



Documentación Proyecto Final

Vehículos en la India

Data Analytics

Sebastian Montenegro M.

# Índice

Índice .....	2
Introducción .....	3
Desarrollo .....	4
ETL Dataset Vehículos en la India .....	4
Bases de datos .....	9
Diagrama E-R .....	12
Conclusión .....	13

## Introducción

Este proyecto consiste en realizar un análisis de datos con el archivo “used\_cars\_dataset.csv”, este dataset trata sobre vehículos usados en la India el cual contiene datos como el tipo de gama, transmisión, combustibles, propietarios, potencia del motor, etc. Se realizará un proceso de ETL el cual implica transformación y limpieza de datos, con el fin de presentar 3 tableros con gráficos, métricas y filtros mediante la herramienta de Power BI, los tableros mostraran información desde la mas específica a la más general respectivamente.

Link del dataset: <https://www.kaggle.com/datasets/ayushparwal2026/cars-dataset>

# Desarrollo

En esta sección se detallarán todos los procedimientos realizados en el dataset.

## ETL Dataset Vehículos en la India

### Parte 1:

Para comenzar, se utiliza el archivo “used\_cars\_data.csv” para estructurar el dataset a utilizar en Power Bi. Se utilizan funciones de Excel y herramientas de Power Query en conjunto para formar al esqueleto del dataset y al modelo de datos, en donde se crean las tablas “Marca”, “Ciudad”, “Transmisión”, “Propietario” y “Combustible” con sus respectivos ID, además se hacen las modificaciones necesarias y se utilizan fórmulas de Excel para vincular el dataset principal con las otras 5 tablas. Para finalizar con la primera parte, el archivo se guarda con el nombre y formato “dataset\_powerbi.xlsx”.

