<div style="background-color:red;color:white;text-align:center">

Microsoft Power Tools for Data Analysis #6

3 Important Power Query Transformations: Merge, Append, UnPivot

Notes from Video:

</div>

Table of Contents:

1. **Example1: Merge 2 Tables to Replace VLOOKUP or Relationship**
   1) **Goal** is to take two related tables from two different files (one is an Excel File and the other is a Text File) and Merge them into one single table using the Merge feature, as seen in this picture:

Example 1: Merge 2 Tables to Replace VLOOKUP or Relationship

   2) **What does Merge feature do?**
      i. When we have a Foreign Key in one table and a Primary Key in the other table, Merge will allow us to pull data from the Primary Key Side to the Foreign Key side, as if were building a helper column with the VLOOKUP function.
      ii. There are other things that Merge can do also, and we will learn about this in the next video.
   3) **The Benefit of using Merge** will be:
      i. In Excel is we can avoid using the VLOOKUP function and thus avoid having many formulas in our spreadsheet solution.
      ii. In the Data Model (in Excel Power Pivot or Power BI Desktop) we can avoid using Relationships and create a more compact Star Schema Data Model.

Steps in Merge:

1) The Two Files we need to import look like this:

006-MSPTDA-dTerritoryTable.txt
006-MSPTDA-dSalesRepTable.xlsx
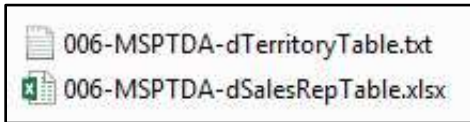
2) The Excel File contains the Excel Table named "dSalesRep", as seen here:

TerritoryID in the dSalesRep Table is the Foreign Key

3) The Tab Delimited Text File looks like this:

TerritoryID in the Tab Delimited Text File (Lookup Table) is the Primary Key

4) **Import Excel File**. To import the Excel File named "006-MSPTDA-StartFilePowerQueryMergeAppendUnPivot.xlsx":
   i.   Click on the Data Ribbon Tab.
   ii. In the Get & Transform group, click the dropdown arrow for Get Data.
   iii. Then point to From File.
   iv. Click on From Workbook, as seen here:
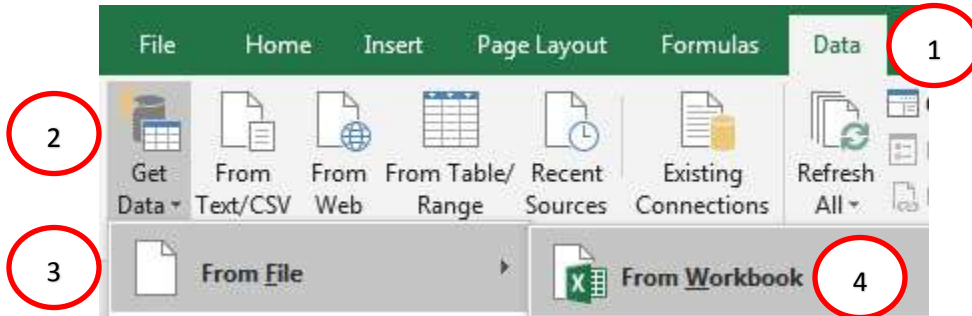
5) Navigate to the Excel File and double-click file to bring it into the Navigation window. In the Navigation window, select the Excel Table named "dSalesRep" on the left and then click the Transform Data button in the lower right, as seen here:
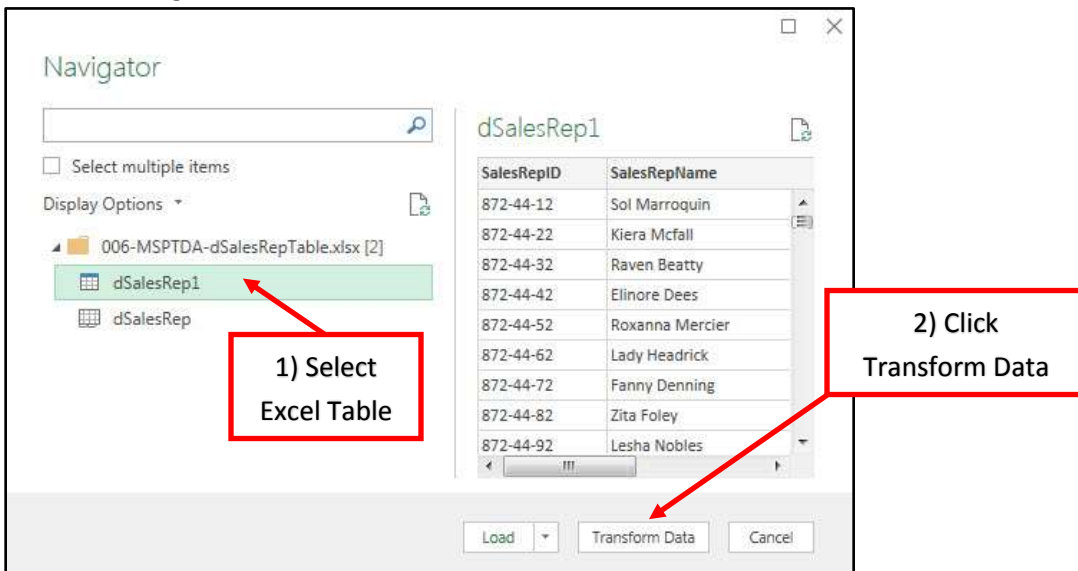
6) In the Power Query Editor window, change the Query Name to "Sales Representative", as seen here:
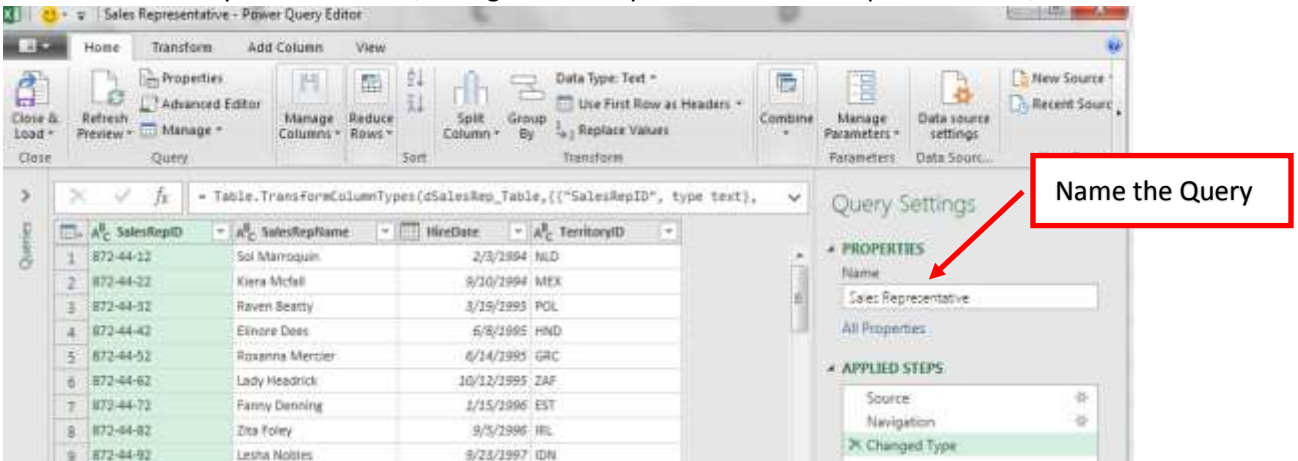
7) In the Home Ribbon Tab, in the Close group, click the dropdown arrow for Close & Load, then click on "Close & Load To…".

