

Trabajo segmentación

Yubar Daniel Marín

Yosel

Mateo

Brayan

16 de mayo de 2018

Introducción

Debido a la heterogeneidad del mercado, dirigir estrategias de marketing generales para todos los clientes se ha convertido en algo inútil; por esta razón a surgido la necesidad de dividir el mercado en grupos cuyos integrantes tengan ciertas características que los asemejen y permitan a la empresa implementar estrategias de mercadotecnia que minimizan costos y que obtener un resultado más satisfactorio.

Método

Para realizar la segmentación se implementará el algoritmo *K-means* el cual es un método de clustering que pertenece a la categoría de aprendizaje no supervisado. Lo que se busca con este método es crear grupos de observaciones con características similares, y así maximizar la variación entre grupos y minimizarla dentro de cada grupo.

Se selecciono este método de clasificación ya que es rápido y con almacenamiento reducido ya que solo se necesitan guardar los K centroides seleccionados.

Procedimiento

Escalar los datos

Lo primero que debemos hacer es escalar las variables con el fin de evitar que algunas tengan más peso que las demás. En este caso utilizaremos el método *Max-Min* representado matemáticamente de la siguiente forma:

$$x'_1 = (max_{objetivo} - min_{objetivo}) * \left[\frac{x_i - min_{valor}}{max_{valor} - min_{valor}} \right] + min_{objetivo}$$

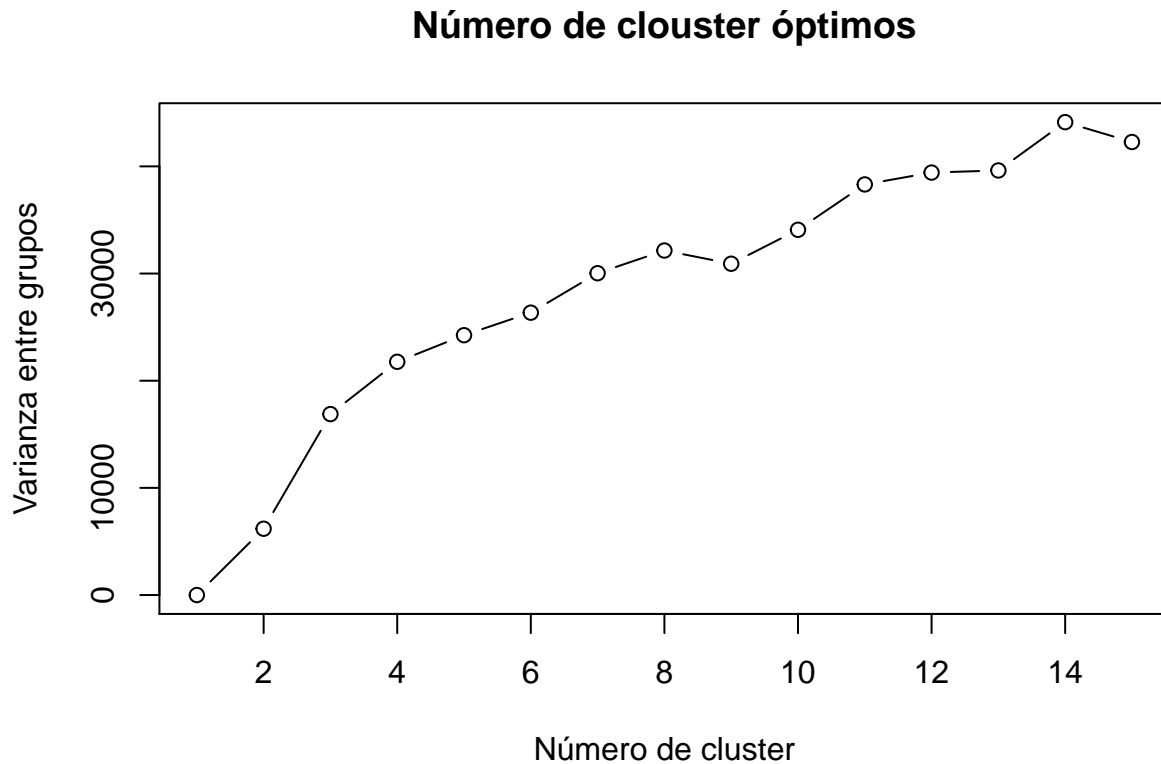
- x'_1 = Entrada normalizada de índice i.
- x_i = Entrada de índice i.
- $max_{objetivo}$ = Valor máximo del rango en el que quiero aplicar las entradas
- $min_{objetivo}$ = Valor mínimo del rango en el que quiero aplicar las entradas.
- max_{valor} = Valor máximo real de las entradas.
- min_{valor} = valor mínimo real de las entradas.

Determinación del número de clúster óptimo

Análisis de la varianza entre grupos

Debemos asignar el número de óptimo clúster para la segmentación por lo que debemos probar desde el 1 hasta el número de clúster que consideramos como un número máximo a permitir en nuestro caso es de 1 a 15, siendo el número máximo de grupos que permitiremos.

En el siguiente grafica podemos observar la relación entre el número de clúster y el valor de la suma de cuadrados Inter grupos teniendo siempre presente que se prefiere una varianza mayor entre los grupos.



La grafica anterior nos permite reconocer puntos de corte, debemos recordar que se prefieren número de clúster con mayores varianzas entre grupos, en este caso un buen punto de corte puede ser 8, ya que se puede observar un decaimiento entre este punto y el siguiente.

Sabemos que el método *K-means* asigna de forma aleatoria los puntos que serán tomadas como las medias de cada subgrupo por lo que cada que se ejecute el siguiente código el cual permite generar las grafica anterior obtendremos un resultado diferente.

```
set.seed(754)
# Verificación de la varianza entre grupos posibles de 1 a 15 clouster
Varianza.en.grupo <- kmeans(escala,centers = 1)$betweenss
for (i in 2:15) entre[i] <- kmeans(escala,centers = i)$betweenss

#grafica de la varianza entre clouster
plot(1:15,entre,type = "b",xlab = "Número de cluster",
     ylab = "Varianza entre grupos", main = "Número de clouster óptimos")
```

El anterior código fue ejecutado 52 veces y se obtuvieron los siguientes resultados.