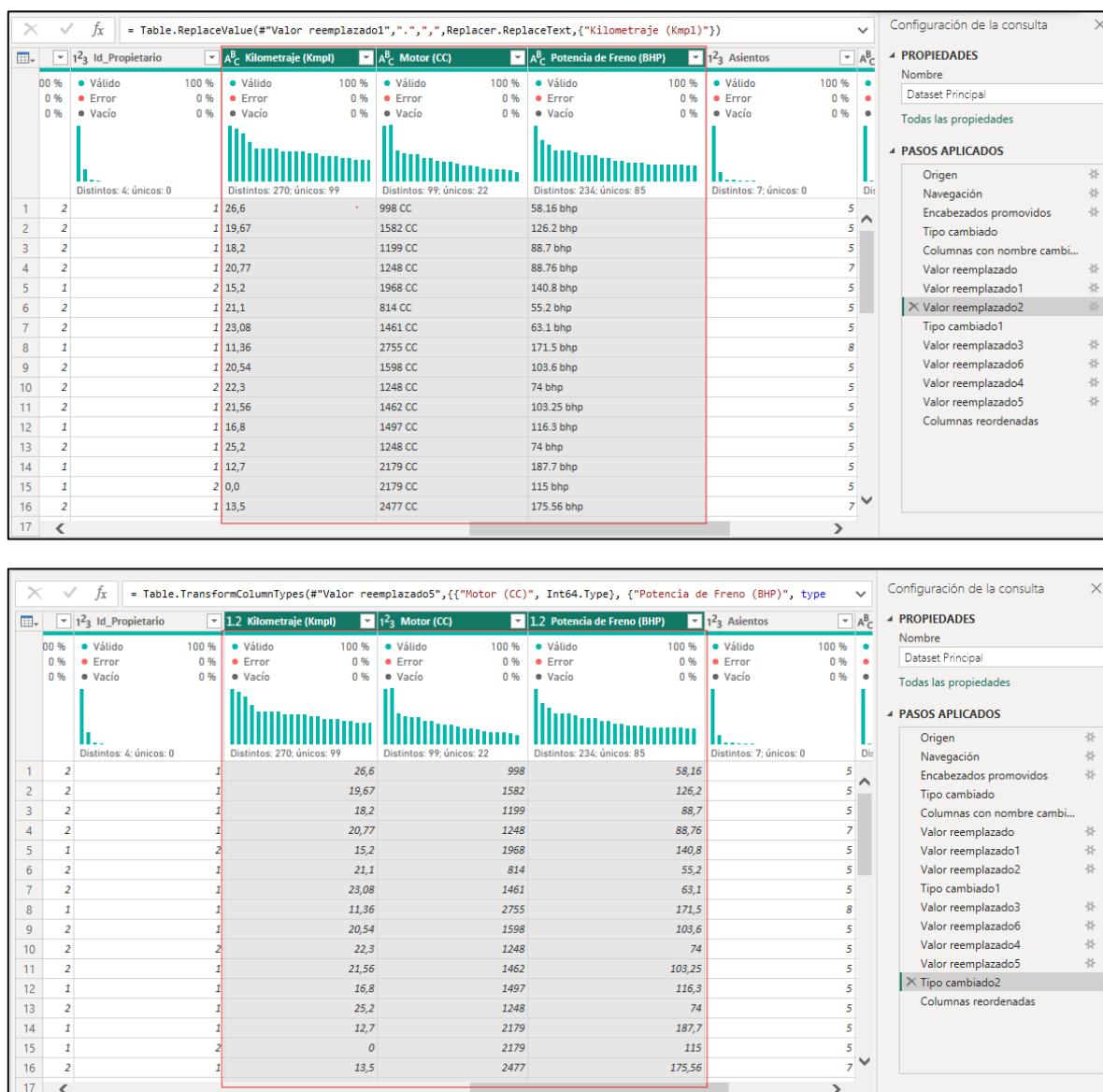


A1	S.No.	Name	Location	Year	Kilometers_Driven	Fuel_Type	Transmission	Owner_Type	Mileage	Engine	Power	Seats	New_Price	Price
1	1	Maruti Wagon R LXI CNG	Mumbai	2010	72000	CNG	Manual	First	26.6 kmpl	938 CC	58.16 bhp	5	1.75	
2	1	Hyundai Creta 1.6 CRDI SX Option	Pune	2015	41000	Diesel	Manual	First	19.67 kmpl	1582 CC	126.2 bhp	5	12.5	
3	2	Honda Jazz V	Chennai	2012	45000	Petrol	Manual	First	18.2 kmpl	1991 CC	88.7 bhp	5	5.61 Lakh	
4	3	Audi A4 2.0 TDI Multitronic	Pune	2012	42000	Diesel	Automatic	Second	19.2 kmpl	2000 CC	140.8 bhp	5	17.74	
5	4	Audi A4 2.0 TDI Multitronic	Chennai	2013	40970	Diesel	Automatic	Second	19.2 kmpl	1998 CC	140.8 bhp	5	17.74	
6	5	Honda EON LPG Era Plus Option	Hyderabad	2012	75000	LPG	Manual	First	21.1 kmpl	934 CC	55.2 bhp	5	2.35	
7	6	Nissan Micra Diesel VX	Jaipur	2013	86393	Diesel	Manual	First	23.03 kmpl	1461 CC	63.1 bhp	5	3.5	
8	7	Tata Indica Vista Crysta 2.8 GX AT	Mumbai	2016	36000	Diesel	Automatic	First	11.36 kmpl	2755 CC	171.5 bhp	8.21 Lakh	17.5	
9	8	Volkswagen Vento Diesel Comfortline	Pune	2013	64430	Diesel	Manual	Second	22.3 kmpl	1248 CC	74 bhp	5	5.195	
10	9	Tata Indica Vista Quadrajet LS	Chennai	2012	65932	Diesel	Manual	Second	22.3 kmpl	1248 CC	74 bhp	5	5.195	
11	10	Maruti Ciaz Zeta Kochi	2019	25692	Petrol	Manual	First	21.56 kmpl	1462 CC	103.25 bhp	5	10.65 Lakh	9.95	
12	11	Honda City 1.5 V AT Sunroof	Kolkata	2012	80000	Petrol	Automatic	First	16.8 kmpl	1497 CC	116.3 bhp	5	4.43	
13	12	Maruti Swift VDI BSIV	Jaipur	2015	64424	Diesel	Manual	First	25.2 kmpl	1248 CC	74 bhp	5	5.6	
14	13	Land Rover Range Rover 2.2L Pure	Pune	2014	72000	Diesel	Automatic	First	12.7 kmpl	2179 CC	187.7 bhp	5	27	
15	14	Land Rover Freelander 2 TD4 SE	Pune	2012	85000	Diesel	Automatic	Second	0.0 kmpl	2179 CC	115 bhp	5	17.5	
