**Working on Multiple Tables**

**Introduction**

In this, you will be working with Power Query! We will walk through the steps of obtaining data, transforming it, then loading it back into Excel.

We will provide you with 4 files: 2017 Order Data, 2018 Order Data, 2019 Order Data, and 2020 Order Data.

1. Download these to your computer.
2. Create a new folder on your computer called “Cookie Company Data.”
3. Place 2017 Order Data, 2018 Order Data, and 2019 Order Data inside the Cookie Company Data folder. DO NOT PLACE 2020 ORDER DATA inside this folder!

\*\*Make sure that you keep the Cookie Company Data folder and the 2020 Order Data file separate for now (at the end of the lab, we will drag the 2020 Order Data file into the Cookie Company Data folder, but hold off on this for now!)

**Step 1: Obtain the Data**

In this analysis, we’ll be using cookie store data. During class, we practiced using Power Query with data from the web and from a range. Now, we’ll practice using Power Query with data from a folder!

1. Open up a new Excel worksheet.
2. Go to the Data tab. Click **Get Data**.
3. Select where you want to get the data from. In this case, you will select the “Cookie Company Data” folder that you downloaded at the start of today’s lab (From File 🡪 From Folder).
4. Select **Combine & Transform Data**. (Note: We could just load the data as we did during class, but transforming the data before combining allows us to modify the data before ever pulling it into Excel!).
5. Click on the **Order Data** worksheet. Click **OK**.
6. The Power Query editor will pop up in a new window.

**Step 2: Transform the Data**

Apply the following transformations to the data using the Power Query editor.

1. Delete the **Source.Name** column.
2. Remove all XXXXXX observations from the data set (you can filter them out of the **Order ID** column).
3. Split the **Customer ID** column into 2 columns: Customer ID and Customer Name.
   1. Select the **Customer ID** column. In the Home tab, click **Split Columns** 🡪 **By Delimiter**.
   2. In the box where you specify the delimiter, Power Query automatically suggested that a hyphen might be the delimiter. Add a space before and after this dash so that we properly split the columns.
   3. Click **OK**.
   4. Make sure that the columns were split properly. If they were not, click the X next to the appropriate Applied Step and retry this step!
   5. Rename **Customer ID.1** to be **Customer ID**.
   6. Rename **Customer ID.2** to be **Customer Name**.
   7. These columns should look like this now:

A screenshot of a computer

Description automatically generated with medium confidence

1. Add a new column called **Profit**.
   1. Profit is calculated by subtracting costs from revenue (Profit = Revenue - Cost).
   2. Select the **Revenue** column, then hit the **Ctrl** key and select the **Cost** column. Now, both **Revenue** and **Cost** should be selected in this order.
   3. Go to the Add Column tab.
   4. Click **Standard** 🡪 **Subtract**.
   5. Change the name of this new column to **Profit**.
   6. Drag the **Profit** column to the left so that it is right after the **Cost** column.
2. Change the data type of **Cost**, **Revenue**, and **Profit** to Currency.
3. Add a new column called **Month Shipped** that contains the month in which the order was shipped. For example, if the **Ship Date** is 12/31/2017, then the **Month Shipped** should be December.
   1. Click on the **Ship Date** column.
   2. Go to the Add Column tab.
   3. Click on the dropdown under **Columns From Examples** 🡪 **From Selection**.
   4. In the green box that pops up under **Column1**, type in **January** and hit Enter.
   5. Click **OK**.
   6. Rename this column to be **Month Shipped**.
   7. Drag this column to the left so that it is right after the **Ship Date** column.
4. We will now add a conditional column called **Order Status (Binary)**. This column will contain a 1 if the order has been shipped and a 0 if the order has not been shipped.
   1. Select the **Order Status** column.
   2. Go to the Add Column tab.
   3. Click **Conditional Column.**
   4. Under New Column Name, type in **Order Status (Binary)**.
   5. A screenshot of a computer

      Description automatically generated with medium confidenceInput the following:
   6. Click **OK**.
5. Remove the **Order Status** column.

**Part 3: Load the Data into Excel**

1. In the Home tab, click **Close and Load** to return to Excel. Your data should look like this:

A screenshot of a computer

Description automatically generated with medium confidence

1. Fortunately, the 2020 data has just arrived! In your File Explorer (or wherever you store your files), drag the **2020 Order Data** file into the **Cookie Company Data** folder.
2. Next, go back to your Excel workbook. Go to the Data tab, and click **Refresh All**. This will allow Excel to recognize that we’ve added a file to our source folder.
3. Now, scroll to the bottom of the sheet. Look—our 2020 data is there, and we didn’t have to do any extra data transformations!