

Informe de Análisis

Exploratorio de Loblolly

Melani Forsythe Matos
Daniela Guerrero Álvarez
Rubén Martínez Rojas

Introducción

El dataset `Loblolly` contiene datos relacionados con el crecimiento de árboles de pino Loblolly (*Pinus taeda*), un tipo de pino comúnmente encontrado en el sureste de los Estados Unidos. El principal objetivo del dataset `Loblolly` es proporcionar información sobre cómo crecen los árboles de pino Loblolly a lo largo del tiempo. Esto permite a los investigadores y analistas explorar las tendencias de crecimiento, identificar factores que afectan el desarrollo de los árboles y crear modelos predictivos del crecimiento forestal. Este dataset cuenta con tres variables : // 1) La variable "Seed" permite identificar de qué semilla proviene cada árbol, lo cual es útil para analizar patrones de crecimiento específicos a semillas individuales. 2) La variable "height" es crucial para evaluar el desarrollo vertical de los árboles. 3) la variable "age" ayuda a entender cómo el tiempo afecta el crecimiento.

Seed

- **Descripción:** Identificador de la semilla.
- **Explicación:** Cada valor en esta columna representa una semilla específica de pino Loblolly, de la cual se han tomado mediciones de crecimiento.
- **Escala:** Nominal.
- **Tipo de variable:** Discreta.

height

- **Descripción:** Altura del árbol.

- **Explicación:** Esta variable representa la altura del árbol medida en pies.
- **Escala:** Razón (ya que tiene un verdadero cero y las diferencias y proporciones son significativas).
- **Tipo de variable:** Continua.

age

- **Descripción:** Edad del árbol.
- **Explicación:** La edad del árbol medida en años. Indica cuántos años tiene el árbol desde que fue plantado.
- **Escala:** Intervalo (aunque la edad tiene un verdadero cero, en este contexto la escala se utiliza principalmente para medir diferencias de tiempo).
- **Tipo de variable:** Discreta.

Análisis de Datos

Distribución Normal

La distribución normal, también conocida como distribución gaussiana o campana de Gauss, es una de las distribuciones de probabilidad más importantes en el análisis de datos y la estadística. Se caracteriza por ser simétrica y tener una forma de campana. A continuación se detallan sus principales características y propiedades:

Características de la Distribución Normal

1. **Simetría:** La distribución normal es simétrica respecto a su media. Esto significa que la mitad izquierda de la curva es una imagen especular de la mitad derecha.
2. **Campana de Gauss:** La curva tiene una forma de campana con una sola cima en el centro. A medida que nos alejamos de la media, la probabilidad disminuye de manera gradual.
3. **Media, Mediana y Moda:** En una distribución normal, la media (promedio), la mediana (valor central) y la moda (valor más frecuente) son iguales y se encuentran en el centro de la distribución.
4. **Desviación Estándar:** La dispersión de la distribución está determinada por la desviación estándar (σ). Aproximadamente el 68% de los datos se encuentran dentro de una desviación estándar de la media, el 95% dentro de dos desviaciones estándar y el 99.7% dentro de tres desviaciones estándar.

Ecuación de la Distribución Normal

La función de densidad de probabilidad de una distribución normal viene dada por la ecuación:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

Donde:

- x es una variable aleatoria continua.
- μ es la media de la distribución.
- σ es la desviación estándar.
- e es el número de Euler (≈ 2.718).

Measure	Value
Mean	13.00000
Median	12.50000
Mode	3.00000
Variance	62.40964

Figure 1: Resultados del análisis de edad.

Conclusiones

Conclusiones del análisis exploratorio.

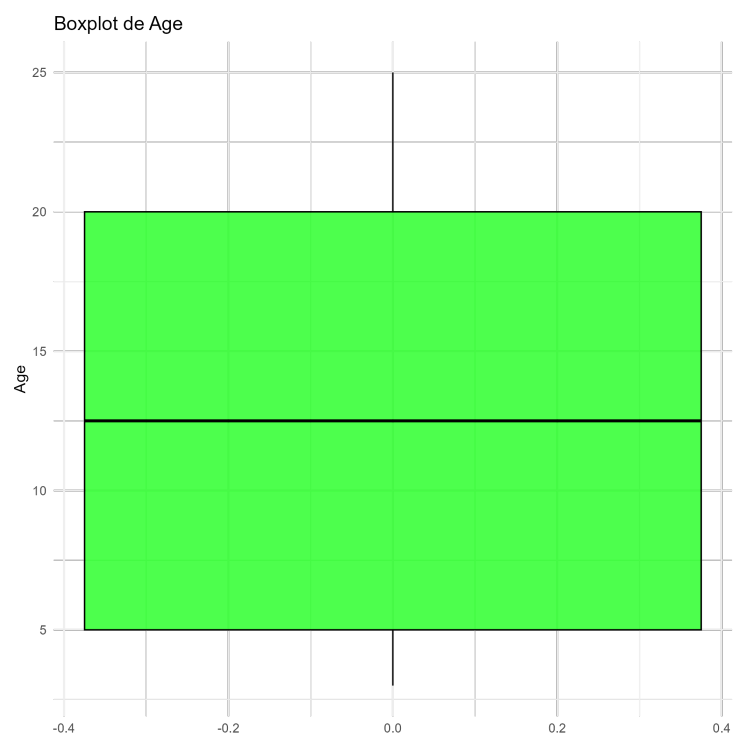


Figure 2: Boxplot de la edad.