16	15	Mitsubishi Pajero Sport 4x4	Bengaluru	2014	100000	Diesel	Manual	First	13.5 kmpl	2477 CC	175.56 bhp	7.32.01 Lakh	15	
17	16	Honda Amaze 1.5 V AT	Kochi	2016	45000	Petrol	Automatic	First	25.8 kmpl	1497 CC	103.25 bhp	5	3.4	
18	17	Maruti Swift DZire 1.0 BS6 VDI	Jaipur	2017	72000	Diesel	Manual	First	28.0 kmpl	1248 CC	62.5 bhp	5	3.4	
19	18	Pearson Delta 1.0 BS6 Petrol	Pune	2014	74560	Petrol	Manual	First	20.45 kmpl	1461 CC	83.18 bhp	5	3.34	
20	19	Mercedes-Benz New C-Class 220 CDI BE Avantgarde	Bangalore	2014	77463	Diesel	Automatic	First	14.94 kmpl	2110 CC	167.62 bhp	5	28	
21	20	BMW 5 Series 320d	Kochi	2014	32382	Diesel	Automatic	First	22.69 kmpl	1995 CC	190 bhp	5	47.87 Lakh	18.55
22	21	Maruti S Cross DZire 200 Alpha	Bangalore	2015	55332	Diesel	Manual	Second	23.65 kmpl	1248 CC	88.5 bhp	5	5.25	
23	22	Hyundai i20 1.2 Magna	Kochi	2010	45807	Petrol	Manual	First	13.53 kmpl	1984 CC	177.01 bhp	5	23.5	
24	23	Hyundai i20 1.2 Magna AT	Kochi	2012	51907	Petrol	Automatic	First	12.9 kmpl	1977 CC	117.7 bhp	5	1.87	
25	24	Volkswagen Vento Petrol Highline AT	Kolkata	2010	33000	Petrol	Automatic	First	14.4 kmpl	1588 CC	103.25 bhp	5	3.25	
26	25	Honda City Corporate Edition	Mumbai	2012	51920	Petrol	Manual	First	23.08 kmpl	1461 CC	63.1 bhp	5	4.25	
27	26	Tata Indica Diesel Comfortline	Pune	2013	64430	Diesel	Manual	Second	20.32 kmpl	1598 CC	67.1 bhp	5	2.75	
28	27	Maruti Alto K10 2010-2014 VXI	Hyderabad	2013	54000	Petrol	Manual	First	17.5 kmpl	1993 CC	88.7 bhp	5	10.57 Lakh	9.9
29	28	Honda WRV i-VTEC VX	Kochi	2016	37430	Petrol	Manual	First	17.5 kmpl	1993 CC	88.7 bhp	5	10.57 Lakh	9.9
30	29	Tata Indica Vista 7-seater	Mumbai	2007	262000	Diesel	Manual	Fourth & Above	12.8 kmpl	2494 CC	102 bhp	7.4		

