

Análisis exploratorio

Ventas Directas en Centroamérica.

Scentia

Maria Fernanda Rodas
Pablo Viana
José Meneses
Daniel Ixcoy
José Martínez
Pablo Viana
Sergio Marchena

Contenido

1. Estado de los datos.
2. Medidas tendencia central de variables más importantes.
3. Mes con mayores pedidos registrados.
4. Productos más pedidos.
5. Variable pronóstico de ventas.
6. Canal de ventas.
7. Líneas más demandadas.
8. Matriz de correlación.
9. Análisis de Componentes Principales.
10. Clustering.

Estado inicial de datos

cada país cuenta con 31 variables:

- 19 generales
- 12 clasificadas como campos por país y periodo

-Se unificó los datos de El Salvador, Guatemala, Honduras y Nicaragua para poder hacer un análisis completo.

-Limpieza de datos

Medidas de tendencia central de variables

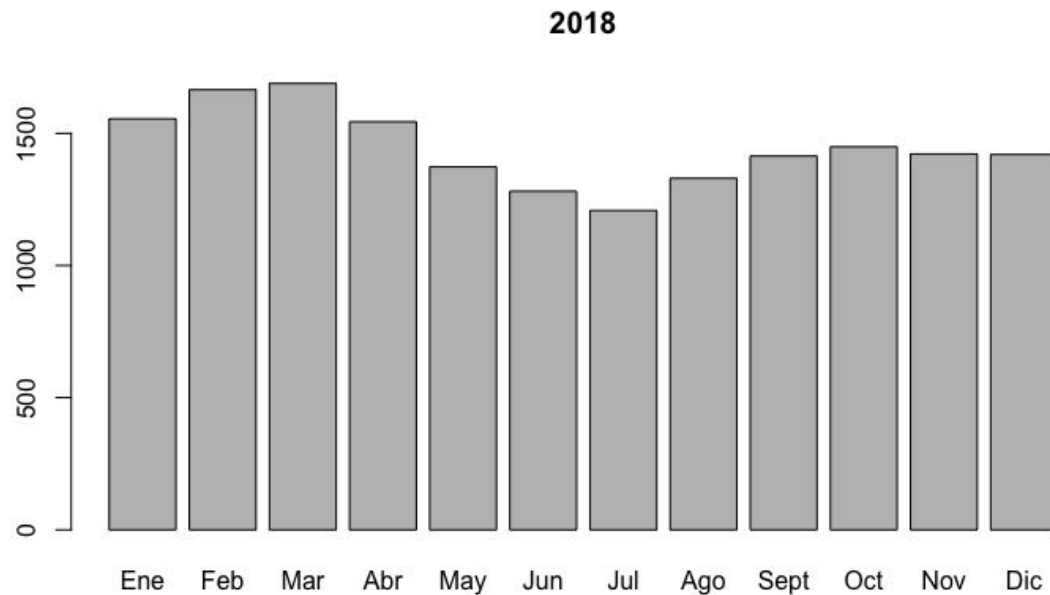
Precio Catalogo	Precio vta s/iva	Pronostico	Unidades Vendidas	Venta Neta s/iva
Min. : 0.00	Min. : 0.000	Min. : 0.0	Min. : -1213.0	Min. : -69528
1st Qu.: 8.95	1st Qu.: 4.925	1st Qu.: 100.0	1st Qu.: 0.0	1st Qu.: 0
Median : 45.95	Median : 26.844	Median : 300.0	Median : 173.0	Median : 8941
Mean : 106.39	Mean : 63.097	Mean : 686.4	Mean : 553.9	Mean : 38876
3rd Qu.: 154.95	3rd Qu.: 91.274	3rd Qu.: 750.0	3rd Qu.: 552.0	3rd Qu.: 33507
Max. : 6819.40	Max. : 4150.939	Max. : 42000.0	Max. : 51957.0	Max. : 2110540
NA's : 1695	NA's : 895			

