

# Matemáticas Computacionales

## Práctica 2: Estudio de una base de datos

Profesor: Ángel Isabel Moreno Saucedo

Alumno: Paola Lizbeth Vázquez Leal

Semestre Febrero - Junio 2021

### Práctica 2

## 1. Introducción

En esta actividad se analizará una Base de datos elegida por el estudiante. De esta se observarán los atributos de cada uno así como también se mostrara graficamente dichas observaciones. Además se analizarán los datos de una sola variable y de dos variables de igual manera. La base de datos elegida es llamada *trees*, que contiene informacion de medidas de un tipo de árbol conocido como *Prunus serotina* [3].

## 2. Base de datos: identificación de árboles de cerezo negro.

La base de datos **trees** contiene 31 observaciones del tipo árbol de cerezo, siendo este el árbol de cerezo negro como nos proporciona [1]. Nos proporciona las mediciones del diametro, la altura y el volumen de la madera de 31 árboles. Dentro de la base de datos nos proporcionan Girth, sin embargo está mal representado ya que este en realidad es una medida hecha de 4 a 6 pies sobre el suelo.

Los atributos de esta base de datos son 3:

1. Girth
2. Height
3. Volume

### 2.1. Estadística de una variable

Para comenzar con la descripción de una variable, contamos con tres tipos de atributos, los cuales han sido mencionados antes. Los datos obtenidos de Girth (diametro) son un mínimo de 8.30 y un máximo de 20.60, en cuanto a su media de 12.90 y mediana de 13.25. Los datos de Height (altura) con un mínimo de 63 y un máximo de 87, en cuanto a su media y mediana son iguales con 76. Por ultimo los datos de Volume (volumen) son un mínimo de 10.20 y un máximo de 77.00. Para ello se encuentran las gráficas de densidad que nos mostraran los datos antes platicados. Además podemos agregar la grafica de caja de bigotes [4], donde podemos apreciar con mayor claritud donde se encuentra la media dada por la diferencia del ranfo intercuartil.

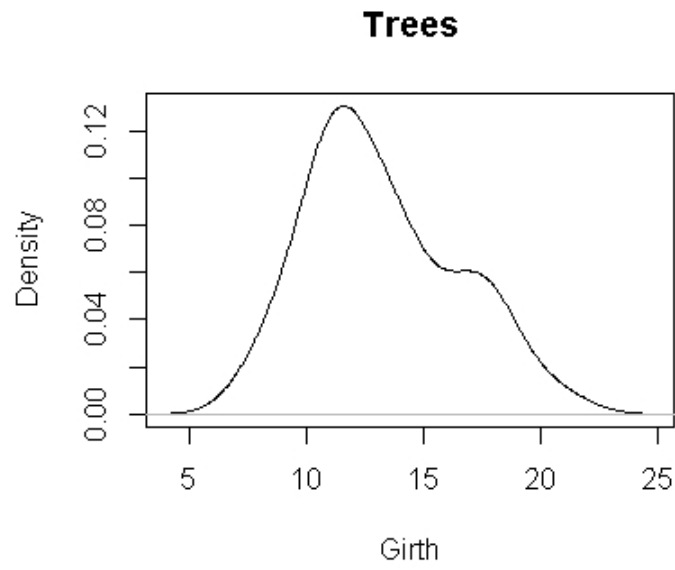


Figura 1: Densidad respecto al atributo Diametro.

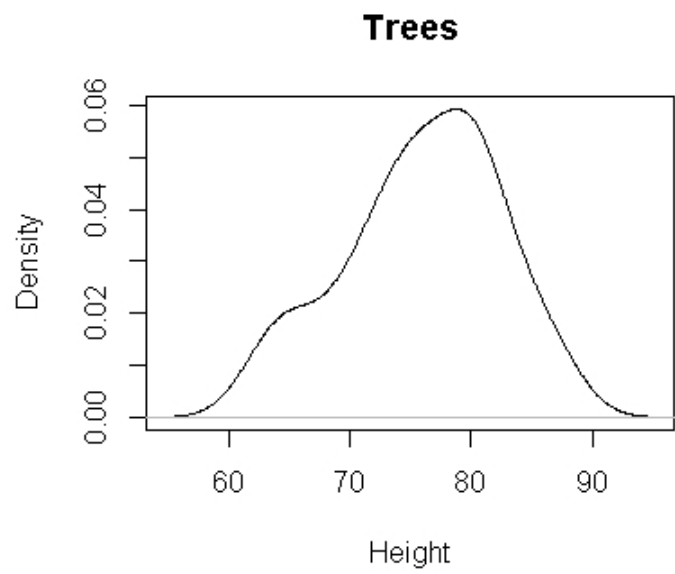


Figura 2: Densidad respecto al atributo Altura.