A1	S.N.	Modelo	Id_Marca	Id_Club	Año	Kilometros_Conducido	Id_Combustible	Id_Transmisi	Id_Propietaria	Kilometri	Mot	Potencia de Fi	Asientos	Precio_Nue	Precio_
1	1	Maruti Wagon R LXI CNG	24	10	2010	72000	1	2	2	1.26.6 kmpl	936 CC	58.16 bhp	5	\$ Sin precio nuevo	17.5
2	1	Creta 1.6 CRDI SX Option	10	11	2015	41000	2	2	2	1.13.67 kmpl	1582 CC	126.2 bhp	5	\$ Sin precio nuevo	12.5
3	2	Jazz V	9	3	2011	46000	4	2	2	1.18.2 kmpl	1199 CC	88.7 bhp	5	\$ 8.61 Lakh	45
4	3	Maruti Ertiga VDI	24	3	2012	87000	2	2	1.20.77 kmpl	1248 CC	88.7 bhp	7	\$ Sin precio nuevo	6	
5	4	A4 New 2.0 TDI Multitronic	10	4	2013	40670	2	1	2.15.2 kmpl	1686 CC	140.8 bhp	5	\$ Sin precio nuevo	177.4	
6	5	EON LPG Era Plus Option	10	6	2012	75000	3	2	1.21.1 kmpl	814 CC	55.2 bhp	5	\$ Sin precio nuevo	23.5	
7	6	Micro Diesel	26	7	2013	86390	2	2	1.23.0 kmpl	1461 CC	63.1 bhp	5	\$ Sin precio nuevo	5.42	
8	7	Indica V2 1.3 G4 AT 8S	26	10	2013	36000	2	2	1.10.54 kmpl	1248 CC	103 bhp	5	\$ 21 Lakh	17.5	
9	8	Vento Diesel Comfortline	27	11	2015	64430	2	2	1.10.54 kmpl	1248 CC	103 bhp	5	\$ Sin precio nuevo	5.42	
10	9	Indica Vista Quadrajet LS	25	3	2012	6532	2	2	1.22.3 kmpl	1248 CC	74 bhp	5	\$ Sin precio nuevo	13.5	
11	10	Maruti Ciaz Zeta	24	8	2018	25892	4	2	1.21.56 kmpl	1462 CC	103.25 bhp	5	\$ 10.65 Lakh	9.95	
12	11	City 1.5 V AT Sunroof	9	9	2012	60000	4	1	1.16.8 kmpl	1437 CC	116.3 bhp	5	\$ Sin precio nuevo	44.9	
13	12	Maruti Swift VDI BSIV	24	7	2015	64424	2	2	1.25.2 kmpl	1248 CC	74 bhp	5	\$ Sin precio nuevo	5.56	
14	13	Range 2.2L Pure	15	5	2014	72000	2	1	1.12.7 kmpl	2179 CC	187.7 bhp	5	\$ Sin precio nuevo	27	
15	14	Freelander 2 TD4 SE	16	11	2012	85000	2	1	2.0.0 kmpl	2179 CC	115 bhp	5	\$ Sin precio nuevo	177.5	
16	15	Pajero Sport 4x4	13	5	2014	100000	2	2	1.13.5 kmpl	2477 CC	175.56 bhp	7.32.01 Lakh	15		
17	16	Amaze S+Dtech	9	8	2016	58950	2	2	1.25.8 kmpl	1498 CC	98.6 bhp	5	\$ Sin precio nuevo	5.4	
18	17	Maruti Swift DZire VDI	24	7	2017	25000	2	2	1.28.4 kmpl	1248 CC	74 bhp	5	\$ Sin precio nuevo	5.99	
19	18	Duster 85PS Diesel Petrol Plus	22	8	2017	77460	2	2	1.14.4 kmpl	1461 CC	103.25 bhp	5	\$ Sin precio nuevo	6.34	
20	19	Maruti Suzuki Celerio 1.0 BS6	17	2	2014	76500	2	1	1.14.4 kmpl	1435 CC	67.1 bhp	5	\$ Sin precio nuevo	2.6	
21	20	Series 200	3	8	2014	32892	2	1	1.22.63 kmpl	1635 CC	180 bhp	5	\$ 47.87 Lakh	185.5	
22	21	Maruti S-Cross DZire 200 Alpha	24	2	2015	55332	2	2	2.23.65 kmpl	1248 CC	88.5 bhp	5	\$ Sin precio nuevo	82.5	
23	22	A6 2011-2015 35 TFSI Technology	1	10	2015	55895	4	1	1.13.53 kmpl	1984 CC	177.01 bhp	5	\$ Sin precio nuevo	23.5	
24	23	i20 1.2 Magna	10	9	2010	45807	4	2	1.18.5 kmpl	1197 CC	60 bhp	5	\$ Sin precio nuevo	16.7	
25	24	Vento Petrol Highline AT	27	9	2010	33000	4	1	1.14.4 kmpl	1598 CC	103.6 bhp	5	\$ Sin precio nuevo	28.5	
26	25	City Corporate Edition	9	10	2012	51920	4	2	1.16.8 kmpl	1437 CC	116.3 bhp	5	\$ Sin precio nuevo	42.5	
27	26	Micro Diesel XV	20	6	2012	54000	2	2	1.23.08 kmpl	1461 CC	63.1 bhp	5	\$ Sin precio nuevo	42.5	
28	27	Maruti Alto K10 2010-2014 VXI	24	6	2013	54000	4	2	2.20.32 kmpl	998 CC	67.1 bhp	5	\$ Sin precio nuevo	27.5	
29	28	WRV i-VTEC VX	9	8	2018	37430	4	2	1.17.5 kmpl	1199 CC	88.7 bhp	5	\$ 10.57 Lakh	9.9	
30	29	Innova 2.5V Diesel 7-seater	26	10	2007	262000	2	2	4.12.8 kmpl	2494 CC	102 bhp	7	\$ Sin precio nuevo	4	

