### OPMT-7750 Data Modelling & Business Intelligence

# SAMPLE MIDTERM PRACTICAL

### WINTER2022

# PRACTICAL

You will complete the following steps in a new Excel workbook you created.

You will start by downloading the MidtermSampleSourceFiles.zip file. Unzip the contents of this file to a folder on your computer. You will then create a new **Excel** file to perform the required steps to come up with your answers.

* 1. **LOAD A TEXT FILE**

Get Data to Edit/Transform the Sales Org.csv file (a text file) into the Power Query Editor in Excel. \*\*\*MAKE SURE YOU GET DATA FROM TEXT/CSV file and not worksheet\*\*\*

* **Promote the first row** values to headers
* **Rename** the **Description** field to **Region**
* **Rename** the **Org field** to **Organization**.

**Load** this as a **Connection Only** and add to the **Data Model**.

*What is the total number of****rows****loaded by this query?* ***4***

---

My Notes

1. Data > Get Data > From Worksheet > from text/csv file
2. Transform/edit
3. Then Promote first row as headers
4. Rename the Description field to Region
5. Rename the Org field to Organization
6. Then Close & Load to…
   1. Only from Create Connection
   2. Add data to data model
   3. When you first create any data model try to always do this step this way – not just as a table

---

* 1. **LOAD MULTIPLE WORKSHEETS**

**Import & Edit/Transform** the Excel file called **Category Product.xlsx** into the Power Query Editor in Excel.

* **Select** both the **Categories**and **Products** sheets from the workbook and **load them as individual queries.**
* Ensure **headers are promoted** and **field types** are selected appropriately
* **Load** these as **Connections** **Only** and add to the **Data Model**.

*What is the* ***combined total number of rows****loaded by the* ***Category****,* ***Products*** *and* ***Sales Org*** *queries?* ***39***

**Edit** the Products query:

* **Change** the data type of the **Price** field to **Currency**
* **Change** the data type of the **Cost** field to **Currency**
* **Replace** blanks in the **Components**field with the value **None**

**Close and load** the query.

**Convert** the Products query to a **Connection Only** (do not load to the data model.)

*How many* ***columns*** *are contained in the* ***Products*** *query?* ***8***

**Import & Edit/Transform** the worksheet in the **Transactions.xlsx**file from the source data folder into the Power Query Editor in Excel.

* **Load** as **Connection** **Only** and add to the **Data Model**.

*What is the* ***name of the query*** *produced by this import?* ***Sales***

*\*\* The name of the query actually takes the sheet name\*\**

---

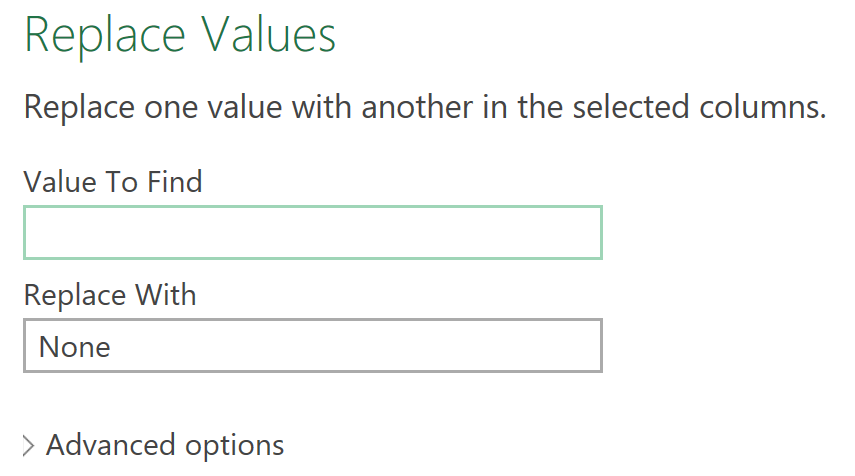
My Notes

1. Data > import Data > Select from Workbook > From Worksheet > Category Product > import & Edit/Transform
2. Select both the Categories and Products sheets from the workbook and load them as individual quieries (use the checkmark part)
3. Promote headers
4. Make sure the right field types are selected

\*\* when you are changing the type value and a notification pops up to ether replace current or add new step it depends:

* the only time when you need to add a new step is when you are converting from text to date. Then you will need to add a new step.

1. Replace the blanks in the Components field with the Value None.
   1. Click on filed under components you want to replace then click replace values. Then you will add the information you need as:



1. Load as connection only and add to the Data Model
2. importing
   1. Data > Get Data > From File > From Workbook > Transactions.xlsx
      1. Load as a connection
      2. Add data to data model

---

* 1. **LOAD FROM A WORKBOOK**

**Edit** the query you just created using the Power Query Editor in Excel.

* **Add** a new **Custom Column** called **Discount Percent** which divides the **Discount** by the **Revenue** fields.
* **Change** the data type of this field to **Percentage**.

*What is the* ***Discount Percent*** *value displayed for* ***TrNum 28*** *of this table? Round to zero decimal places and include the % sign.* ***3%***

In the same query, **Create**a new **Conditional Column** called **Revenue Multi** which returns 1.14 if the Currency field contains EUR, and otherwise returns 1.

*What is the* ***Revenue Multi*** *value displayed for* ***TrNum 14*** *of this table?* ***1.14***

***\*\* when it asks you for TrNum make sure you are looking in that column not just the rows that are autolabelled on the left. If it is asking for the row number then you can look on just the left hand side which was generated by excel automatically.***

**Add** a new **Custom Column** called **USD Revenue** which multiplies **Revenue** by **Revenue Multi**. **Change** the data type of this field to **Currency**.

*What is the value for* ***USD Revenue*** *displayed for* ***TrNum 16****of this table? Round to two decimals places.* ***$8,025.59***

(While still in the Power Query editor) **merge** the **Products** query into this query (Left Outer Join) using the **ProductID** field.

* **Expand** the **Products** field and select only the **Price, Cost,** **Product description** and **Category** fields.

