Exercise 1 - Managing Data Sources and Visualisation

Objective: To explore Tableau Desktop and Data Sources

Steps:

Opening an Excel File

1. Download the zip ***ManagingDataSources*** from Moodle. When downloaded, unzip this file.
2. Open ***OrderDetails*** in Tableau.
    a. Once open, browse through the data.
    b. Note the number of rows
    c. To begin a visualisation, Click on the ***Sheet 1*** tab (bottom left).
    d. To return to the data, Click on ***Data Source*** tab (bottom left).

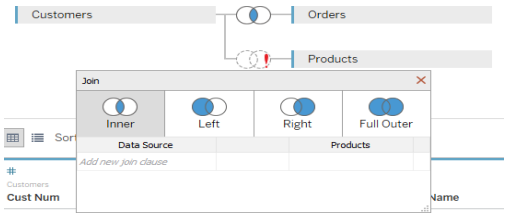
Join Related Data Sources (with Common Names)

1. First, open ***JoinExamples*** in Excel.
    a. Note the three worksheets and their structure.
    b. Also, notice how columns in some sheets are named the same as columns in other sheets. This will allow us to join the data sources.
    c. Close Excel.
2. Open ***JoinExamples*** in Tableau.
    a. Go to **data** tab and select, **New Data Source**. When the dialog opens, chose the Excel option to open the file, ***JoinExamples***.
    b. When open, on the LHS under sheets, double click the ***Customers*** sheet and the ***Orders*** sheets.
3. Notice how Tableau creates a join automatically.

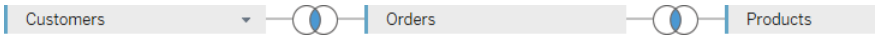
4. If you hover over the join symbol, Tableau will indicate that there is an ***Inner Join*** based on ***Cust Num***.
5. If we now click on the ***Sheet1*** tab, notice how we have access to ***Customers*** and ***Orders*** under the ***Dimensions*** section.
6. To create a very basic visualisation, drag ***Cust Full Name*** to the sheet 1 area. Also, from the ***Measures*** section, drag ***Number of Records*** to the sheet area. This will create a table (list) of customers and the number of orders they have. Notice how this data spans two different tables and requires the join to bring this data together.

Join Related Fields (with Inconsistent Names)

1. When it comes to joining data sources, we may need to join on fields that **don't** have the same name.
2. Open ***JoinExamples*** in Excel and note the three sheets. On the ***Orders*** sheet, note the ***ProductID*** column. Also, on the ***Products*** sheet, note the ***SKU*** column. We need to make a *join* based on these two columns.
3. Open ***JoinExamples*** in Tableau.
    a. Open the three sheets in Tableau. Note how Tableau automatically creates a join (as before).
    b. When we add the Products sheet, the Join Dialog appears. Tableau doesn't know how to join the information – so we must tell it.

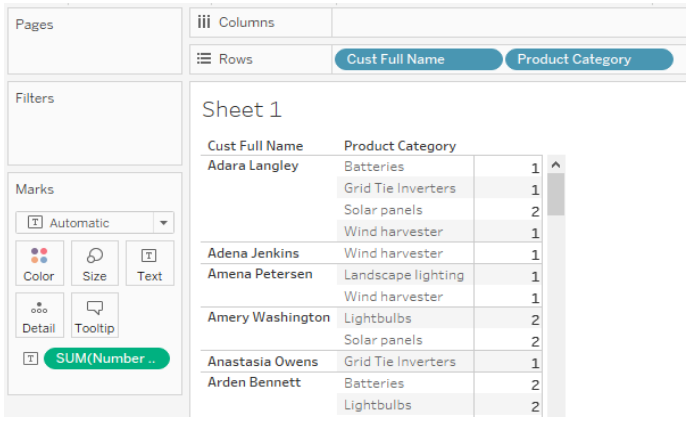


    c. Under ***Data Source***, select ***ProductID***
    d. Under ***Products***, select ***SKU***



    e. All three tables are now joined together.

4. Open the ***Sheet 1*** tab, all fields are now visible under dimensions.
    a. Drag out ***Customer Full Name***, note how it sits on the ***Row Shelf.***
    b. Drag ***Product Category*** to the Row Shelf – note how we get the items bought per customer.
    c. Now drag ***Number of Records*** (from the measures section like before) to the Sheet area. This displays the number of products that was bought per customer.



