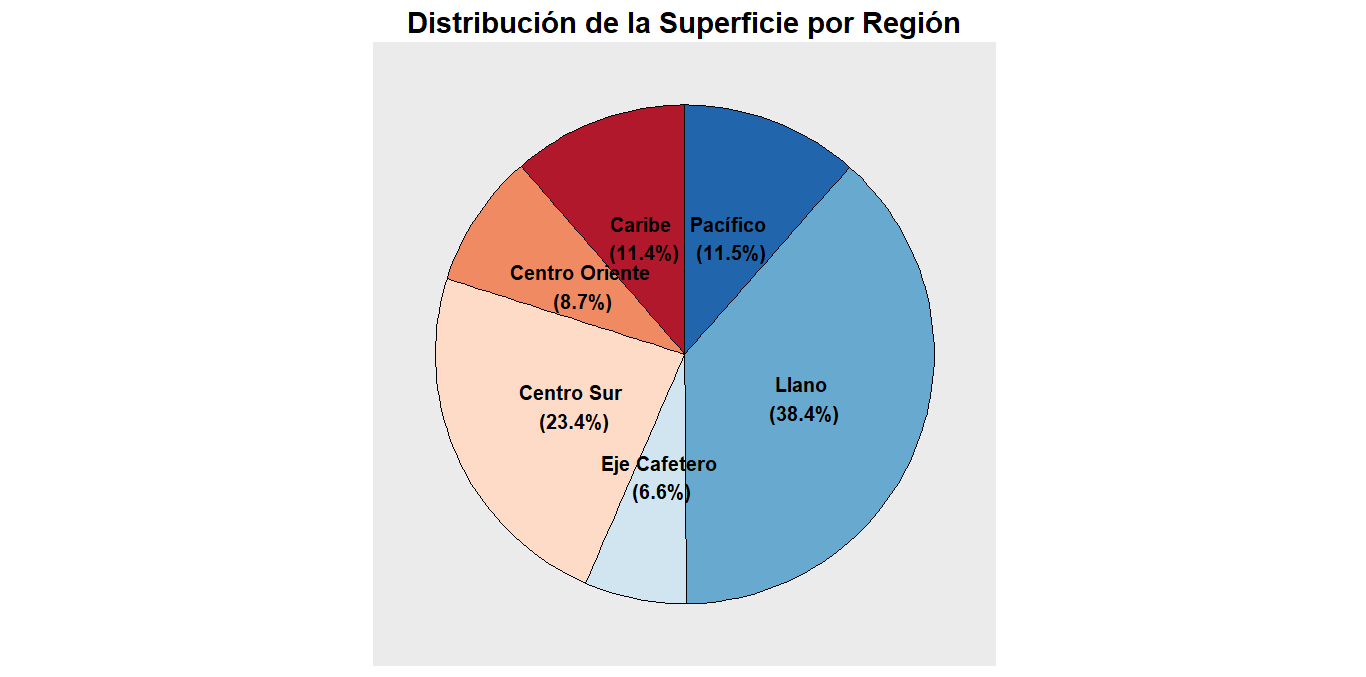
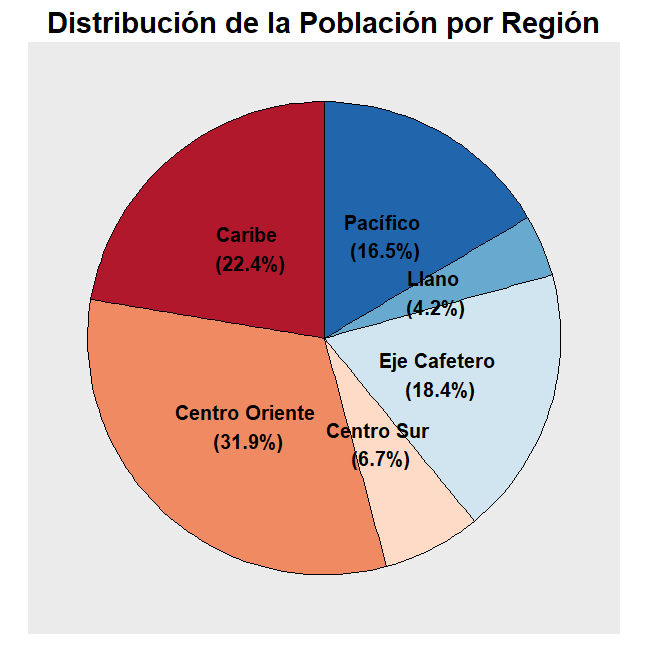
En el siguiente documento se encuentra el paso a paso de la ejecución del código.

