How Matching Works

Our strategy is to compare each U.S. model to every CM model. Then we subtract the numeric values of the product characteristics to create a "Dif-Code." If your product characteristic variables are properly aligned (left-to-right in order of importance), then the Dif-Code can be looked at as a giant number, reading it from left-to-right without the hyphens. The smaller the Dif-Code is, the smaller the physical differences between the U.S. model and the CM model.

US Model	US Char	CM Char	CM Model	Dif-Code
A	01-01	01-01	A	00-00
A	01-01	01-02	B	00-01
B	01-02	01-01	A	00-01
B	01-02	01-02	B	00-00
C	02-01	01-01	A	01-00
C	02-01	01-02	B	01-01

We sort by Dif-Codes from smallest to largest, bringing the smallest differences to the top (first line) for each U.S. model. Then we create a variable (see "Choice" below) that ranks the choices by counting from the first line of data for each U.S. model to the last. The best match will appear on the first line for each U.S. model and will have the smallest choice number (*i.e.*, 1).

US Model	CM Model	Difcode (sorted)	Choice	Match
A	A	00-00	1	√
A	B	00-01	2	
B	B	00-00	1	√
B	A	00-01	2	
C	A	01-00	1	√
C	B	01-01	2	

Not only is the order of the product characteristics (left-to-right in order of importance) crucial but the order and spacing of the number valuations within an individual product characteristic are important in matching, too. Let's look at an example of a product with just one determining product characteristic--wooden dowels for craft stores. They are all made of the same wood and are all the same length. The only difference from one model to another is the diameter of the dowels. As you can see from the chart below, US Model D is closest to CM Model C in size, with the next closest being B, then A and then, at far last, E.

CM Model	DIAMETER (mm)	US Model
A	10	
B	11	
C	12	
	13	D
E	25	

If the product characteristic were reported in a way that just numbered the models from smallest to largest without spacing the numbers based on actual size (*i.e.*, A=1, B=2, C=3, D=4, E=5), we could get US Model D finding CM Models C & E to be equally similar—an incorrect result.

US Model	US Char	CM Model	CM Char	Dif-Code
D	4	A	1	3
D	4	B	2	2
D	4	C	3	1
D	4	E	5	1

One solution would be to use the actual diameters as the value of the product characteristic. This would work fine for wooden dowels but may not work well for other products with characteristics that are not defined by usable numbers. In such cases, you can adjust the spacing of the product characteristic values to ensure that matching works properly. In the case above, we want the US Model D to match to CM models in this order: C, B, A and then E. To achieve this result, you need to adjust the value of the product characteristic for Model E such that it generates the largest Dif-Code, as so:

US Model	US Char	CM Model	CM Char	Dif-Code
D	4	A	1	3
D	4	B	2	2
D	4	C	3	1
D	4	*E*	*8*	*4*