*What is the* ***Product*** *description displayed in* ***TrNum******22****of the table?* ***T-Shirt***

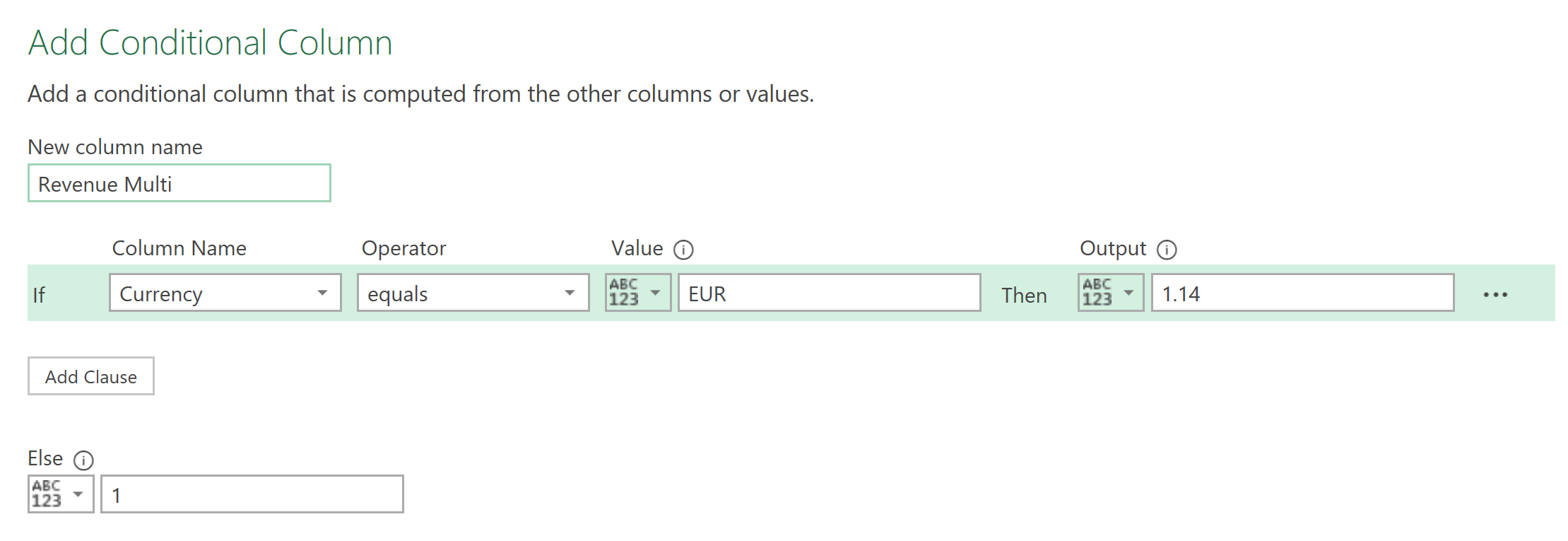
**Close and Load** the query to Excel as a **Connection only** and add to the **Data Model**.

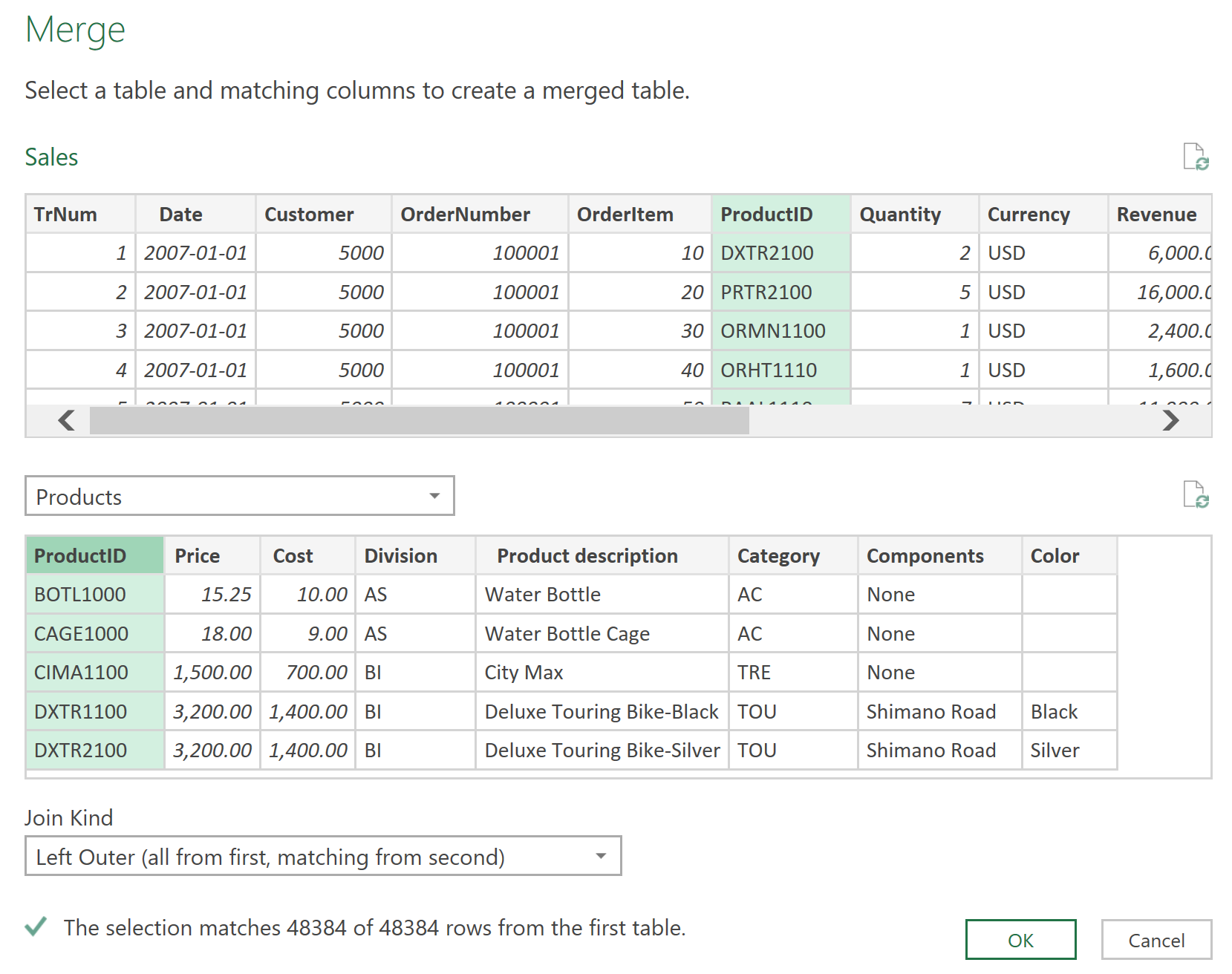
*How many* ***rows*** *were loaded by the* ***Sales****query?* ***48,384***