## Parte 2:

En esta parte se realizan modificaciones a los datos mediante Power Query, para empezar se renombran las columnas “Kilometraje”, “Motor” y “Potencia de Freno” añadiendo “Kmpl”, “CC” y “BHP” respectivamente. En la columna “Kilometraje (Kmpl)” se reemplazan los puntos por comas en todos los registros y se cambia a formato decimal, en las columnas “Motor (CC)” y “Potencia de Freno (BHP)” se eliminan los caracteres “CC” y “bhp” respectivamente de todos los registros y se reemplazan los puntos. Para terminar, el formato de la columna “Motor (CC)” es cambiado a numero entero y decimal en la columna “Potencia de Freno (BHP)”.



### Parte 3:

En esta parte se explicará cómo se dio forma a la tabla “Calendario” y las 3 medidas calculadas dentro de la tabla “Dataset Principal”.

- Calendario:

Mediante la función DAX CALENDAR se crea un calendario desde el 01-01-1998 al 31-12-2019.

The screenshot shows the Power BI Data Editor interface. At the top, there is a formula bar with two lines of DAX code: 1. `Calendario = CALENDAR(DATE(1998, 1, 1), DATE(2019, 12, 31))` and 2. An empty line below it. Below the formula bar is a table preview with three columns: "Fecha", "Año", and "Mes". The table contains 7 rows of data, all corresponding to January 1998. The "Fecha" column shows dates from "01-01-1998" to "07-01-1998". The "Año" column shows the year "1998" for all rows. The "Mes" column shows the month "Enero" for all rows.

Fecha	Año	Mes
01-01-1998	1998	Enero
02-01-1998	1998	Enero
03-01-1998	1998	Enero
04-01-1998	1998	Enero
05-01-1998	1998	Enero
06-01-1998	1998	Enero
07-01-1998	1998	Enero

Con la función DAX YEAR se extrae el año desde la columna “Fecha” de la tabla “Calendario”.

The screenshot shows the Power BI Data Editor interface. At the top, there is a formula bar with one line of DAX code: 1. `Año = YEAR(Calendario[Fecha])`. Below the formula bar is a table preview with three columns: "Fecha", "Año", and "Mes". The table contains 10 rows of data, starting from January 1998 and ending at January 1999. The "Fecha" column shows dates from "01-01-1998" to "10-01-1998". The "Año" column shows the year "1998" for all rows. The "Mes" column shows the month "Enero" for all rows. The "Año" column is highlighted with a red border.

