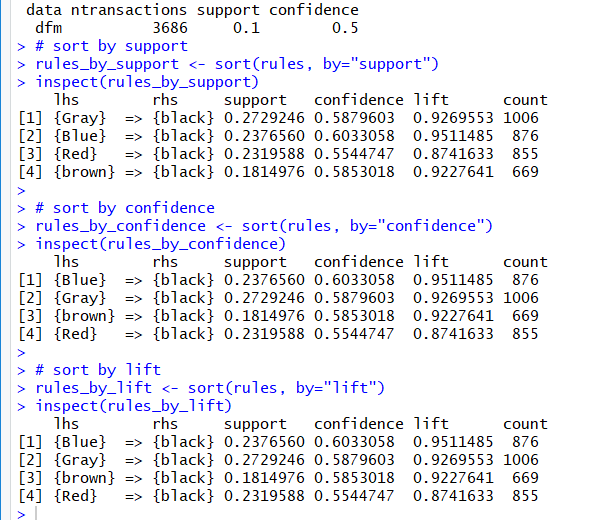
Apriori Algorithm:

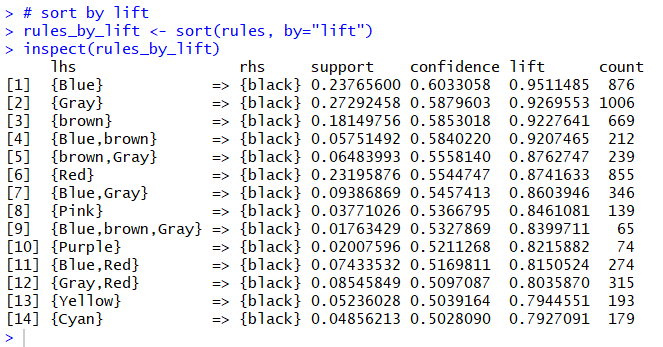
* rules <- apriori(dfm, parameter = list(supp = 0.1, conf = 0.5, minlen = 2))

when increasing support and confidence it results in all rules implies black, and this is logical as black is the most frequent color in the dataset



* rules <- apriori(dfm, parameter = list(supp = 0.01, conf = 0.5, minlen = 2))

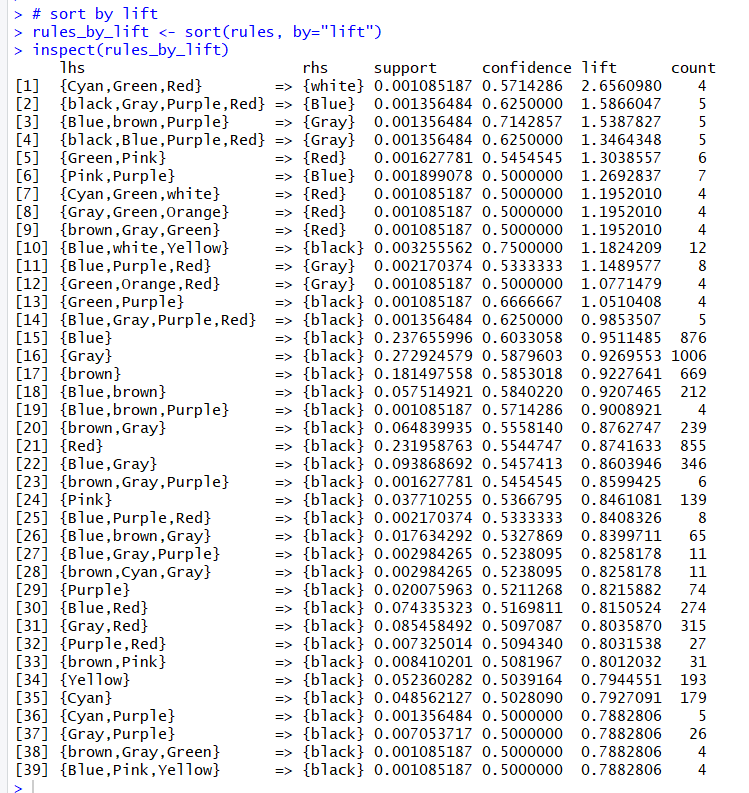
Increased confidence but less support. This will get more colors and more rules as the colors that are found less will be included. But will also implies black as we’re using increased confidence