---

My Notes

1. Sales Query
   1. Add custom column called discount percent
      1. Add Column > Custom Column
      2. [Discount]/[Revenue]
      3. Change the data type to Percentage
   2. Add conditional column
      1. Called Revenue Multi
      2. If currency = EUR then display 1.14
      3. Else return 1



* 1. Add custom Column
     1. Call it USD Revenue
     2. Revenue x Revenu Multi
     3. Change the data field to Currency
  2. Merge
     1. Home > Merge Queries
     2. First table will be Sales then the other will be products.
     3. Highlight the ProductID to connect the tables.
     4. Now click on that column that was created called Products
        1. Expand it
        2. Uncheckmark Use original column name as prefix
        3. Only select the **Price, Cost,** **Product description** and **Category** fields.
  3. Close & Load
     1. Connection only
     2. Add data to data model

---

* 1. **LOAD FROM A FOLDER**

**Import & Edit/Transform** the file **Cust Header Conversion.xlsx** into the Power Query Editor in Excel.

* Ensure **headers are promoted**
* **Rename** this query **Conversion Table**
* Load as a **Connection Only** (do not add to the data model)

**Import** the contents of the **Customers** folder into the Power Query Editor in Excel.

* Use the **Combine and Edit/Transform**command to combined your mismatched data.
* **Select** the **US Customers** as the sample file and select **Sheet1** under **Parameter1**

**Edit** the **Transform Sample File** using the methods from Assignment 3 (Chapter 4) as a reference.

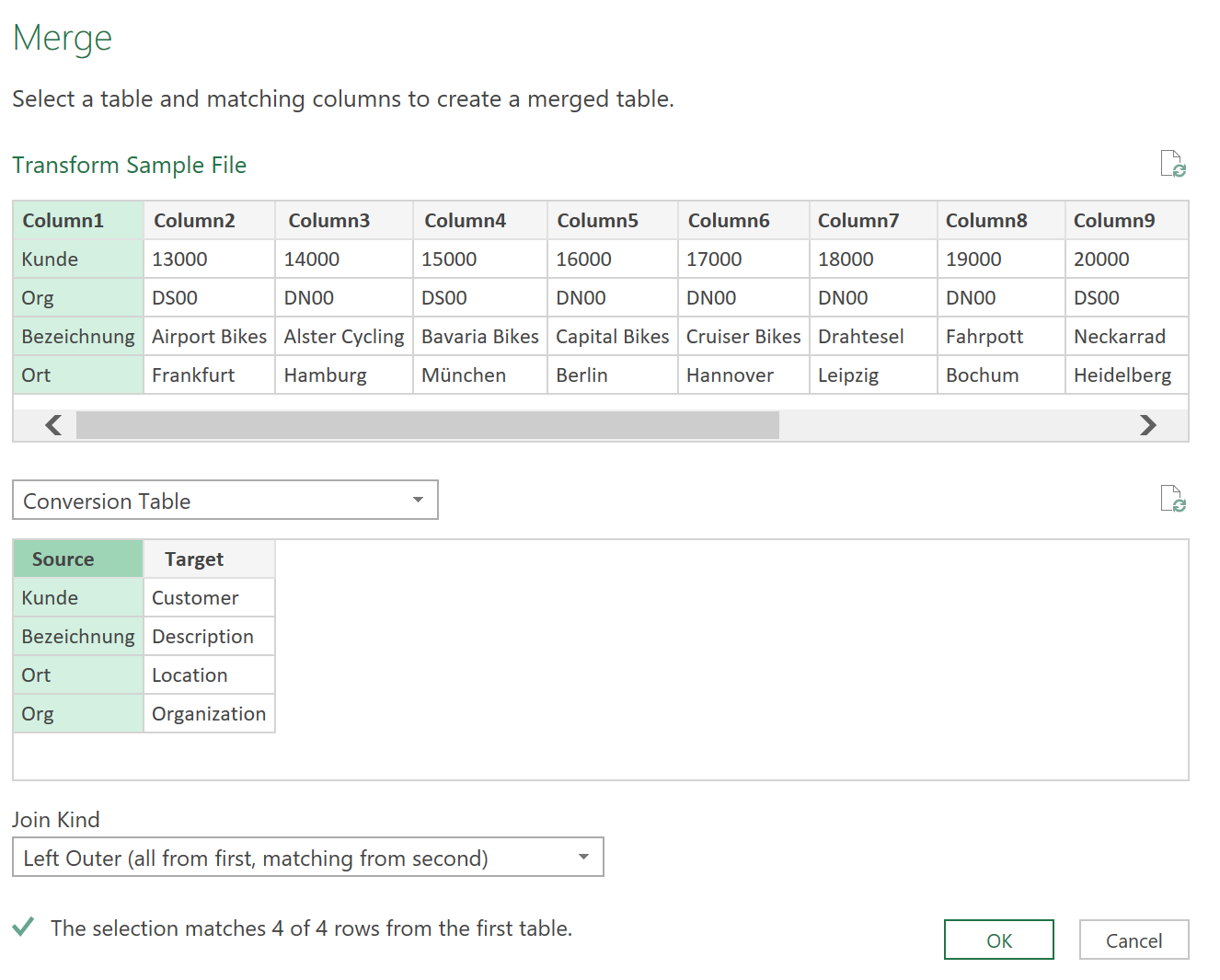
* **Merge** to the **Conversion Table** query and perform the steps necessary to replace the headers in your sample file with the **Target** column
* **Load** this query as **Connection** **Only** and add to the **Data Model**