Fecha	Año	Mes
01-01-1998	1998	Enero
02-01-1998	1998	Enero
03-01-1998	1998	Enero
04-01-1998	1998	Enero
05-01-1998	1998	Enero
06-01-1998	1998	Enero
07-01-1998	1998	Enero
08-01-1998	1998	Enero
09-01-1998	1998	Enero
10-01-1998	1998	Enero

La columna “Mes” está compuesta por 2 funciones DAX:

Fecha	Año	Mes
01-01-1998	1998	Enero
02-01-1998	1998	Enero
03-01-1998	1998	Enero
04-01-1998	1998	Enero
05-01-1998	1998	Enero
06-01-1998	1998	Enero
07-01-1998	1998	Enero
08-01-1998	1998	Enero
09-01-1998	1998	Enero

-Con la función “**UPPER(LEFT(FORMAT(Calendar[Fecha], "MMMM"), 1))**” se transforma la primera letra del mes en mayúscula.

-Con la función “**LOWER(MID(FORMAT('Calendar'[Fecha], "MMMM"), 2, LEN(FORMAT(Calendar[Fecha], "MMMM"))))**” se transforma las demás letras del nombre del mes en minúsculas.

- **Medidas Calculadas**

Las 3 medidas calculadas se encuentran dentro de la tabla “Dataset Principal” y fueron hechas de la siguiente forma:

-Promedio de kilometrajes: Se utiliza la función DAX AVERAGE en la columna “Kilometraje (Kmlp)” para obtener esta medida.

```
1 Promedio de kilometrajes = AVERAGE('Dataset Principal'[Kilometraje (Kmlp)])
```

- Promedio de kilómetros recorridos: Se utiliza la función DAX AVERAGE en la columna “Kilómetros Conducidos” para obtener esta medida.

```
1 Promedio de kilometros recorridos = AVERAGE('Dataset Principal'[Kilometros Conducidos])
```

- Promedio de potencia de freno: Se utiliza la función DAX AVERAGE en la columna “Potencia de Freno (BHP)” para obtener esta medida.

```
1 Promedio de potencia de freno = AVERAGE('Dataset Principal'[Potencia de Freno (BHP)])
```

Al terminar con las medidas calculadas, dentro de la tabla “Dataset Principal” saldrán los siguientes cálculos.

✓  Dataset Principal
<input type="checkbox"/> Año
<input type="checkbox"/> Asientos
<input type="checkbox"/> Id_Ciudad
<input type="checkbox"/> Id_Combustible
<input type="checkbox"/> Id_Marca
<input type="checkbox"/> Id_Propietario
<input type="checkbox"/> Id_Transmision
<input type="checkbox"/> $\sum$ Kilometraje (Kmpl)
<input type="checkbox"/> $\sum$ Kilometros Conducidos
<input type="checkbox"/> Modelo
<input type="checkbox"/> $\sum$ Motor (CC)
<input type="checkbox"/> $\sum$ Potencia de Freno (BHP)
<input type="checkbox"/> $\sum$ Precio
<input type="checkbox"/> Precio Nuevo
<input type="checkbox"/> Promedio de kilometrajes
<input type="checkbox"/> Promedio de kilometros recorridos
<input type="checkbox"/> Promedio de potencia de freno
<input type="checkbox"/> S.No.

## Bases de datos

### Listado de tablas

Dataset Principal: Esta tabla contiene las características principales de cada vehículo tales como, nombre del vehículo, número de serie, kilómetros recorridos y centímetros cúbicos del motor.

- PK: Número de Serie
- Modelo
- FK: Id\_Marca
- FK: Id\_Ciudad
- Año
- Kilómetros Conducidos
- FK: Id\_Combustible
- FK: Id\_Transmision
- FK: Id\_Propietario
- Kilometraje (Kmpl)
- Potencia de freno (BHP)
- Motor (CC)
- Asientos
- Nuevo Precio
- Precio

Combustible: La tabla Combustible indica cual es el tipo de combustible que utiliza el vehículo tales como, petróleo, diesel o eléctrico.

- PK: Id\_Combustible
- Tipo de Combustible

Transmisión: Esta tabla indica cual es el tipo de transmisión que utiliza el vehículo para funcionar las cuales son automático y manual.

- PK: Id\_Transmision
- Tipo de Transmisión

Ciudad: La tabla ciudad indica donde está ubicado geográficamente el vehículo.

