

# Matemáticas Computacionales

## Práctica 2: Estudio de una Base de Datos

Arely Anahi Alvarado Solis  
1904628

### 1. Introducción

En esta práctica se estudiara un base de datos con la estadística descriptiva básica. Se analizara los atributos, tipos de datos, indicadores básicos de una y de dos variables. Se muestran algunos de los primeros pasos para luego implementar un preprocesamiento de datos para luego utilizarla en alguna experimentación o trabajo a tratar.

### 2. Base de Datos: cars

Dada esta base de datos identificada como cars son los datos que dan la velocidad de los automóviles y las distancias que se toman para detenerse. Tomemos en cuenta que los datos que se registraron fueron en la década de 1920. Esto se implementa en [3] y [2]

#### 2.0.1. Descripción del conjunto de datos

El datasets cars tiene 50 observaciones y aparte cuenta con 2 atributos que son:

1. Velocidad ("mph": milla por hora)
2. Distancia de frenado ("ft": pies)

#### 2.1. Estadística descriptiva de una variable

Los datos de velocidad marca con un valor mínimo de 4.00, un máximo de 25.00, una mediana de 15.00 y media de 15.9; los datos de la distancia de frenado contiene un valor mínimo de 2.00, un valor máximo de 120.00, una mediana de 36.00 y una media de 42.98, esto se ve en la Figura (1).

Con 12.00 sea el primer cuartil, el tercer cuartil de 19.00 y la mediana, que comentamos anteriormente, de 15.00, son datos de *velocidad* y se ven en la Figura (2) ademas, se observa donde se encuentra el valor mínimo y máximo.

Aparte los datos de *distanciade frenado*, tenemos con el primer cuartil 26.00, el tercero con 56.00 y la mediana, que también comentamos anteriormente, de 36.00; esto se ve en la gráfica de la Figura (3) ademas, se ve donde se encuentra el valor mínimo y máximo.

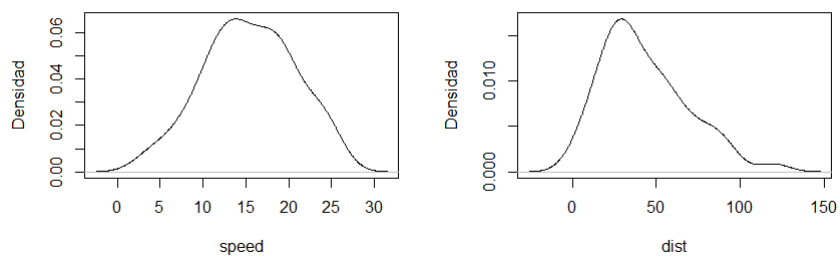


Figura 1: Gráfica de densidad de los atributos.

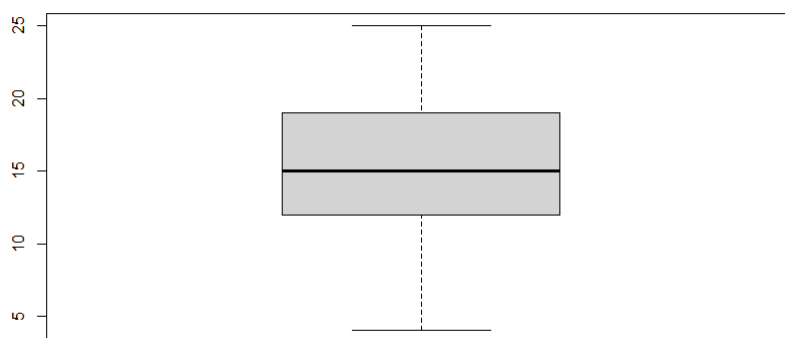


Figura 2: Gráfica de caja bigotes de la velocidad.

