

# ACTIVIDAD USANDO PYTHON EN R CON RETICULATE

El dataset seleccionado contiene información de 1000 usuarios, con información sobre su edad, estado civil y más. También incluye si compraron una bicicleta o no.

## Configuramos el entorno virtual.

```
# Configurar el Entorno de Python
message("--- Configurando entorno de Python con reticulate ---")

# Ruta del entorno de python
python_env_path <- "C:/EnviromentR1"

# Ejecutar el entorno virtual
tryCatch({
  use_virtualenv(virtualenv = python_env_path, required = TRUE)
  message(paste("Usando el entorno de Python:", py_config()$python))

  # Opcional: Instalar paquetes Python si no están en el entorno
  # Este paso solo es necesario la primera vez para asegurar que tienes los paquetes Python necesarios
  py_install(packages = c("pandas", "numpy", "scikit-learn"), envname = python_env_path)
  message("Paquetes Python (pandas, numpy, scikit-learn) verificados/instalados.")
}, error = function(e) {
  message("Error al inicializar el entorno Python:")
  message(e$message)
  message("Por favor, asegúrate de que la ruta al entorno virtual sea correcta y que esté funcional.")
  stop("El script se detiene debido a un error de configuración de Python.")
})
```

## Importamos los módulos de Python necesarios y convertimos nuestro dataframe de R a Python.

```
#Importar módulos de Python
pd <- import("pandas")
np <- import("numpy")
sns <- import("seaborn")
plt <- import("matplotlib.pyplot")

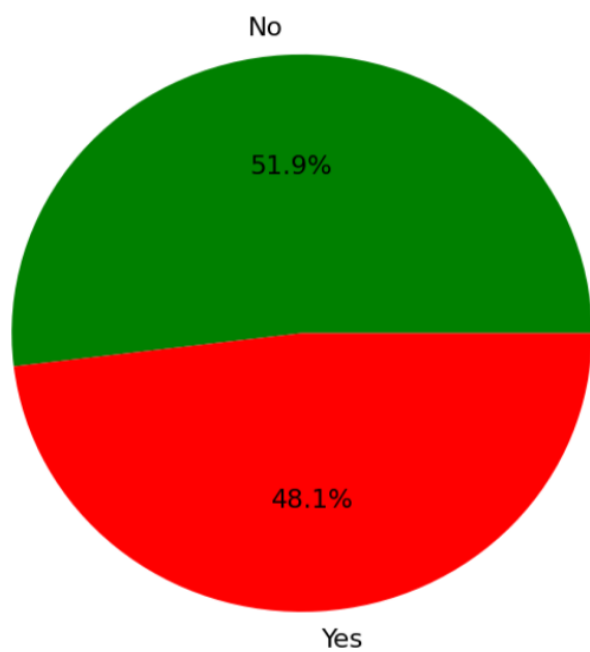
# Leemos el Dataset en R
df <- read_csv("C:/EnviromentR1/bike_buyers.csv")

#Convertir el data frame de R a un DataFrame de Pandas
df_py <- r_to_py(df)
message("DataFrame de R convertido a Pandas DataFrame.")

# Primeras 5 filas
```