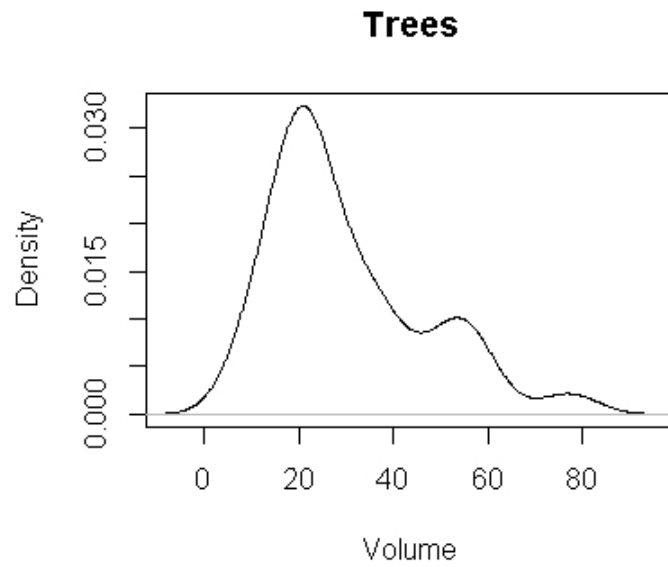


Figura 3: Densidad respecto al atributo Volumen.

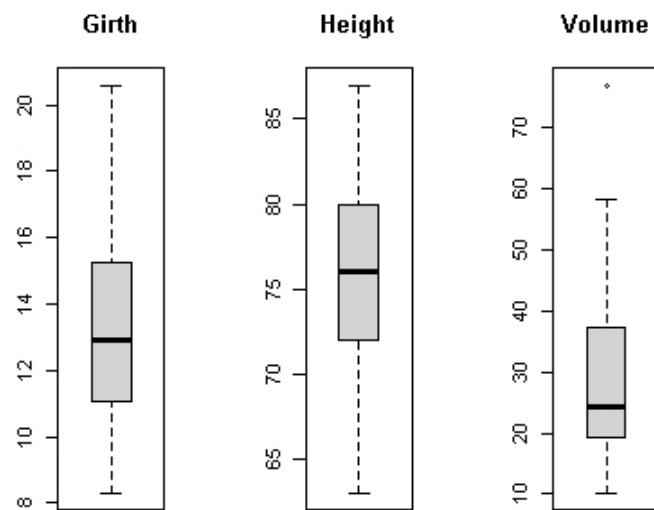


Figura 4: Muestra de la media junto al primer y tercer intercuartil.

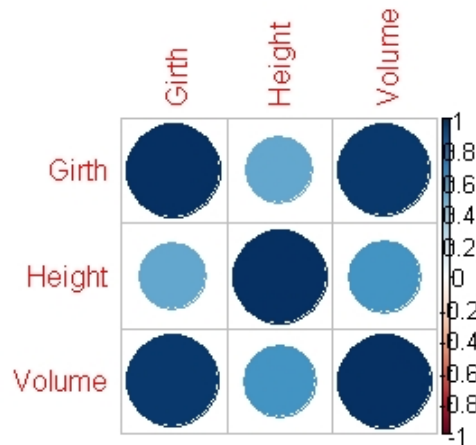


Figura 5

## 2.2. Estadística de dos variables

De la figura [5] podemos notar como las variables de Girth y Volume se encuentran fuertemente correlacionadas la una con la otra, esto significa que si una de las variables disminuye su tamaño la otra también lo hará, así como también pueden aumentar. En este caso donde la variable (llamémoslas 1 a Girth y 2 a Volume respectivamente) 1 por lo que, si una de las variables disminuye, lo hará la otra. Mientras que la altura (height) está altamente correlacionada con el volumen como se muestra en la figura [8] mientras que con el diámetro (Girth) está medianamente relacionada como se ve en la figura [7]. [4]

