



**BOOTCAMP DE
DATA SCIENCE**

BIT

**BUILD
INNOVATE
TRANSFORM**

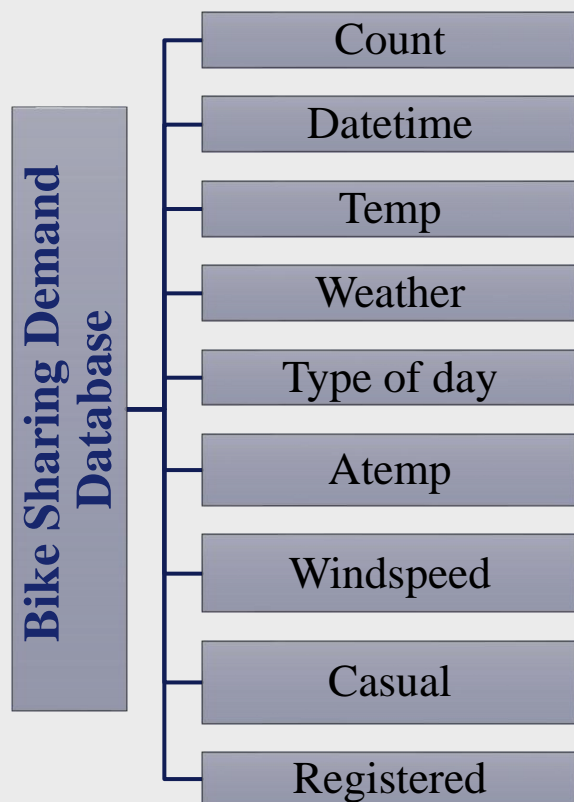
IMPLEMENTACIÓN DE ALGORITMOS PARA LA PREDICCIÓN DE DEMANDA DE ALQUILER DE BICICLETAS EN WASHINGTON D.C.

PRESENTADO POR:

PAOLA CASTRO



INFORMACIÓN DE LA BASE DE DATOS



División de la base de datos

- **Train:** 10886 filas y 15 columnas
- **Validation:** 6493 filas y 9 columnas



OBJETIVO DEL PROYECTO

Motivación

- Mejorar el servicio a los clientes.

Audiencia interesada

- Directivos de Capital Bikeshare.