*In the* ***Customer Query,*** *what is the the* ***Location****displayed for****Customer 5000?***

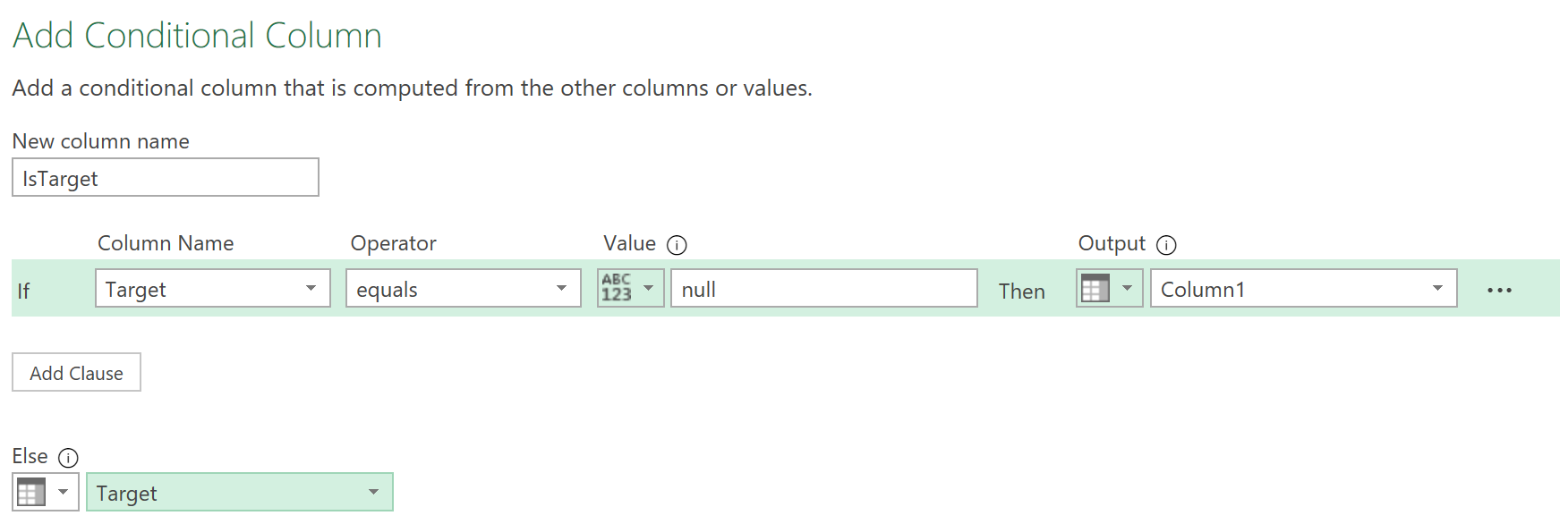
---

My Notes

1. Import worksheet
   1. Data > Get Data > from workbook > from worksheet > Cust Header Conversion.xlsx
2. Promote first row as headers
3. Rename this query to Conversion Table
4. Close & Load
   1. Load as connection Only (Do not add to data model)
   2. We don’t want to connect this data to the data model because we will merge this with another table which will load into the data model
5. Import Folder
   1. Customer Folder
6. Transform Sample Data
   1. Remove promote headers from Applied steps
   2. Transform > Transpose
   3. Merge Conversion
      1. Home > Merge Queries



* + 1. In the converstion table title column
       1. Click on the arrows then remove source and only leave target selected.
       2. Uncheckmark Use original column name as prefix
       3. Call this new column Target
  1. Add conditional column



* + 1. Call this new column IsTarget
    2. Move this new conditional column to the beginning
  1. Remove
     1. Column1 and Target
  2. Transform > Transpose
  3. Promote first row as headers

1. Go to Customers Table
   1. Remove the Changed Type from Applied Steps
   2. Make sure Data Types are properly selected for each column
   3. Remove the Source.Name column
2. Close & Load to > Load too..
   1. Create connection
   2. Add data to data model

---

* 1. **BUILD THE RELATIONAL MODEL**

*How many* ***relationship joins*** *did you need to create?* ***3***

*Which is the* ***FACT*** *table in this model?* ***Sales***

---

My Notes

1. Data > Manage Data Model
2. Diagram view
   1. Select the organization in the Customers table to the organization in the Sales Org
   2. Select Category from the Categories table to the Categories from the Sales table
   3. Select Customer from Sales to Customer in Customers table
3. Then close

---

* 1. **PIVOT TABLE - PRODUCTS IN CATEGORY**

**Add a Pivot Table** on a new worksheet called **Products in Category**.

* **Display** the **Distinct Count of ProductID** in each **Category Description** in the Pivot Table.

*How many* ***Trend Bikes****products does the company have?* ***2***

*How many distinct Products are sold* ***in all Categories****?* ***28***

**Filter** the **Pivot Table** to hide the **Accessories** and **Wearable**category rows.

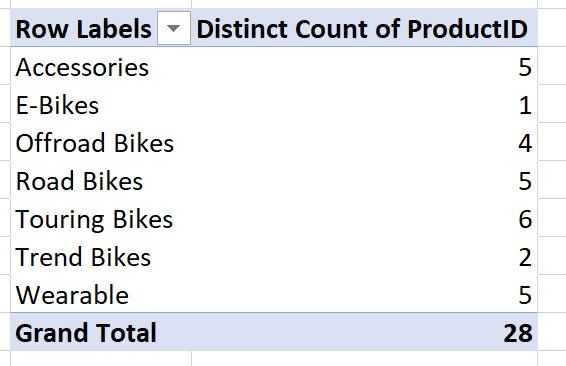
* **Rename** your **field headers** to communicate concisely and clearly
* **Sort** the value field in **descending** order
* **Apply** White, Pivot Style White Medium 1 **PivotTable Style**

