PART 1: Connecting & Shaping the Data

*Open a new Power BI Desktop file, and complete the following steps:*

**1)** Update your Power BI options and settings as follows:

* Deselect the "*Autodetect new relationships after data is loaded*" option in the **Data Load** tab
* Make sure that Locale for import is set to "*English (United States)*" in the **Regional Settings** tab

**2)** Connect to the **MavenMarket\_Customers** csv file

* Name the table "**Customers**", and make sure that headers have been promoted
* Confirm that data types are accurate (**Note:** "*customer\_id*" should be whole numbers, and both "*customer\_acct\_num*" and "*customer\_postal\_code*" should be text)
* Add a new column named "*full\_name"*to merge the the "*first\_name"* and "*last\_name"* columns, separated by a space
* Create a new column named "*birth\_year"*to extract the year from the "*birthdate"*column, and format as text
* Create a **conditional column**named "*has\_children"*which equals "**N**" if "*total\_children"* = 0, otherwise "**Y**"

**3)** Connect to the **MavenMarket\_Products** csv file

* Name the table "**Products**" and make sure that headers have been promoted
* Confirm that data types are accurate (**Note:** "*product\_id*" should be whole numbers, "*product\_sku*" should be text), "*product\_retail\_price*" and "*product\_cost*" should be decimal numbers)
* Use the statistics tools to return the number of distinct product brands, followed by distinct product names
  + ***Spot check:****You should see****111****brands and****1,560****product names*
* Add a calculated column named "*discount\_price*", equal to 90% of the original retail price
  + Format as a fixed decimal number, and then use the rounding tool to round to 2 digits
* Select "*product\_brand*" and use the **Group By** option to calculate the average retail price by brand, and name the new column "*Avg Retail Price*"
  + ***Spot check:****You should see an average retail price of****$2.18****for Washington products, and****$2.21****for Green Ribbon*
* Delete the last applied step to return the table to its pre-grouped state
* Replace "*null*" values with zeros in both the "*recyclable*" and "*low-fat*" columns

**4)** Connect to the **MavenMarket\_Stores** csv file

* Name the table "**Stores**" and make sure that headers have been promoted
* Confirm that data types are accurate (**Note:** "*store\_id*" and "*region\_id*" should be whole numbers)
* Add a calculated column named "*full\_address*", by merging "*store\_city*", "*store\_state*", and "*store\_country*", separated by a comma and space (***hint:****use a custom separator*)
* Add a calculated column named "*area\_code*", by extracting the characters before the dash ("-") in the "*store\_phone*" field

**5)** Connect to the **MavenMarket\_Regions** csv file

* Name the table "**Regions**" and make sure that headers have been promoted
* Confirm that data types are accurate (**Note:** "*region\_id*" should be whole numbers)

**6)** Connect to the **MavenMarket\_Calendar** csv file

* Name the table "**Calendar**" and make sure that headers have been promoted
* Use the date tools in the query editor to add the following columns:
  + *Start of Week (starting Sunday*
  + *Name of Day*
  + *Start of Month*
  + *Name of Month*
  + *Quarter of Year*
  + *Year*

**7)** Connect to the **MavenMarket\_Returns** csv file

* Name the table "**Return\_Data**" and make sure that headers have been promoted
* Confirm that data types are accurate (all ID columns and *quantity* should be whole numbers)

**8)** Add a new folder on your desktop (or in your documents) named "**MavenMarket Transactions**", containing both the **MavenMarket\_Transactions\_1997** and **MavenMarket\_Transactions\_1998**csv files

* Connect to the folder path, and choose "Edit" (*vs. Combine and Edit*)
* Click the "*Content*" column header (double arrow icon) to combine the files, then remove the "*Source.Name*" column
* Name the table "**Transaction\_Data**", and confirm that headers have been promoted
* Confirm that data types are accurate (all ID columns and *quantity* should be whole numbers)
  + ***Spot check:****You should see data from 1/1/1997 through 12/30/1998 in the "transaction\_date" column*

**9)** With the exception of the two data tables, disable "*Include in Report Refresh*", then **Close & Apply**

* Confirm that all 7 tables are now accessible within both the **RELATIONSHIPS** view and the **DATA** view

**10)** Save your .pbix file (*i.e. "****MavenMarket\_Report****"*)

***Solution screenshot (for reference):***