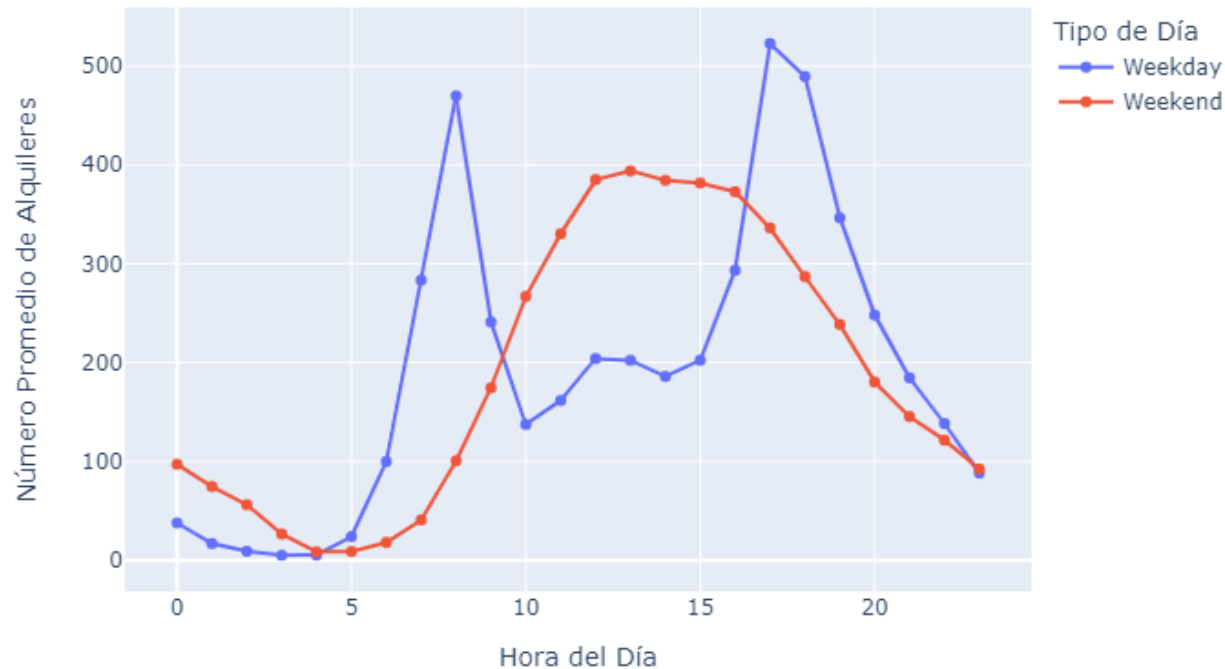
Contexto comercial

- Aumentar alcance y la calidad del servicio



ANÁLISIS EXPLORATORIO DE LOS DATOS

Demanda Promedio por Hora: Días Laborables vs. Fines de Semana



Horas pico

- **Weekday:** 8 a.m. y 5 p.m.
- **Weekend:** Entre las 9 a.m. y 5 p.m.