8) In the Import Data dialog box, select the dialog button for "Connection Only", as seen here:
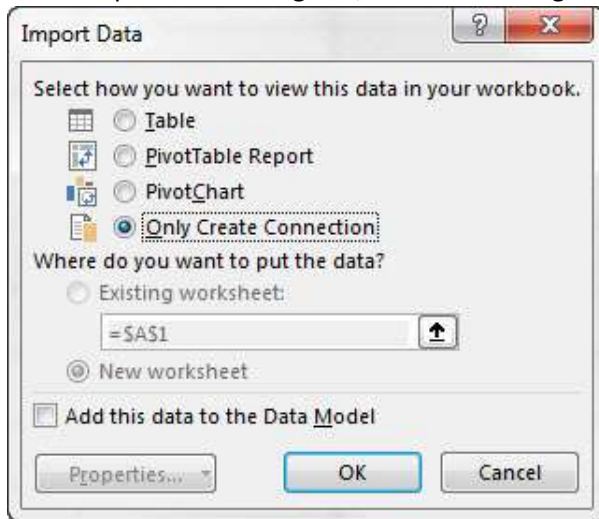
9) In the Queries and Connections Task Pane, you can see our first Query loaded as a Connection, as seen here:

10) **Import Text File**. To import the Text File named "006-MSPTDA-dTerritoryTable.txt":
   i. Click the Data Ribbon Tab.
   ii. In the Get & Transform group, click the dropdown arrow for Get Data.
   iii. Then point to From File.
   iv. Click on From Text File, as seen here:
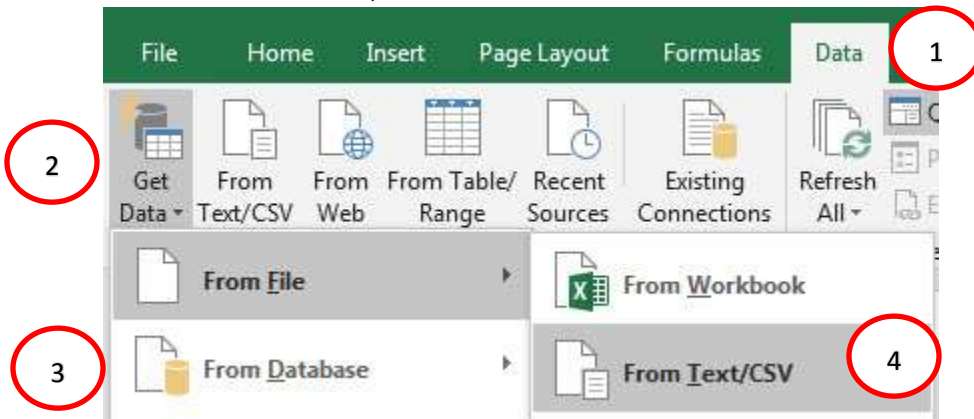
11) Navigate to the Text File and double-click file to bring it into the next window. In the next window, verify that the Delimiter is Tab. Then click the Transform Data button in the lower right, as seen here:
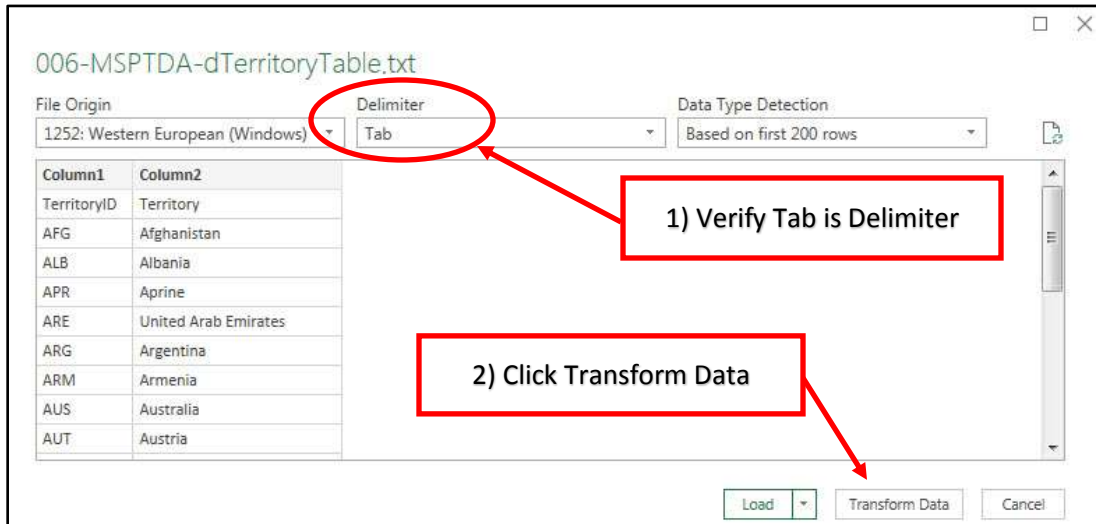