- PK: Id\_Ciudad
- Latitud
- Longitud
- Nombre de Ciudad

**Marca:** Esta tabla indica a que marca pertenecen los vehículos.

- PK: Id\_Marca
- Nombre Marca

**Propietario:** Con esta tabla podremos saber a qué tipo de dueño pertenece un vehículo, nos indicara si es el primer dueño, segundo u otro.

- PK: Id\_Propietario
- Tipo de Propietario

**Calendario:** Esta tabla nos indica de que año es cada vehículo.

- Año
- Fecha
- Mes

#### Listado de columnas por tablas

#### **Dataset Principal**

Campo	Tipo de Campo	Tipo de Clave
Número de Serie	INTEGER	PK
Modelo	CHAR	-
Id_Marca	INTEGER	FK
Id_Ciudad	INTEGER	FK
Año	INTEGER	-
Kilómetros Conducidos	INTEGER	-
Id_Combustible	INTEGER	FK
Id_Transmision	INTEGER	FK
Id_Propietario	INTEGER	FK
Kilometraje (Kmpl)	INTEGER	-
Potencia de Freno (BHP)	INTEGER	-
Motor (CC)	INTEGER	-
Asientos	INTEGER	-
Nuevo Precio	VARCHAR	-
Precio	INTEGER	-

### **Combustible**

<b>Campo</b>	<b>Tipo de Campo</b>	<b>Tipo de Clave</b>
<b>Id_Combustible</b>	INTEGER	PK
<b>Tipo de Combustible</b>	CHAR	-

### **Transmisión**

<b>Campo</b>	<b>Tipo de Campo</b>	<b>Tipo de Clave</b>
<b>Id_Transmision</b>	INTEGER	PK
<b>Tipo de Transmisión</b>	CHAR	-

### **Marca**

<b>Campo</b>	<b>Tipo de Campo</b>	<b>Tipo de Clave</b>
<b>Id_Marca</b>	INTEGER	PK
<b>Nombre Marca</b>	CHAR	-

### **Propietario**

<b>Campo</b>	<b>Tipo de Campo</b>	<b>Tipo de Clave</b>
<b>Id_Propietario</b>	INTEGER	PK
<b>Tipo de Propietario</b>	CHAR	-

