1. (5) What does RFM stand for? \_\_\_Recency, Frequency, Monetary\_\_\_\_\_\_\_\_\_\_\_\_\_
2. (30)

|  |  |
| --- | --- |
|  | Given these 5 receipts, calculate the following for Coke => Milk |

* 1. Support

2/5

* 1. Confidence

1/2

* 1. Lift

1. (10) What is the name of the R package to perform Market Basket Analysis? \_\_\_\_arules\_\_\_\_\_
2. (10) If a grocery item has a support of 0.002, what does this mean?

It means that this product is relatively less preferred than other products. In other words, 2 out of 1000 purchases have this item.

1. (20) Give an example application Market Basket Analysis that does not appear in the classroom slides?

Web page mining

1. (30) Explain in detail what the following code does:

retail %>%

group\_by(InvoiceNo) %>%

summarize(n\_items = n()) %>%

ggplot(aes(x=n\_items))+

geom\_histogram(fill="indianred", bins = 100000) +

coord\_cartesian(xlim=c(0,80))

We are trying to draw the histogram of the number of items in an order.