12) In the Power Query Editor, name the query "TerritoryLookupTable". Then Delete the Change Type Step in the Applied Steps List.
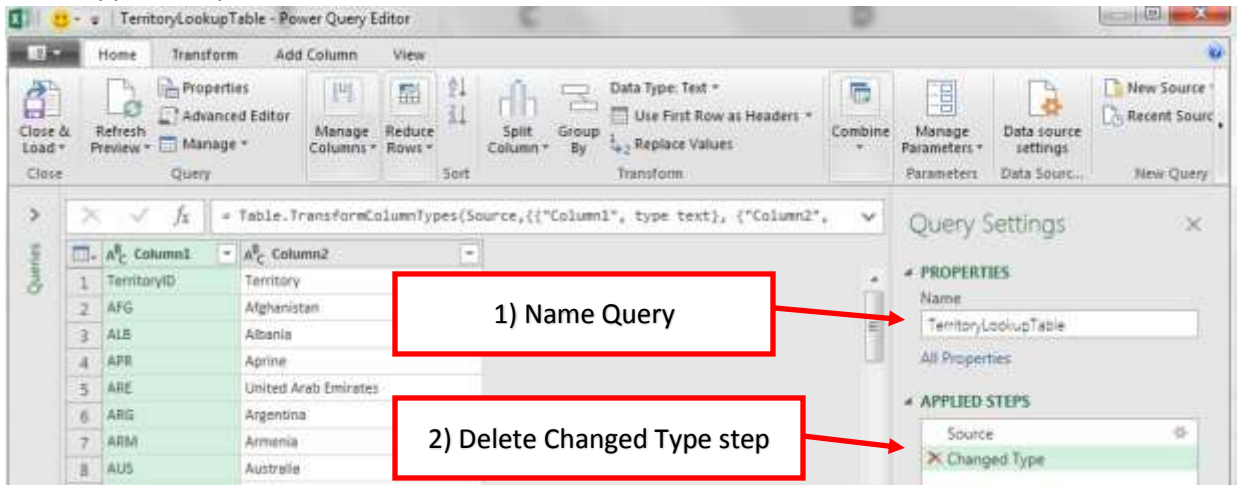
13) **Promote Headers**. Power Query did not interpret the first row of text values in the Tab Delimited File as Field Names. In the above picture, you can see that the Column Headers (Field Names) have the un-helpful names of "Column1" and "Column2". So, we must "Use First Row as Headers". To do this:
   i.   Click on the dropdown arrow in the Table Icon in the tables upper left corner
   ii. Then from the list select "Use First Row as Headers", as seen here:
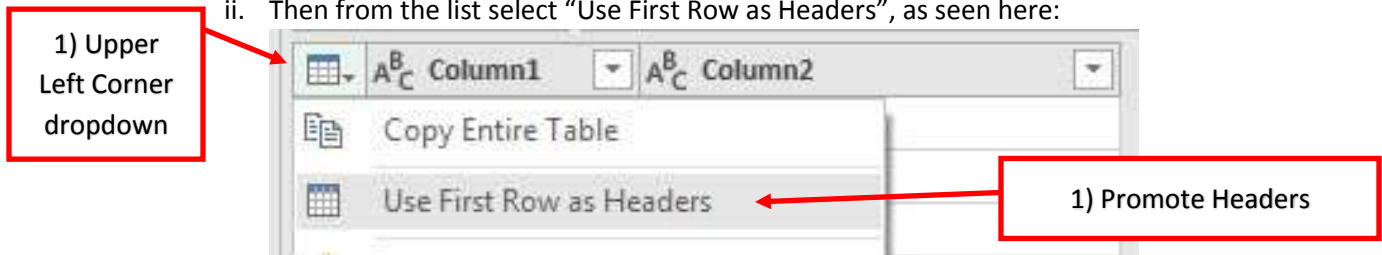
14) After we "Use First Row as Headers", two steps are added to the Applied Steps Take Pane: 1) Promote Headers, and 2) Change Type, as seen here:
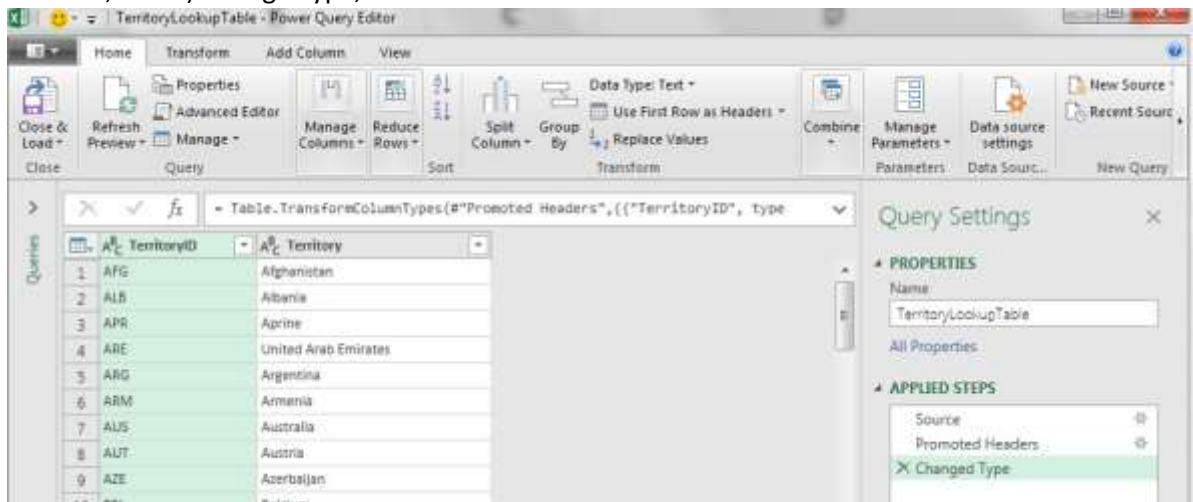
15) Close the Query and Load it as a Connection Only. Now in the Queries and Connections Task Pane, you can see our first two Queries loaded as a Connection, as seen here:
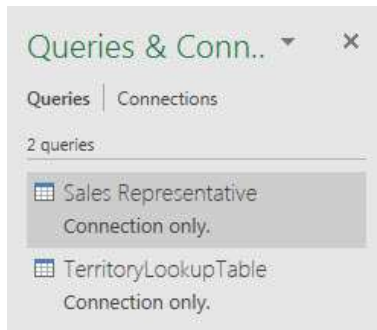