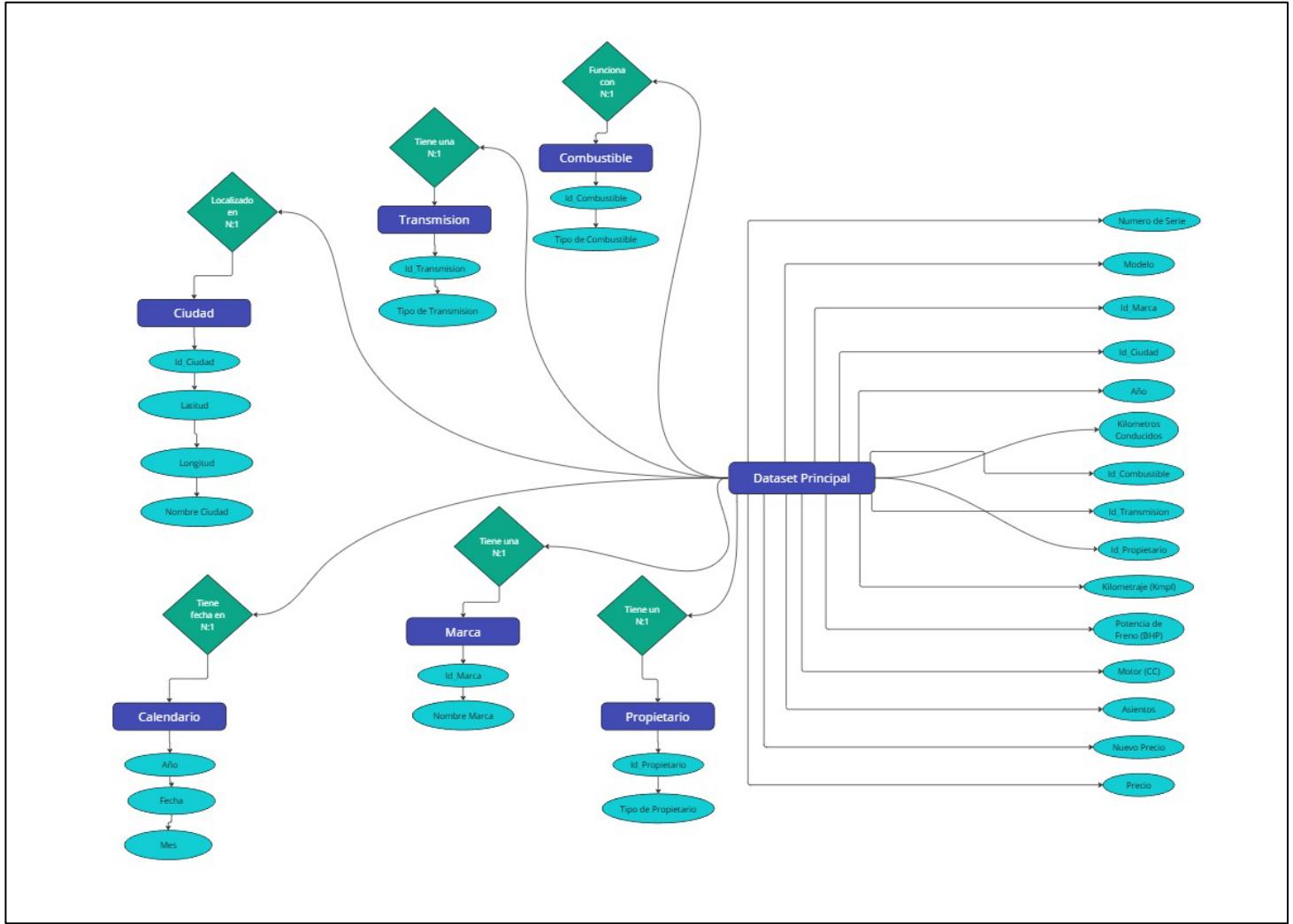
### **Ciudad**

<b>Campo</b>	<b>Tipo de Campo</b>	<b>Tipo de Clave</b>
<b>Id_Ciudad</b>	INTEGER	PK
<b>Latitud</b>	CHAR	-
<b>Longitud</b>	CHAR	-
<b>Nombre Ciudad</b>	CHAR	-

### **Calendario**

<b>Campo</b>	<b>Tipo de Campo</b>	<b>Tipo de Clave</b>
<b>Año</b>	INTEGER	-
<b>Fecha</b>	INTEGER	-
<b>Mes</b>	CHAR	-

## Diagrama E-R



Link del diagrama:

[https://miro.com/app/board/uXjVLI76ZBc=/?share\\_link\\_id=344177486797](https://miro.com/app/board/uXjVLI76ZBc=/?share_link_id=344177486797)

## Conclusión

En conclusión, se puede identificar información muy interesante respecto a los vehículos en la India con la ayuda del dataset y los tableros de Power Bi. Para analizar los tableros, en el primero se puede identificar que las 3 marcas que lideran en cuanto a promedio de precios son Porsche, Land Rover y Audi, además los vehículos con mayores kilómetros recorridos funcionan con Diesel y en la ciudad de Coimbatore se ubican la mayor cantidad de vehículos de menor precio liderando las marcas Hyundai y Susuki.

Referente al segundo tablero, se logra observar en primer lugar que las marcas con los motores más potentes Bentley, Mercedes-Benz y Lamborghini, además los vehículos con mas potencia de freno serian a Diesel y Petróleo y la mayor gama en marcas se encuentra en Nueva Delhi.

En el tercer y último tablero se puede observar de manera más general, que las marcas que tiene mayor variedad de modelos son Susuki, Hyundai y Honda respectivamente, también las marcas mas escasas en cuanto a su cantidad son Lamborghini, Bentley e Isuzu además los vehículos de transmisión manual son los que más kilómetros recorridos en promedio tienen.