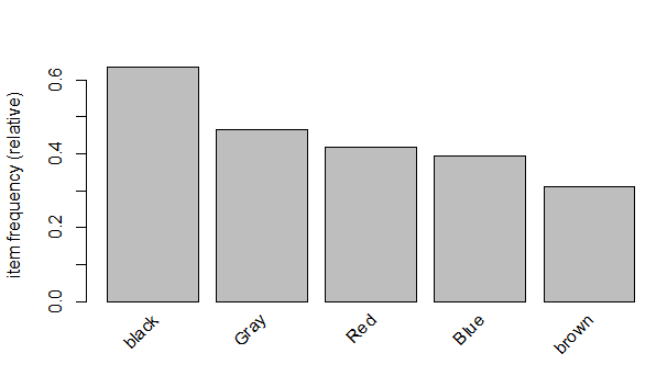


Note that in the 2 above cases, the lift is less than one in all rules and this makes sense. Because implying black happens due to having black as the most frequent item and this means that this is happening by coincendence.

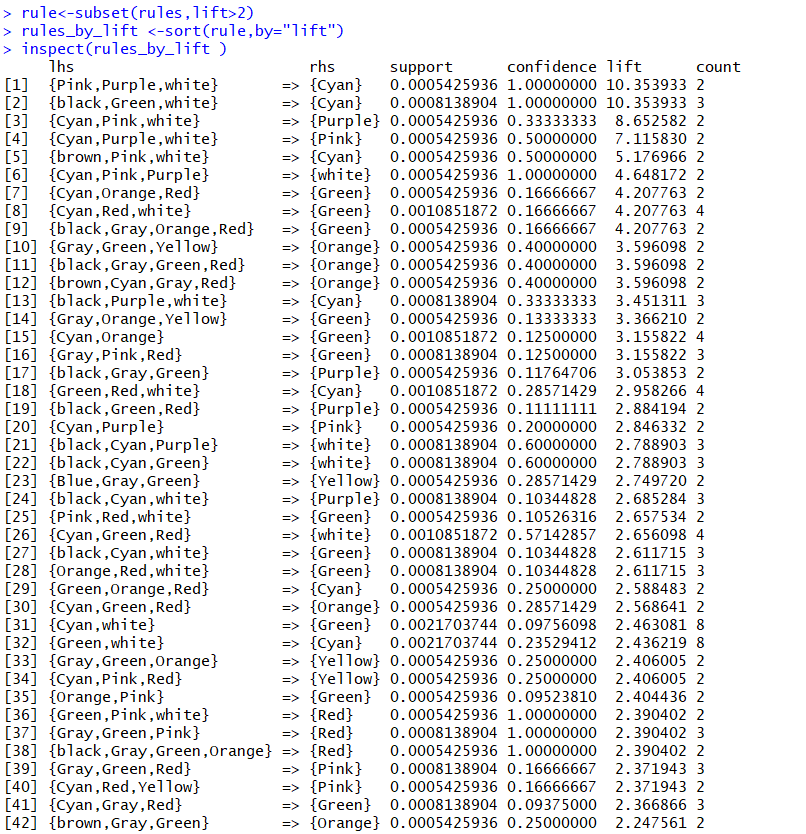
* rules <- apriori(dfm, parameter = list(supp = 0.001, conf = 0.5, minlen = 2, maxlen = 5))

By decreasing the support and leaving the confidence as it is. It begins to show rules that implies colors other than black. These rules have lift bigger than one and this shows that they are positively correlated and correct rules. The black rules’s lift are less than one so they are by coinsendence.