16) **Merge Two Queries**. To Merge the two queries
   i.   Click the Data Ribbon Tab.
   ii.   In the Get & Transform group, click the dropdown arrow for Get Data.
   iii. Then point to Combine Queries.
   iv.   Click on Merge, as seen here:
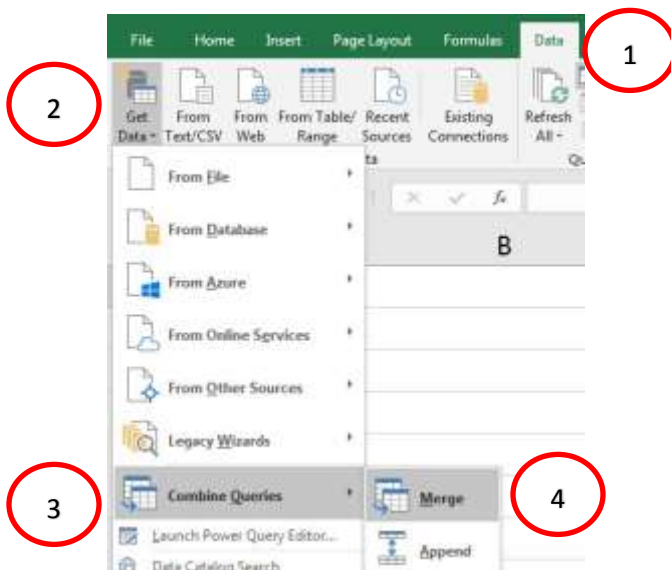
17) In the Merge dialog box:
   i.   For the first table select "Sales Representative".
   ii. Click TerritoryID (Foreign Key).
   iii. For the second table select "TerritoryLookupTable".
   iv. Click TerritoryID (Primary Key in Lookup Table).
   v.   Join Kind = Left Outer.
   vi. Click OK.
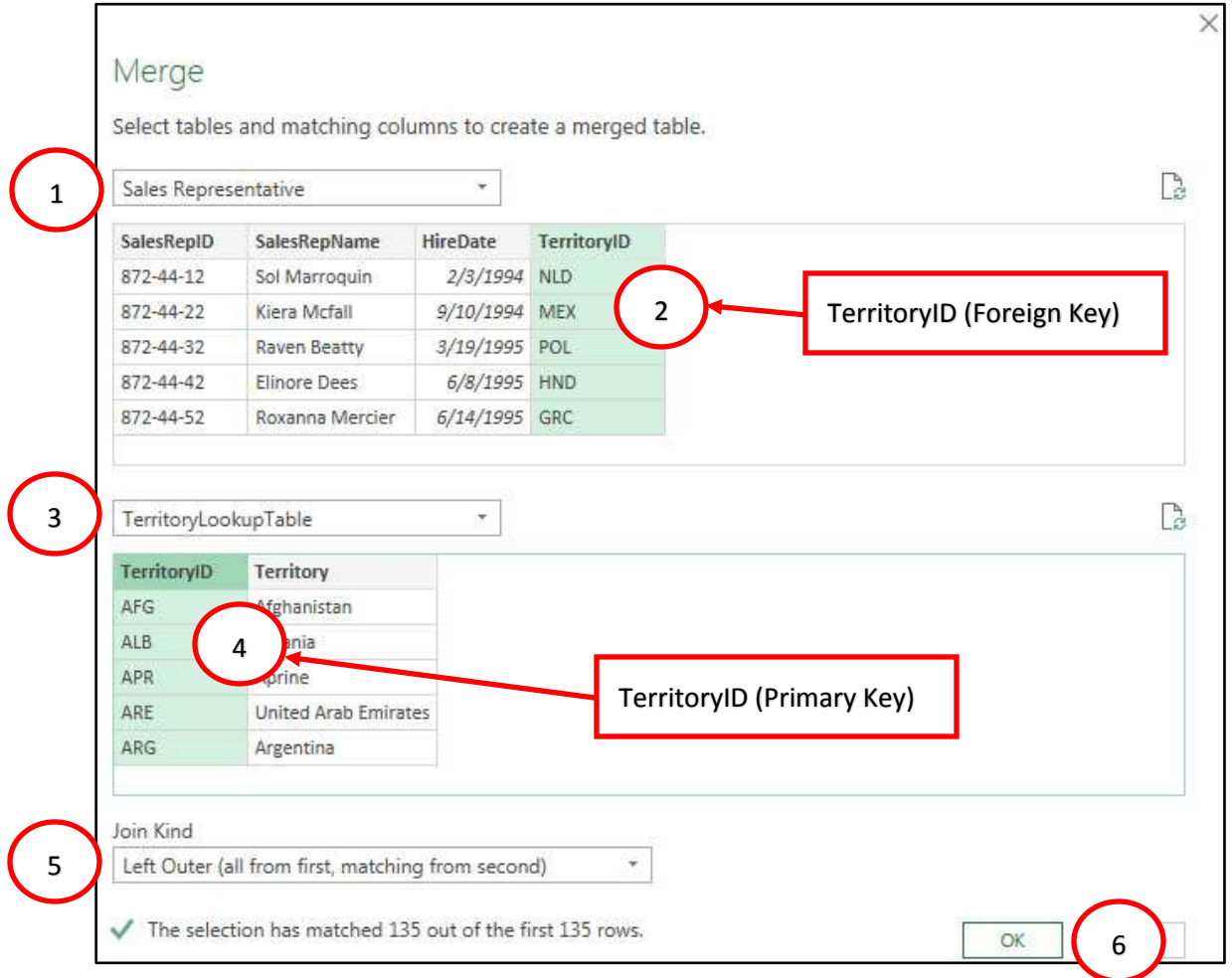
18) The merge process should look like the picture below. If you click off to the side of the word "Table" in the TerritoryLookupTable Column, you can see the related record from the Primary Key side of the Merge in the left-lower ares of the Power Query Editor, as seen in the picture below:
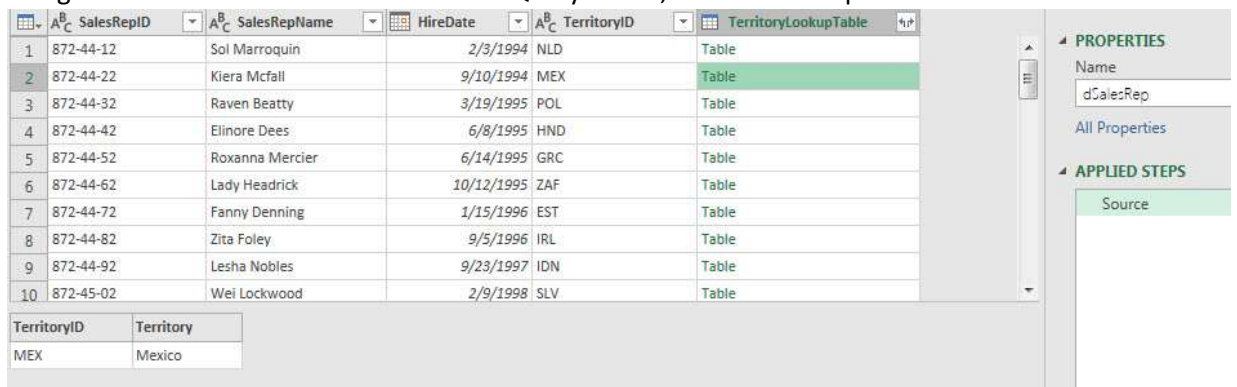
19) To pull the Territory Name into the dSalesRep Table, click the expand button in the TerritoryLookupTable Column, then check the Territory Column name as seen in the picture below:
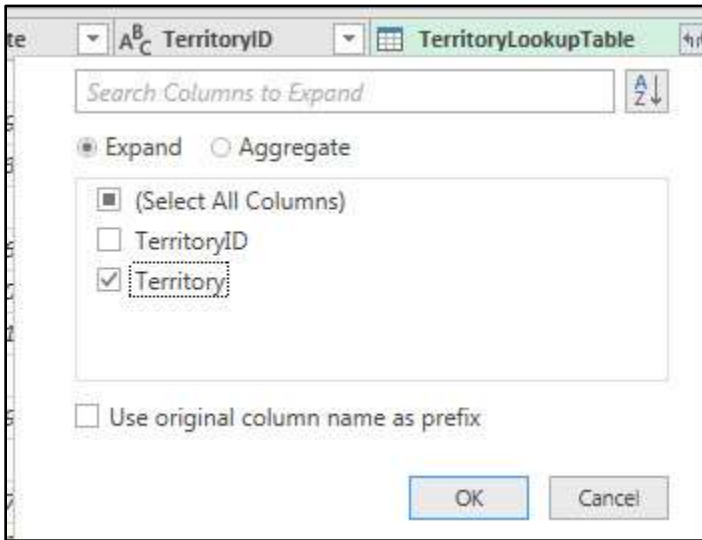
20) Then "Close & Load To…" the Worksheet named "Append", as seen in this picture.
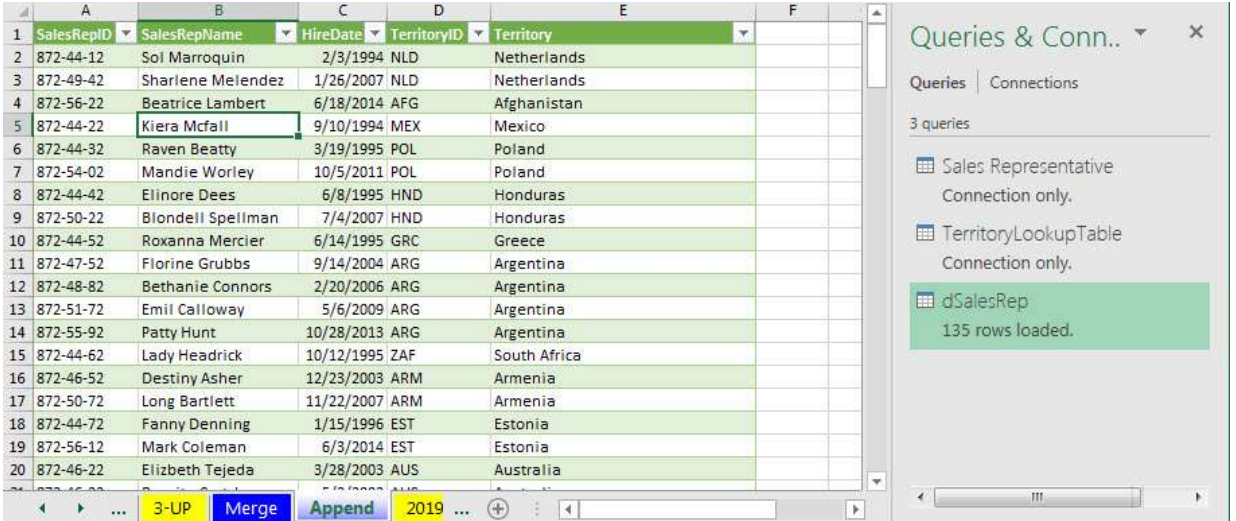
21) **Change Data & Refresh**. To see that the Merge of the two tables between two different files is dynamic, open the Excel File named "006-MSPTDA-dSalesRepTable.xlsx" and add a new record, as seen below:

SalesRepID	SalesRepName	HireDate	TerritoryID
872-57-52	Sioux Radcoolinator	7/12/2018	USA

22) Now when you go back to the Query and refresh, you will see that this new record is added to the Merged Table Power Query Output, as seen here:

	A	B	C	D
1	SalesRepID	SalesRepName	HireDate	Territory
136	872-57-52	Sioux Radcoolinator	7/12/2018	United States

2. **Example 2: Transform 3 Non-Uniform Sales Tables Then Append into Single Proper Data Set**:
    1) Goal is to take three Sales Tables that do not have the same number of columns or types of columns, and append them into a single Proper Data Set with the correct sales from all three tables using a number of Power Query features including the Append feature, as seen in this picture:

Example 2: Transform 3 Non-Uniform Sales Tables Then Append into Single Proper Data Set

    2) **What does Append feature do?** :
        i. The Append feature will take tables with the same Field Names and Data Types and stack then on top of each other to create a single table.
        ii. In Order to have a single Proper Data Set result from the Append process you must:
            1. Have the same number of columns.
                a. If you have an extra column in one of the tables, then the resulting table with have one extra column that contains data from the originating table and null values for the tables that did not contain the column.
            2. Each table must have consistent Field Names that are spelled the same in each table.
            3. The Data Types for each column that will be appended, for example the Revenue columns, must have the same Data Types.
                a. If you do not have the same Data Types, the column may result in a column without a Data Type and for columns with numbers you may not be able to make calculations on those columns.

3) **Source Excel Table Data**. If we look at the three tables that we are trying to Append, we can see that each table has a different set of columns:

   i. In the 2017 table, we can see the three columns: Units, NetCostEquivalent and Price. The other two tables do not have these columns. The other two tables have a column name Revenue. This means that in the transformation process we will have to take the columns Units, NetCostEquivalent and Price and perform multiplication to create a new column called Revenue.

	A	B	C	D	E	F
1	Date	Product	Region	Units	NetCostEqivalent	Price
2	1/1/2017	Fast Catch	East	120	0.91	58
3	1/1/2017	Darnell	South	72	0.91	15
4	1/1/2017	Quad	West	132	0.98	43
5	1/1/2017	Quad	West	72	0.9	43
6	1/1/2017	Fast Catch	West	180	0.97	58
7	1/1/2017	Quad	South	144	0.94	43

3-UP | Merge | Append | 2019 | 2018 | **2017** | UnPivot

   ii. In the 2018 table, we can see the correct four columns.

	A	B	C	D	E	F	G
1	Date	Product	Region	Revenue			
2	1/1/2018	Darnell	South	$1,965.60			
3	1/1/2018	Fast Catch	West	$6,960.00			
4	1/1/2018	Quad	East	$5,221.92			
5	1/1/2018	Darnell	West	$2,052.00			
6	1/1/2018	Aspen	East	$2,138.40			
7	1/1/2018	Yanaki	East	$2,257.20			

3-UP | Merge | Append | 2019 | **2018** | 2017 | UnPivot

   iii. In the 2019 table, we can see the correct four columns and one extra column named Color. This means that in the transformation process, we will have to delete the Colors Column.

	A	B	C	D	E	F	G
1	Date	Product	Region	Color	Revenue		
2	1/1/2019	Carlota	South	7	$1,482.00		
3	1/1/2019	Darnell	East	1	$1,411.20		
4	1/1/2019	Quad	West	10	$5,335.44		
5	1/1/2019	Aspen	South	1	$1,943.04		
6	1/1/2019	Tri Fly	West	10	$441.60		
7	1/1/2019	Yanaki	East	9	$820.80		