*Copy and paste your* ***pivot table*** *into the answer box****.***

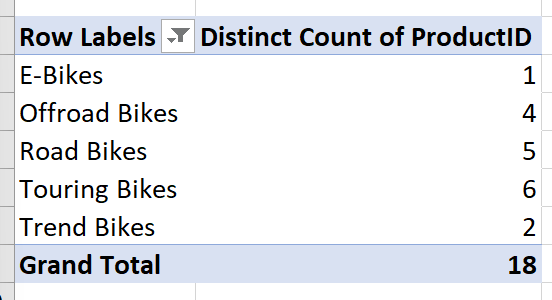
---

My Notes

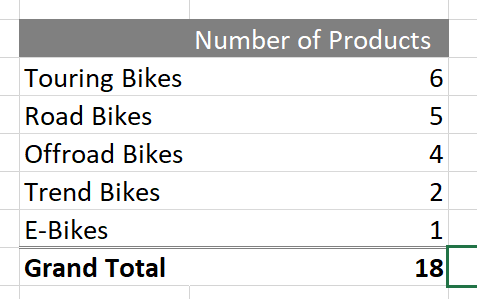
1. Label the new worksheet as Products In Category
2. Insert > PivotTable > Data from Data Model
3. Drag Fields
   1. Add ProductID to values and change the count to Distinct Count
   2. Then go to the Category Table and move over the Description to the rows



* Remove accessories and wearables



1. Click on the Row Labels dropdown and remove the Accessories and Wearables Category.
   1. Rename field headers
   2. Sort values by decending
   3. **Apply** White, Pivot Style White Medium 1 **PivotTable Style**



---

* 1. **MODIFY THE SALES QUERY**

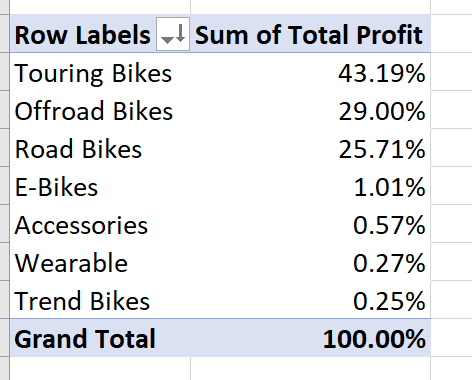
**Edit** the **Sales** query to add Custom Columns for **Total Cost**, **Total Sales** and **Total Profit** using the correct fields to calculate the values.

* **Format** these columns as **Currency**.

*What is the* ***Total Cost*** *displayed for* ***TrNum 16****of the Sales table?* ***$3,000***

**Add a Pivot Table** on a new worksheet called **Category Profit** to enable you to answer the following question.

*What is the* ***Description*** *of the* ***Category****that contributes the smallest percentage of****Total Profit****?* ***Trend Bikes***

**

*What percentage does the category E-Bikes represent of Total Profit? Round your answer to 2 decimal places.* ***1.01%***

---

My Notes

1. Click on the Sales Query
2. Create Custom Columns
   1. Total Sales
      1. [Quantity] \* [Price]
      2. Then change the data type to currency
   2. Total Cost
      1. [Quantity] \* [Cost]
      2. Change it to currency
   3. Total Profit
      1. Minus the Total Sales by Total Cost
3. Close & Load… > Load To
   1. As connection
   2. As data from data model
4. Insert > PivotTable
   1. Description from category table to Rows
   2. Sum of Total Profit to Values from the Sales table
5. Right click the area where the numbers are and show value as > % of Grand Total

---

**2.1 LOAD MULTIPLE SHEETS**

**Import & Edit/Transform** all the worksheets from **Competitive Data.xlsx** to combine the data from different regions into one query.

*If you need help please refer to Assignment 4 (Chapter 5.) Hint: Right click on the workbook name to edit/transform.*

* **Remove** any columns with **null** values
* Use **Region** as the field name
* **Rename** your query to **Competitive Data**
* **Load** as a **Connection Only** and add to **Data Model.**

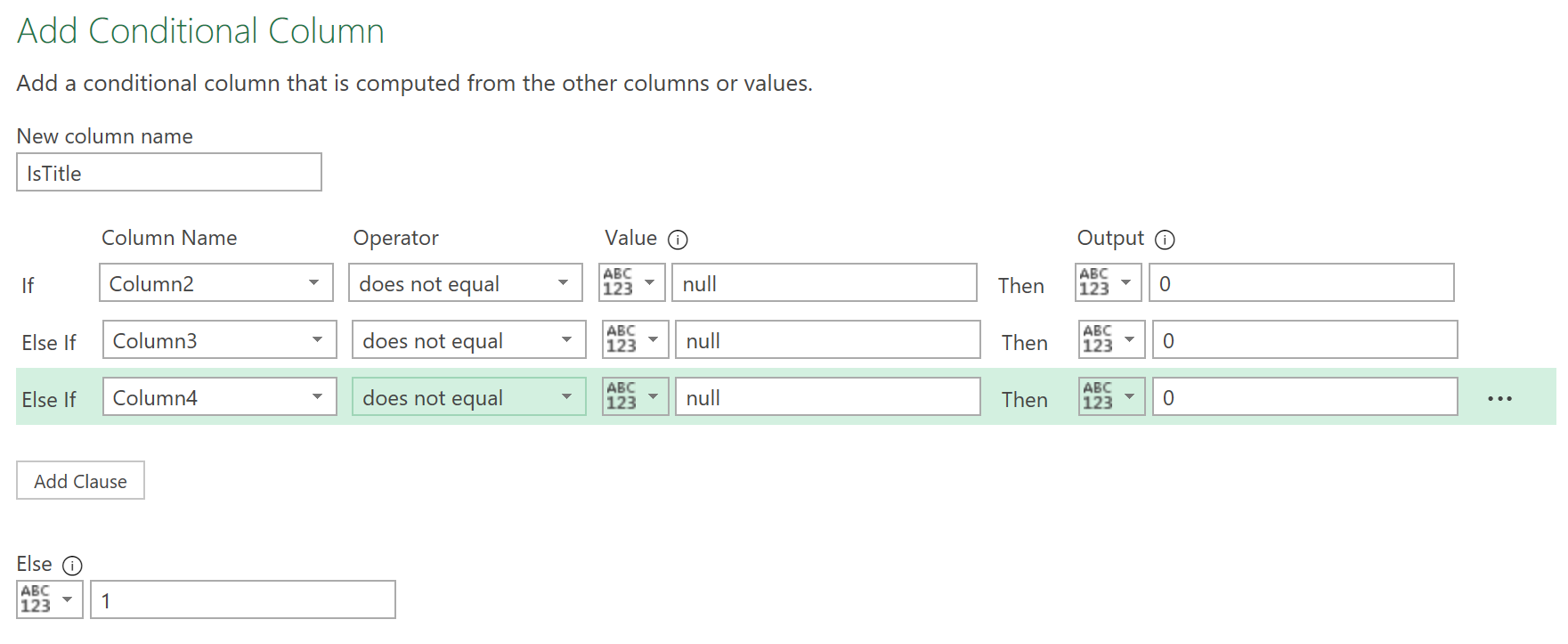
*Create a new pivot table on a new worksheet* ***Competition*** *to show the competitive****Region****that sold the highest quantity of****Off Road Helmets****?* ***Asia***