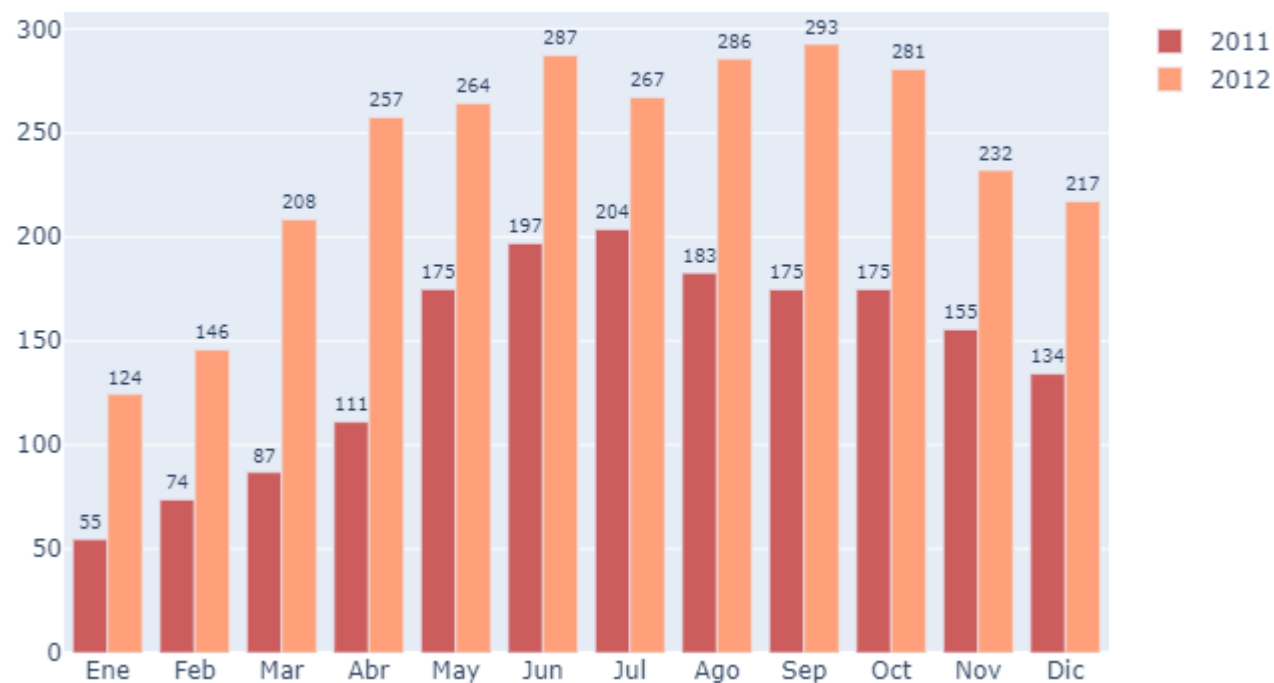


ANÁLISIS EXPLORATORIO DE LOS DATOS

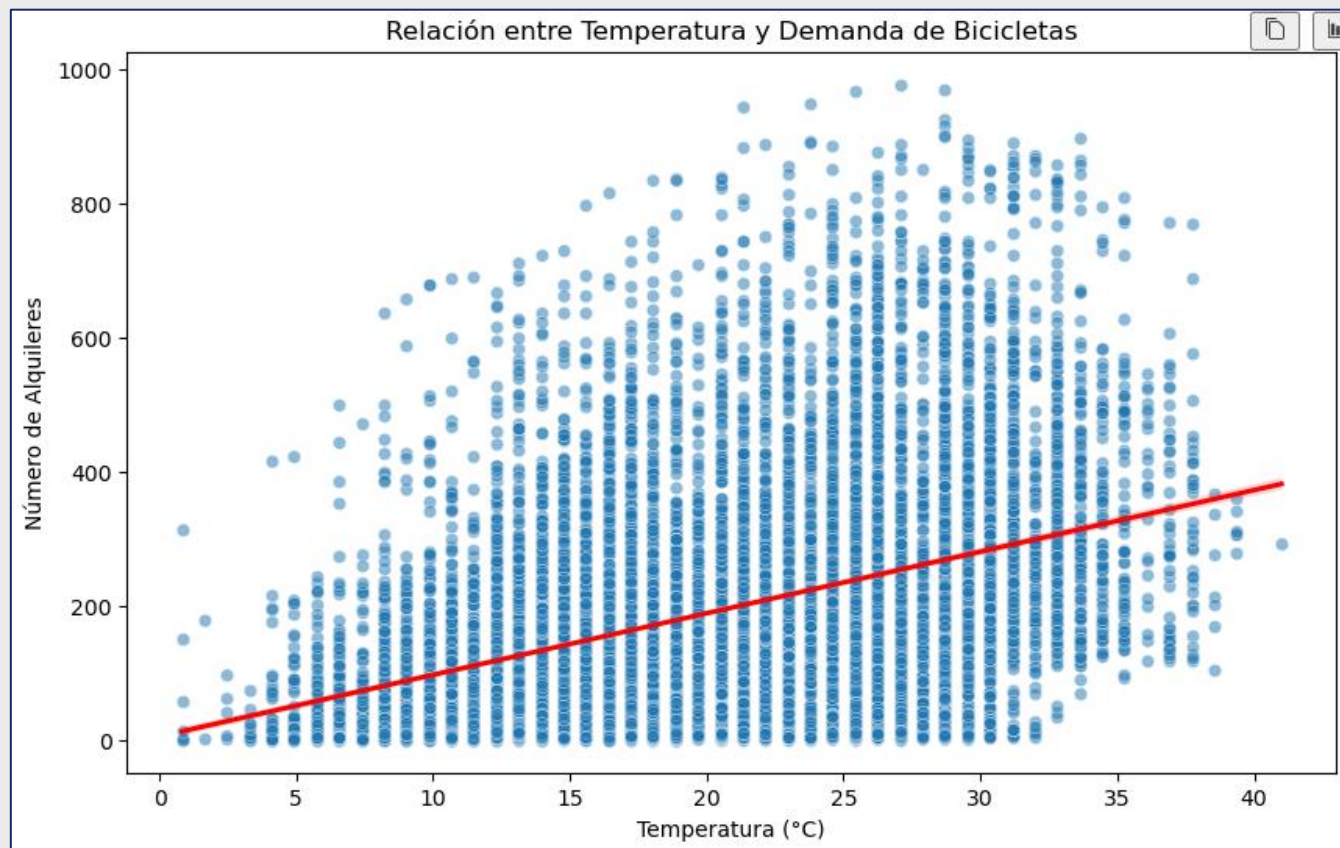
**Sube la demanda de
bicicletas
significativamente**



Demanda Promedio Mensual por Año (2011 vs 2012)