Costo	utilidad	Margen	Pedido Real	Ratio
Min. : 0.00	Min. : -451731.8	Min. : -1997.0000	Min. : 0	Min. : -0.17413
1st Qu.: 0.65	1st Qu.: 279.8	1st Qu.: 0.1482	1st Qu.: 0	1st Qu.: 0.00000
Median : 5.92	Median : 6353.9	Median : 0.6001	Median : 7764	Median : 0.01358
Mean : 17.47	Mean : 30730.1	Mean : 0.5256	Mean : 6828	Mean : 0.05557
3rd Qu.: 27.69	3rd Qu.: 23978.8	3rd Qu.: 0.9768	3rd Qu.: 10540	3rd Qu.: 0.05275
Max. : 431.94	Max. : 2110539.5	Max. : 754.0161	Max. : 14527	Max. : 6.75959
NA's : 1262	NA's : 817		NA's : 800	

Meses con más pedidos

2018

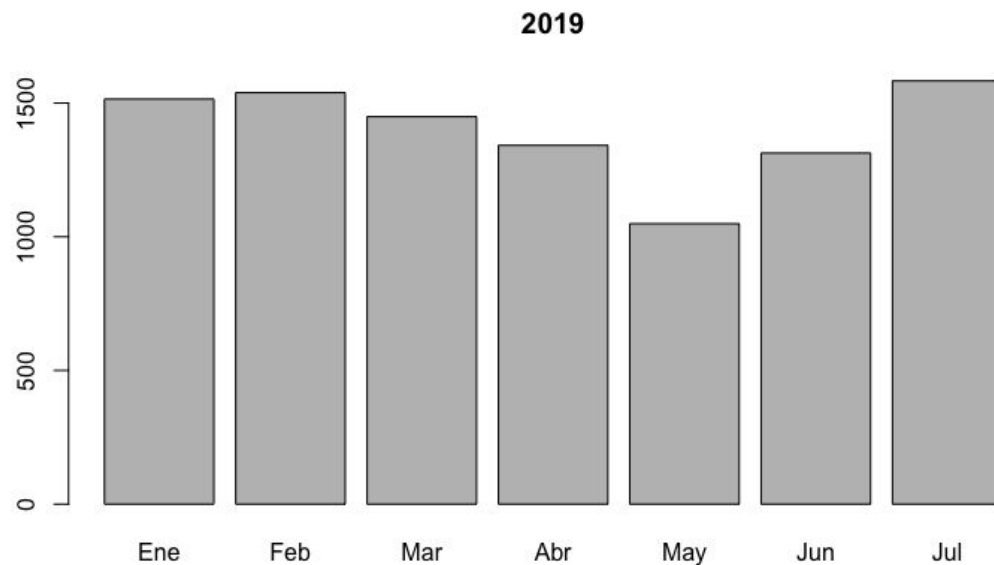
Marzo, con 1689.