*Use the same pivot table to show how many* ***Water Bottles*** *were sold in* ***Canada****?* ***675***

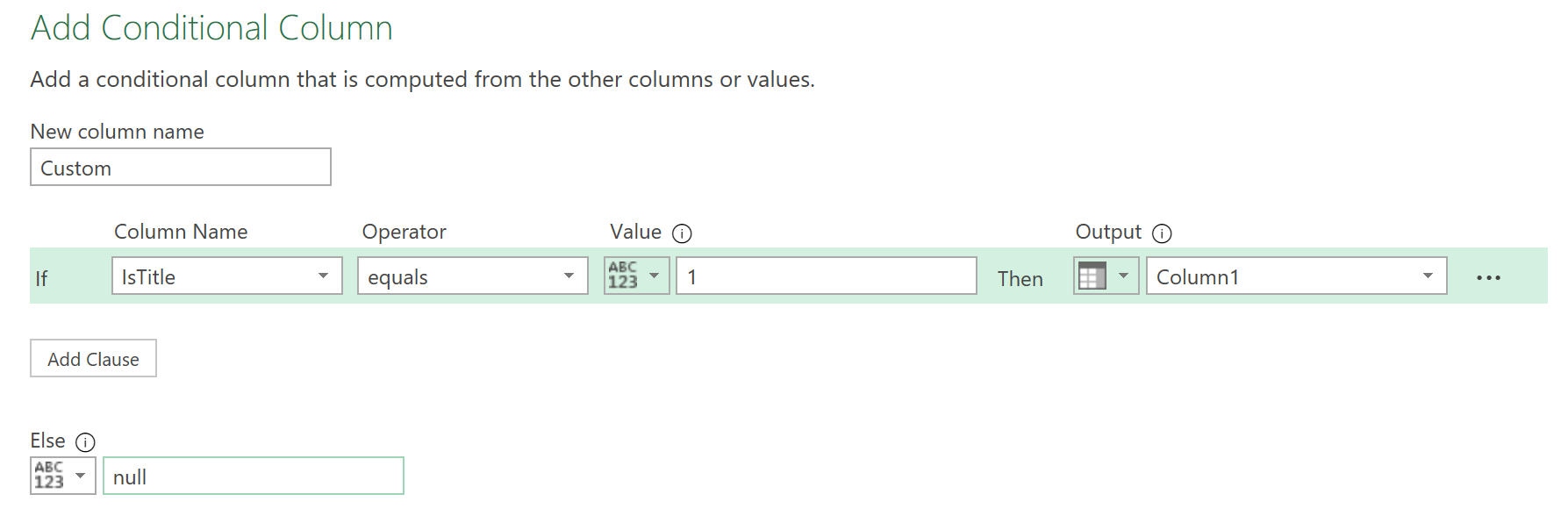
---

My Notes

1. Data > Get Data > From Workbook > From Worksheet > Competitive Data.xlsx
2. Stay in the Competitive Data Query not Transform Sample
   1. Right click the data column and click remove other columns
   2. Data > Expand information
      1. Uncheckmark Use original column name as prefix
   3. Remove all the the columns that have nulls
   4. Remove null rows by clicking column1 dropdown and uncheckmarking
   5. Add Column > Conditional Column

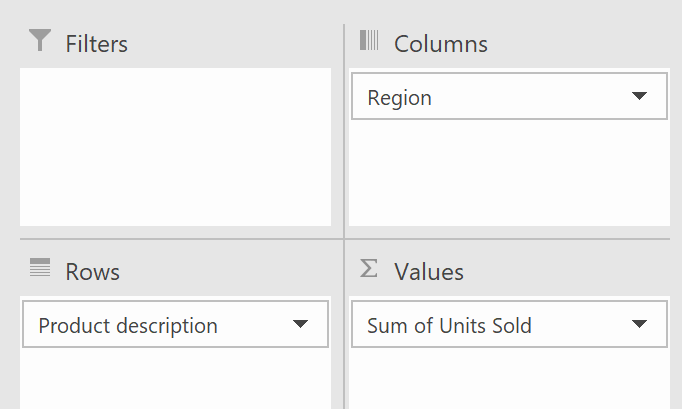


* 1. Add Column > Conditional Column



* 1. Filter out the 1 in the isTitle Column
  2. Then promote first row as headers
  3. Remove that column that has 0’s
  4. In the Products Description filter out the Product Description
  5. Change data types
  6. Close and Load > Load to..
     1. Connection only
     2. Data as data model

1. Insert > PivotTable > From Data Model
   1. Then place the items in the fields



---

**Add a Pivot Table** on a new worksheet called **Product Avg Price by Region** using the **Competitive Data** query to illustrate **Average Unit Price** by **Product** (rows) by **Region**(columns).