Número de clouster	Total de apariciones
6	2

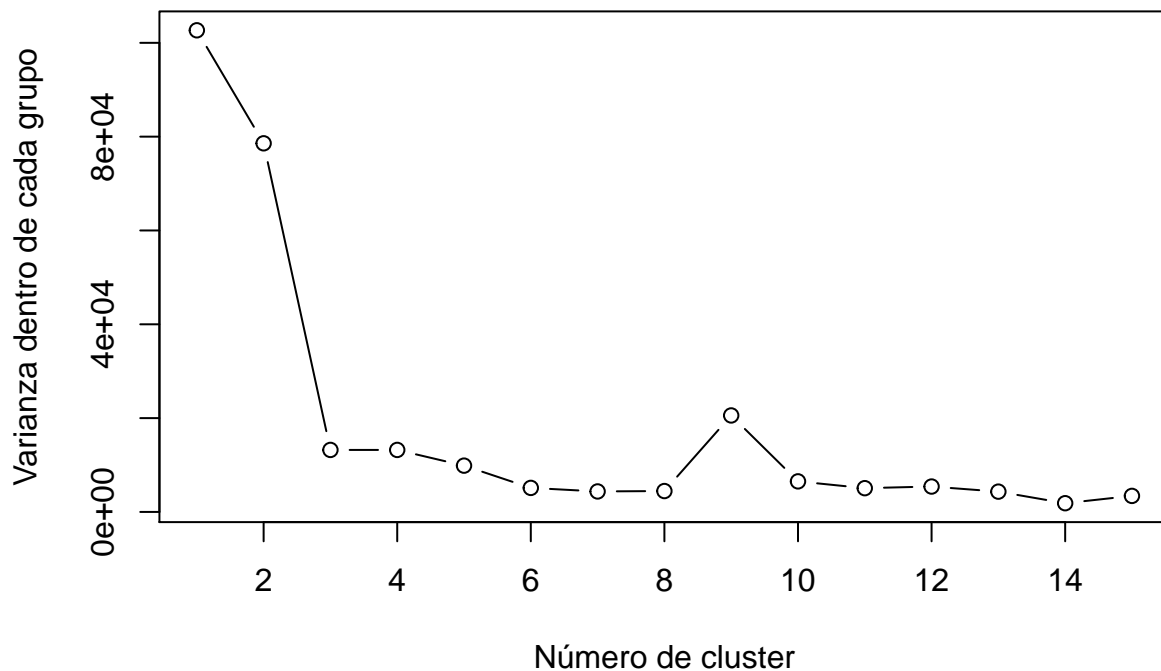
Número de clouster	Total de apariciones
7	6
8	10
9	10
10	9
11	5
12	4
13	3
14	2

De la anterior tabla podemos observar que el número de veces que se seleccionaron 8 clúster y 9 clúster como número es óptimo es de 10, otra posible selección será 10 clúster con un total de aprisiones de 9.

Análisis de la varianza dentro de cada grupo

Ahora analizares el grafico recursivo de la varianza entre los elementos de cada grupo. En este caso utilizaremos un código similar al anterior, la única diferencia es que se remplaza `betienses` por `withinss`

Número de clouster óptimos



La idea es tener un criterio de selección de numero de clúster óptimo, debemos recordar que se prefieren un numero de clúster en el cual la varianza dentro del grupo sea menor, en la gráfica anterior se puede observar que hay un incremento obvio en la varianza tenemos 8 grupos y 9 clúster.

Así por el criterio anterior se selección como número óptimo de clúster 8 ya que se observar un comportamiento decreciente para la varianza cuando tenemos un número inferior de clúster.

Finalmente se debe aclarar que para un número inferior al seleccionado la varianza es decreciente, pero por el criterio de la primera grafica se rechazaron y para un número superior de clúster como 9 y 10 aunque podrían

ser aceptados, la varianza en las gráficas probadas con el criterio anterior nunca llegó a ser decrecientes hasta estos cortes.

Creación de los subgrupos

Las medidas obtenidas con la técnica *K-means* con un total de 8 grupos son:

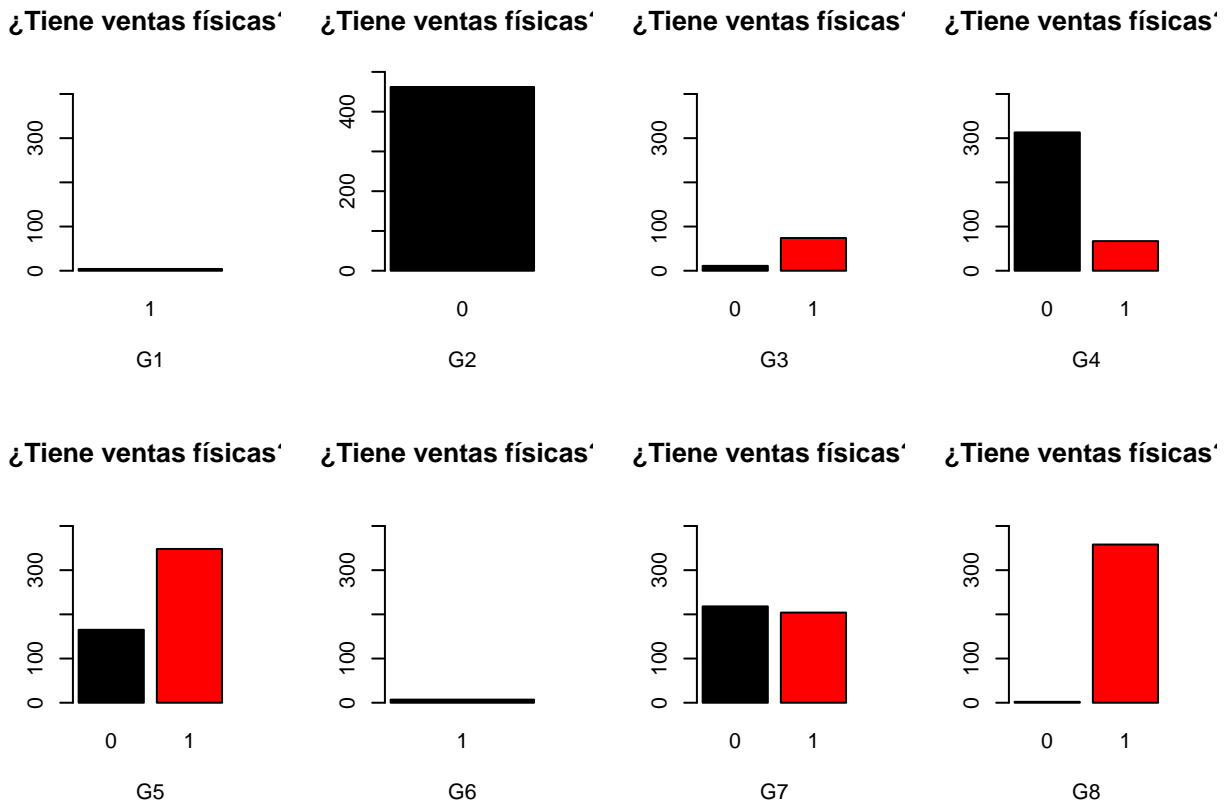
- La inercia total es de 102672
- La inercia entre grupos es de 31989.67
- la inercia dentro cada uno de los 8 grupos es de :