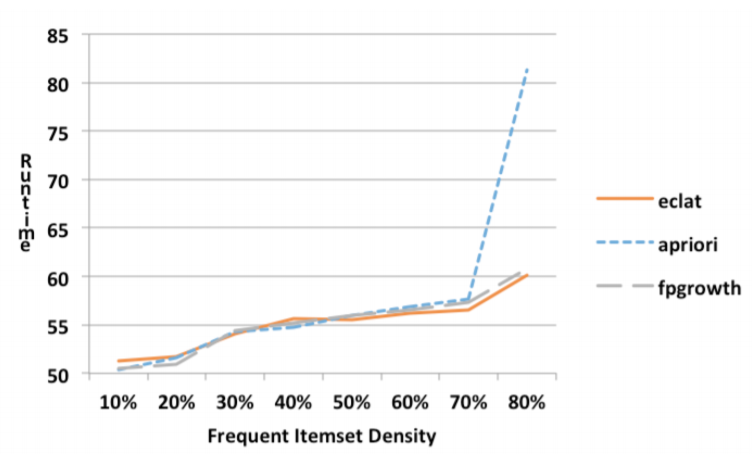


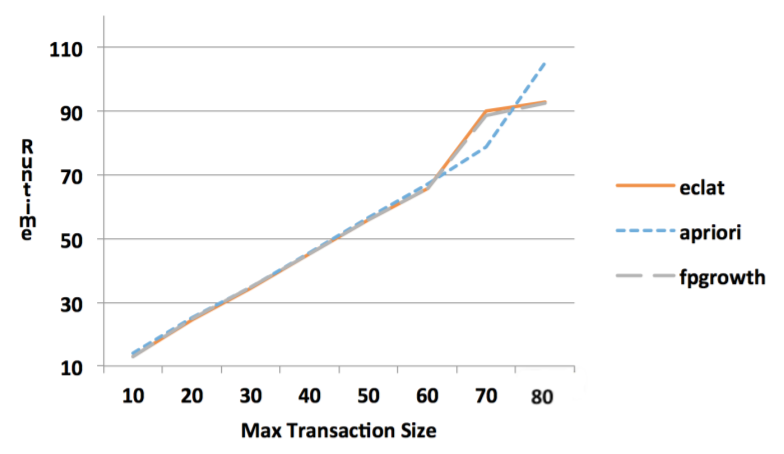


* rules <- apriori(dfm, parameter = list(supp = 0.0005, conf = 0.03, minlen = 2, maxlen = 5))h



**Difference between Apriori, eclat and fp growth:**



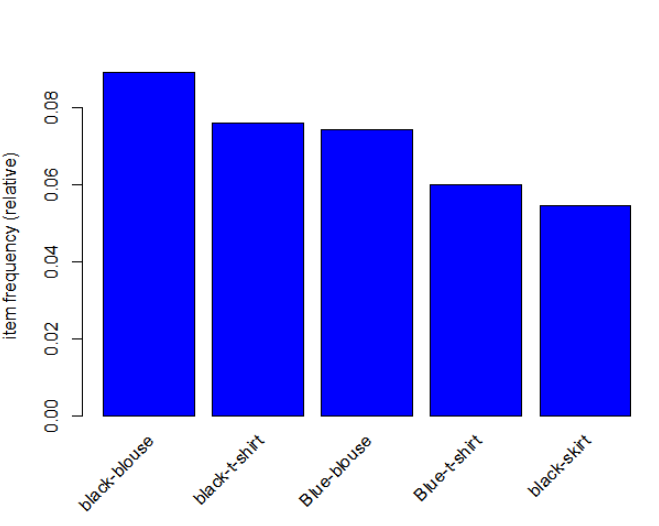


Apriori is an easily understandable frequent itemset mining algorithm. Because of this, Apriori is a popular starting point for frequent itemset study. However, Apriori has serious scalability issues and exhausts available memory much faster than Eclat and FP-Growth. Because of this Apriori should not be used for large datasets. Most frequent itemset applications should consider using either FP-Growth or Eclat.

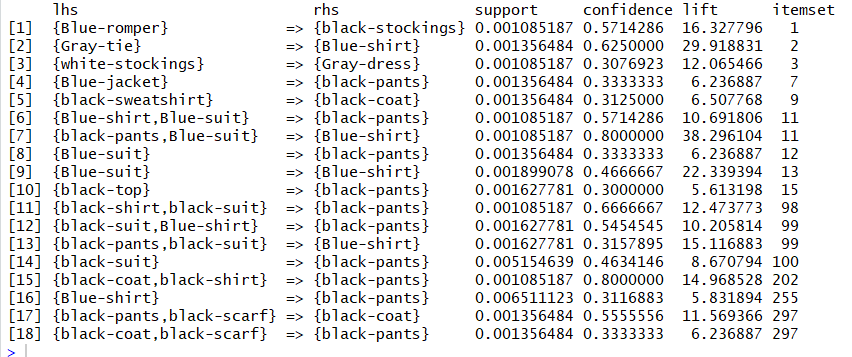
Final:

**Clothes Only (Color-Piece)**

Most Frequent Color-Piece:



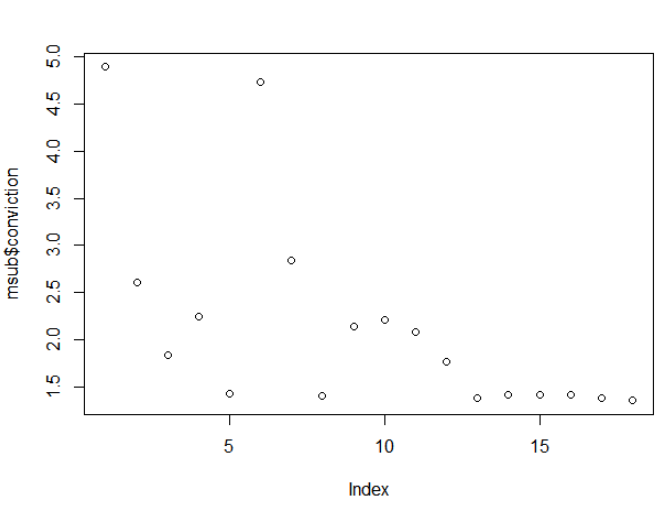
Rules generated using Eclat Algorithm with minimum support = 0.001 and minimum confidence = 0.3



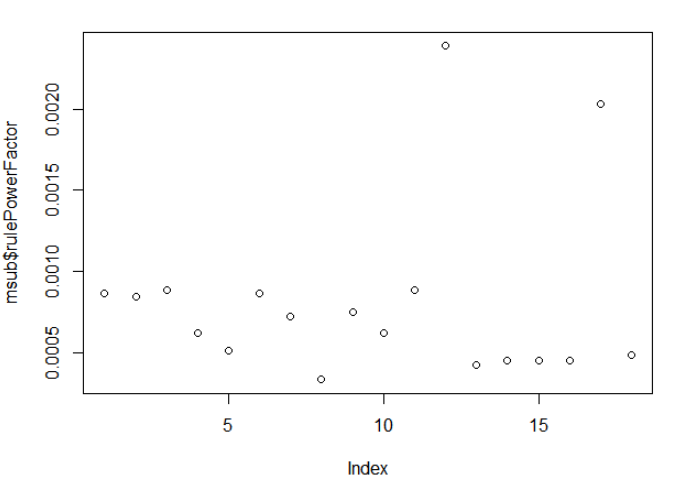
* Rules with lift near to 1 (By coincidence) from the above rules:

None

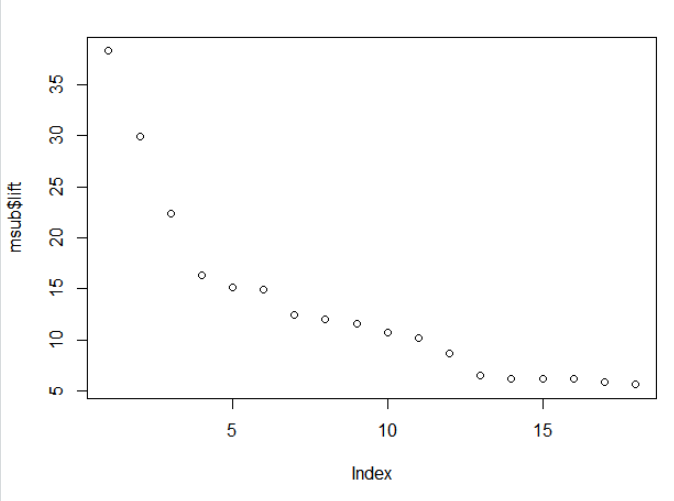
* Plotting rules with conviction:



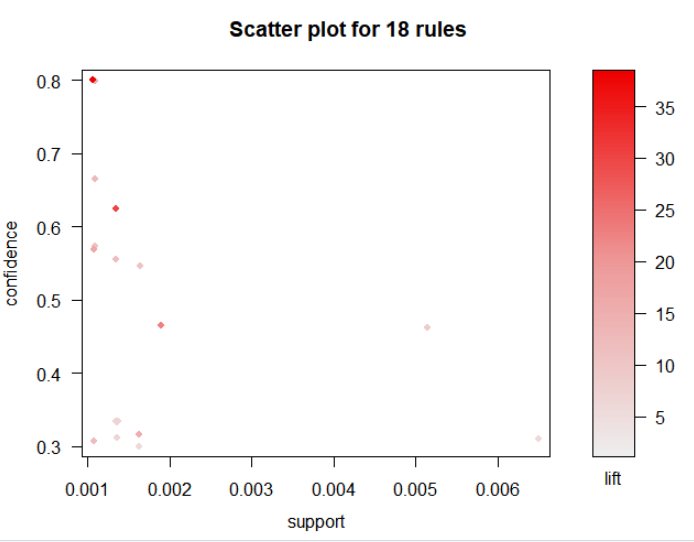
* Plotting rules with RPF:



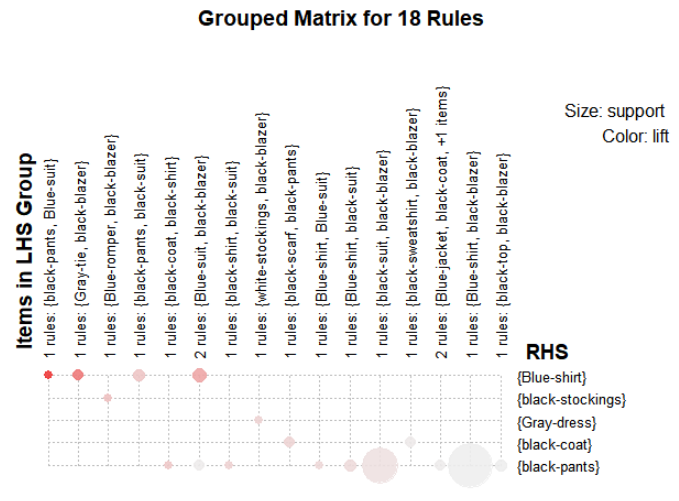
* Plotting Rules with Lift:



* Scatter Plot showing the Support, Confidence and Lift of the Rules:



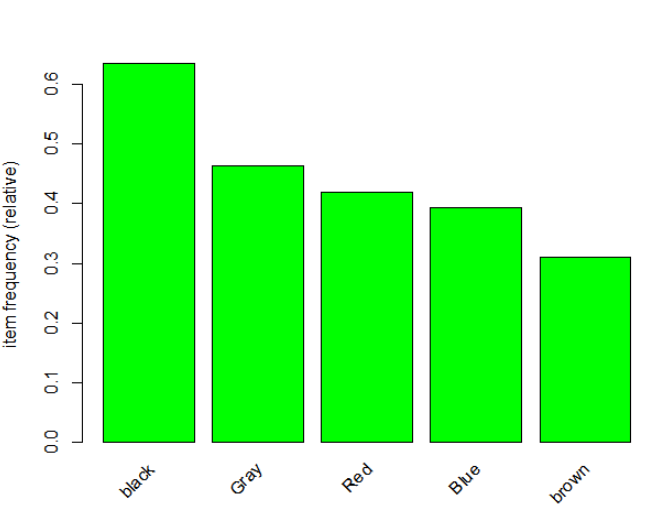
* Grouped Matrix plot which shows the rules itself with the lift and support to make it more clear