Meses con más pedidos

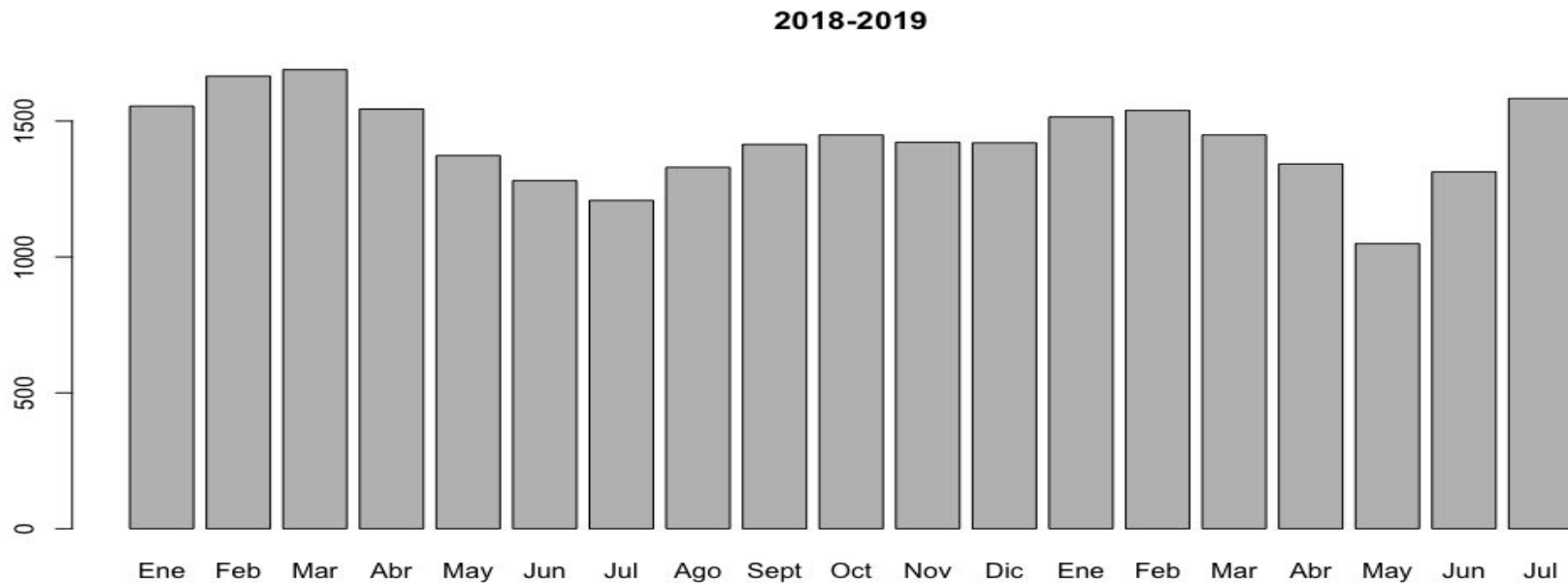
2019

El mes de Julio, con 1583 pedidos.