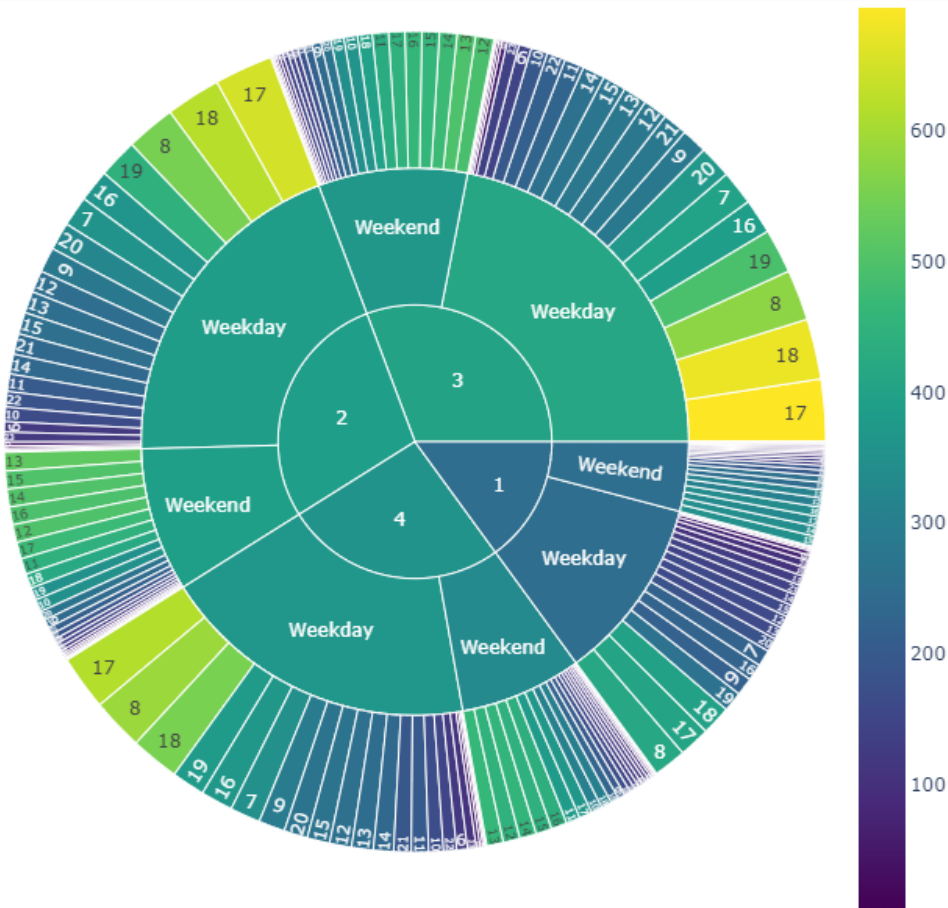


ANÁLISIS EXPLORATORIO DE LOS DATOS



- **Pendiente positiva**
- **Coefficiente de correlación de 0.39**

ANÁLISIS EXPLORATORIO DE LOS DATOS



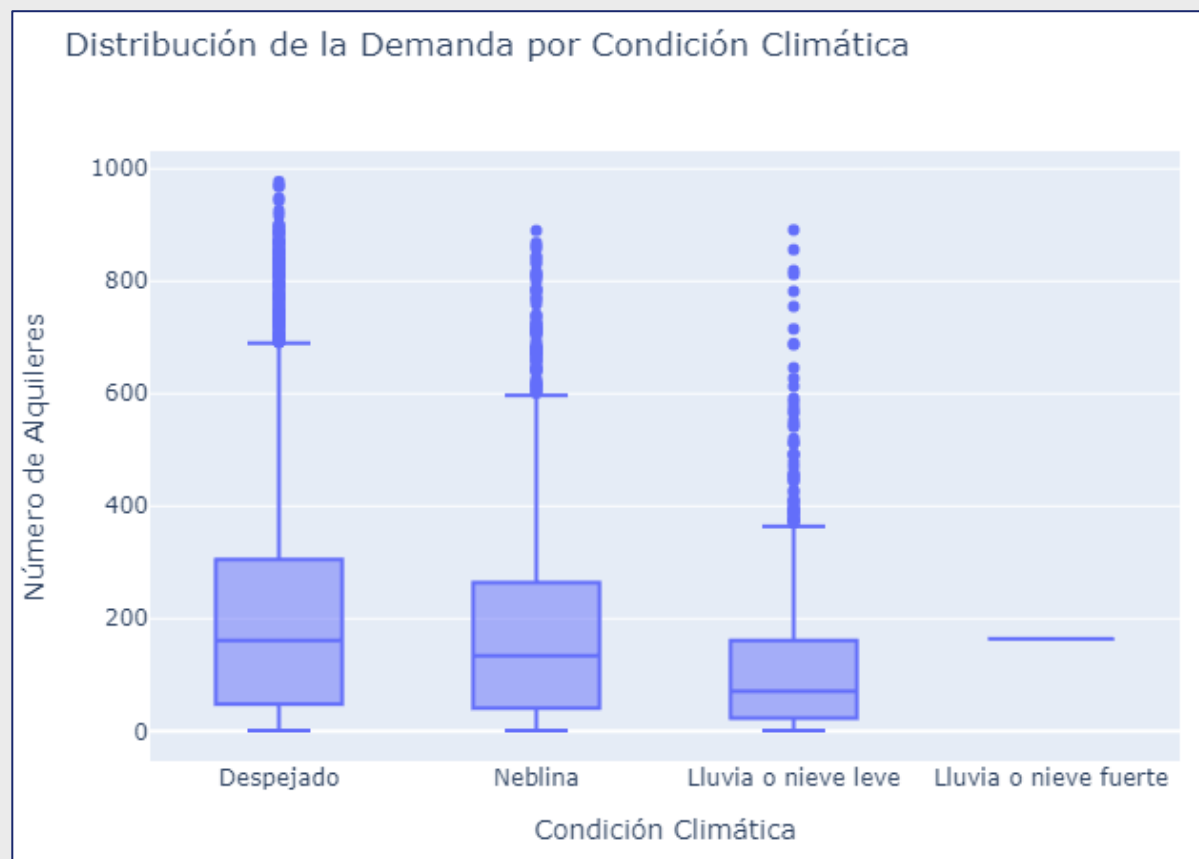
Importancia de la gráfica:

- Estaciones del año.
- Tipo de día.
- Hora del día.