En una primera parte se importan las librerías y documentos necesarios (Municipios y Prestadores). En la exploración de los datos vemos que el documento de **Municipios** necesita una limpieza exhaustiva, posteriormente se revisa la presencia de **missings** y se encuentra que hay uno en superficie, se encuentra la información de la superficie de **Mapiripana** (6457km) continuamos con el archivo de **Prestadores** y se procede a eliminar las variables que tenían más del 50% de **missings** y otras no tan relevantes.

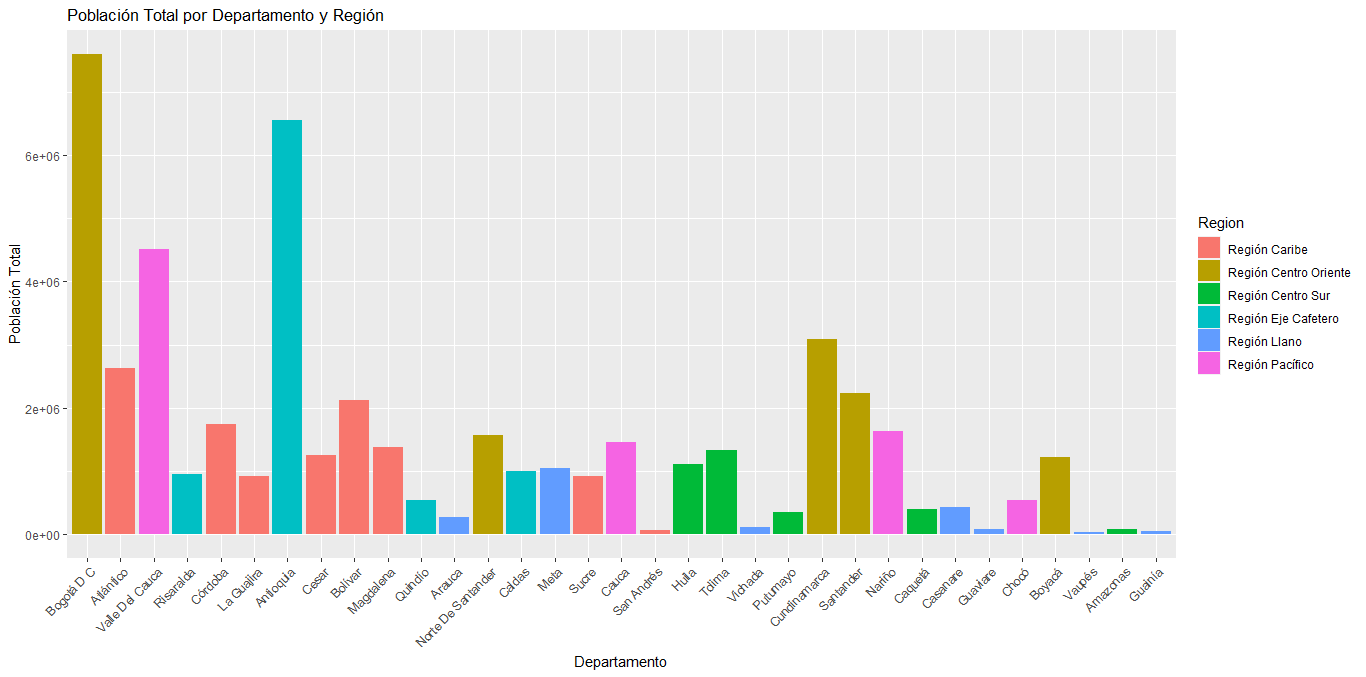
En el análisis descriptivo se importa el archivo **mismacros.txt** para ejecutar la función **mytable** con esta función creamos dos tablas de frecuencia. En la primera vemos la cantidad de municipios por departamento, donde en el primer lugar tenemos Antioquia con 125 municipios (11.18%) en segunda, tenemos la cantidad de prestadores por departamentos donde Bogotá ocupa el primer lugar con 14728 (24.17%).