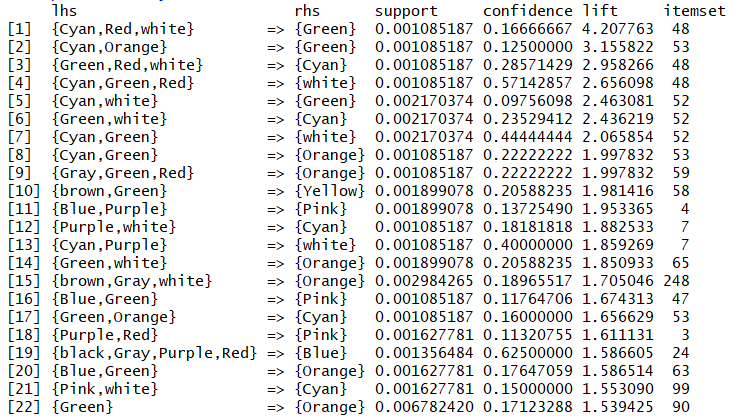
**Colors Analysis Only:**

**Colors Analysis on All Pieces:**

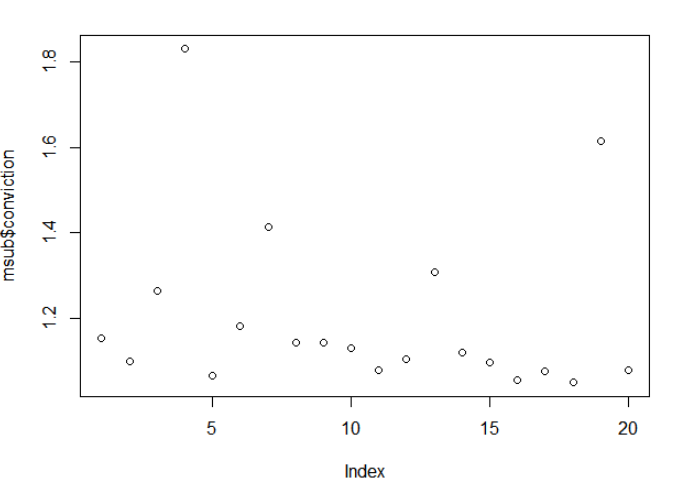
* Most Frequent Colors



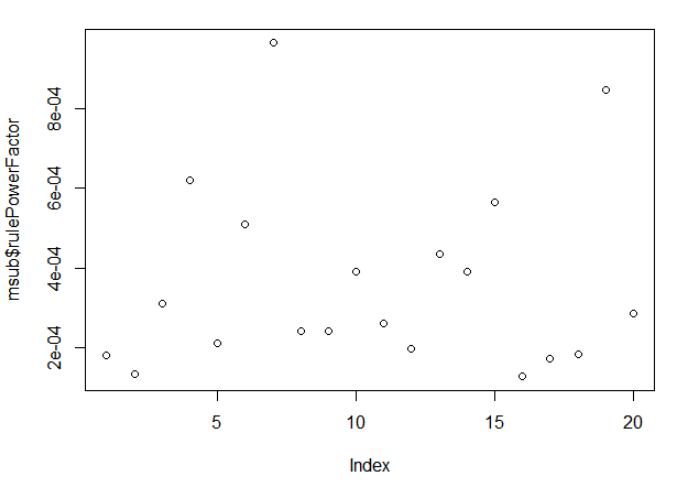
Some of the colors rules which are having lift more than 1.5:



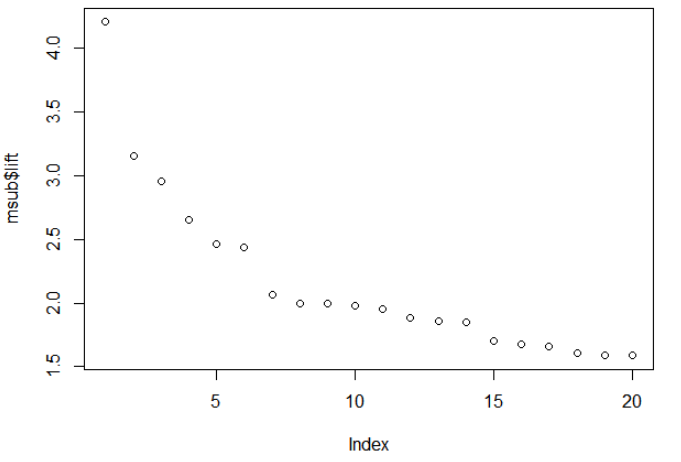
* Plotting Rules with Conviction



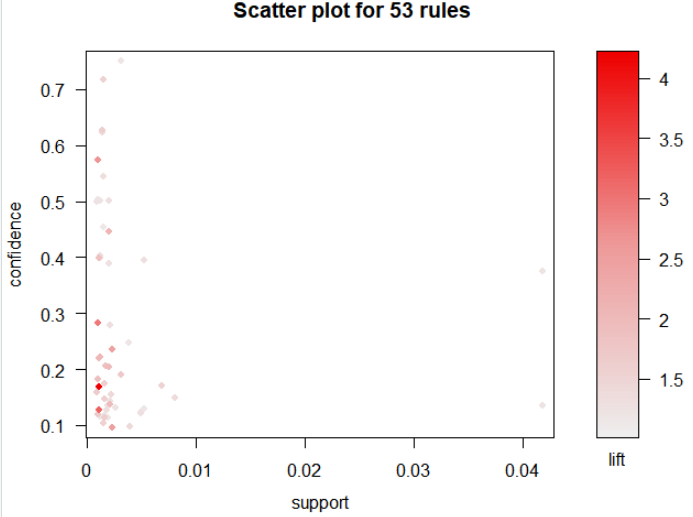
* Plotting Rules with RPF



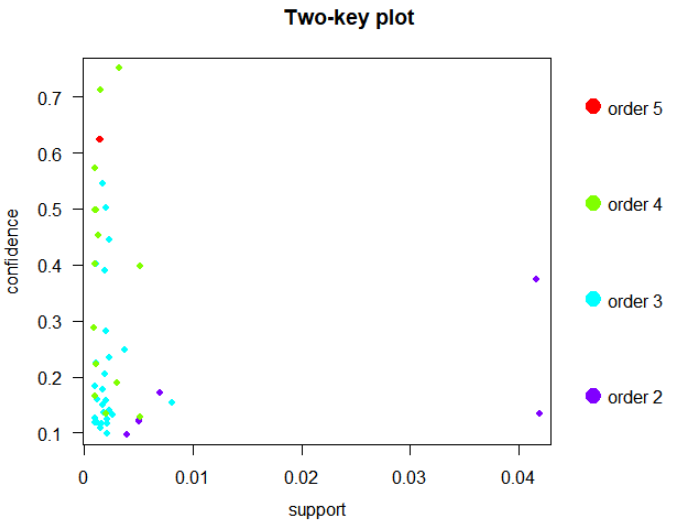
* Plotting Rules with Lift

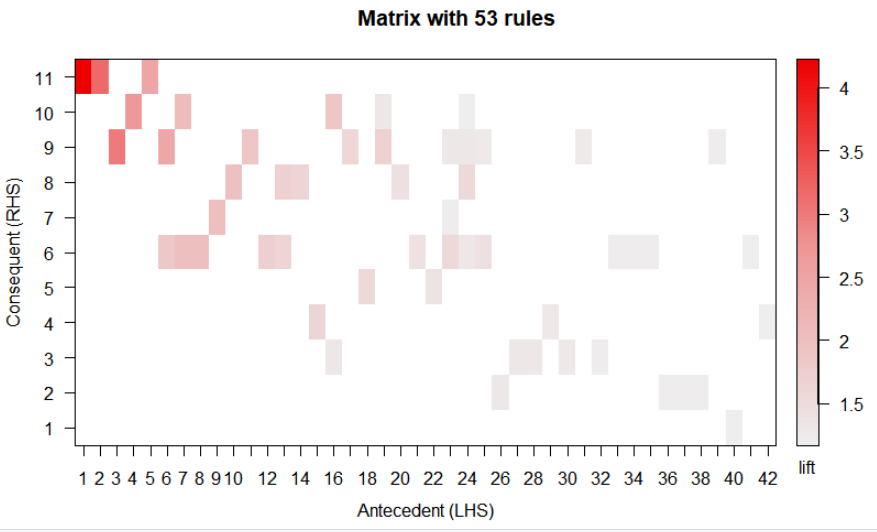


Scatter Plot:

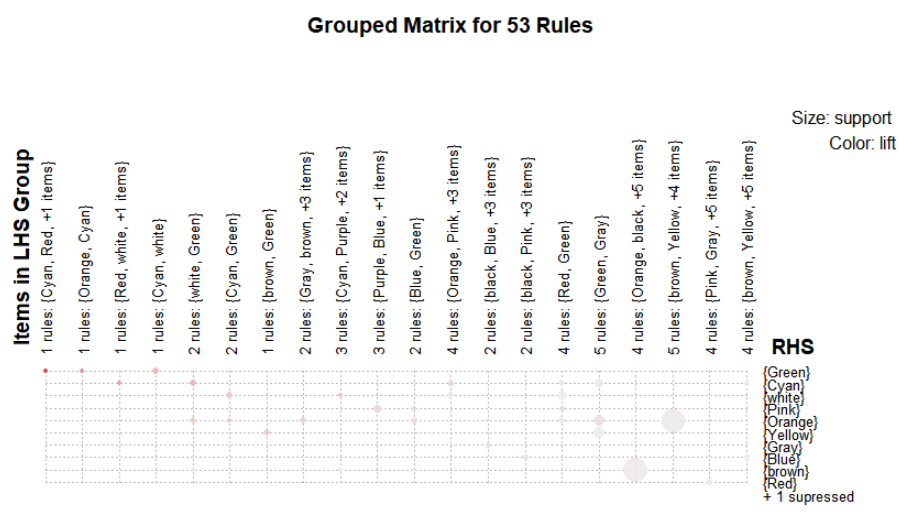


Two-Key Point:

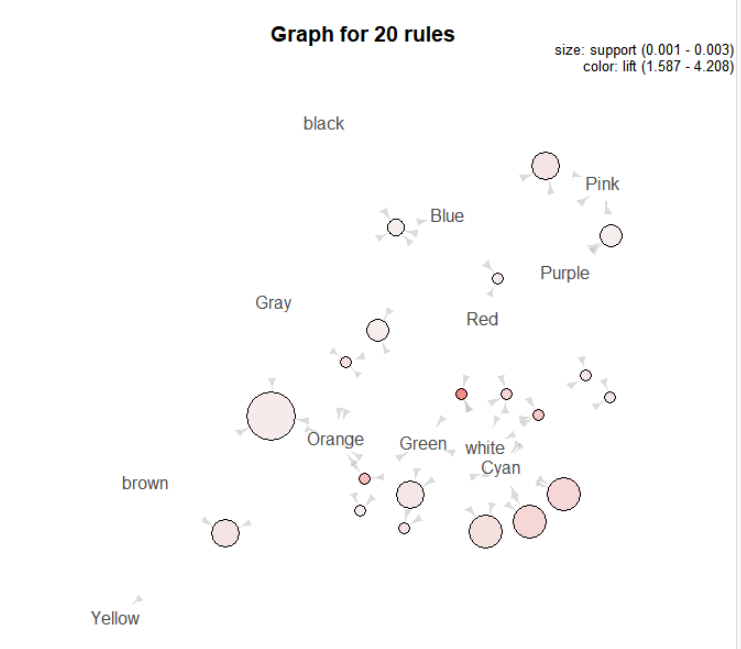




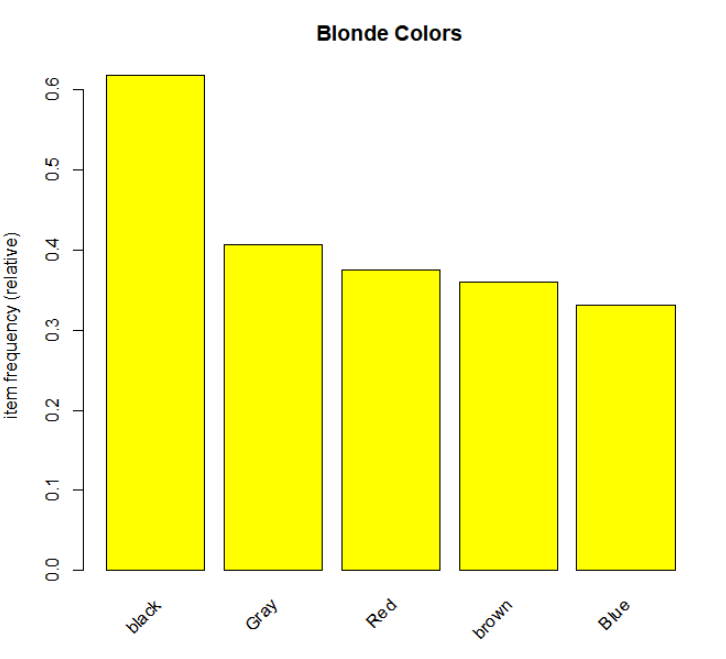
Grouped Matrix:



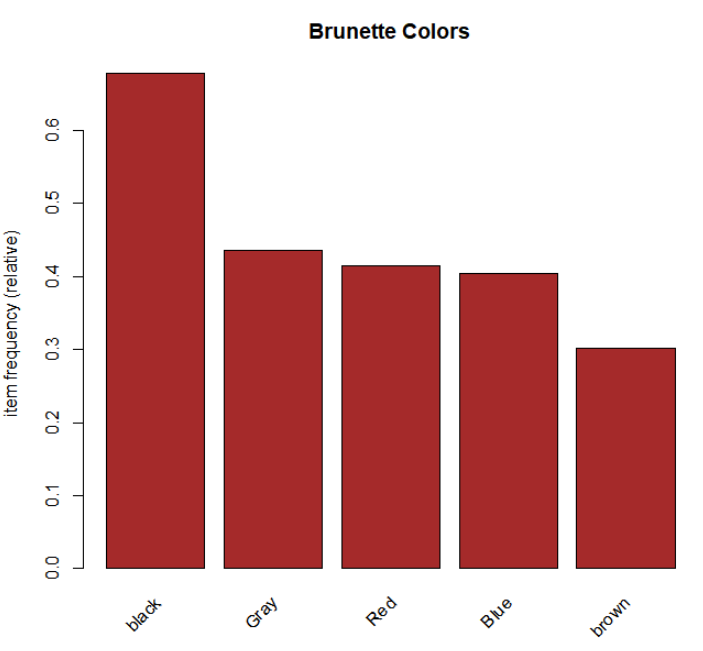
Graph for 20 Rules:



Blonde Frequent Colors:



Brunette Colors:



References:

[https://en.wikipedia.org/wiki/Association\_rule\_learning](https://en.wikipedia.org/wiki/Association_rule_learning#Conviction)

<https://en.wikipedia.org/wiki/Unsupervised_learning>

<https://cran.r-project.org/web/packages/arulesViz/vignettes/arulesViz.pdf>

<https://rdrr.io/cran/arules/man/interestMeasure.html>

<https://arxiv.org/pdf/1701.09042.pdf>