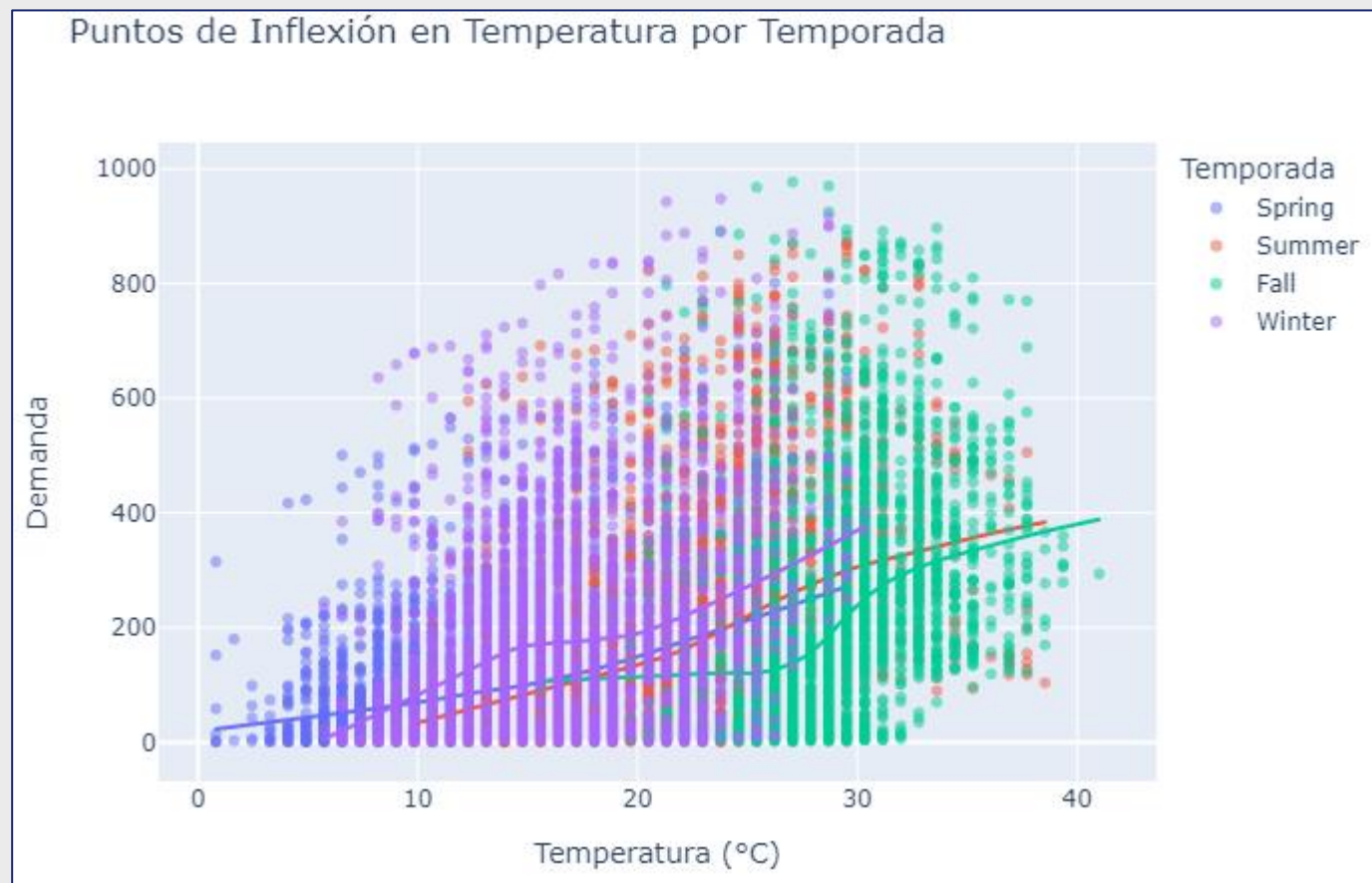
ANÁLISIS EXPLORATORIO DE LOS DATOS

Cuando la condición climática es favorable, mayor número de alquileres.



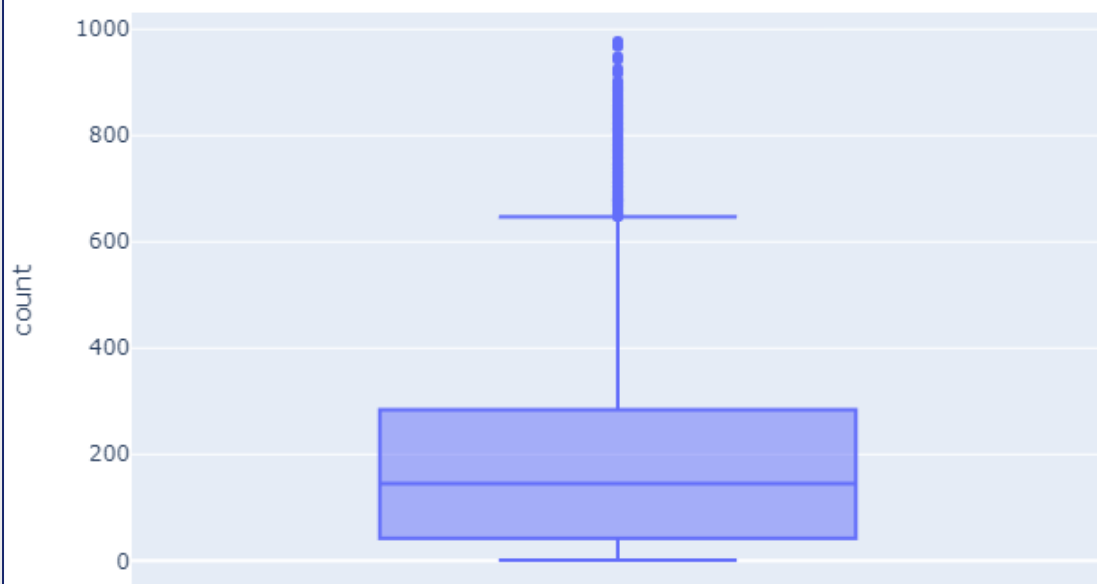
ANÁLISIS EXPLORATORIO DE LOS DATOS

Asegurar disponibilidad de bicicletas a temperaturas más altas (entre los 20°C y 30°C).