Continuamos con las consultas de SQLite, en la primera vemos ver la cantidad de municipios por departamento además del promedio, mínimo y máximo de superficie, en la segunda tenemos la cantidad de prestadores por departamentos además el porcentaje de persona natural y jurídica.

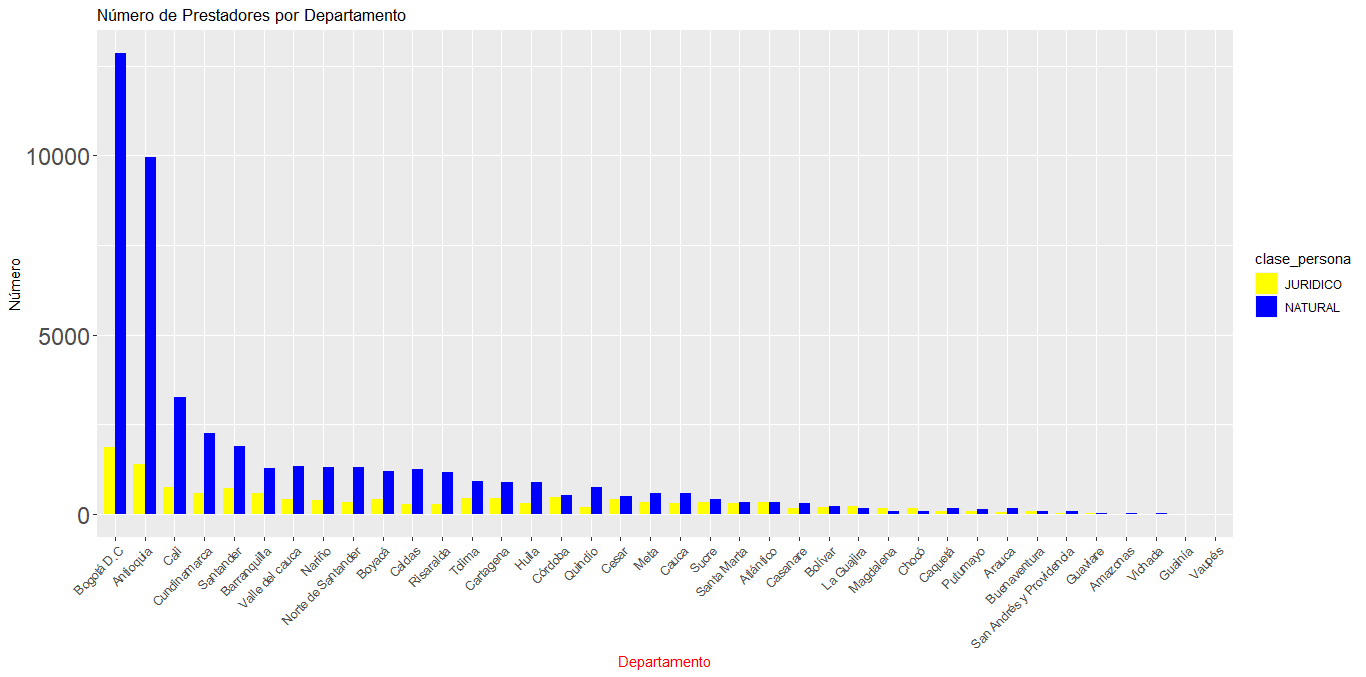
En la visualización de los datos en una primera parte tenemos los siguientes graficos.



Acá se puede apreciar como el Llano es la región con mayor superficie y Centro Oriente la región con mayor población.