1. 881.617
2. 5117.989
3. 16094.385
4. 10536.707
5. 9616.878
6. 8665.118
7. 7700.730
8. 7068.905

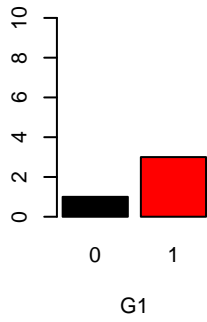
Utilidad de la segmentación para la empresa

Ahora se presenta la utilidad del procedimiento de segmentación propuesto, para esto presentan graficas que permiten diferenciar por colores los elementos de cada uno de los grupos.

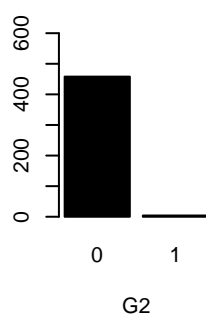
```
seg <- data.frame(base[1],clu8$cluster)
```



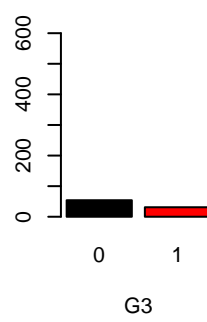
Ventas electronicas



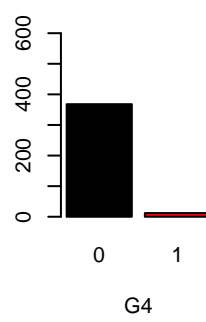
Ventas electronicas



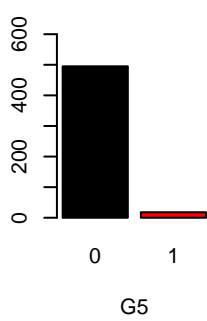
Ventas electronicas



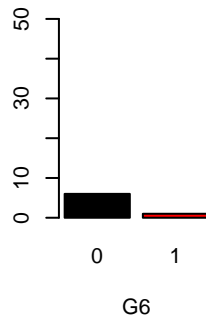
Ventas electronicas



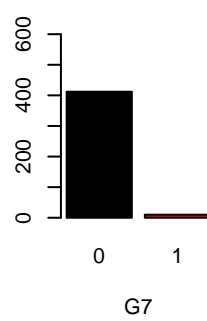
Ventas electronicas



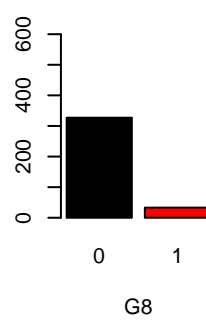
Ventas electronicas



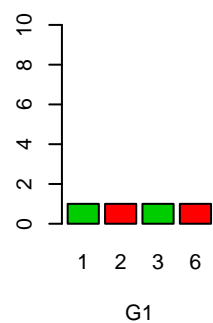
Ventas electronicas



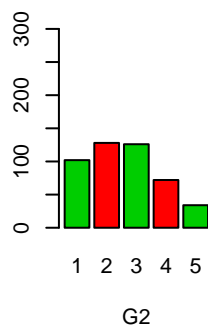
Ventas electronicas



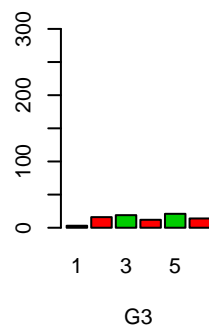
Ciclo financiero



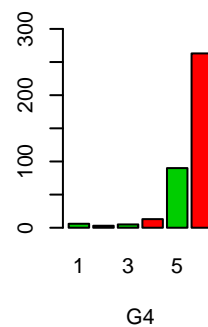
Ciclo finanziario



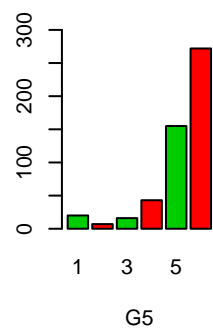
Ciclo finanziario



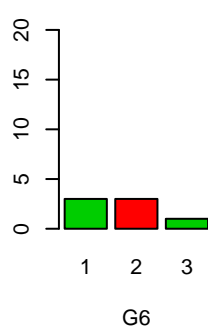
Ciclo finanziario



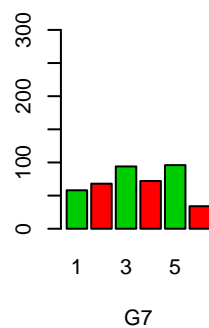
Ciclo finanziario



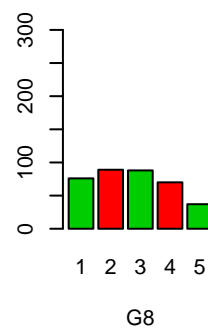
Ciclo finanziario

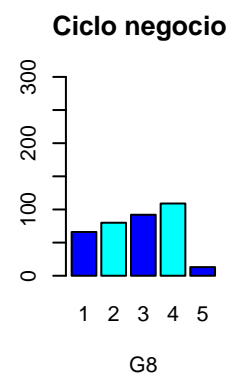
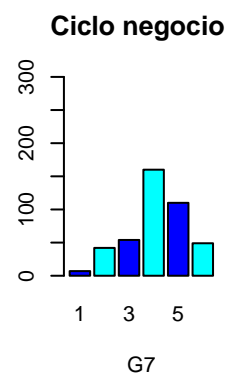
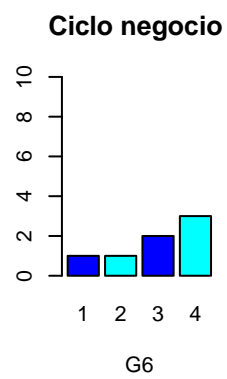
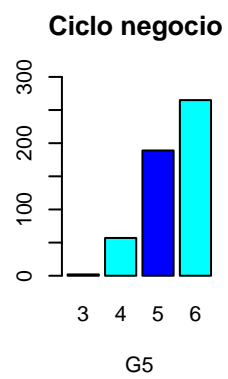
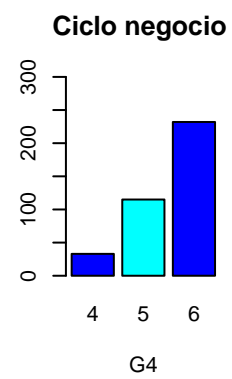
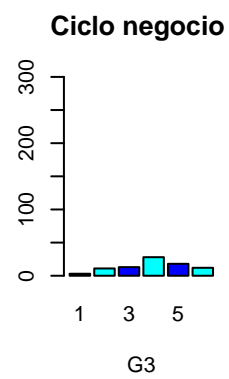
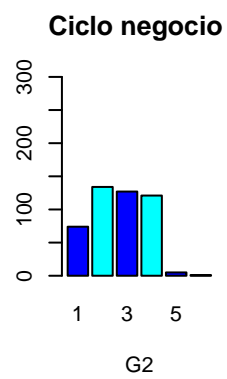
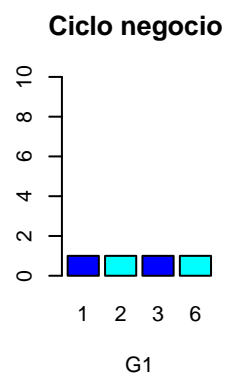


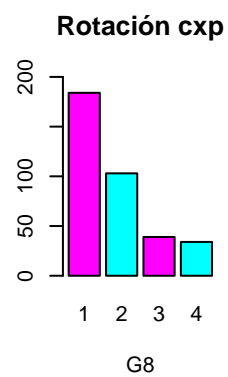
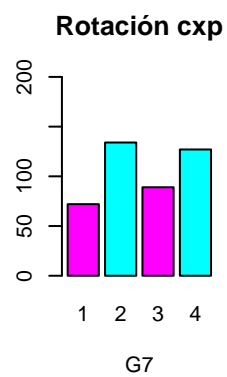
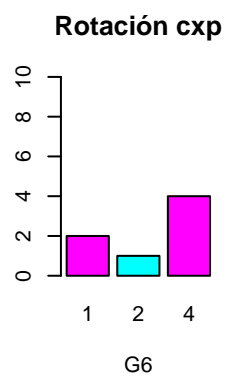
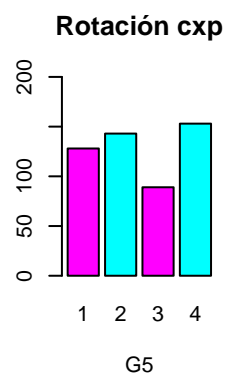
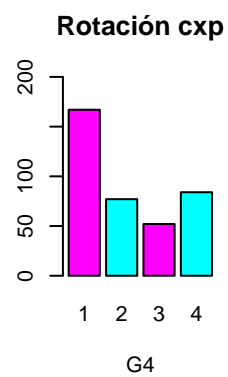
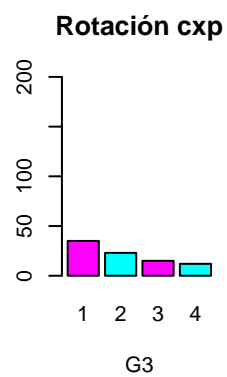
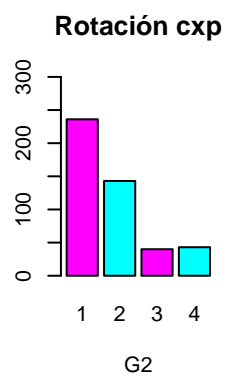
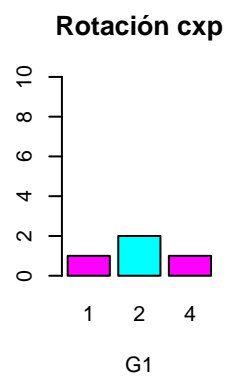
Ciclo finanziario