Cleaning Source Data

1. With Tableau open, open the Excel file – ***CleanData***.
2. Drag Order Details to the Sheets area and note how the data is not formatted properly. What errors/inconsistencies can you notice?
    a. Null Errors
    b. Vague Column Names
    c. First row of data may be headers

OrderDetails Sales Summary fo...	OrderDeta... F2	OrderDetails F3	OrderDetails F4	OrderDetails F5	OrderDetails F6	OrderDeta... F7	OrderDetails F8	OrderDetail F9
null	*null*	CustFullName	*null*	SKU	ProductName	*null*	*null*	*null*
1	77	Graham Frederick	01/01/2017	KE625	K-Eco 625	1	1,201.23	1,201
2	42	Jael Rice	01/01/2017	KI9K	K-Invert 9K	10	19,806.00	198,060
3	5	Nolan Slater	05/01/2017	KI7K	K-Invert 7K	2	10,714.00	21,428
4	49	Blaine Ashley	15/01/2017	KE12Li	K-Eco 12-volt Lithium-...	8	4,504.00	36,032

3. To clean the data, we will use ***Data Interpreter.***
    a. On the LHS under Sheets, click the check box for ***Use Data Interpreter***.
    b. Note how the data gets cleaned up.

Visualisation using Show Me

1. With Tableau open, open file – **ShowMe.twb**.
    a. To go ***file*** -> ***open*** and then select the required file.
2. To start, under the ***Measures*** section, highlight ***Total Price.***
3. Once highlighted, click ***Show Me*** – which is located at the top RHS.
    a. This will show us the available Visualisations.
4. Drag ***Total Price*** into the sheet area. The number that appears is the Total Price for everything. The only other options we have is a bar chart or histogram. Select the Bar Chart.
    a. Notice that you only get a Single Bar Line. To get more detail, we need to add more dimensions. Drag ***ProductName*** to the ***Row Shelf***. Notice how we get a break down by product name now.
    b. We can swap the ***ProductName*** and ***Total Price*** around to get a different graph. Drag ***ProductName*** from the *Row Shelf* to the *Column Shelf* and vice-versa for the ***Total Price***. Note the result.

Change the order of the fields in a Visualisation.

1. With Tableau open, open file – **Reorder.twb**.
    a. To go ***file*** -> ***open*** and then select the required file.
2. Let's create a visualisation based on the Cust Full Name, Product Name and Total Price. To achieve this:
    a. ***Visualisation 1:*** Drag ***Cust Full Names*** to Columns, then drag ***Product Name*** to Rows, finally, drag ***Total Price*** to the data area. Notice how the visualisation is not very informative.
    b. Let's change it around

c. ***Visualisation 2:*** Drag ***Cust Full Names*** down to Rows to the left of ***ProductName***. Note how with the same data, the visualisation is more informative now.
d. Therefore, if you feel that you are not getting as much from your data as you think, try changing things about.

Change the summary data operation.

1. With Tableau open, open file – **Summary.twb**.
    a. To go ***file*** -> ***open*** and then select the required file.
2. Notice, that in the table, we have a ***Total Price*** – this is automatically summed for us (total price per month). What if we wanted something different – such as the average (average price per month)?
3. Under the ***Marks*** section, select ***Total Price***, click on the ***drop menu option***. From the menu, select ***Measure*** …. Note the different options available and then select ***average***. Note the changes.
4. Now try and determine the ***Count*** of orders for each month. (Use CTRL+C to undo)

5. To add a second value, drag another copy of ***Total Price*** to the Marks section and drop it under the existing entry.
    a. Set one to ***SUM*** and the other to ***AVERAGE***.
    b. Now change the new entry to ***Text***, by selecting the small box beside it and the select ***Text*** from the available options.
    c. Verify that we now have two entries in the table, a total price and an average price for the month.
    d. To remove it, just drag it from ***Marks*** to ***Measures***.

Split text into multiple Columns

1. Open the Excel file ***SplitData*** in Tableau. (Go to Data Tab)
2. Notice how we have the full name in one column. To split this, hover over the column and select the drop down icon when it appears.
3. From the list select, ***Split***.
    a. Notice how we still have a full name, but we also have a split 1 and split 2 entry.
    b. To rename that column, double click it and then rename.
    c. Rename the split columns to FirstName and LastName.
4. To split a column using Delimiters, hover over the column and select the drop down icon when it appears. Then select ***Custom Spilt***. This will allow you to use delimiters.

Present data using Storylines.

1. In addition to data analysis and visualisation, it is important to communicate findings effectively. Tableau uses Storyline to build a narrative.
2. Open the ***Storyline.twb*** file in Tableau. Note the three sheets.

3. To create a storyline, we select the ***New Story*** Icon found at the bottom of the window.
4. We can now add elements to the story:
    a. Drag in Categories, add the caption ***Sales by Category for 2017 and 2018***
    b. Try dragging on another text box to add more captions.
    c. We can create a new story point, by selecting blank on the LHS. Note the navigation bar that appears at the top once a story point is added.
    d. Drag ***ByQuarter*** into the new story line. Create the following caption, ***Total Sales by Quarter***.
    e. Add a third story point and drag map summary into it. Add the caption***,*** ***Map Summary***.
    f. We can now navigate between the different story points as we require.
    g. We can rename the storyline by double clicking on the tab.

------- End of Exercise -------