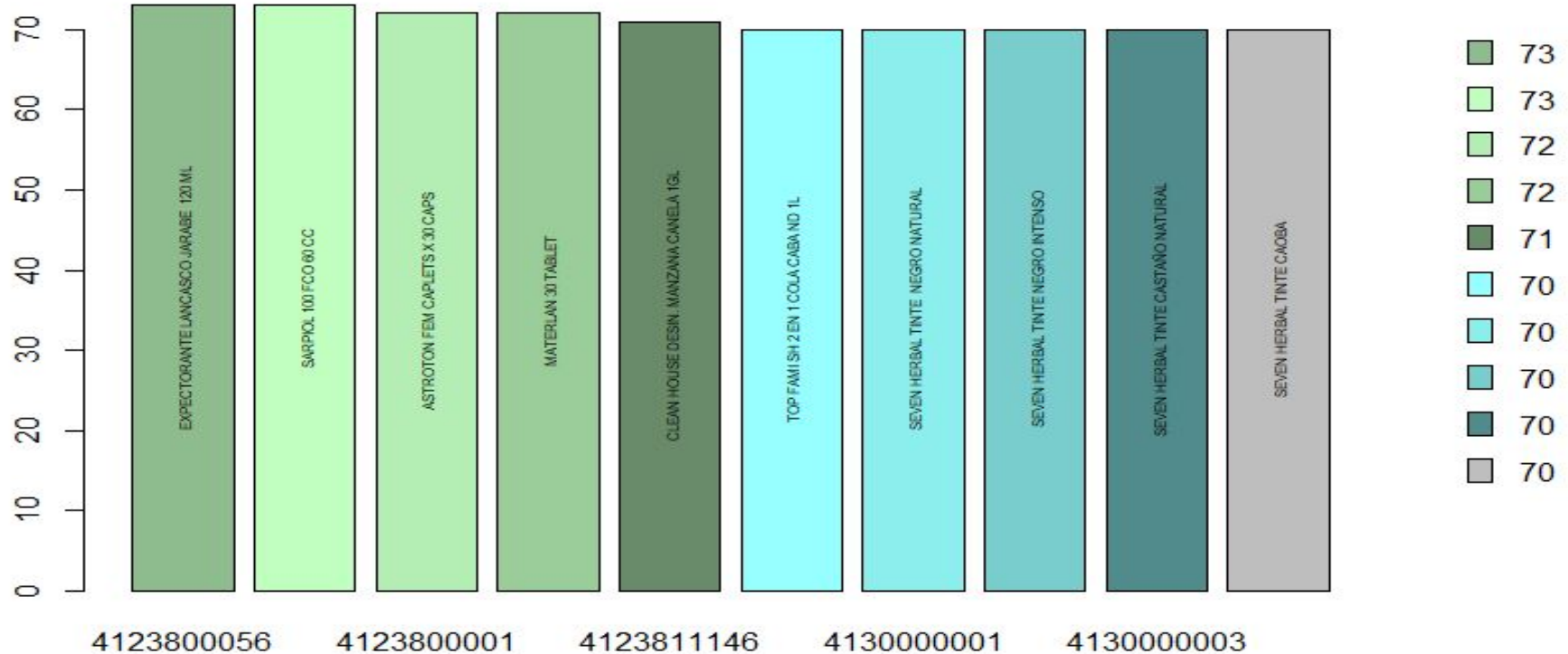
Meses con más pedidos

Total registrados → Marzo, con 1689.

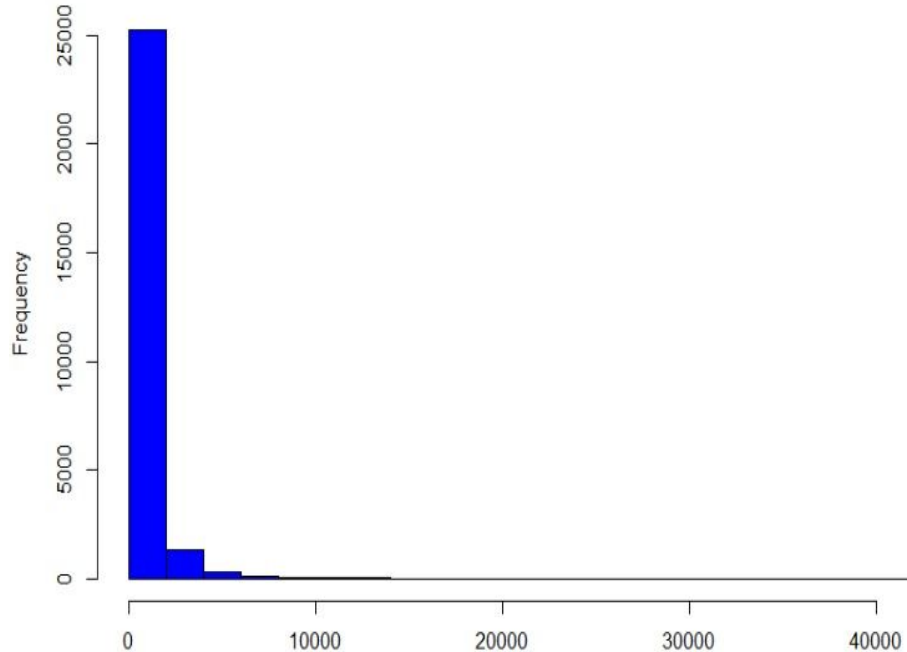


Productos más pedidos

1. “EXPECTORANTE LANCASCO JARABE 120 ML”
2. “SARPIOL 100 FCO 60 CC”

