**Objective**

For every “entry” in the Excel file (i.e. every row that contains nutrition information), five columns need to be created. To understand the information that needs to go into the five columns, one needs to understand the “hierarchy” of the information as well as how the “serving size” information may be displayed. Both are described below. To illustrate, an Excel file with a few examples of how the five columns would look like for a few entries has been created (“sample column split.exe”).

The original PDF file from the book as well as the Excel file that was created after OCR reading are provided.

**Brand Hierarchy**

1. Product Category (column 1)
   1. Listed in bold
   2. Sorted alphabetically(a-z)

Ex. **Cake, mix**

* 1. Applies to every entry (cell) below it until the a cell containing a bolded word appears
  2. Note: every entry MUST to have information in the Product Category column

1. Product Description (column 2)
   1. Description ends with a semicolon.
   2. Sorted alphabetically (a-z). Ex. angel food:
   3. It applies until a new product description appears, or b) a new product category appears.
   4. It is possible that a product category does not have product descriptions (i.e. the product category is fine enough to describe all brand names – see below- in it).
   5. Notes: Those entries that do not have information for the Product Description column can have a NA written in this column. Also, note that the product description is often, but not always, listed separately in a row. But sometimes the product description may be listed in the same row as the product category (in non-bolded letters, after a comma). In the example below, “Angel-hair pasta entree” is the Product Category (since it is in bold) and “frozen” is the Product Description (more on the serving size “1 pkg” later).



1. Type (column 3)
   1. This is additional information about the product that may or may not exist for an entry. When it does it appear, it does so in two ways:
      1. Immediately before the parentheses containing the brand name (see below for brand name), in the same cell as the brand name. In the example below, Type would be “Asian”.



* + 1. In the rows following the different variants of the brand. In the example below, the brand “*Pacific Organic*” has four “types”: “original”, “unsweetened”, “vanilla”, “vanilla, unsweetened”



* 1. Note: those entries that do not have information for Type can have a NA written in this column.

1. Brand name (column 4)
   1. This information is contained in parentheses and usually (not always) all words inside parentheses are in Italics. Note: “(cont.)” appears frequently but is not a brand name, although it is sometimes italicized
   2. Applies to every cell below it until either: a) a new brand appears, b) a new product description appears, or c) a new product category appears.
   3. Sorted alphabetically (a-z). Ex. (*Betty Crocker*)
   4. Note: not all entries will contain a “brand name”. That is, there will be entries that will have nutrition information but for which there is no brand attached to it. This is the case for fresh produce, for instance, such as “apple”, “lettuce”, etc.
   5. Note: those entries that do not have a brand name can have a NA written in this column.

**Serving size** (column 5)

1. It is listed in a variety of ways:
   * In the same cell as Product Category, after the bolded name of the product category. Ex: **Bagel chips/crisps**, 1 oz. This is the “default” product size for all brands in that product category. Sometimes the cell of the Product Category would explicitly state the default nature of the serving size right after (e.g. **Beef entree, microwave**, 10-oz. cont., except as noted)
   * In the same cell as the Product Description. Ex: ground, 1/3 cup: . This also is a default serving size for all brands in that product description. As in the Product Category case, sometimes the “except as noted” appears to denote the default nature.
   * In the same cell as the brand name: ex. (*Pillsbury Funfetti*), 1/12 pkg.
2. Comes in various units such as:
   * pkg.
   * pcs.
   * tbsp.
   * oz.
3. There MUST be serving size information for **every** entry
   * If not stated in the brand name cell, pick the default serving size in the Product Description. Failing either, there should be a product size in the Product Category.