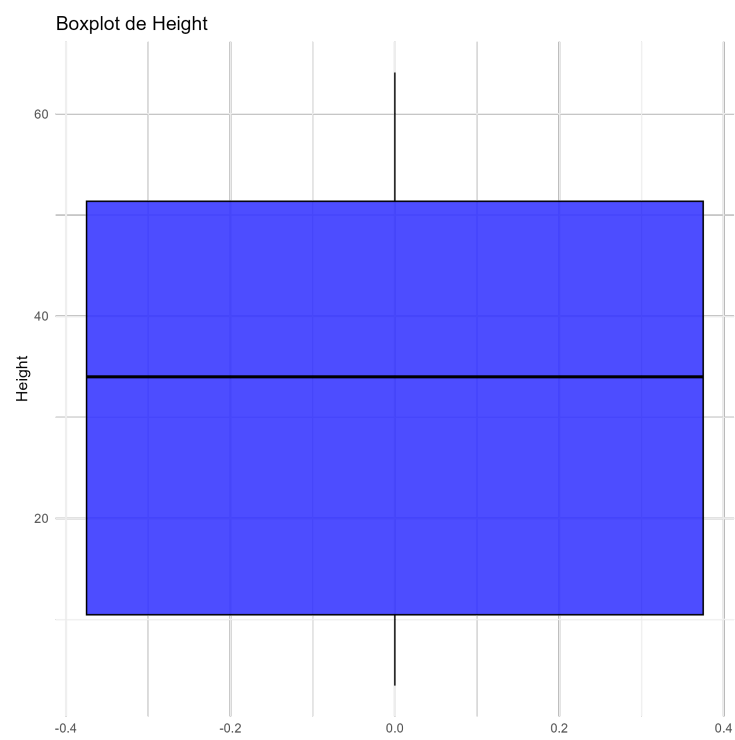


Figure 3: Boxplot de la altura.

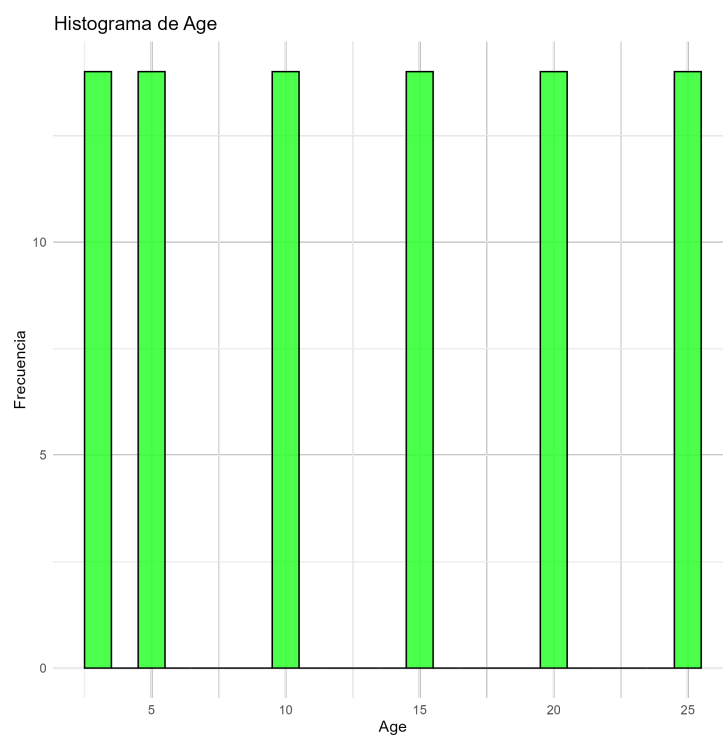


Figure 4: Histograma de la edad.

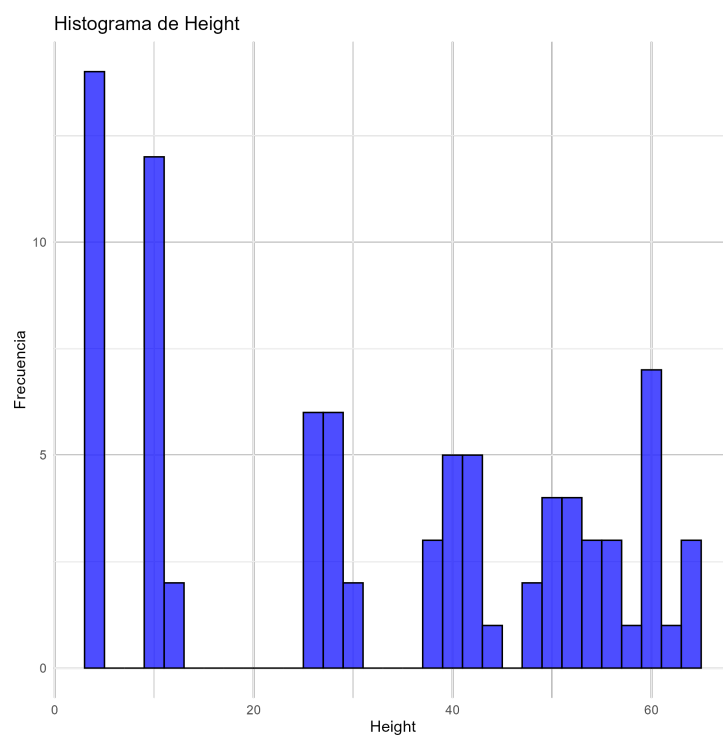


Figure 5: Histograma de la altura.

Measure	Value
Mean	32.3644
Median	34.0000
Mode	3.4600
Variance	427.3979

Figure 6: Resultados del análisis de altura.

Measure	Value
Mode	329

Figure 7: Resultados de la semilla.