Activities in a Store

Gabriel Martins Dias

07/05/2015

# Summary

In this report we will study the activities in a store.

# Data Cleaning

First, we clean the data using the following code:

ds1.json <- lapply(ds1.json, function(list.item){   
 data.frame(colour.name = list.item$colour$name,   
 colour.code = list.item$colour$code)  
 } )  
ds1.df <- do.call(rbind, ds1.json)  
  
# add a new column  
ds1.df$productid <- ds1$productid  
ds1.df$skuid <- ds1$skuid  
  
rownames(ds1.df) <- 1:nrow(ds1)  
ds1 <- ds1.df

The cleaned data is the following set of rows:

## colour.name colour.code productid skuid  
## 1 NEGRO #2c2b31 6103880034 8433815102928  
## 2 AVELLANA #E1D8D1 6103870033 8433815103376  
## 3 AVELLANA #E1D8D1 1813659542 8433815080622  
## 4 VERDE MANZANA #65db80 1333989684 8433815097019  
## 5 MARENGO #595f6b 1783619976 8433815103338  
## 6 CARMIN #fe6557 1333989684 8433815097064  
## 7 NEGRO #2c2b31 6103940009 8433815102881  
## 8 ROJO #ff0040 6103830027 8433815103581  
## 9 VIOLETA #B072ED 6470618819 8433815111081  
## 10 ROJO #ff0040 6412780171 8433815112354  
## 11 PÚRPURA #db86dd 5860669440 8433815090973  
## 12 AVELLANA #E1D8D1 1527589489 8433815074980  
## 13 CARMIN #fe6557 1290599722 8433815098610  
## 14 CARMIN #fe6557 1208948963 8433815081490  
## 15 ROJO #ff0040 6103840028 8433815103611  
## 16 MARRÓN #73605C 6103890035 8433815103352  
## 17 MARRÓN #73605C 6412790172 8433815112316  
## 18 PLATA #F0F0F2 6412849830 8433815115690  
## 19 NEGRO #2c2b31 1333969688 8433815096890  
## 20 AZUL MARINO #363D5A 1527638952 8433815082312

# Which Product is the Most Wanted?

To calculate this we used the following code:

# the number of sales per product  
tries <- tapply(ds1$skuid, ds1$productid, length)  
  
# the product with more tries  
tries[which(tries == max(tries))]

## 5860769906   
## 35

The most tried product is 35.

# Which Combination is the Most Tried?

To calculate this we used the following code:

# number of sales per product and color  
tries <- tapply(ds1$skuid, ds1[, c("productid", "colour.code")], length)  
  
# the product with more sales  
ind <- which(tries == max(tries, na.rm = TRUE), arr.ind=TRUE)

One of the most tried combinations is the product 1018109000 with color #FF00FF.