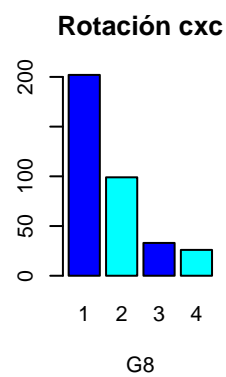
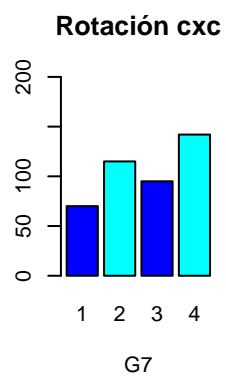
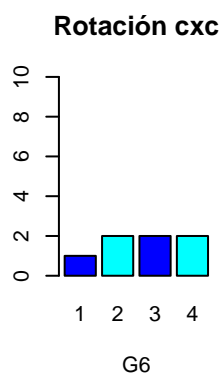
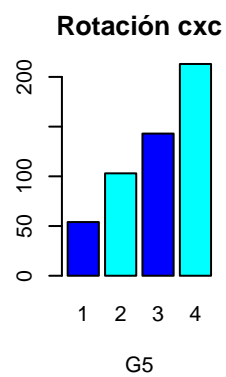
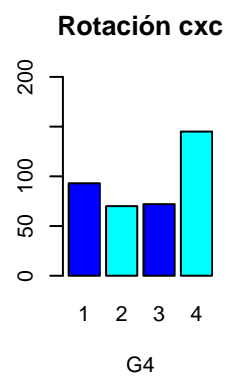
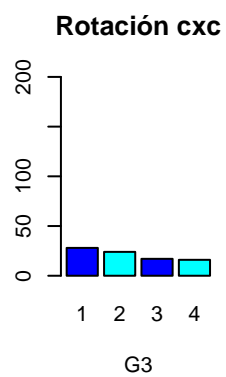
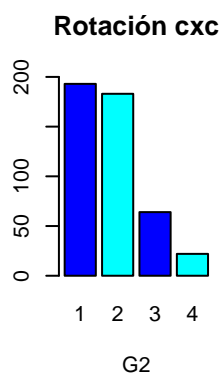
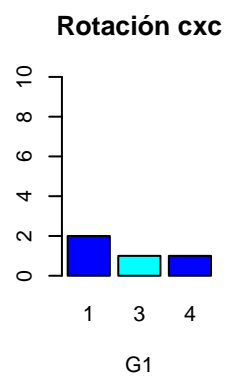