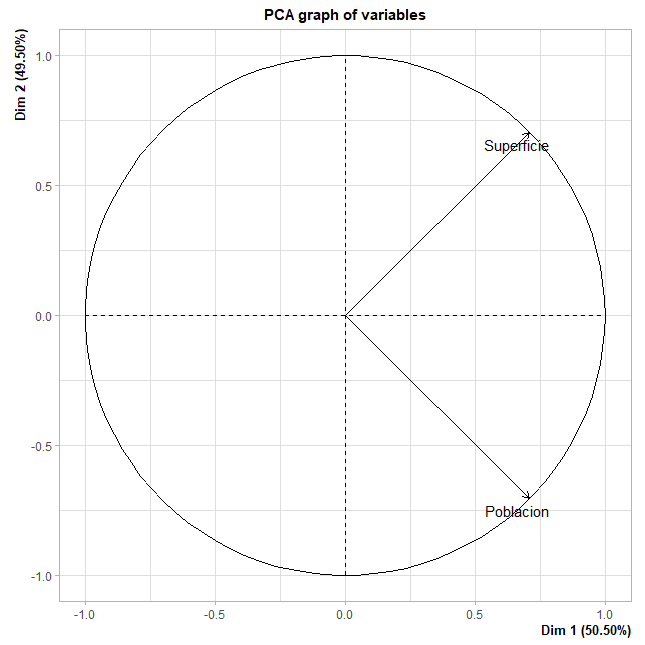


En este grafico podemos apreciar como bogotá es el departamento con mayor población.



Es claro que Bogotá en el departamento con más prestadore además de que la clase de persona en general más común es la Natural.

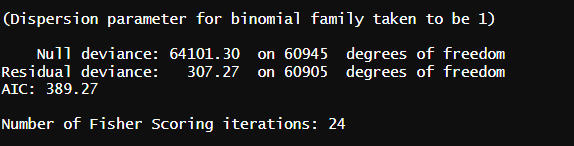
En la parte de modelamiento iniciamos un análisis de componentes principales, en este caso se puede apreciar en el grafico que forman un angulo de 90% por tanto estas variables serían independientes entre sí.



En una segunda parte se ajusta un modelo mixto teniendo en cuenta la agrupación por departamento. Se encontró que en promedio, por cada unidad adicional de superficie, se espera un aumento de aproximadamente 2,422 unidades en la población del municipio, además de encontrarse datos atípicos que podrían sesgar los resultados.

|  | **StudRes** | **Hat** | **CookD** |
| --- | --- | --- | --- |
| 1 | 18.9218377 | 0.008002144 | 1.4440834 |
| 149 | 5.6790974 | 0.990602105 | 1699.7979784 |
| 1002 | 16.7408611 | 0.023804983 | 3.4170936 |
| 1118 | -0.5687665 | 0.527602668 | 0.1806498 |

Continuando con el modelamiento se realiza un modelo de regresión logística para el archivo de **prestadores** con diferentes variables dejando el modelo con mejor AIC. Acá se puede apreciar como las dos variables tienen una influencia significativa en pertenecer a la clase persona “Juridica”, tanto los departamentos de una manera positivas como clpr\_codigo de una manera negativa.

****

Además, se puede apreciar que el modelo convergió después de 24 interacciones, además de tener un valor predictivo significativo en comparación al modelo nulo, es decir, tiene una mejor bondad de ajuste.

Por último, se ajusta un modelo de supervivencia donde se toma la diferencia de tiempo entre las fechas de radicación y vencimiento como el tiempo de supervivencia, aquí se aprecia como la variable clase\_persona tendría un impacto significativo en este tiempo, es de notar algunos departamentos cuentan con coeficientes significativos Bogotá con un coeficiente negativo (-0.4)  
lo que surgiere menor tiempo de supervivencia, es decir, menor la brecha entre la fecha de radicación y la fecha de vencimiento, por otro lado chocó sería el caso contrario con un coeficiente positivo (0.39)

Texto

Descripción generada automáticamente

¡¡Gracias por la atención prestada cualquier duda o inquietud quedo super atento!!