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) Save your .pbix file (*i.e. "**Retail Report**"*)

3) Creating a **Staging Area** or **Staging Layer** so that raw data is accessible in its original form

4) Connect to the **Raw_Customers** csv file in the Staging Area and use Reference option to make a copy. Right click to create a New Group named **Data Model.** Uncheck the **Enable Load** option

- 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**"

5) Connect to the **Raw_Products** csv file. Repeat the process of using Reference, moving the query to the Data Model area and unchecking Enable Load

- 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 data profiling 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

6) Connect to the **Raw_Stores** csv file. Repeat the process of using Reference, moving the query to the Data Model area and unchecking Enable Load

- 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

7) Connect to the **Raw_Regions** csv file. Repeat the process of using Reference, moving the query to the Data Model area and unchecking Enable Load

- 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)

8) Connect to the **Raw_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*

9) Connect to the **Raw_Returns** csv file. Repeat the process of using Reference, moving the query to the Data Model area and unchecking Enable Load

- 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)

10) Add a new folder on your desktop (or in your documents) named "**Transactions**", containing both the **Raw_Transactions_1997** and **Raw_Transactions_1998** csv files

- Connect to the folder path, and choose "Edit" (*vs. Combine and Transform Data*)
- Repeat the process of using Reference, moving the query to the Data Model area and unchecking **Enable Load.**
- Verify that both files have been combined correctly and delete 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*

11) 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