Ciclo finanziario

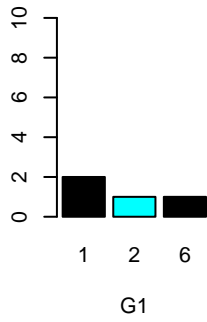




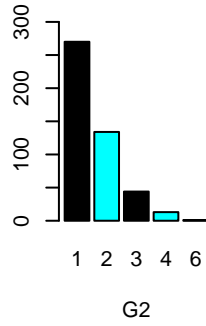




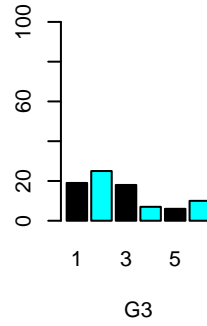
Rotación inventario



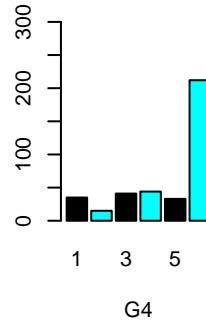
Rotación inventario



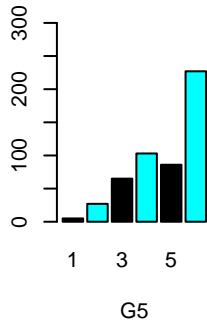
Rotación inventario



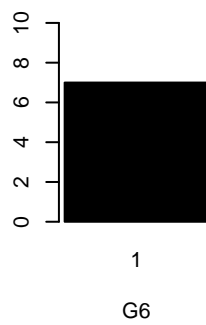
Rotación inventario



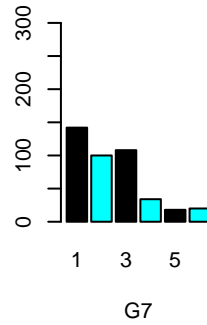
Rotación inventario



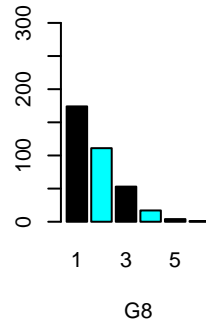
Rotación inventario



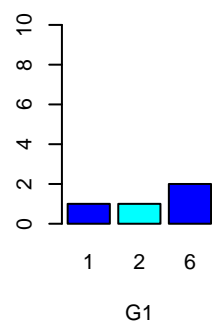
Rotación inventario



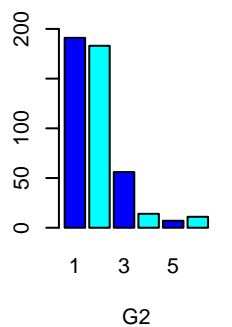
Rotación inventario



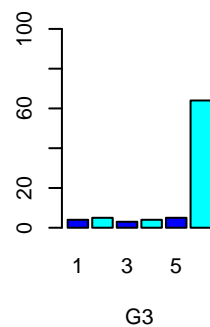
Total inventario



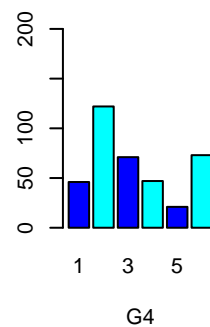
Total inventario



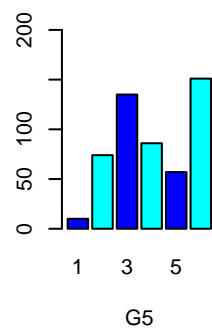
Total inventario



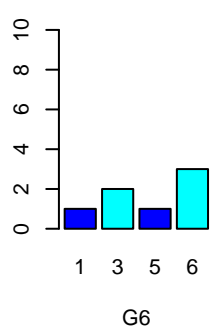
Total inventario



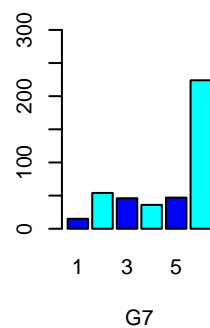
Total inventario



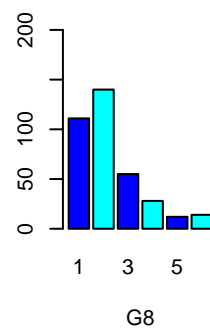
Total inventario