3-UP | Merge | Append | **2019** | 2018 | 2017 | UnPivot

4) **Import the 2017 Excel Table**. Using the "From Table/Range" button in the Get & Transform group in the Data Ribbon Tab, import the Excel Table named "Sales2017" from the Worksheet named 2017 into the Power Query Editor. Then change the Data Type for the Date Field to Date. The result should look like this:
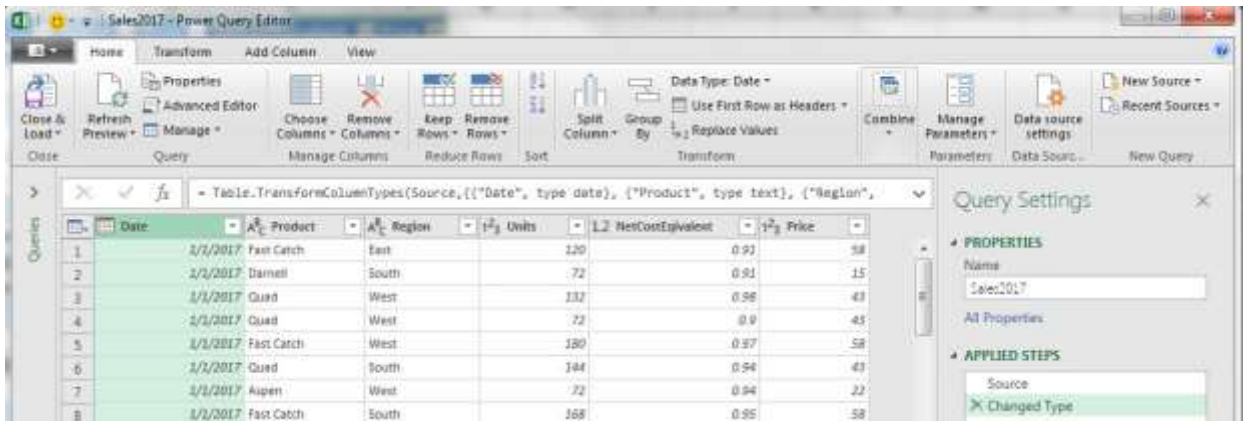
5) **Add Custom Column for Multiplication**. In order to add a new column and multiply the columns Units, NetCostEquivalent and Price, select the three columns, then:
   i.   Click on the Add Column Ribbon Tab.
   ii. In the From Number group, click the dropdown arrow in the Standard button, then from the dropdown menu, click on Multiply, as seen here:
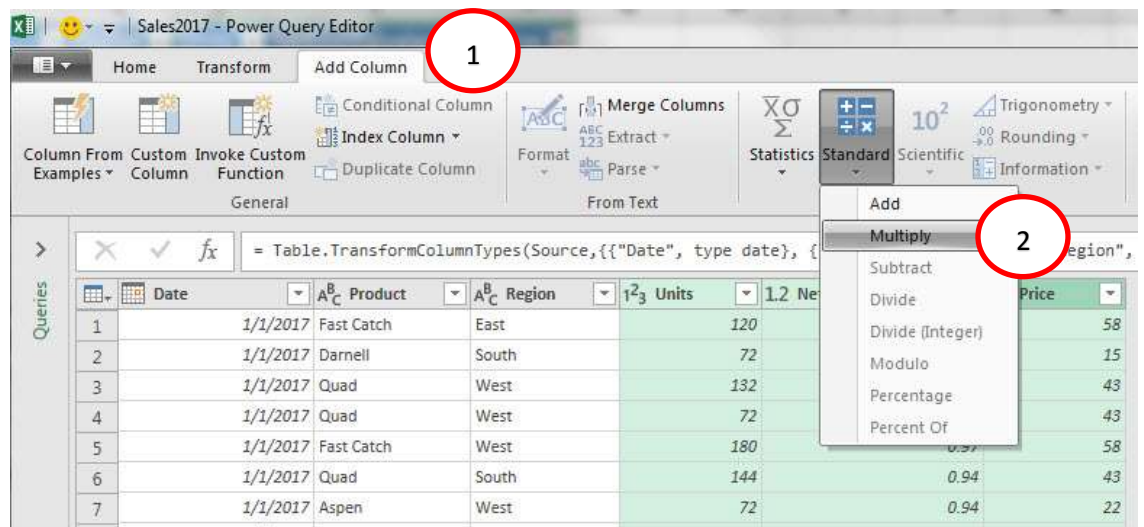
6) In the below picture you can see that the step "Inserted Multiplication" was added to the Applied Steps list, a new Column was added and you can see the formula in the formula bar, as seen here:

Formula can be seen in Formula Bar.

New Column

Inserted Multiplication step
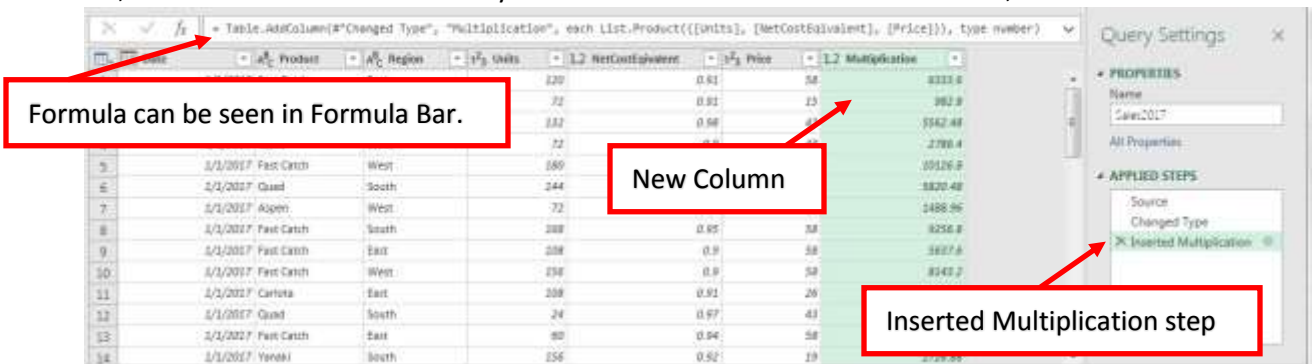
7) In the Formula Bar, double click the word "Multiplication" and then replace it with the new Field Name "Revenue". The result is seen here:

| | 1) Field Name "Revenue" is changed in the Formula Bar | | 2) New Field Name appears here |

8) **Taking a closer look at the Table.AddColumn Function**:

```
= Table.AddColumn(#"Changed Type", "Revenue", each List.Product({[Units], [NetCostEqivalent], [Price]}), type number)
```

   i.   In the first argument we see #"Changed Type". This argument lists the previous step that the function is acting on.
   ii. In the second argument we see "Revenue". This is the name of the new column.
   iii. The third argument lists the formula that will be calculated for each row in the table.
   iv. The fourth argument defines the Data Type for the column and we can see that it defines it as "number", which is the Decimal Data Type.

9) **Taking a closer look at the List.Product Function**:

```
List.Product({[Units], [NetCostEqivalent], [Price]})
```

   i.   List.Product function will take items from a list and multiply them to yield a product.
   ii. Inside the List.Product function was see the Curly Brackets being used to house a list of Field Names that will be multiplied.
   iii. For each one of the Field Names, the Lookup Operator (Square Brackets [and]) are being used so that for each row in the table, the formula can lookup the correct number for the Units, NetCostEquivalent and Price Columns.

10) Next we need to remove the Units, NetCostEquivalent and Price Columns. To do this, select the three columns and then right-click any one of the columns and from the menu click on Remove Columns, as seen here:
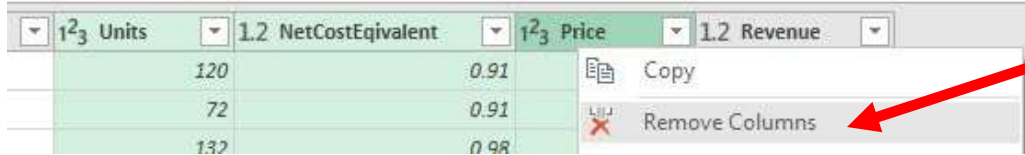