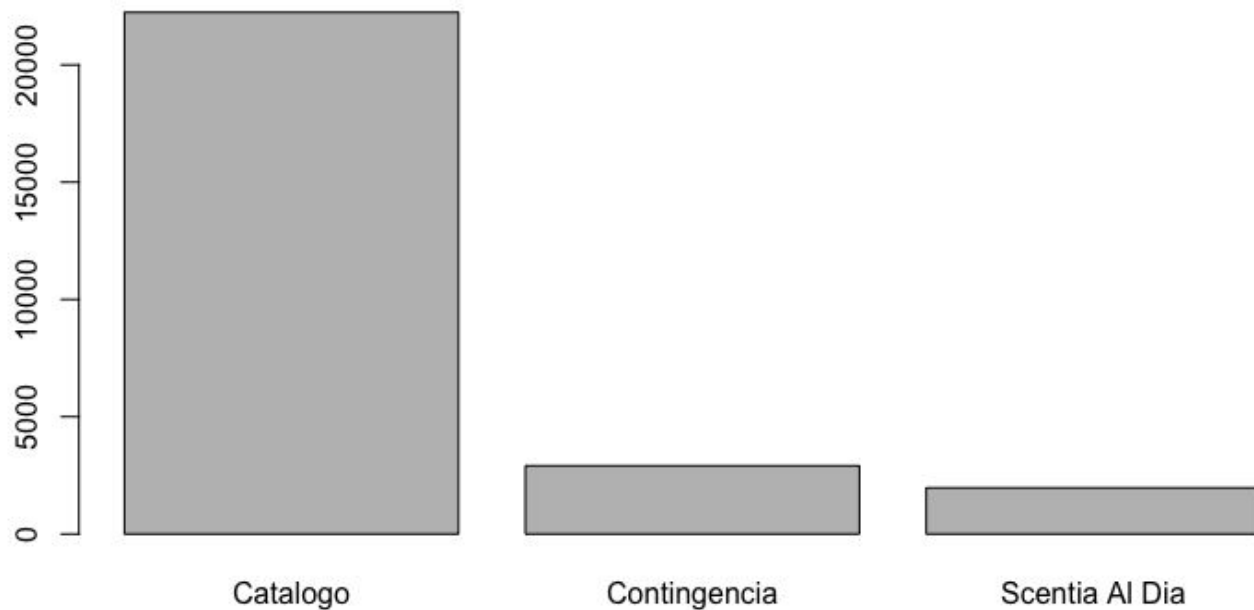


Variable pronóstico de ventas

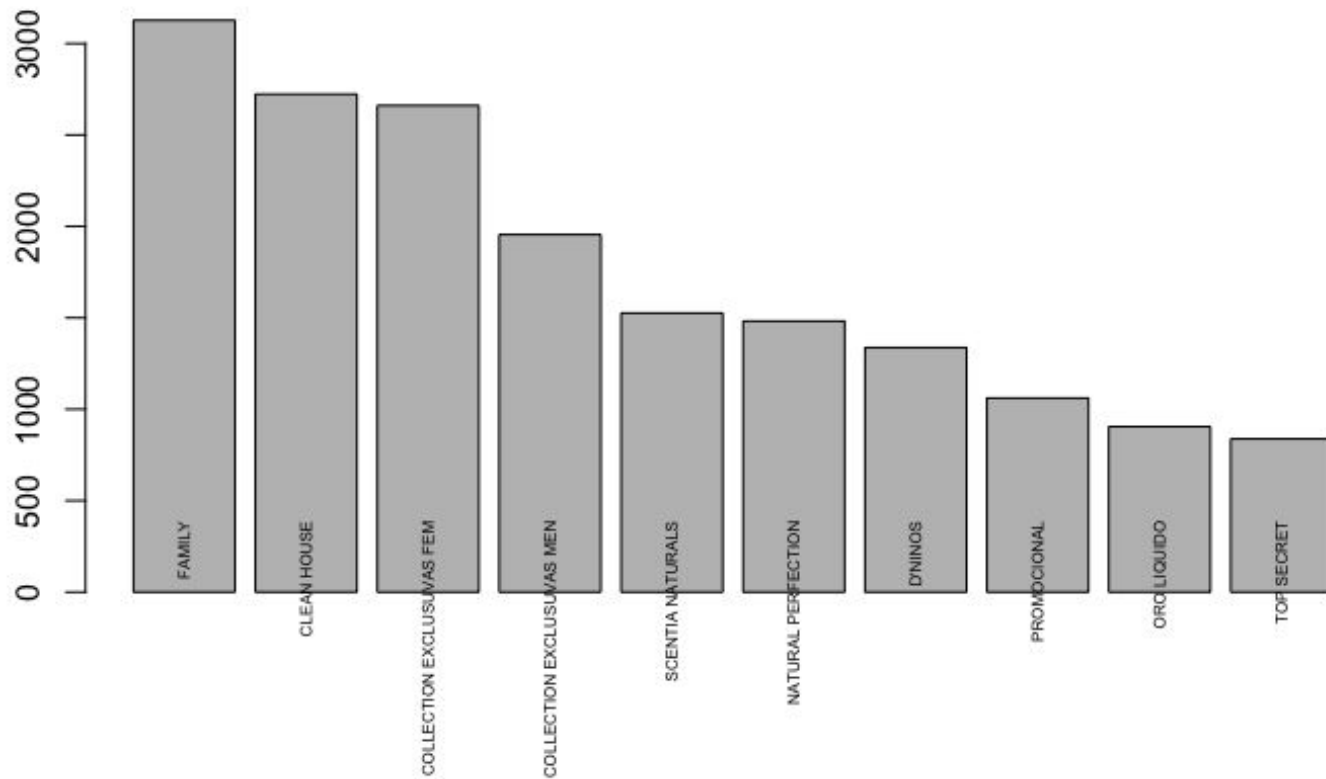


- Es claro que no hay normalidad en los datos.
- Casi todos los pronósticos se encuentran dentro de la misma clase.
- Pronósticos no varían, lo que indica que no son ajustados a cada temporada.

Canal de ventas → catálogo



Top Líneas → Familia, Clean House y Collection Exclusive FEM.

