

PROYECTO DE IRONHACK POR: CARLOS CUERDA

CALCULADORA DE PRECIOS

Estimación del precio de tu vehiculo en el mercado actual





ÍNDICE GENERAL

- 01** MI PROYECTO
- 02** PUNTOS CLAVES
- 03** DATOS GENERALES
- 04** ESTADÍSTICAS
- 05** INFORMACIÓN TÉCNICA
- 06** ASPECTOS A MEJORAR
- 07** AGRADECIMIENTOS Y BIBLIOGRAFÍA

NUESTRO PROCESO DE ESTIMACIÓN ONLINE

1



Entra en nuestra web:

www.calculadoracoches.es



2



Contesta a 5 sencillas
preguntas.



3



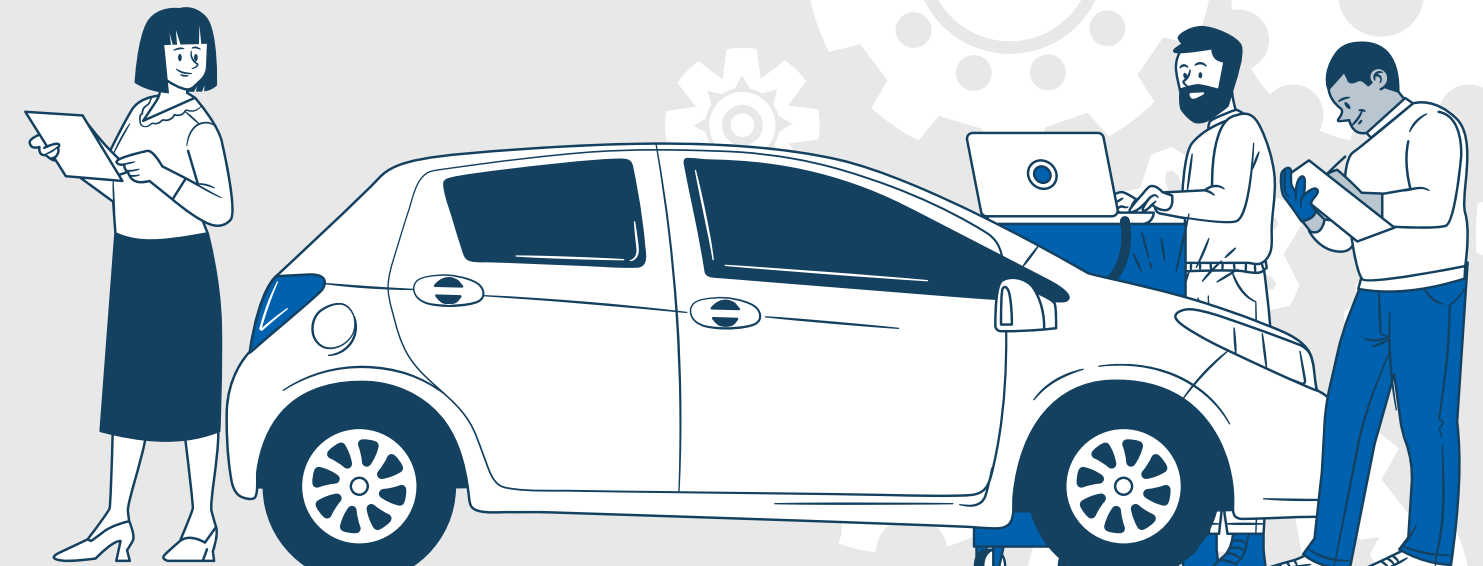
Recibe una estimación
directa del precio de tu
coche en el mercado
actual.