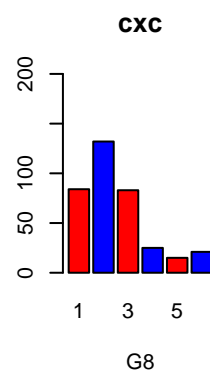
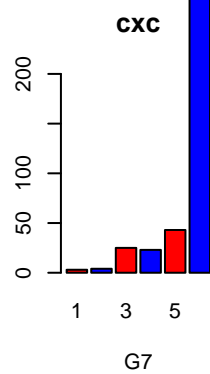
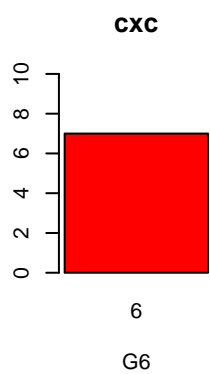
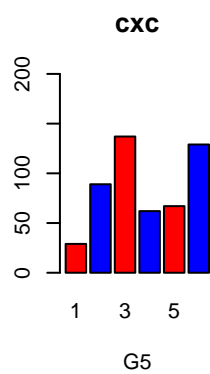
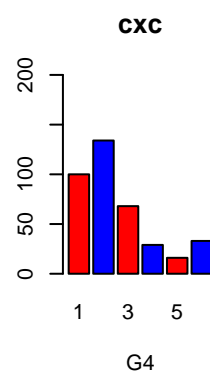
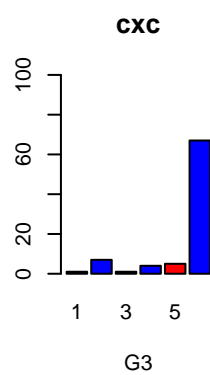
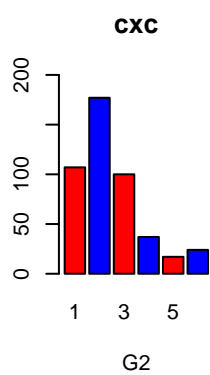
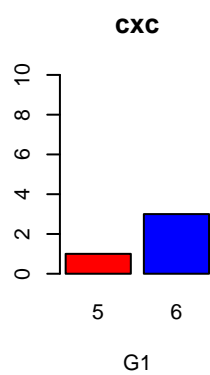


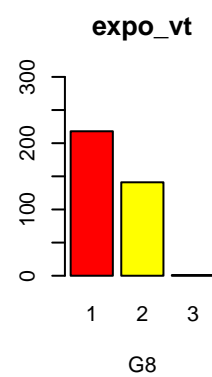
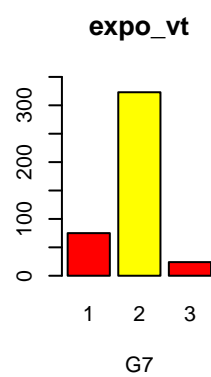
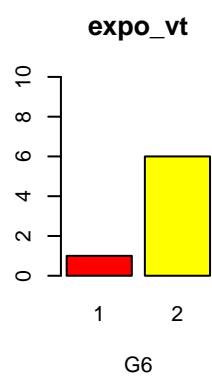
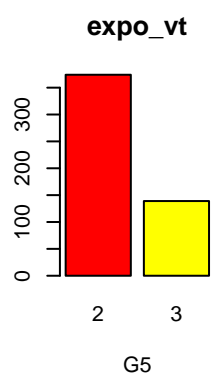
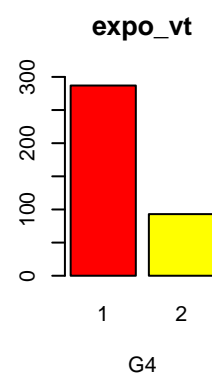
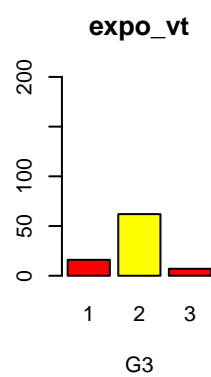
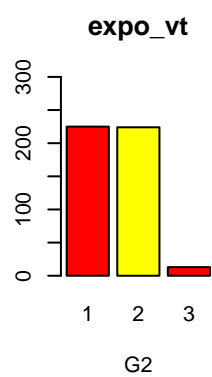
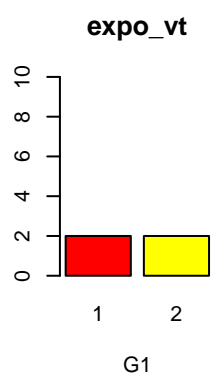
Total inventario

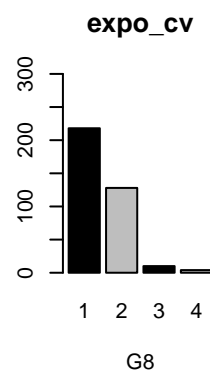
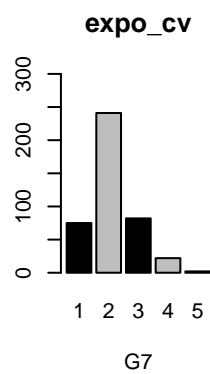
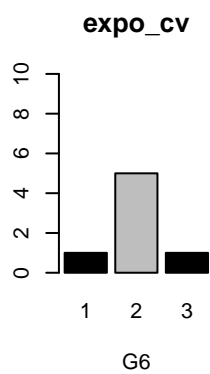
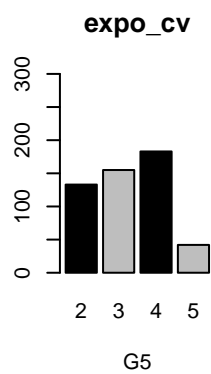
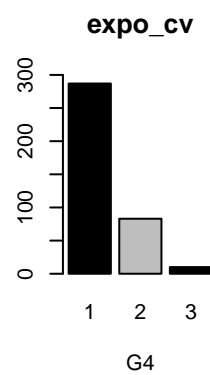
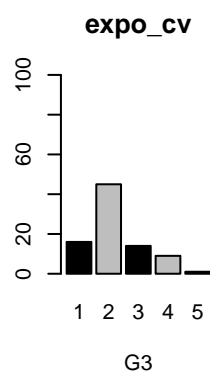
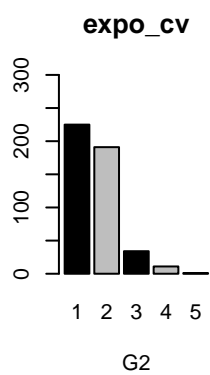
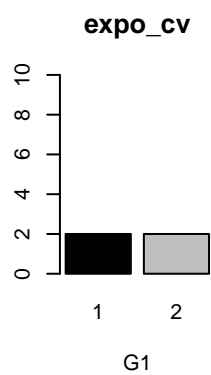