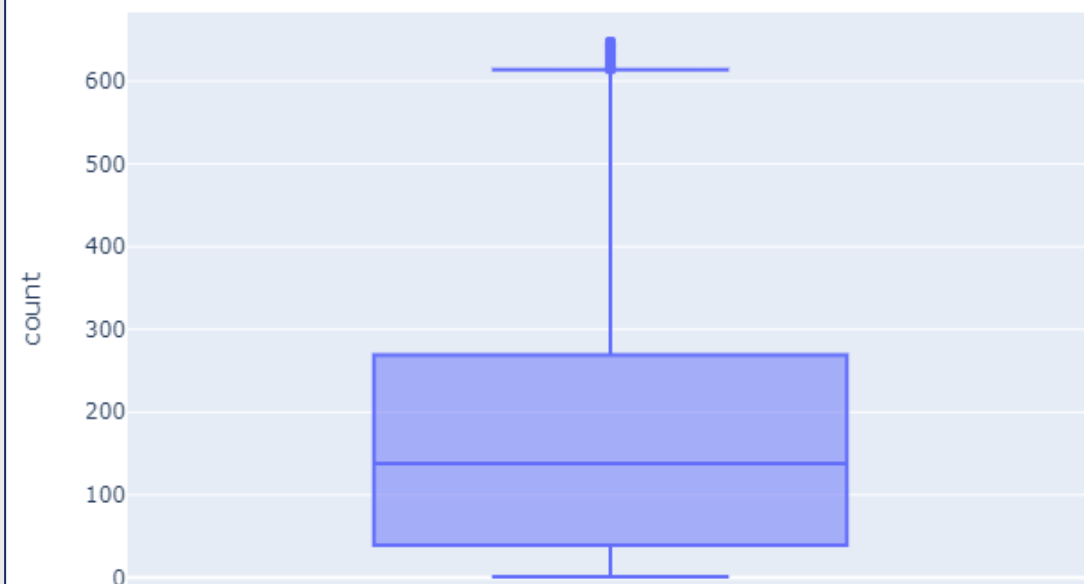


LIMPIEZA DE DATOS

Boxplot de Count para Identificar Outliers

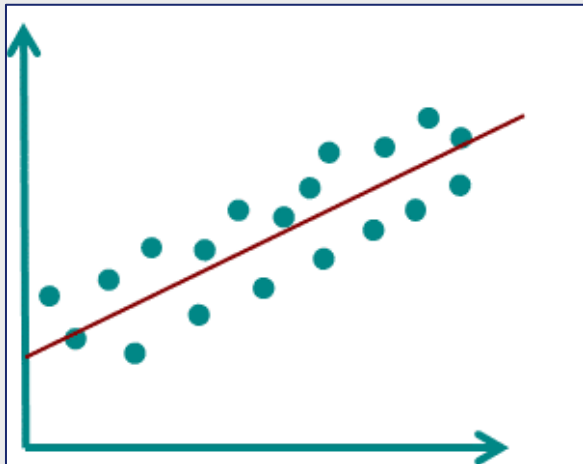


Boxplot de Count en Dataset Limpio

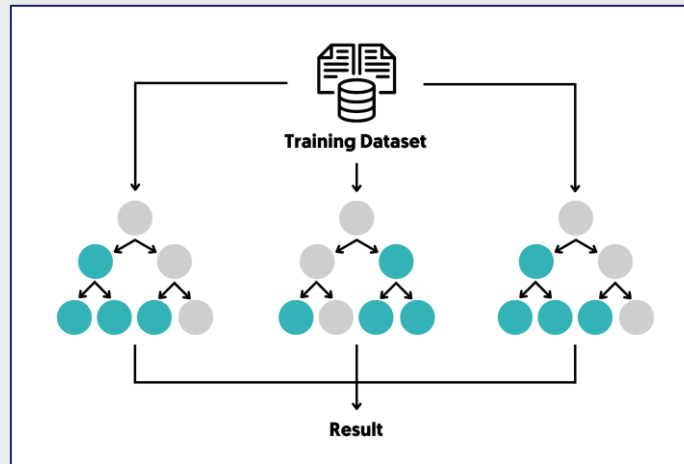


