**Check of SPA**

When checking the frequencies in the dataset “NETS2014\_cat\_loc\_final.txt”

For SPA – finding there are 0 coded as SPA

Cumulative Cumulative

SPA Frequency Percent Frequency Percent

ƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒ

0 78201615 100.00 78201615 100.00

When checking only the SIC codes that should be coded as spa (without including the word search with it) it should be:

Cumulative Cumulative

SPAs Frequency Percent Frequency Percent

ƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒƒ

0 78147007 99.93 78147007 99.93

1 54608 0.07 78201615 100.00

I’m thinking I must have messed up the code for this. Here’s what the logic should be to code them:

If BEH\_SIC = (79910103, 72990203) then spa = yes (so everything in these 2 SIC codes should be included.)

Else if BEH\_SIC in (72990200-72990202 or 80499912) and contains (“SPA”) then spa = yes (so if falls in the SIC code ranges above, then should be included.)

The other note is that it should only contain SPA as a separate word, not if the letter combination SPA falls within a word (so for example don’t include SPACE)

Here’s the code that I sent. Can you figure out where I messed up?

"SPA" : {

"sic\_exclusive" : [80499912],

"sic\_range" : [79910103,

72990203],

"sic\_range\_2": [72990200,

72990202 ],

"name" : ["SPA"],

"conditional" : 11

}

"SPA" : {

"sic\_exclusive" : [79910103,

72990203],

"sic\_exclusive\_2" : [80499912],

"sic\_range": [72990200,

72990202],

"name" : [" SPA", “SPA ”, “ SPA ”],

"conditional" : 11

}