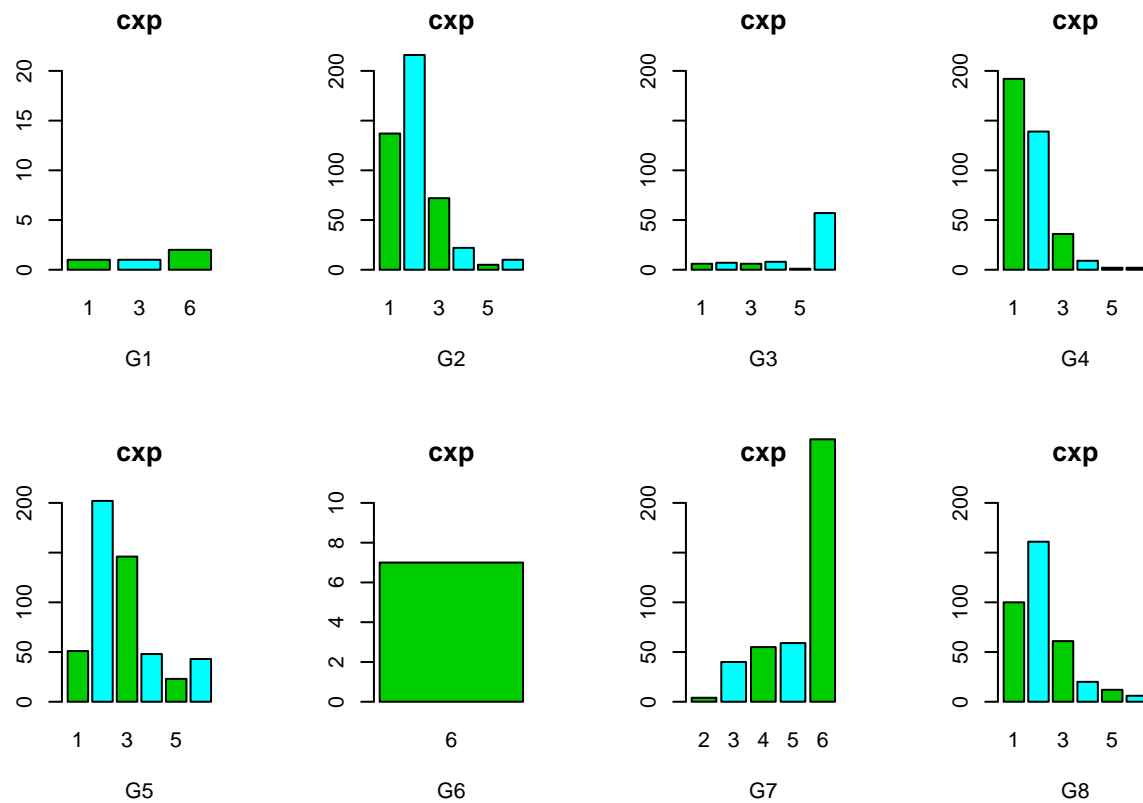
Total inventario











```
##          promedio.pagos_pn promedio.pagos_pj
## clouster1      0.7125000      0.2875000
## clouster2      0.4226537      0.5600390
## clouster3      0.6528941      0.3471059
## clouster4      0.3461211      0.6065263
## clouster5      0.3841618      0.6139006
## clouster6      0.6035714      0.3964286
## clouster7      0.5342536      0.4254668
## clouster8      0.4777278      0.5195028
```

```
##          promedio.recaudos_pn promedio.recaudos_pj
## clouster1      0.715500000      0.034500000
## clouster2      0.003006494      0.001322511
## clouster3      0.214435294      0.185576471
## clouster4      0.046363158      0.003636842
## clouster5      0.032465887      0.024064327
## clouster6      0.769571429      0.087571429
## clouster7      0.089381517      0.033841232
## clouster8      0.103000000      0.033111111
```

```
##          Promedio_en_canales_y_servicios
## clouster1      70281348814
## clouster2      133982345
## clouster3      4747157848
## clouster4      119640853
## clouster5      182533666
```

## clouster6	8611396796	
## clouster7	841946156	
## clouster8	208754394	
##	media.productos.grupo1	media.productos.grupo2
## en_vm_canal1	1.189040e+11	5.100066e+08
## en_vm_canal2	2.318742e+11	1.524908e+09
## en_vm_canal3	3.391351e+10	6.028937e+07
## en_vm_canal4	1.027012e+11	0.000000e+00
## en_vm_canal5	3.977824e+11	3.061291e+08
## en_vm_canal6	5.352109e+07	1.989460e+07
## en_vm_canal7	2.495363e+11	3.270518e+07
## en_vm_canal8	2.555703e+08	5.954614e+06
## en_vm_canal9	2.002516e+09	3.903916e+06
## en_vm_canal10	7.500000e+04	1.328544e+04
## en_vm_otros	2.197480e+10	1.871412e+06
## en_tx_canal1	4.490352e+04	8.550034e+01
## en_tx_canal2	1.027687e+03	3.588825e+01
## en_tx_canal3	1.238180e+05	6.510018e+01
## en_tx_canal4	8.788651e+05	0.000000e+00
## en_tx_canal5	5.729167e+01	2.512858e+00
## en_tx_canal6	2.114583e+01	4.519760e+00
## en_tx_canal7	9.131250e+01	6.047652e-01
## en_tx_canal8	4.009708e+02	2.694504e+00
## en_tx_canal9	1.129028e+03	2.806318e+00
## en_tx_canal10	2.500000e-01	1.298701e-02
## en_tx_otros	3.121198e+04	1.381189e+00
## sal_vm_canal5	3.998111e+11	2.796549e+08
## sal_vm_canal2	5.496288e+11	1.273052e+09
## sal_vm_canal8	1.507391e+06	6.591834e+05
## sal_vm_otros	0.000000e+00	4.276255e+05
## sal_tx_canal5	5.777083e+01	2.576953e+00
## sal_tx_canal2	1.633552e+04	4.624917e+02
## sal_tx_canal8	5.000000e-01	5.554654e-01
## sal_tx_otros	0.000000e+00	6.953847e-01

