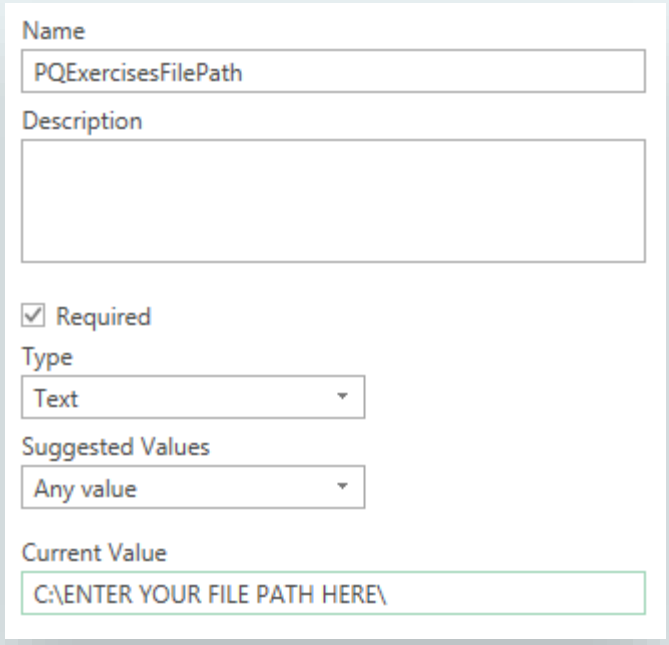
- **PQExercisesFilePath = C:\Users\CFI\Desktop\CFI PQ Student Folder\Data**

We can change the file path of all the data at the same time.

This is an easier way to modify queries without accessing the editor.

Power Query Exercise 4D: Errors with File Locations

Query Steps

Steps		Directions
1	Create a new query to load the CSV data from **File 1A**. Rename it.	
2	Create a new parameter called PQExercisesFilePath.	Home > Manage parameters > New 
3	Click on the Source step of your query.	
4	Replace the current file reference with: **PQExercisesFilePath & "PQ Exercise 1a.csv"**	
5	The final source code line should look like:	

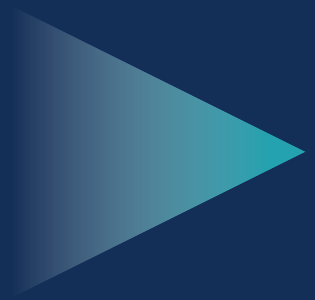
```
let
    Source = Csv.Document(File.Contents(PQExercisesFilePath & "PQ Exercise 1a.csv"),[Delimiter=",", Encoding=1252, QuoteStyle=QuoteStyle.None])
in
    Source
```

Power Query Exercise 4Z: Chapter Review

Estimated time: 10 minutes

1. Transform data from file 4Z.
2. Create a conditional Category column to identify categories by the following names: Alcohol, Tobacco, Food, Luxury, Perfume
3. In addition, remove all error entries for the 6th ghost category that shouldn't be present.
4. Change the date column to a date type and deal with any errors.

Date	Sales	Cat
10/01/2020	326.7212	ALC
10/01/2020	265.4218	TOB
10/01/2020	252.3812	FD
10/01/2020	310.4698	LXY
10/01/2020	270.8957	PER
10/01/2020		#REF!
11/01/2020	170.0706	ALC
11/01/2020	203.1826	TOB
11/01/2020	169.4264	FD



Date	Sales	Cat	Category
10/01/2020	326.7212148	ALC	Alcohol
10/01/2020	265.4218139	TOB	Tobacco
10/01/2020	252.3812408	FD	Food
10/01/2020	310.4697612	LXY	Luxury
10/01/2020	270.8957147	PER	Perfume
11/01/2020	170.0705617	ALC	Alcohol
11/01/2020	203.1825873	TOB	Tobacco
11/01/2020	169.4264092	FD	Food

Power Query Exercise 4Z: Dealing with Errors Review

Clues

Clue

You will need more than one clause in your conditional column