Tableau at UNH: Using Data Interpreter with Tableau 9 and Other Data Cleanup you can do with Excel Data Sources

1) Info on Tableau’s website about how to use Tableau Data Interpreter in Tableau Desktop: http://onlinehelp.tableau.com/v9.0/pro/online/windows/en-us/help.htm#qs_excel_cleaning.html

2) Info on Tableau’s website about Excel data source options you can set in Tableau Desktop before building a view: http://onlinehelp.tableau.com/v9.0/pro/online/windows/en-us/help.htm#examples_excel.html

Set Excel table options: In the join area, hover over the table name until the gear icon displays. Click the icon and specify if the data includes field names in the first row. If column names are not included, they will be automatically generated by Tableau. You can rename fields later.

Turn on Data Interpreter: Sometimes the format of your Excel data is difficult to analyze in Tableau. It might have hierarchical headers, extraneous headers/footers or empty rows & columns. Data Interpreter can remove extraneous info to help prepare your Tableau data source. After you set up the data source, if Tableau detects unique formatting or extraneous info, the Data Interpreter option becomes available. The Data Interpreter makes changes only to the Tableau data source; it never modifies underlying data.

To turn on Data Interpreter and review results:
1. After you have set up your Excel data source, on the data source page, click Turn on.
2. Click Review results. A copy of your data source opens in Excel on the Key tab. (See below screenshot for an example of the information it provides).
3. Review the annotation key to find out how to read the results. Click subsequent tabs to review how the Data Interpreter interpreted the data source. You can also review Data Interpreter results in the preview area below the Data Interpreter. If Data Interpreter does not provide expected results, you can turn it off and use the original data source.
4. To turn off the Data Interpreter in Tableau Desktop, click Turn off.
Preview data, pivot, and split: The first 10,000 rows in the data source are automatically displayed below the join area. If you add/remove tables, or make changes to join conditions, the preview updates.

You can also do the following in the preview area:
- Change data type or geographical role of a column by clicking the data type icon
- Rename or hide a column by clicking the column header drop-down menu and selecting the respective option
- Pivot fields to transform data in a crosstab format into columnar format. This is only available for non-legacy connection types. For more info, see Pivot Data (from Columns to Rows).
- Split a string field into multiple columns of string fields. This is only available for non-legacy connection types. For more info, see Split a Field into Multiple Fields.
3) Info on Tableau’s website about two issues you may encounter with Data Interpreter:


a) Data Interpreter Option Not Available - The Data Interpreter option may not be available if:

- The data source is already in a format Tableau can interpret: if Tableau doesn't need help from the Data Interpreter to handle unique formatting or extraneous info, Data Interpreter will not be available.
- Many rows or columns: Data Interpreter is not be available when your Excel data has the following attributes:
  - Excel data contains more than 2000 columns
  - Excel data contains more than 3000 rows and more than 150 columns
- Data source not supported: Data Interpreter is only available for Excel data sources in XLS and XLSX formats. Excel data in CSV formats is not supported.

b) Tables Excluded - If the Data Interpreter detects more than one table in an Excel worksheet, it will exclude all but the first table. You can review what the Data Interpreter has included & excluded in Data Interpreter results from the Data Source page.

To resolve this issue, consider modifying the underlying data in one of two ways:

- Make sure that each table is in its own Excel worksheet
- Create named ranges for each table