Analysis

In my analysis I found, through trial and error, that a support value of 0.3-0.4 and a confidence value of 0.5-0.6 is the best for producing meaningful and interesting rules. From testing with the support and confidence values that were given to me, I found that a support value any higher than 0.4 produced no rules, and any support value lower than 0.3 produced a large number of rules that had lower confidence values. I tried to aim for a confidence level of at least 0.65 when producing rules, and I was able to produce that by setting my confidence to 0.6 and my support to 0.3. A lot of values in the data set ended up having low support (0.29 and under) which is why my output changed so drastically between my first run (with support = 0.2 and confidence = 0.5) and subsequent runs.

As I was typing this, I decided to run a few more tests and found that with a lower support threshold and a higher confidence threshold (0.1 and 0.7) that I produced more rules with high confidence and lift which I found to be very interesting. I have a hunch it relates to how often certain items appear in the data, and how often they appear together.