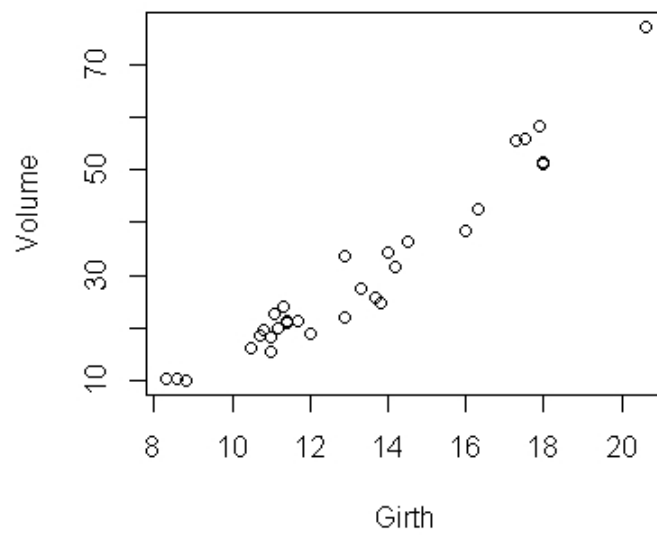


Figura 6: Dispersión existente entre diametro y altura

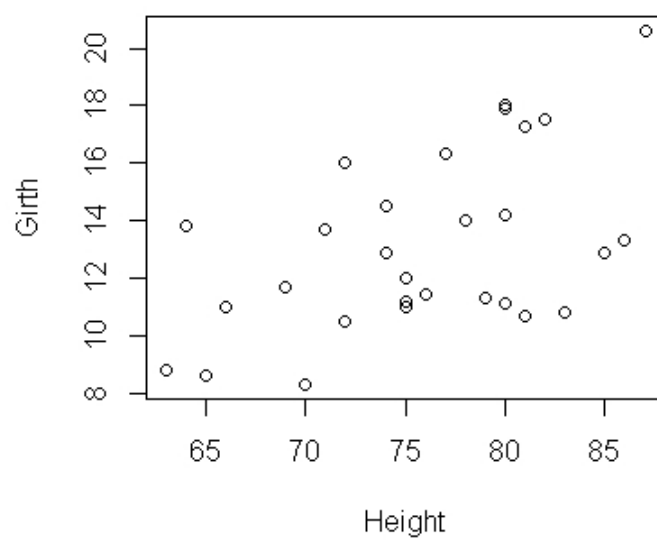


Figura 7: Dispersión existente entre altura y diametro

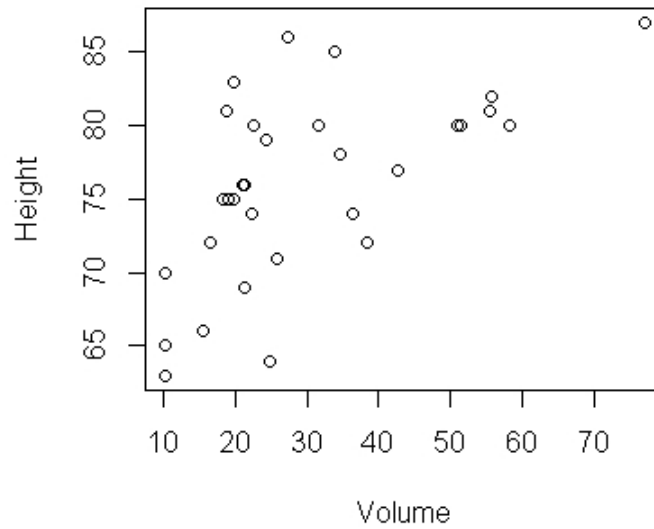


Figura 8: Dispersión existente entre volumen y altura

### 3. Conclusión

La base de datos de trees o bien conocida como trees black cherry, se usa como referencia a las mediciones de dichos árboles, llevando un listado de cada árbol que fue medido y relacionado a este tipo. Como se pudo observar los datos que esta base contenía fueron de ayuda al momento de buscar la información que era pedida para realizar la actividad. En ella se pudo observar que los tres atributos estaban correlacionados entre sí, aunque unos más que otros como fue el caso de Girth y Volume, mediante las graficas esta información pudo ser comprobada al igual que con los demás casos. En cuanto al momento de realizar el código, como se podrá observar en él, la relación de porcentaje con la frecuencia decidí obtener esta individualmente, ya que la base no tenía ningún clase factor. [2]

## Referencias

- [1] A.C. Atkinson. *Plots, Transformations and Regression*. Oxford University Press, 1985.
- [2] Thomas Arthur Ryan, Brian L. Joiner, and Barbara F. Ryan. *The Minitab Student Handbook*. Duxbury Press, 1976.
- [3] Natural Resources Conservation Service. *Prunus serotina* ehrh, black cherry.
- [4] Paola Vázquez. Repositorio de Github. <https://github.com/Li-vzz/MatematicasComputacionales>, 2021.