



PROYECTO FINAL DATA SCIENCE I

NESTOR IVAN MILLAN NUÑEZ

FECHA: 20/07/2024

INTRODUCCION

- El conjunto de Datos brinda información sobre las Ventas de vehículos realizadas durante los Años 2022 y 2023.
- El Dataset brinda información útil que puede ser analizada por la Empresa, por ejemplo, tenemos la Fecha de venta, también se detalla información sobre la marca de Vehículo que se adquirió, los Ingresos Anuales de cada comprador, el género, etc. Todo esto nos podrá brindar información sobre preferencias al momento de escoger un modelo de vehículo en particular.
- . Con éste Dataset se buscará encontrar cualquier tipo de relación e información que motive la compra o adquisición de una marca o un vehículo en particular, a los fines de generar información valiosa para la Empresa y los vendedores

PREVISUALIZACION

- Realizamos una previsualización de los datos a los fines de encontrar errores, datos nulos, variables sin nombre, faltantes, etc

```
df.head()
```

	Car_id	Date	Customer Name	Gender	Annual Income	Dealer_Name	Company	Model	Engine	Transmission
0	C_CND_000001	2022-01-02	Geraldine	Male	13500	Buddy Storbeck's Diesel Service Inc	For			
1	C_CND_000002	2022-01-02	Gia	Male	1480000	C & M Motors Inc	Dodge			
2	C_CND_000003	2022-01-02	Gianna	Male	1035000	Capitol KIA	Cadilla			

```
df.columns
```

```
Index(['Car_id', 'Date', 'Customer Name', 'Gender', 'Annual Income', 'Dealer_Name', 'Company', 'Model', 'Engine', 'Transmission', 'Color', 'Price ($)', 'Dealer_No ', 'Body Style', 'Phone', 'Dealer_Region'], dtype='object')
```

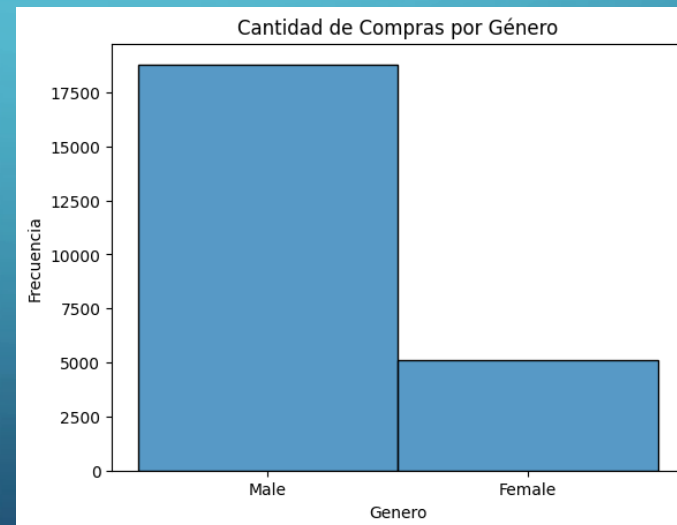
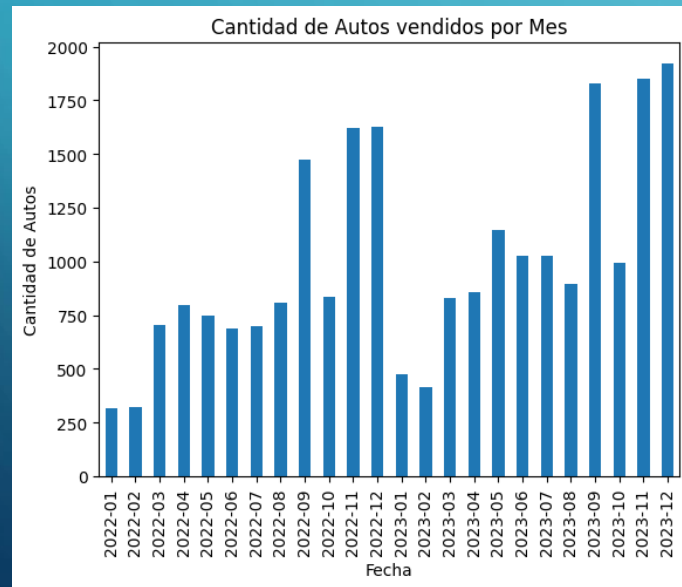
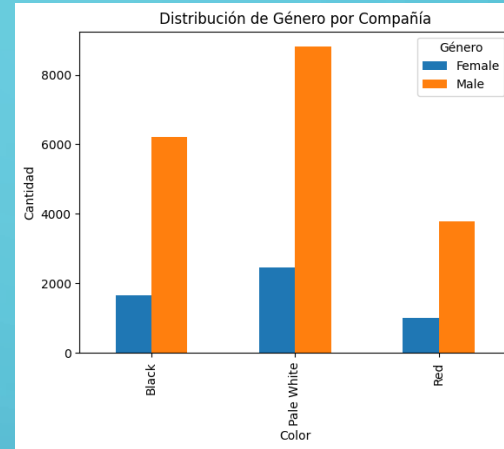
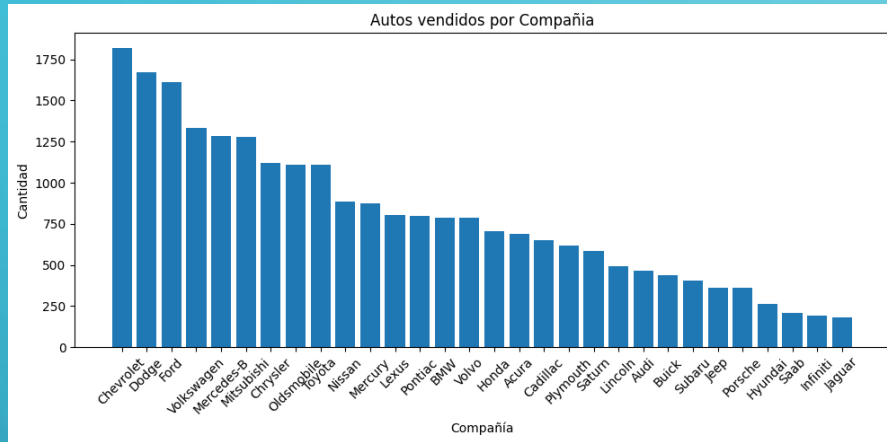
```
df.columns = ["Identificacion_auto", "Fecha", "Nombre_Cliente", "Genero", "Ingreso_Anual", "Nombre"]  
print(df)
```