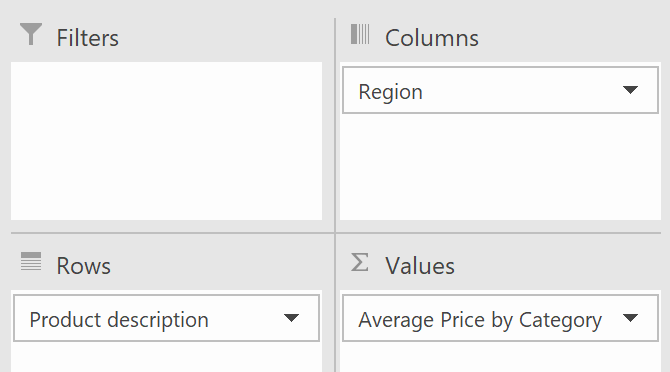
* **Rename** columns for clarity
* **Do not show Field Headers**
* **Format** values as **Currency** with zero decimal places
* **Remove** **Grand Total** columns & rows
* **Apply** PivotTable Style **White, Pivot Style Medium 1**
* **Apply** conditional formatting **Solid Fill Blue Data Bar**

*Copy and paste your* ***pivot table*** *in the answer box****.***

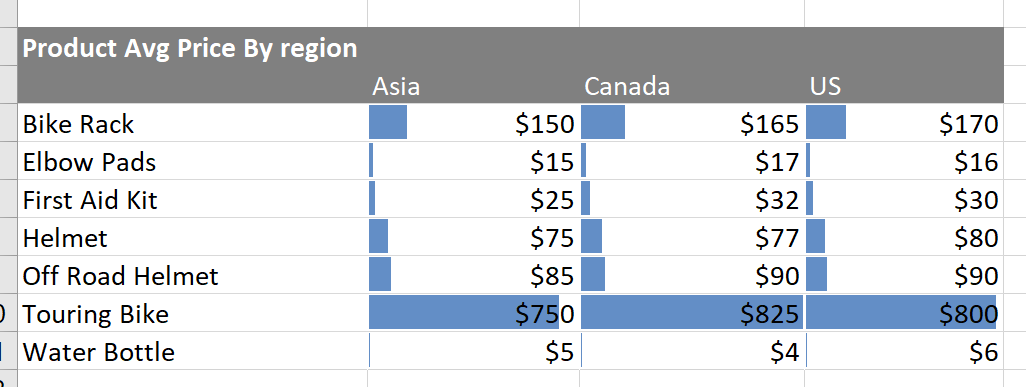
*---*

*My Notes:*

* + - 1. *Insert > PivotTable > From Data Model*
      2. *Move the fields to:*

**

* + - 1. *Add this formatting*
         1. **Rename** columns for clarity
         2. **Do not show Field Headers**
         3. **Format** values as **Currency** with zero decimal places
         4. **Remove** **Grand Total** columns & rows
         5. **Apply** PivotTable Style **White, Pivot Style Medium 1**
         6. **Apply** conditional formatting **Solid Fill Blue Data Bar**

**

*---*

**2.2 APPEND NEW DATA**

**Load** the Excel file called **Mexico.xlsx** into the Power Query Editor in Excel.

* **Add** a **Custom Column** named **Region** with the value "**Mexico**".
* **Rename** your query **Mexico**.
* **Close & Load** the query as a **Connection Only** (do not load to the data model.)

Open the **Competitive Data** query and edit in Power Query Editor in Excel.

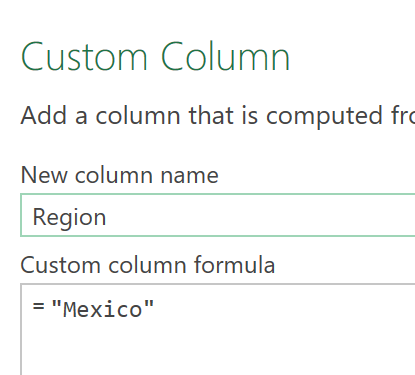
* **Append** the **Mexico** query, making adjustments as required to ensure the data appends correctly

*How many* ***columns****now appear in your* ***Competitive Data*** *query****? 5***

---

My Notes

1. Data > Get Data > From File > From Workbook
2. Add Column
   1. Custom column



1. Rename query to Mexico
2. Close & Load
   1. Load too..
   2. Only create connection
   3. DO NOT add the data to the Data Model
3. Append data
   1. Home > Append Query
   2. Select Mexico
   3. If the columsn do not match up just go back into the Mexico folder and change the headings to say the right thing. Make sure to remove the append query in the competitive data query

---

**Close and load** your **Competitive Data** query.

* **Refresh** your **Average Price** by **Product** & **Region** Pivot Table.

***What is the Average Price of Elbow Pads in Mexico.*** *Round to zero decimal points.* ***$15***

---

My Notes

* + - 1. *Close and Load*
      2. *Refresh the table to see the update by right clicking the query on the right side and right click and select refresh.*

*---*

**Edit** the **Competitive Data**query to add **Custom Columns** for **Total Sales, Total Cost, Total Profit $** and **Total Profit %**using the correct fields to calculate the values.

* **Format** these new columns appropriately

**Add a Pivot Table** on a new worksheet called **Avg Profit By Regio**n using the**Competitive Data** query to illustrate **Average Total Profit %** by **Region.**

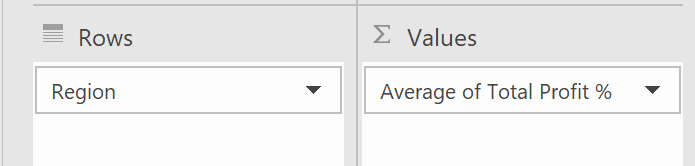
* **Rename** columns for clarity
* **Do not show Field Headers**
* **Format**values as **Percentage**with two decimal places
* **Apply** PivotTable Style **White, Pivot Style Medium 1**
* **Sort**in descending value