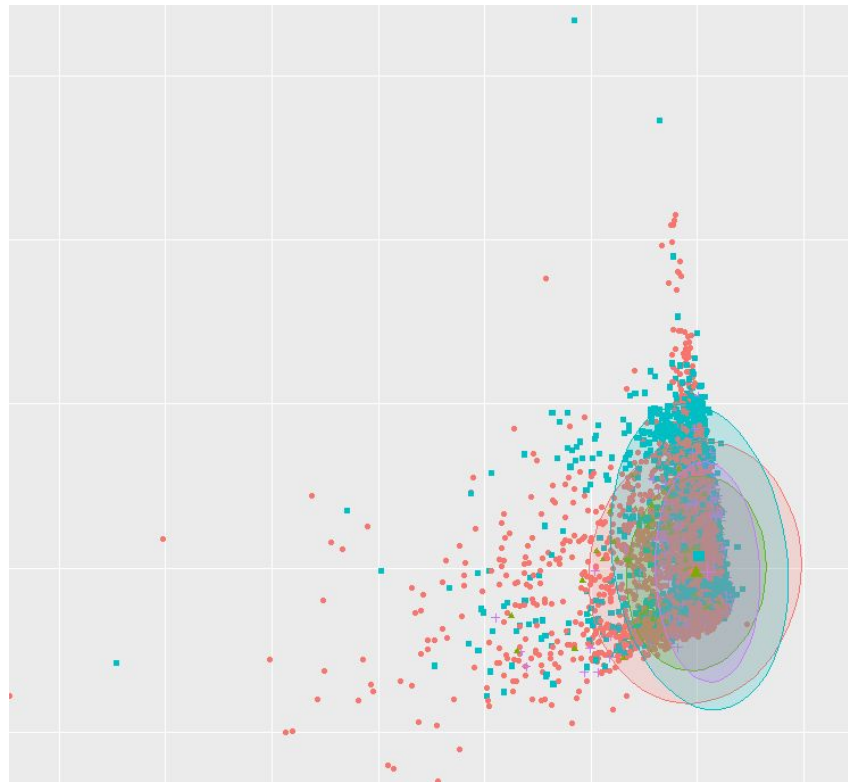


Matriz de correlación

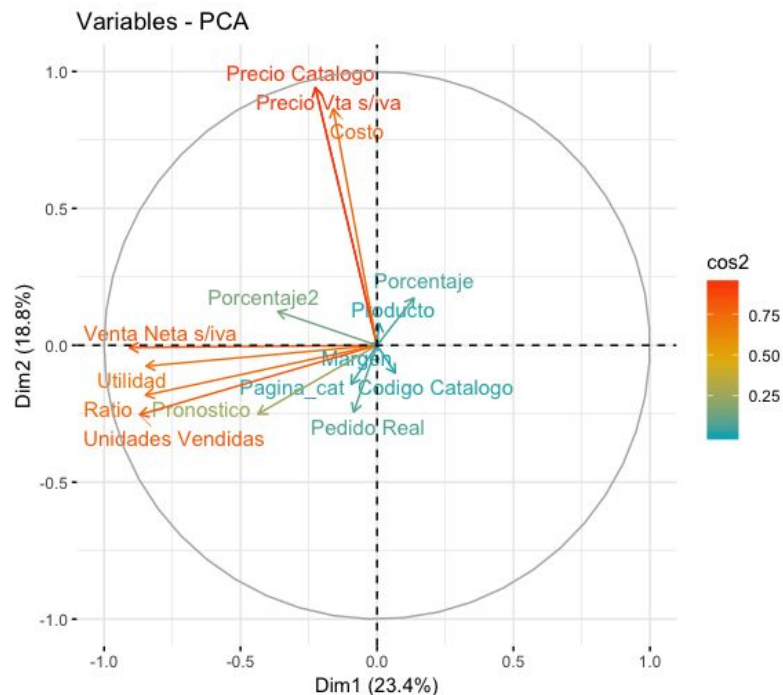
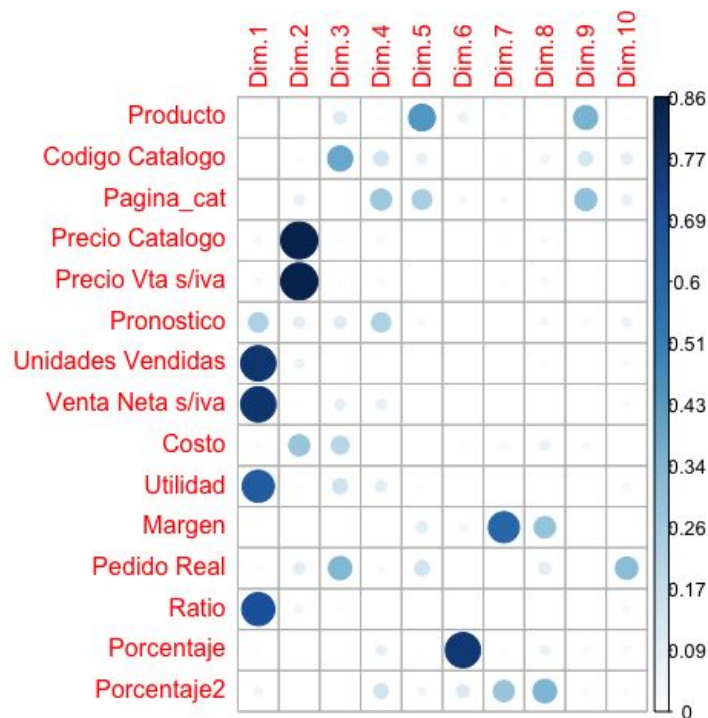


- Tonos azules correlaciones más altas.
- Tonos rojos correlaciones más bajas.
- Más adelante se podría eliminar variables en función de sus altas correlaciones para el modelo.

Clustering



Análisis de Componentes Principales (PCA)



Análisis de Componentes Principales (PCA)

```
Importance of components:
      PC1  PC2  PC3  PC4  PC5  PC6  PC7  PC8  PC9  PC10  PC11  PC12  PC13  PC14
Standard deviation    1.872 1.681 1.1704 1.0314 1.017 1.0030 0.9891 0.9132 0.8930 0.8242 0.6977 0.5212 0.33897 0.18504
Proportion of Variance 0.234 0.188 0.0913 0.0709 0.069 0.0671 0.0652 0.0556 0.0532 0.0453 0.0324 0.0181 0.00766 0.00228
Cumulative Proportion 0.234 0.422 0.5132 0.5842 0.653 0.7202 0.7854 0.8410 0.8942 0.9395 0.9719 0.9901 0.99772 1.00000

      PC15
Standard deviation    0.00329
Proportion of Variance 0.00000
Cumulative Proportion 1.00000
```

Componentes:

Producto, Código Catálogo, Pagina_cat, Precio Catálogo, Precio Vta s/iva, Pronóstico, Unidades Vendidas, Venta Neta s/iva, Costo y Utilidad