	Identificacion_auto	Fecha	Nombre_Cliente	Genero	Ingreso_Anual	\
0	C_CND_000001	2022-01-02	Geraldine	Male	13500	
1	C_CND_000002	2022-01-02	Gia			
2	C_CND_000003	2022-01-02	Gianna			

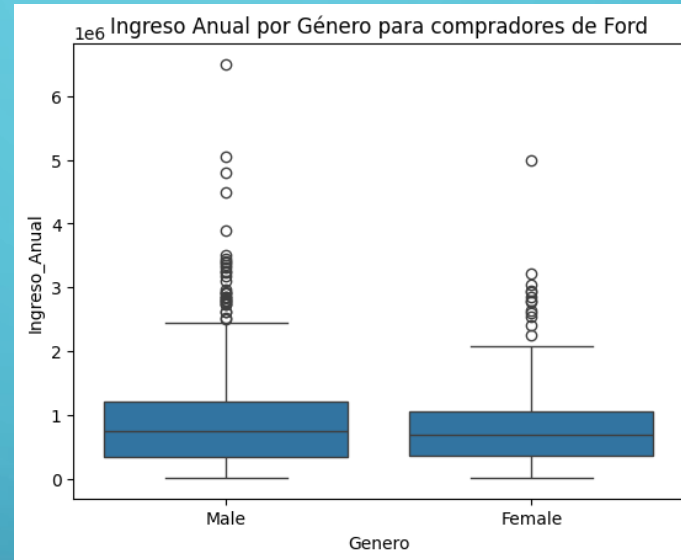
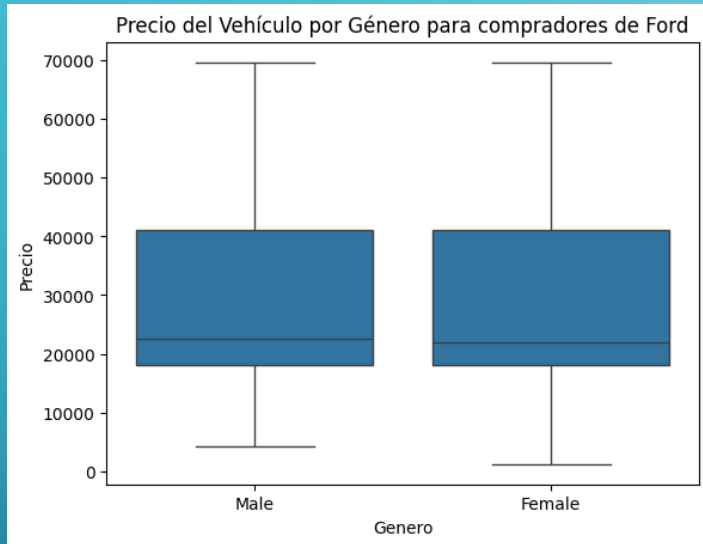
```
df.describe()
```

	Fecha	Ingreso_Anual	Precio
count	23906	23906	23906
mean	2023-03-01 14:28:10.822387456	830840	28090
min	2022-01-02 00:00:00	10080	1200

GRÁFICOS



MACHINE LEARNING



Accuracy: 0.13747908533184608
Precision: 0.1371888717654097
Recall: 0.13747908533184608
F1-Score: 0.1360041456723252

El modelo demuestra tener una precisión del 13,72% siendo bajo, con lo cual se determina que el Ingreso Anual, el Precio del Vehículo o el Género no influyen al momento de la compra de vehículos, ya que no hay una predisposición clara en estos aspectos