Though we want a smaller sample of data, we have to ensure that most of information from original data will not lose due to sampling process, so we choose to draw our sample data proportionally based on the strata of State variable, since its representativeness is the most significant and it has obvious levels.

First, we decided to draw a dataset with size of 7280 rows, because we have 72864 rows in the original dataset. Then, we used table() function to see the proportion of State variable, what we got was not very satisfying, like 1158 for Alabama State, to simplify the process, we can take for level i,

We used ceiling function because are not integers, we take ceiling values for to protect integrality of levels with few rows. After this, “Sampling” package is installed and loaded, strata() function is used to get stratified sample named as food.