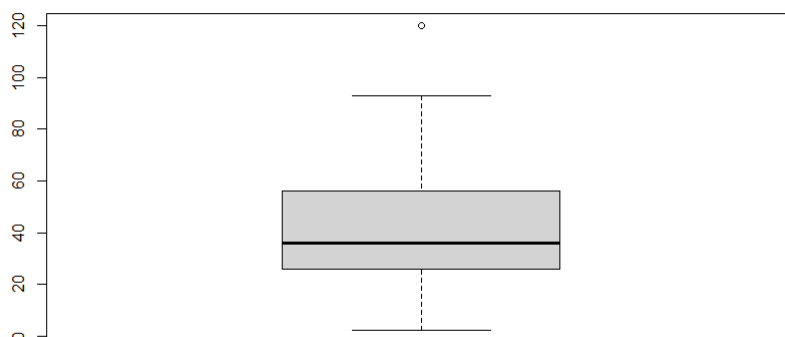


Figura 3: Gráfica de caja bigotes de la distancia de frenado.

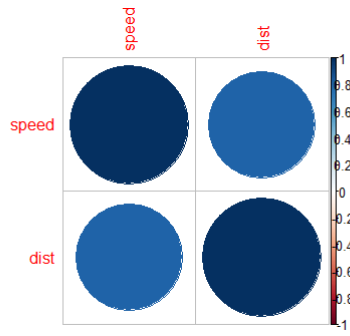


Figura 4: Gráfica de correlaciones entre los atributos.

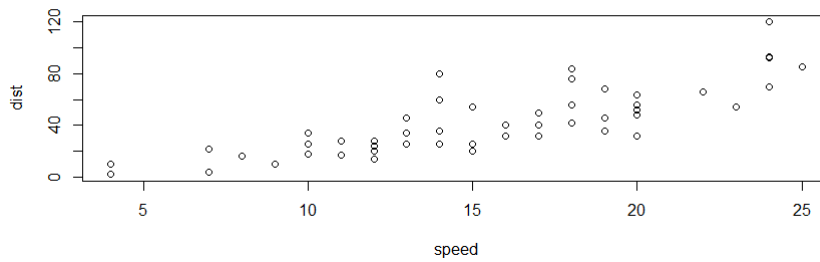


Figura 5: Gráfica de dispersión 1 (scatterplot).

## 2.2. Estadística descriptiva de dos variables

En la Figura (4) se muestra la correlación entre los únicos dos atributos utilizando el coeficiente de correlación de Pearson, donde se ve que dado los únicos dos atributos no hay mucho que decir entre estos, pues como va en un rango de  $[-1,1]$ , se ve que ambos están fuertemente correlacionados ya que ambos se acercan al 1 positivo.

En las gráficas de las Figuras (5) y (6) se ven los dos atributos correlacionados, donde se ve que indica una relación positiva entre ellas.

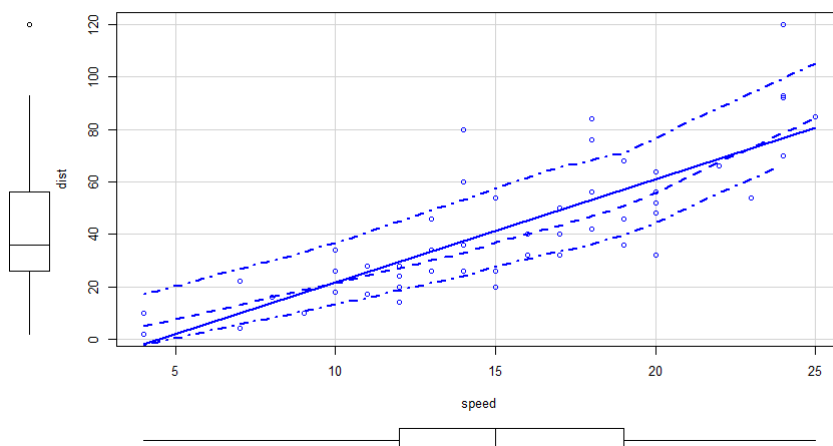


Figura 6: Gráfica de dispersión 2 (scatterplot).

Puede encontrar los códigos en el repositorio [1]

## Referencias

- [1] Arely Alvarado. Repositorio de Github. <https://github.com/ArelyAnahiaA/MateComputacional>, 2021.
- [2] M. Ezekiel. *Methods of Correlation Analysis*. Wiley, 1930.
- [3] D. R. McNeil. *Interactive Data Analysis*. Wiley, 1977.