##	media.productos.grupo3	media.productos.grupo4
## en_vm_canal1	1.377389e+10	3.880588e+08
## en_vm_canal2	3.814404e+10	1.103101e+09
## en_vm_canal3	8.713120e+08	8.093572e+06
## en_vm_canal4	3.334103e+09	1.826998e+07
## en_vm_canal5	1.781388e+10	4.332578e+08
## en_vm_canal6	1.420558e+09	4.963423e+07
## en_vm_canal7	9.270971e+09	1.376001e+07
## en_vm_canal8	3.131938e+08	5.399821e+06
## en_vm_canal9	4.241712e+07	3.690027e+06
## en_vm_canal10	1.052247e+07	6.723363e+05
## en_vm_otros	1.538317e+09	3.844710e+06
## en_tx_canal1	2.754944e+03	5.652262e+01
## en_tx_canal2	8.199833e+02	2.046187e+01
## en_tx_canal3	2.749670e+03	7.910478e+00
## en_tx_canal4	9.947437e+03	7.950550e+01
## en_tx_canal5	4.768579e+01	3.294621e+00

## en_tx_canal6	1.445620e+02	4.682444e+00
## en_tx_canal7	5.564902e+01	4.887531e-01
## en_tx_canal8	9.756980e+01	3.449557e+00
## en_tx_canal9	3.407884e+01	2.209599e+00
## en_tx_canal10	1.339599e+00	1.025585e-01
## en_tx_otros	1.780119e+03	1.315217e+00
## sal_vm_canal5	1.797760e+10	4.097589e+08
## sal_vm_canal2	3.757611e+10	1.151306e+09
## sal_vm_canal8	3.205542e+08	1.830010e+05
## sal_vm_otros	7.225371e+06	1.950267e+05
## sal_tx_canal5	4.947431e+01	3.176463e+00
## sal_tx_canal2	5.676065e+03	1.857388e+02
## sal_tx_canal8	5.694543e+01	1.714474e-01
## sal_tx_otros	1.118175e+01	3.605988e-01

##	media.productos.grupo5	media.productos.grupo6
## en_vm_canal1	1.014571e+09	1.535867e+10
## en_vm_canal2	2.000984e+09	7.932408e+10
## en_vm_canal3	1.967738e+07	7.377184e+09
## en_vm_canal4	1.625569e+08	1.855395e+09
## en_vm_canal5	4.755144e+08	1.110358e+10
## en_vm_canal6	8.618395e+07	2.494769e+07
## en_vm_canal7	5.308703e+07	4.060867e+08
## en_vm_canal8	2.129361e+07	2.897248e+08
## en_vm_canal9	8.428669e+06	6.075266e+07
## en_vm_canal10	5.314528e+05	0.000000e+00
## en_vm_otros	1.999586e+06	3.975980e+10
## en_tx_canal1	2.282636e+02	6.814024e+04
## en_tx_canal2	1.219583e+02	4.822976e+02
## en_tx_canal3	2.280323e+01	8.594301e+04
## en_tx_canal4	8.112239e+02	7.572564e+03
## en_tx_canal5	4.260836e+00	5.059524e+00
## en_tx_canal6	2.287878e+01	1.186111e+01
## en_tx_canal7	5.445500e+00	3.892857e+01
## en_tx_canal8	1.104102e+01	2.208468e+03
## en_tx_canal9	6.893018e+00	4.938135e+02
## en_tx_canal10	1.301829e-01	0.000000e+00
## en_tx_otros	5.161738e+00	1.541199e+05
## sal_vm_canal5	4.719541e+08	1.109735e+10
## sal_vm_canal2	1.158552e+09	9.168363e+10
## sal_vm_canal8	4.219654e+05	3.724827e+05
## sal_vm_otros	2.526321e+05	0.000000e+00
## sal_tx_canal5	4.176709e+00	4.226190e+00
## sal_tx_canal2	2.918338e+02	3.076893e+03
## sal_tx_canal8	3.976165e-01	1.904762e-01
## sal_tx_otros	4.633946e-01	0.000000e+00

##	media.productos.grupo7	media.productos.grupo8
## en_vm_canal1	2.786384e+09	8.477394e+08
## en_vm_canal2	9.200367e+09	1.584227e+09
## en_vm_canal3	2.656071e+08	7.688327e+07
## en_vm_canal4	1.902406e+08	2.707033e+08

## en_vm_canal5	1.949597e+09	4.964943e+08
## en_vm_canal6	1.159005e+08	9.291979e+07
## en_vm_canal7	8.793906e+08	4.816786e+08
## en_vm_canal8	1.956137e+07	3.008263e+07
## en_vm_canal9	1.340452e+07	7.299903e+06
## en_vm_canal10	7.254148e+05	1.298077e+06
## en_vm_otros	1.469956e+08	3.277305e+07
## en_tx_canal1	5.856618e+02	2.056289e+02
## en_tx_canal2	1.727861e+02	9.839340e+01
## en_tx_canal3	1.074018e+03	2.772861e+02
## en_tx_canal4	8.787188e+02	1.742831e+03
## en_tx_canal5	5.387800e+00	4.252113e+00
## en_tx_canal6	2.130586e+01	1.426388e+01
## en_tx_canal7	7.101770e+00	1.623199e+01
## en_tx_canal8	9.985665e+00	2.431189e+01
## en_tx_canal9	1.154907e+01	5.968545e+00
## en_tx_canal10	1.297393e-01	2.050926e-01
## en_tx_otros	2.429256e+02	9.667478e+01
## sal_vm_canal5	1.939348e+09	4.845295e+08
## sal_vm_canal2	7.749930e+09	1.853756e+09
## sal_vm_canal8	2.880237e+05	2.018662e+05
## sal_vm_otros	6.404770e+05	2.042610e+06
## sal_tx_canal5	5.174524e+00	4.286869e+00
## sal_tx_canal2	1.216374e+03	5.253885e+02
## sal_tx_canal8	1.943523e-01	1.880688e-01
## sal_tx_otros	3.137543e-01	6.564881e-01