***Copy and paste****your pivot table in the box below.*

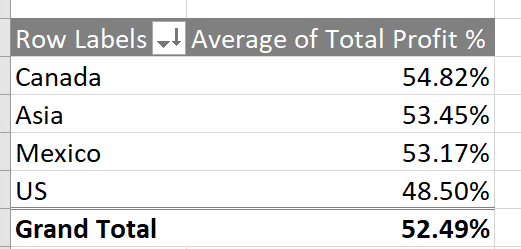
---

My Notes

1. Add Custom Columns for:
   1. Total Sales
      1. Units Sold \* Unit Price
   2. Total Cost
      1. Units Sold \* Unit Cost
   3. Total Profit
      1. Total Sales – Total Cost
   4. Total Profit %
      1. Total Profit / Total Sales
      2. Make sure this is set to percentage
2. Close & Load
3. Insert > PivotTable
   1. Make it like this:
   2. Make sure profit is average
   3. Make it percentage and to two decimal places



* 1. **Apply** PivotTable Style **White, Pivot Style Medium 1**
  2. **Sort**in descending value



---

**3.1 UNPIVOT**

**Import & Edit/Transform** the **ORDERS** sheet in **DrinkDelivery.xlsx**. The data on the worksheet is formatted as a 2x2 table so you will need to unpivot it to add it to the data model.

*If you need help please refer to textbook activities 6-5 and 6-6.*

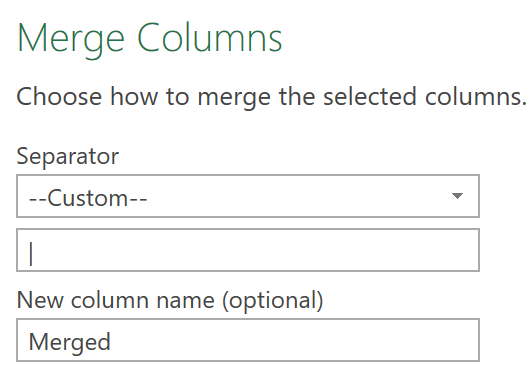
* **Load** as a **Connection Only** and add to **Data Model.**

*How many* ***rows*** *does the* ***ORDERS*** *query load to the data model?* ***72***

---

My Notes

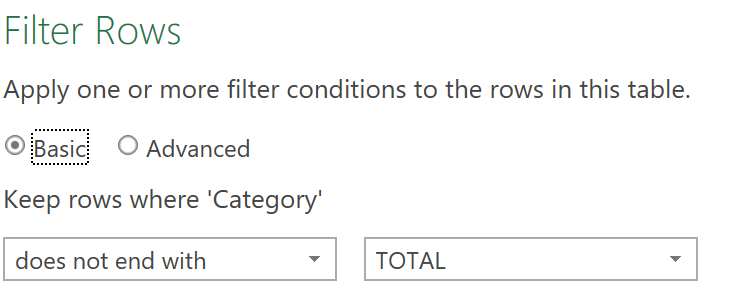
1. Import Data > Workbook > Worksheet > DrinkDelivery.xlsx
2. Load as a connection only and add to Data Model
3. Delete the last two step sfrom the Applied Steps to make sure the header is depromoted
4. Remove Rows
   1. Home > Remove Rows > Remove Bottom Row
   2. Enter in 1
5. Delete last Grand Total Column
6. Select Column1
   1. Transform > Fill > Fill Down
7. Merge Columns
   1. Select Column 1 and Column2
   2. Transform > Merge Columns



1. Transform > Transpose
2. Fill
   1. With column 1 selected Transform > Fill > Fill down
3. Promote first row as headers
4. Select Column 1 and Column2 then go to Unpivote Other Column
5. Format to all Attribute values to uppercase
6. Right click the Attribute > Split Column > By Delimeter
7. Now we can rename the columns



1. Remove The totals by
   1. Click the arrow dropdown on Category and select
   2. Text Filters > Does Not End With..
   3. Click the arrow dropdown on Region and select
   4. Text Filters > Does not end with…



---

Create a new **pivot table** based on the **ORDERS** table. Your pivot table should illustrate the breakdown of **Wine, Beer, Spirits by Region** (by percentage.)

*What is the* ***percentage of overall orders*** *were for* ***Wine*** *in the* ***Okanagan****?* ***13%***

Ensure your **Pivot Table** for the **ORDERS** query in the previous step shows Region (rows) and Category (columns) and Orders as a percent of grand total.

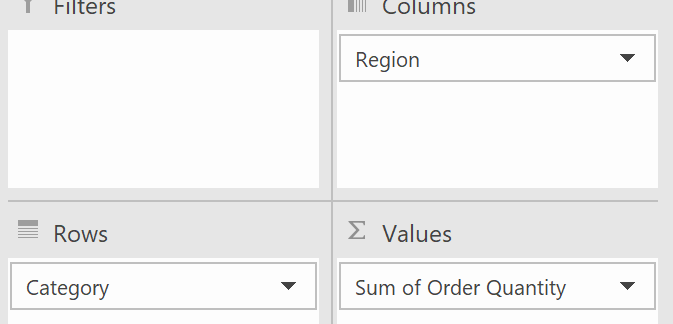
* **Format** values as **Percentage** with zero decimal places
* **Apply** PivotTable Style **White, Pivot Style Medium 1**

*Copy and paste your* ***pivot table*** *in the answer box****.***

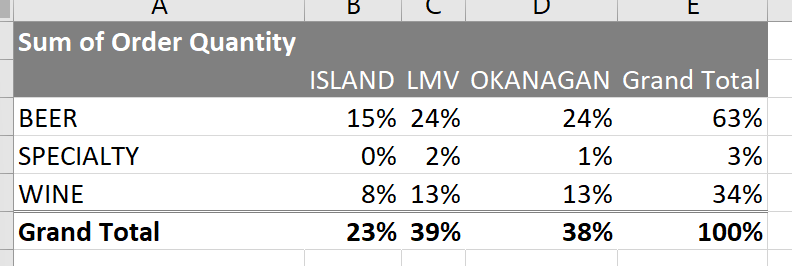
---

My Notes:

1. Create a table



1. Formate values as percentage by clicking drop down for Sum of Order Quantity > Value Field Setting > Number format > Percentage and set to zero decimal places
2. **Apply** PivotTable Style **White, Pivot Style Medium 1**



---