11) **Import the 2018 Excel Table**.
   i. Using the "From Table/Range" button in the Get & Transform group in the Data Ribbon Tab, import the Excel Table named "Sales2018" from the Worksheet named 2018 into the Power Query Editor.
   ii. Then change the Data Type for the Date Field to Date.
   iii. Then Close and Load as a Connection Only.
12) **Import the 2019 Excel Table**.
   i. Using the "From Table/Range" button in the Get & Transform group in the Data Ribbon Tab, import the Excel Table named "Sales2019" from the Worksheet named 2019 into the Power Query Editor.
   ii. Then change the Data Type for the Date Field to Date.
   iii. Then Remove the Color Column.
   iv. Then Close and Load as a Connection Only.
13) After you have imported and transformed the three tables, the Queries & Connections Task Pane should look like this:
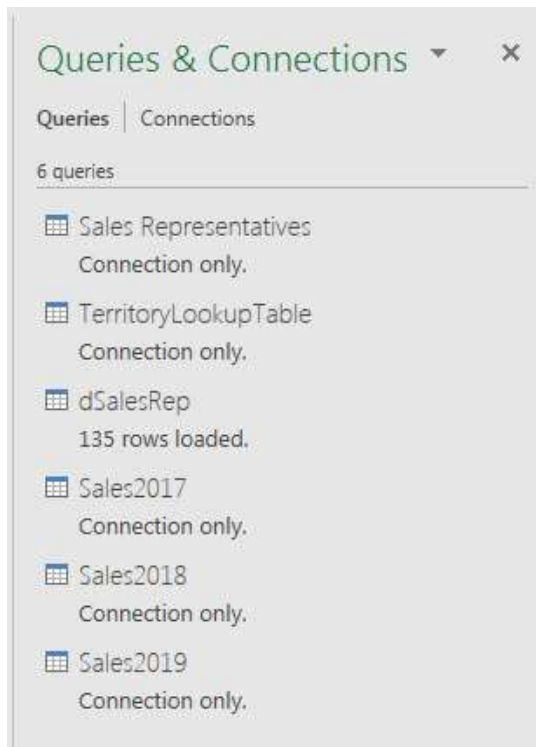
14) **Append feature**. To Append the three Sales tables,
    i.   Click the Data Ribbon Tab.
    ii.   In the Get & Transform group, click the dropdown arrow for Get Data.
    iii.   Then point to Combine Queries.
    iv.   Then click on Append, as seen here:
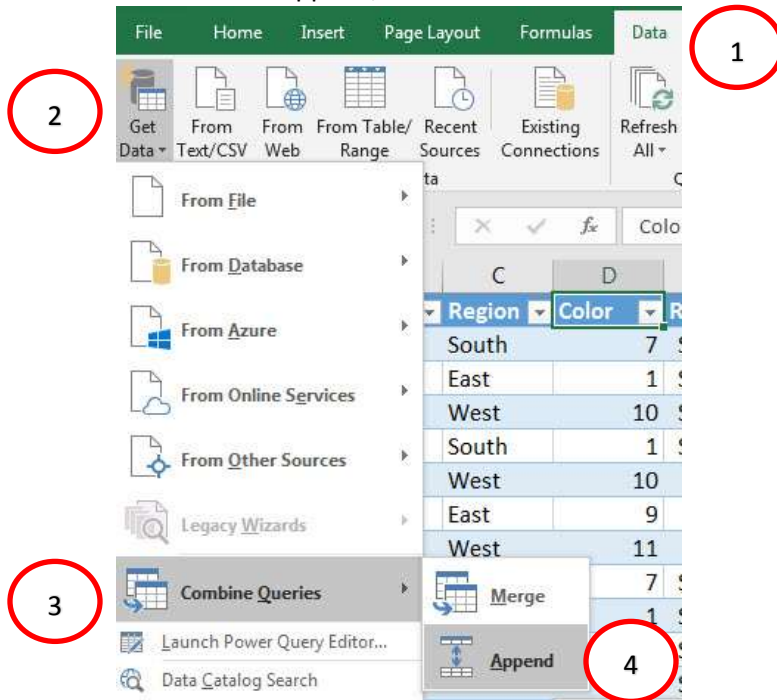
15) In the Append dialog box:
    i.   Click dialog button for "Three or more tables".
    ii.   On the left select the three Sales tables.
    iii.   Click the Add button.
    iv.   Click OK.
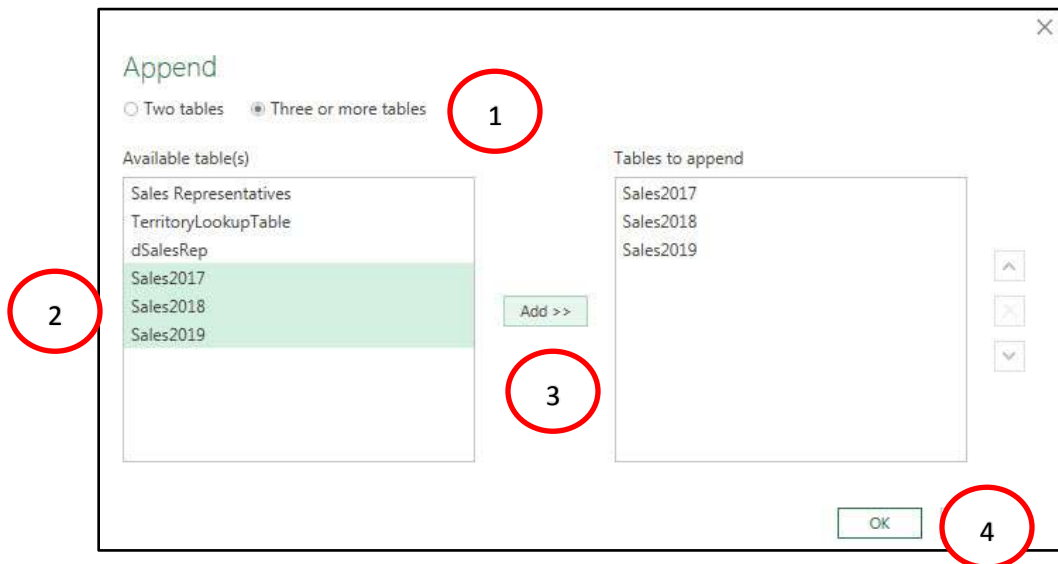
16) The final Appended Proper Data Set can be seen in the picture below. Be sure to name this query "ThreeYearsData. Notice the Function, Table.Combine which uses a List that contains the names of the three queries used to import and transform the tables.

17) Finally, Close and Load the Proper data Set to the Worksheet named Append.

3. **Example 3: UnPivot to Convert Cross Tabulated Data Tables into Proper Data Set**:

1) Goal is to take a table of cross tabulated data and convert it to a Proper Data Set using the UnPivot Feature, as seen in this picture:

Example 3: UnPivot to Convert Cross Tabulated Data Tables into Proper Data Set

2) **Idea Behind UnPivot**.

   i. In the below Cross Tabulated Data Table, the first column contains a unique list of student names.

   ii. Columns 2 to 8 contain the grades for each student for a particular class.

   iii. For columns 2 to 8, the Column Headers are the names for each class.

   iv. The problem with the below data set is that it is not a Proper Data Set, and so we cannot easily perform tasks like sorting.

   v. Our goal is to transform the below table into a three-column table with the Field Names: Student, Class and Grade.

   vi. To UnPivot this non-Proper Data Set, we must tell Power Query:

   1. To use the first column for a new column named Students.

   2. Take the Column Names for columns 2 to 8 and use those values for a second column named Classes.

   3. Take all the values for grades and use those values for a third column named Grades.

vii. To Further see how we need to "UnPivot" this table, look at the first grade of 3.7 in the table in the below picture:

Grades:

Student/Quantitative Class	Busn 216	Busn135	Busn218
Abdi Hyde	3.7	3	3.9
Tyrone Lord	3.6	3.4	4

viii. Then imagine it as a record in a three-column table, as seen here:

Student	Class	Grade
Abdi Hyde	Busn 216	3.7

ix. Now look at the second grade of 3 in the table in the below picture:

Grades:

Student/Quantitative Class	Busn 216	Busn135	Busn218
Abdi Hyde	3.7	3	3.9
Tyrone Lord	3.6	3.4	4

x. Then imagine it as a record in a three-column table, as seen here:

Student	Class	Grade
Abdi Hyde	Busn 216	3.7
Abdi Hyde	Busn135	3

xi. The Process of Unpivoting a table in Power Query will be simple because all we will have to do is select the first column and then click the "UnPivot Other Columns" button.