ALGORITMOS

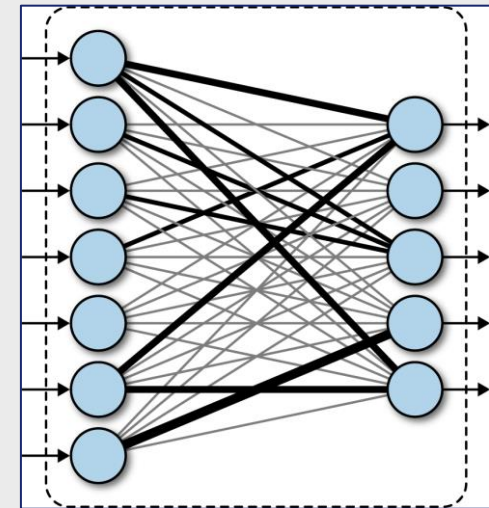
Linear Regression



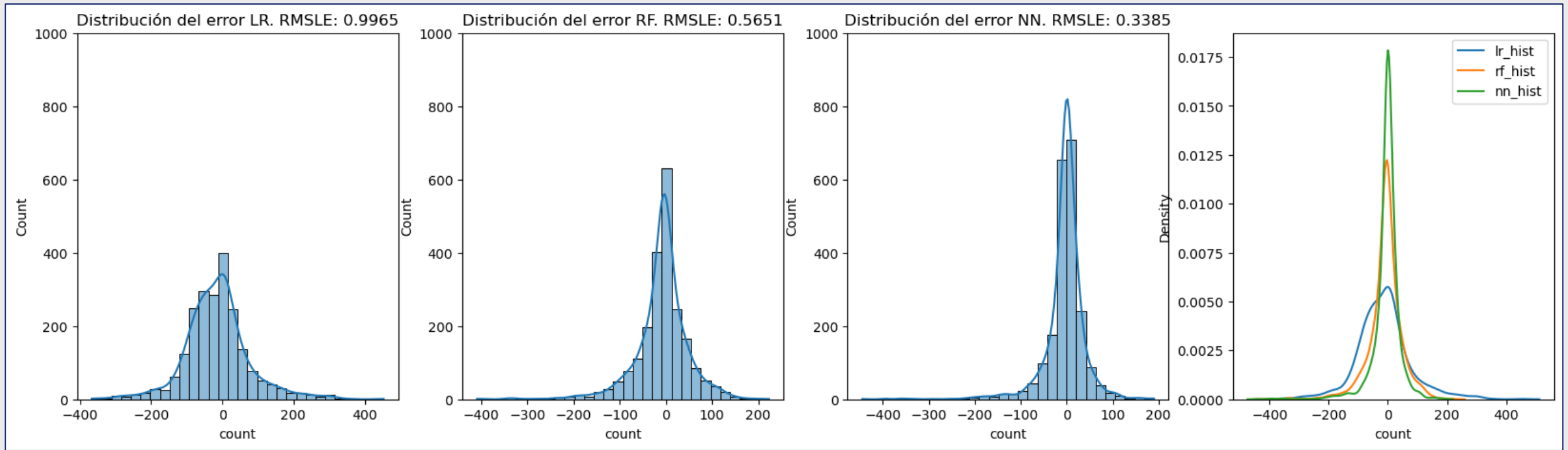
Random Forest



**Fully connected
Neural Network**



RESULTADOS



PRUEBA DEL MODELO

