

Obtención de estadísticas descriptivas

Una introducción al conjunto de datos ¿Qué es? ¿De dónde se obtuvo? ¿Qué representa?

Los datos que se utilizaron en esta actividad fueron obtenidos de choques de automóviles. Cuentan con una descripción del evento y datos relevantes sobre el ambiente, duración, fecha y lugar del suceso. Cada registro representa un accidente y recopila las variables que lo acompañaron.

Cantidad de datos que tienes, las variables que contiene cada vector de datos y el tipo de variables.

La base de datos cuenta con 499 registros, los vectores contienen 47 variables de tipo entero, flotante, objeto y booleano.

Para las dos variables que escogiste:

-Los rangos de las variables que escogiste

Pressure(in) tiene un rango de valores de 28.37 a 30.62

Visibility(mi) tiene un rango de valores de 0.5 a 20.0

-Basándote en la media, mediana y desviación estándar de cada variable, ¿qué conclusiones o asunciones puedes obtener de los datos? ¿Parecen muy dispersos? ¿Alcanzas a vislumbrar algún patrón? ¿Parecen relacionados?

En el caso de la presión los valores no varían mucho si nos basamos en la desviación estándar y a su acotado rango. La mayoría de los accidentes suceden cuando la presión está cerca de los 29.9 in de acuerdo a la media y mediana.

En cuanto a la visibilidad, la dispersión es mucho mayor. Sin embargo, se repiten muchos valores a comparación de la columna de presión. Juzgando por la media y mediana la mayoría de accidentes suceden cuando la visibilidad es mayor.

Entre estas dos variables no parece haber una relación clara.