Steps to UnPivot Table

1) **Import the Cross Tabulated Data Table**. Using the "From Table/Range" button in the Get & Transform group in the Data Ribbon Tab, import the Excel Table named "GradeTable" from the Worksheet named UnPivot into the Power Query Editor.
2) Name the query "FinalProperDataSetGradeTable".
3) Delete the automatically created "Changed Type" step from Applied Steps list.
4) **Unpivot Other Columns**. Select the first column, and right-click, then from the menu, click on "Unpivot Other Columns", as seen here:
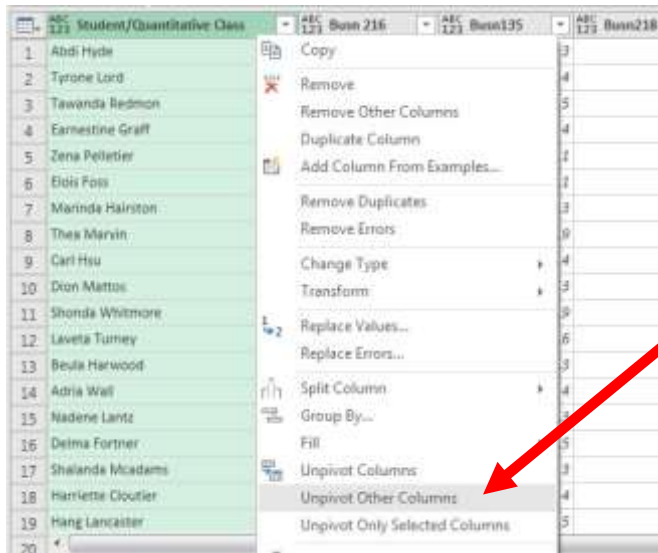
5) **Table.UnpivotOtherColumns function**. The Three-Column Table result can be seen in the below picture. Notice the "Unpivoted Other Columns" step in the Applied Steps list, and notice the Table.UnpivotOtherColumns function in the Formula Bar.

6) To Finalize the three-column grade table, Change the Field Names and Data Types, as seen in the picture below:

Student	Class	Grade
1 Abdi Hyde	Busn 216	3.7
2 Abdi Hyde	Busn135	3
3 Abdi Hyde	Busn218	3.9
4 Abdi Hyde	BI 348	3.6
5 Abdi Hyde	Math 148	2.4
6 Abdi Hyde	Math 111	3
7 Abdi Hyde	Busn 210	2.6
8 Tyrone Lord	Busn 216	3.6
9 Tyrone Lord	Busn135	3.4
10 Tyrone Lord	Busn218	4
11 Tyrone Lord	BI 348	3.6

PROPERTIES

Name

FinalProperDataGetGradeTable

All Properties

APPLIED STEPS

Source
Unpivoted Other Columns
Renamed Columns
✗ Changed Type

7) **Load to Worksheet**. Load the transformed grade table to cell K4 on the sheet named UnPivot, as seen below:

	A	B	C	D	E	F	G	H	I	J	K	L	M
1													
2		Grades:											
3													
4		Student/Quantitative Class	Busn 216	Busn135	Busn218	BI 348	Math 148	Math 111	Busn 210		Student	Class	Grade
5		Abdi Hyde	3.7	3	3.9	3.6	2.4	3	2.6		Gigi Espinal	BI 348	3.9
6		Tyrone Lord	3.6	3.4	4	3.6	3.4	4	3.1		Tawanda Redmon	BI 348	3.8
7		Tawanda Redmon	3.8	3.5	3.7	3.8	3.8	3.6	3.6		Abdi Hyde	BI 348	3.6
8		Earnestine Graff	3.9	4	1.7	2.9	3.1	2.6	3		Tyrone Lord	BI 348	3.6
9		Zena Pelletier	2.1	1.1	2.3	0.5	3.2	1.9	0.8		Delma Fortner	BI 348	3.5
10		Elois Foss	3.8	3.1	2.8	3	0.1	2.6	1.6		Kenton Reardon	BI 348	3.5
11		Marinda Hairston	2.9	0.3	3.3	3	2.7	2.8	3.1		Luciano Neville	BI 348	3.5
12		Thea Marvin	1.8	1.9	0.6	2.2	1.7	3.4	3.1		Lizabeth Solis	BI 348	3.4
13		Carl Hsu	3.5	3.4	3.1	3.3	2.6	2.4	2.8		Carl Hsu	BI 348	3.3
14		Dion Mattos	3.7	3	2	3.2	4	2.9	2.1		Dion Mattos	BI 348	3.2
15		Shonda Whitmore	2.7	2.9	3.2	2.5	4	0.7	2.7		Elois Foss	BI 348	3
16		Laveta Turney	3.3	3.6	0.5	2.9	2.4	2	3.1		Marinda Hairston	BI 348	3
17		Beula Harwood	2.6	3.3	3	2.4	2.5	2.2	3.5		Earnestine Graff	BI 348	2.9
18		Adria Wall	2.8	4	3	0.9	2.9	3	3.4		Laveta Turney	BI 348	2.9
19		Nadene Lantz	2	3.3	0.5	2.4	3	3.7	2.2		Alethea Rodrigues	BI 348	2.7
20		Delma Fortner	3.4	3.5	2.6	3.5	3.1	3.2	0.1		Veta Mejia	BI 348	2.6
21		Shalanda Mcadams	3.1	3	2.7	0.6	2.5	0.5	2.8		Shonda Whitmore	BI 348	2.5
22		Harriette Cloutier	2.2	4	3	2.4	2.5	2.2	2.1		Beula Harwood	BI 348	2.4
23		Hang Lancaster	2.6	2.5	3.6	0.9	3	3	1.1		Nadene Lantz	BI 348	2.4
24		Kenton Reardon	3.7	3	2.6	3.5	3.3	3	2.8		Harriette Cloutier	BI 348	2.4
25		Lizabeth Solis	2.6	2.8	0	3.4	2.4	1.5	3.8		Thea Marvin	BI 348	2.2
26		Luciano Neville	3.7	1.1	3.6	3.5	0.1	2.7	2.8		Adria Wall	BI 348	0.9
27		Gigi Espinal	1.3	2.7	1.8	3.9	3.9	2.9	2		Hang Lancaster	BI 348	0.9
28		Alethea Rodrigues	2.6	3.2	2.8	2.7	0.7	1.2	0		Shalanda Mcadams	BI 348	0.6
29		Veta Mejia	2.3	3.2	3.5	2.6	3.7	3.4	2.6		Zena Pelletier	BI 348	0.5
30											Lizabeth Solis	Busn 210	3.8
31											Tawanda Redmon	Busn 210	3.6
32											Beula Harwood	Busn 210	3.5