

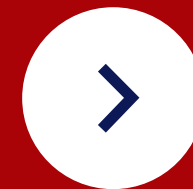
---

# ANÁLISIS Y CLASIFICACIÓN DE MOVIMIENTOS DE PRECIOS DE ACCIONES DE TSLA

PREPARED BY:

**GABRIEL KOHEN**

---



TESLA

# INTRODUCCIÓN

En este proyecto, analizaremos datos históricos del periodo 2020-2022 de Tesla, aplicando técnicas de preprocesamiento, exploración visual y entrenamiento de diversos modelos. Nos ayudará a mostrar a los inversionistas si les conviene invertir en Tesla o no. Mi objetivo es clasificar la variación de precios en la columna "Price Variation" para determinar si, en un día dado, la acción sube o baja.

# PREPROCESAMIENTO DEL DATASET

Columnas Originales:

- Date, Close, Open, High, Low, y Unnamed 8.

Nuevas columnas :

- Open Comparison, Price Variation, Day, Year, Month

Eliminación de la primera fila :

Date	Close	Open	Open Comparison	Price Variation	High	Low	Volumen	unnamed 8
1/9/2020	31.88	32.12	32.90	SUBE	32.33	31.58	194.65M	NaN

# DATAFRAME

**El dataframe consta de 997 filas y 10 columnas.**

COLUMNA	DESCRIPCIÓN
CLOSE	Precio de cierre del activo financiero al final del día.
OPEN	Precio de apertura del activo financiero al inicio del día.
OPEN COMPARATION	Tiene la misma informacion que la Columna Open con una dato menos
PRICE VARIATION	Indica si el precio SUBE o BAJA.
HIGH	Representa el precio más alto de las acciones para el día de la transacción correspondiente.
LOW	Representa el precio más bajo de las acciones para el día de la transacción respectiva.
VOLUMEN	Representa el volumen de transacciones para el día de la transacción correspondiente.
DAY	Representa Día del mes.
MONTH	Representa el Mes.
YEAR	Representa el Año.

# PREPROCESAMIENTO DE DATOS

Ingeniería de Características de la columna Date:

- Extracción de Día, Mes y Año de la Fecha.

Limpieza y Conversión de Datos:

- Eliminación de la Columna unnamed 8.
- Eliminación de la Columna Date.

Manejo de Datos Categóricos:

- Conversión de Variación de Price Variation.

# ANÁLISIS EXPLORATORIO DE DATOS (EDA)

Análisis de valores duplicados:

- Identificación y evaluación de la presencia de valores duplicados en el conjunto de datos.

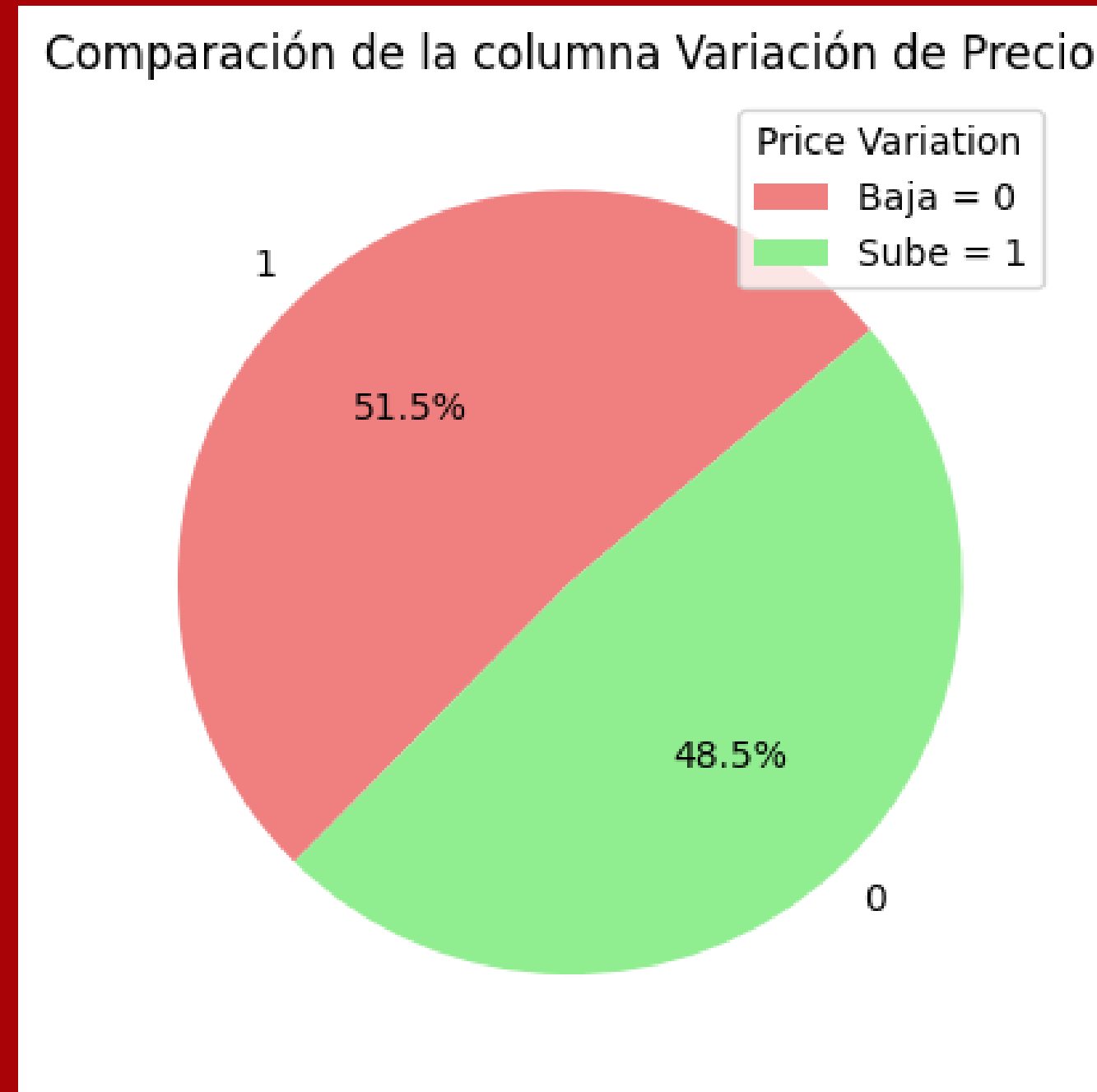
Análisis de valores nulos:

- Evaluación de la existencia de valores nulos o faltantes en las columnas.

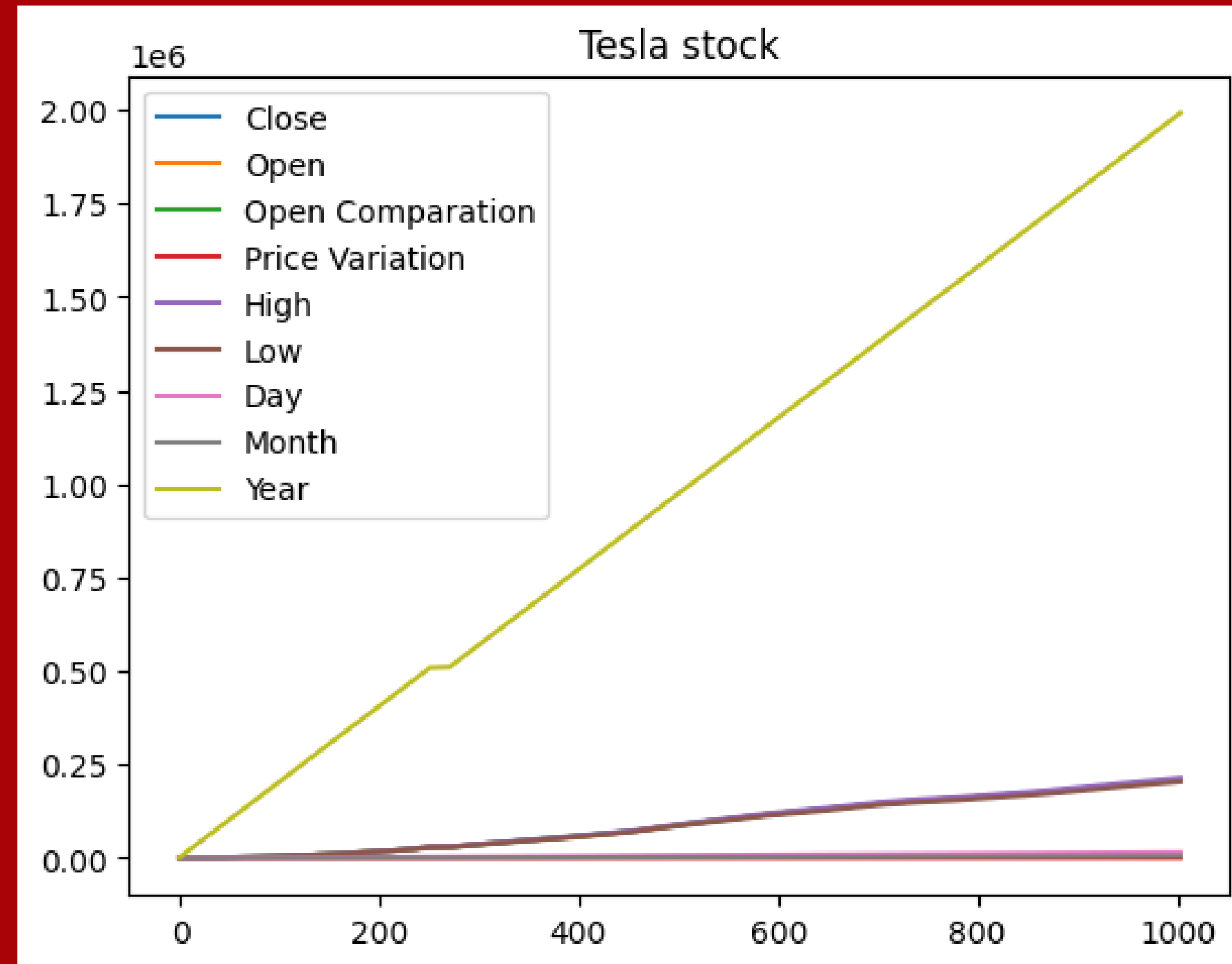
Visualización gráfica:

- Generación de 3 gráficos para mejorar la comprensión de las relaciones entre variables y aspectos significativos del conjunto de datos.

# VISUALIZACIÓN DE LA COLUMNA PRICE VARITION

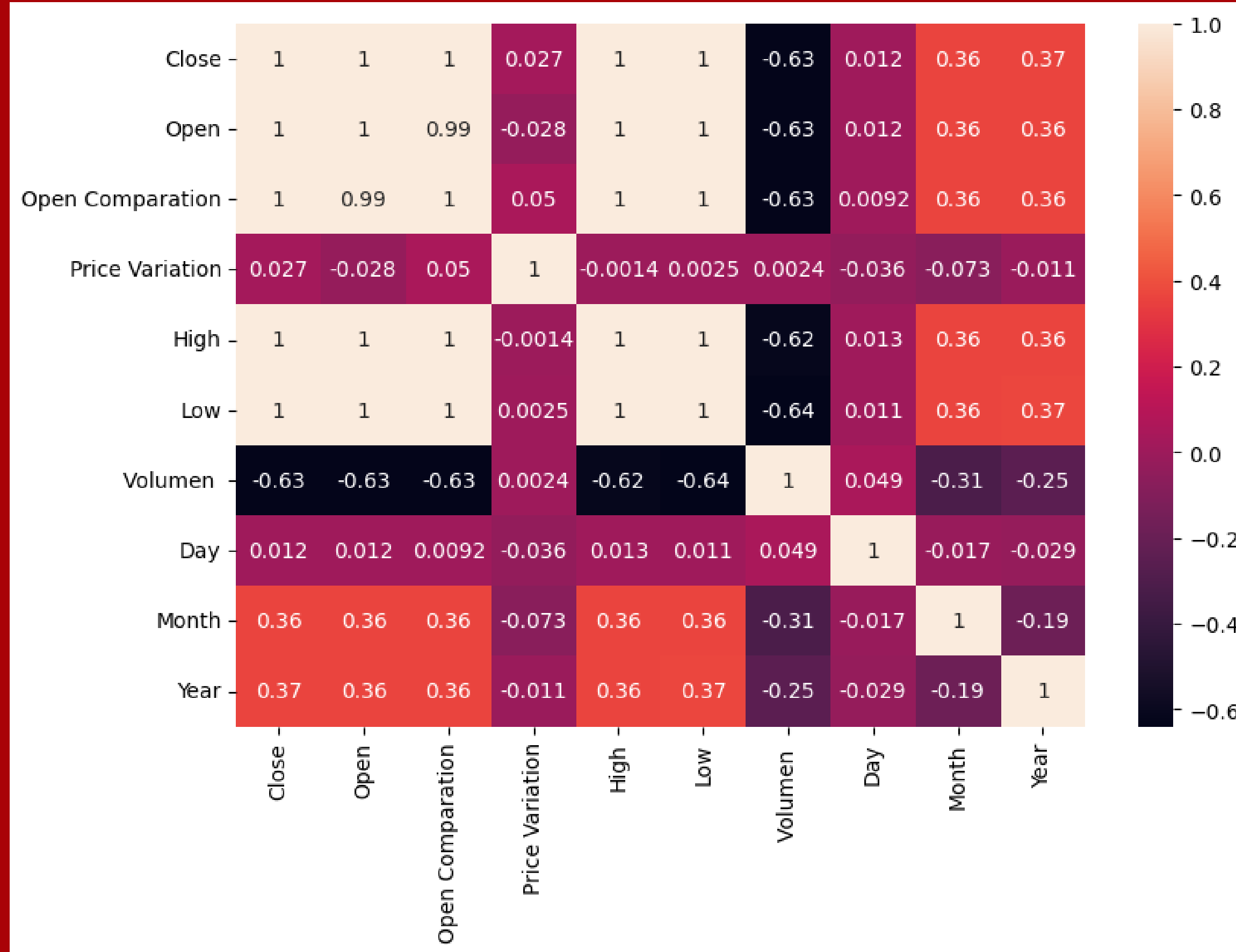


# VISUALIZACIÓN DE TODAS LAS COLUMNAS





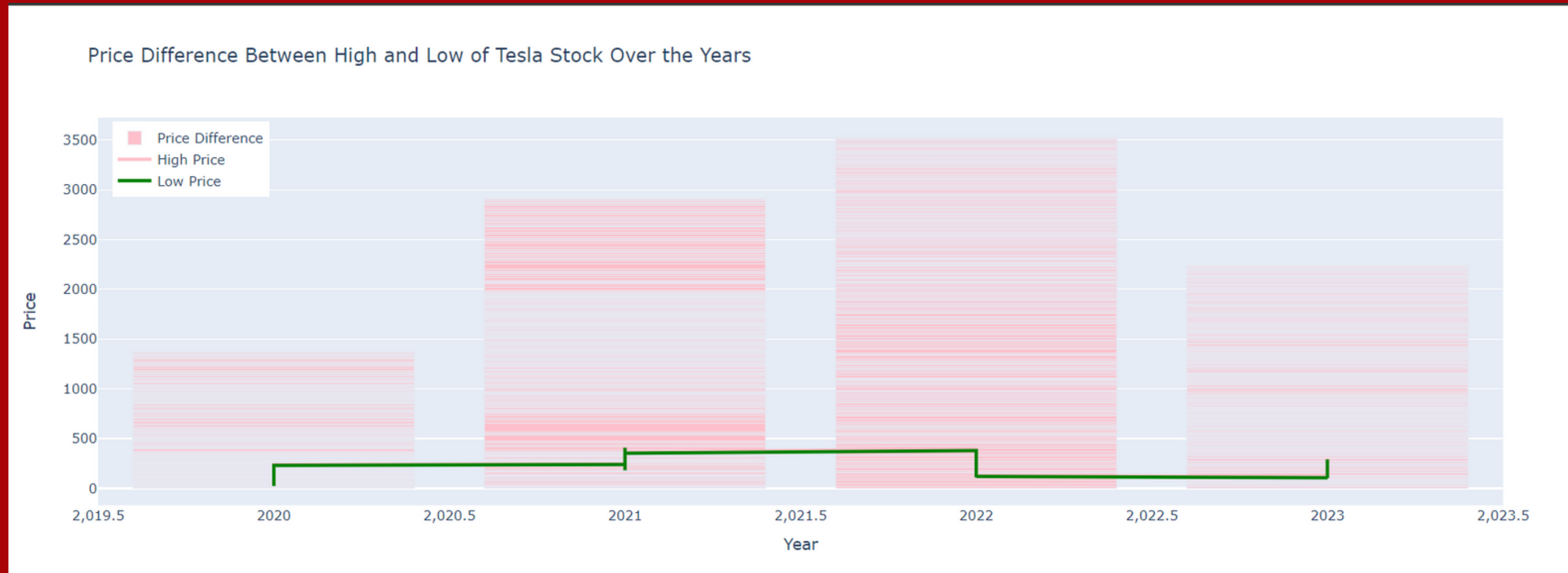
# VISUALIZACIÓN DE MAPA DE CALOR



# VISUALIZACIÓN DE LAS COLUMNAS OPEN Y CLOSE



# VISUALIZACIÓN DE LAS COLUMNAS LOW Y HIGH



# MODELOS Y MÉTRICAS

MODELOS	METRICAS	GRIDSEARCH
LogisticRegression	Accuracy: 0.81   Recall: 0.8069 Precision:0.827 f1_score:0.8069	Accuracy: 0.8818 Recall: 0.8808 Precision: 0.8829 f1_score: 0.8814
KNeighborsClassifier	Accuracy : 0.4764   Recall : 0.4768 Precision : 0.4768 f1_score : 0.4717	Accuracy: 0.6554   Recall: 0.6557 Precision: 0.6555 f1_score: 0.6553

# CONCLUSIÓN

**MUCHAS GRACIAS!**

**Alguna consulta?**