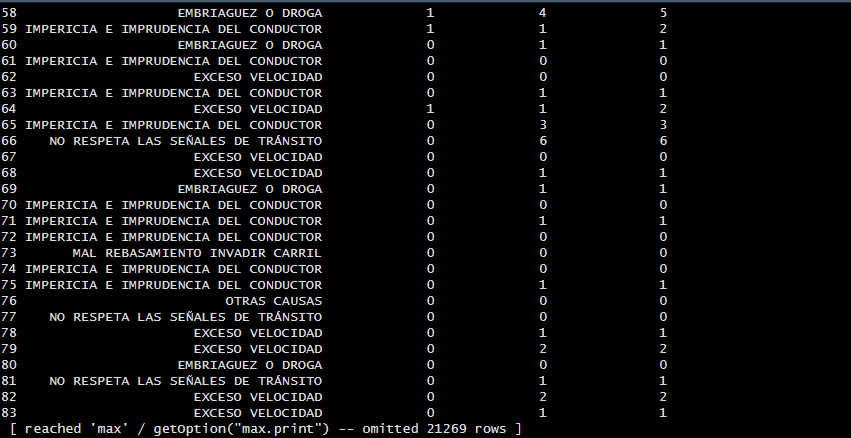
**Universidad técnica particular de Loja.**

**Nombre: Roberto Bustamante**

**Fecha: 25/04/2025**

1. **Examinar los datos (estructura, dimensionalidad, etc)**

**Print.**



Head.

Interfaz de usuario gráfica, Texto

El contenido generado por IA puede ser incorrecto.

Str.

Interfaz de usuario gráfica, Texto

El contenido generado por IA puede ser incorrecto.

Summary.

Interfaz de usuario gráfica

El contenido generado por IA puede ser incorrecto.

Dim.

Texto

El contenido generado por IA puede ser incorrecto.

Class.

Una captura de pantalla de un celular con texto e imagen

El contenido generado por IA puede ser incorrecto.

Mode.

Texto

El contenido generado por IA puede ser incorrecto.

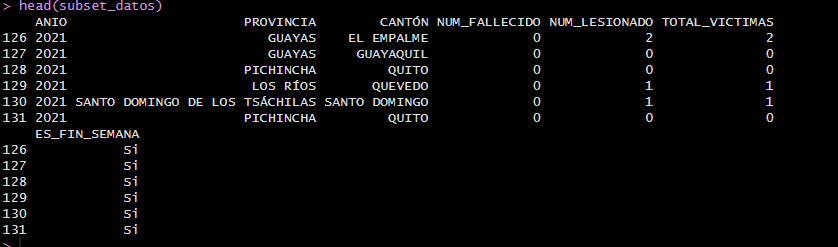
1. **Agregar al menos dos atributos derivados**

datossini$GRAVEDAD <- ifelse(as.numeric(datossini$NUM\_FALLECIDO) > 0, "Grave", "No Grave")  
crea una nueva columna llamada gravedad la cual analiza si el numero de muertos es mayor a 0 lo clasifica como grave, caso contrario lo pone como no grave.

Interfaz de usuario gráfica, Texto

El contenido generado por IA puede ser incorrecto.

datossini$ES\_FIN\_SEMANA <- ifelse(datossini$DIA %in% c("SÁBADO", "DOMINGO"), "Si", "No")



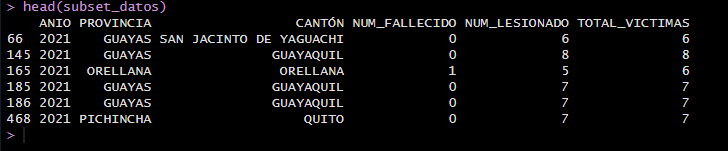
crea una nueva columna llamada es\_fin\_semana la cual se encarga de recopilar si el día en el que ocurrió el siniestro fue un sábado o domingo, con lo cual podemos sacar patrones para analizar cuando es que se producen más siniestros.

1. **Extraer un subconjunto de datos (filtrando tanto filas como columnas)**

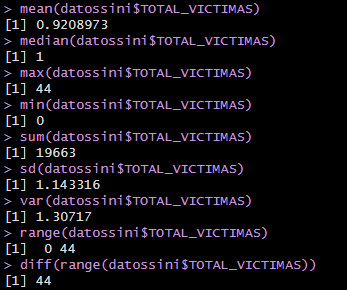
subset\_datos <- subset(datossini,(TOTAL\_VICTIMAS) >= 6,select = c("ANIO",

"PROVINCIA", "CANTÓN", "NUM\_FALLECIDO", "NUM\_LESIONADO", "TOTAL\_VICTIMAS"))

Se encarga de filtrar los datos para que los que se muestren sean los siniestros donde las victimas totales hayan sido mayores o iguales a 6.



1. **Para datos numéricos obtenga estadísticas globales (media, mediana, moda, varianza, desviación tipica, rango)**



1. **Obtenga estadísticas mediante agrupamientos (Investigar).**

install.packages("dplyr")

library(dplyr)

datossini %>%

group\_by(PROVINCIA) %>%

summarise(

promedio\_victimas = mean(as.numeric(TOTAL\_VICTIMAS), na.rm = TRUE),

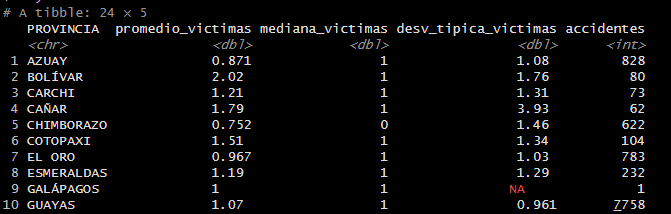
mediana\_victimas = median(as.numeric(TOTAL\_VICTIMAS), na.rm = TRUE),

desv\_tipica\_victimas = sd(as.numeric(TOTAL\_VICTIMAS), na.rm = TRUE),

accidentes = n()

)

Se instala la librería dplyr y luego se la ejecuta, luego se procede a armar la consulta por agrupamiento de los datos por provincia con la cual obtenemos los siguientes resultados.



Se usa el as.numeric para que transforme los datos que se encuentren en tipo texto a tipo numérico y así poder realizar operaciones sin problemas