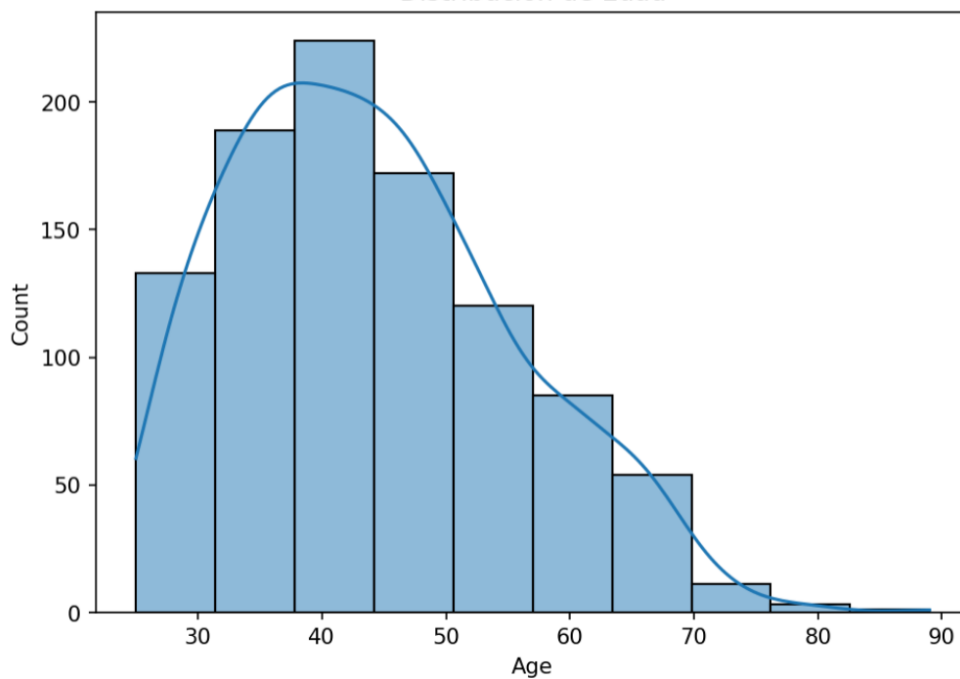
## Distribución de variable objetivo

Distribución de Compra de Bicicleta

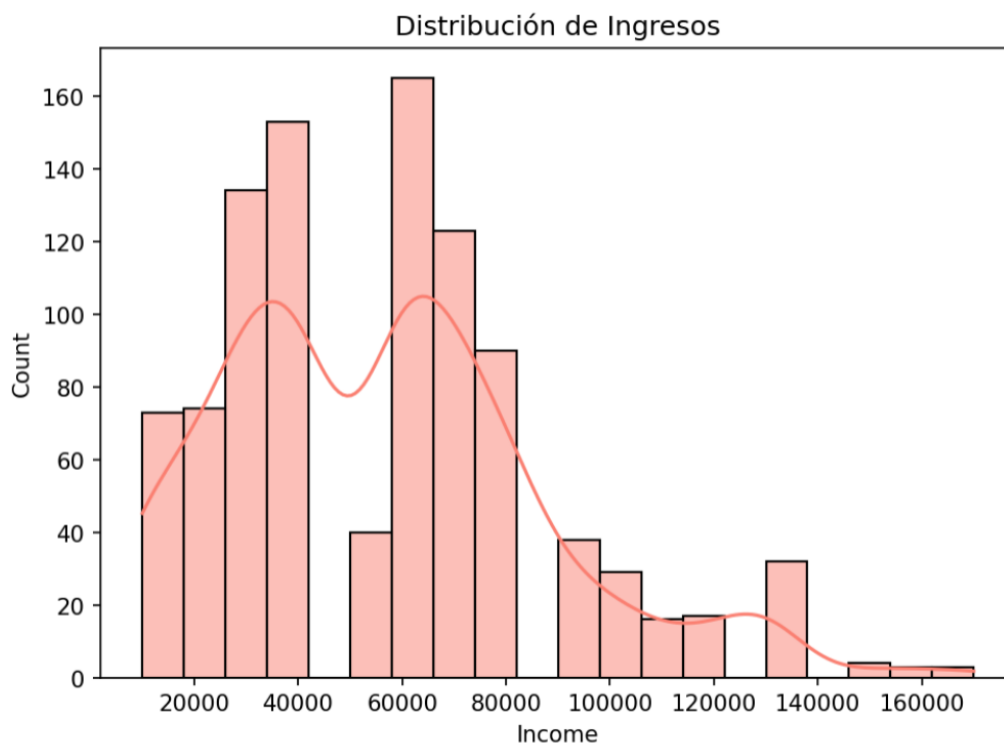


## Histograma de edades

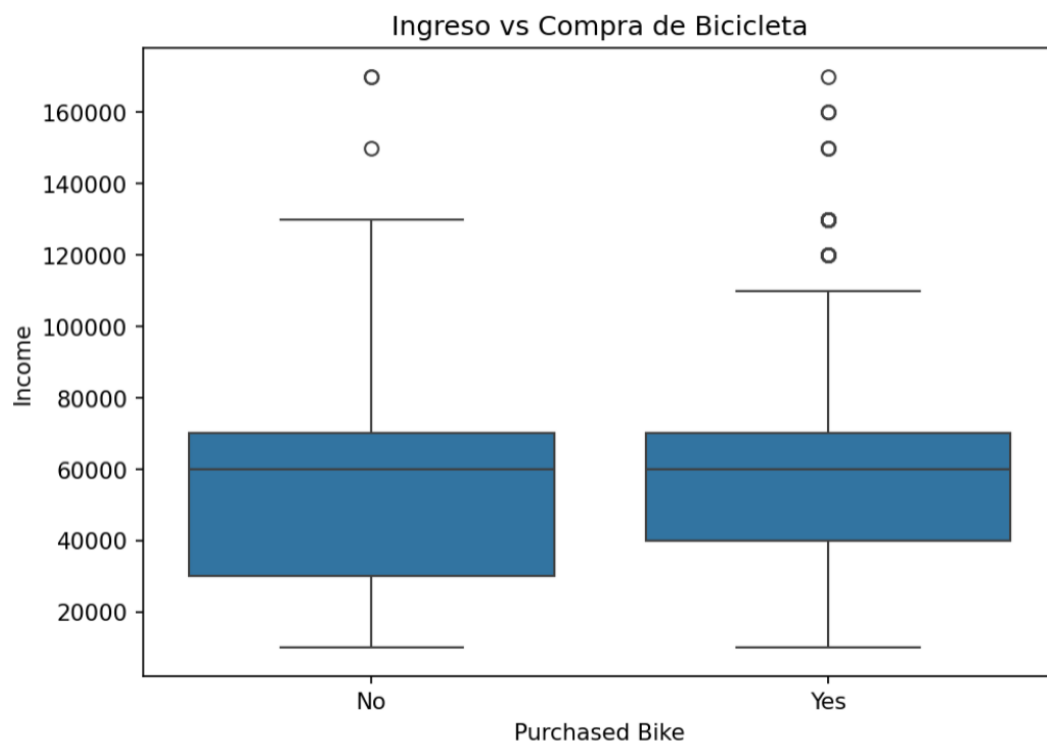
Distribución de Edad



## Histograma de ingresos



## Boxplot de ingreso por compra de bicicleta



## **Conclusiones**

Compras de bicicleta: La mayoría no compró bicicleta.

Ingresos: Hay una gran variabilidad. Quienes compraron bicicleta tienden a tener ingresos más altos (según el boxplot).

Edad: Las compras se concentran entre los 30 y 50 años. Las bicicletas parecen ser más populares entre adultos de mediana edad, con ingresos medios, pocas responsabilidades familiares (0-2 hijos) y que poseen 1 o 2 autos. por lo cual la compra de bicicleta es para ejercicio, recreación y no para sustituir su transporte principal.

Para el siguiente paso se puede tomar en cuenta las variables más relevantes, realizar un mejor tratamiento con los datos nulos y crear un modelo predictivo para saber si un cliente comprará o no.