No pierdas tiempo
No hay que darse de alta
No necesitamos conocer tu matrícula



WWW.CALCULADORACOCHESES.ES

PUNTOS CLAVE



1 ACCESIBILIDAD
PARA EL USUARIO.

2 VELOCIDAD DE
ESTIMACION DEL
PRECIO.

3 5 PREGUNTAS

4 OPCIONES DE
MEJORA.

Datos Generales

SOBRE EL DATASET

558837 filas x 16 columnas
No tiene valores duplicados pero si Nan
5 columnas numéricas 11 object
Target: Price

YEAR

Año de fabricación del vehículo

CAR YEAR

Años de uso del vehículo

MARKET PRICE

Precio de mercado del vehículo

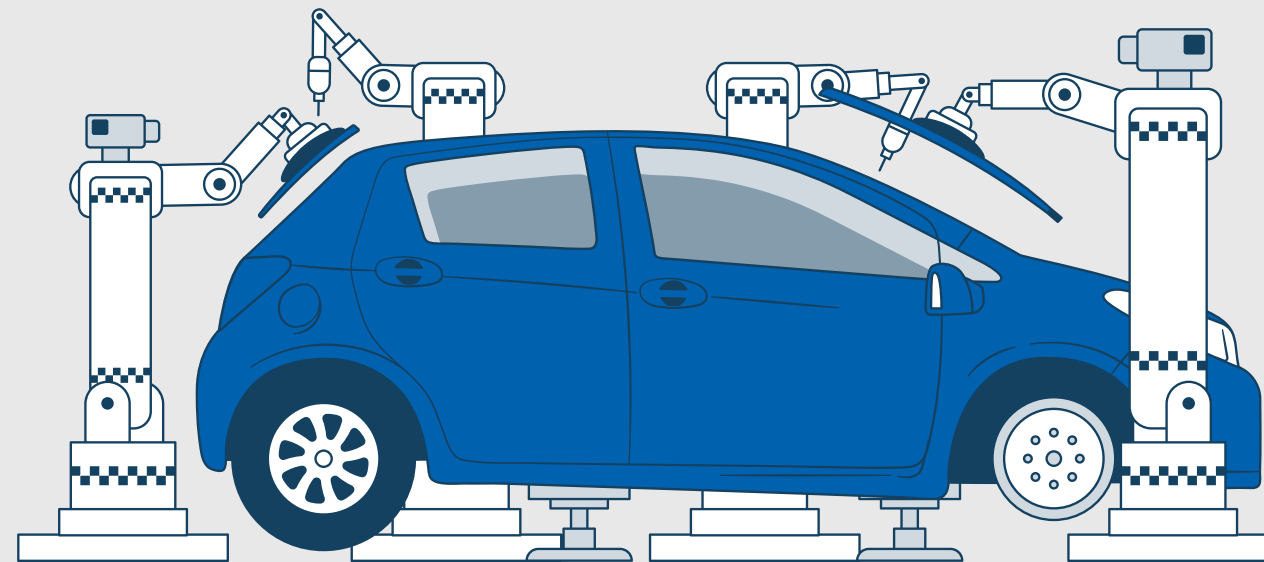
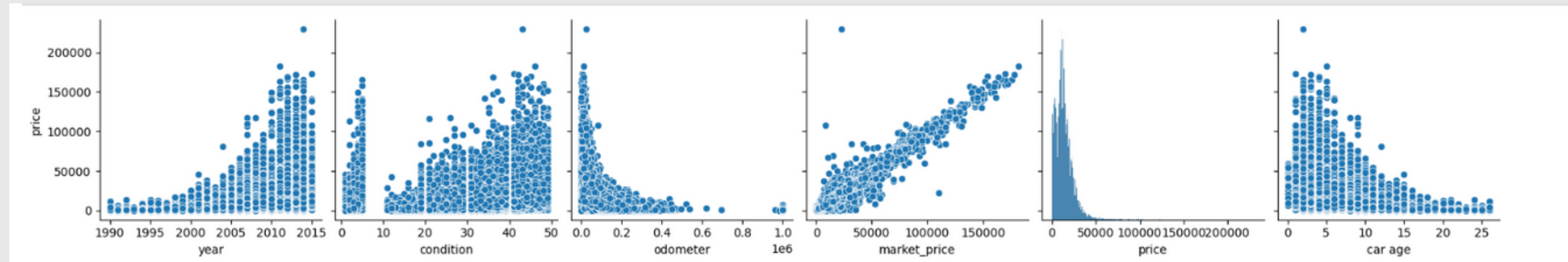
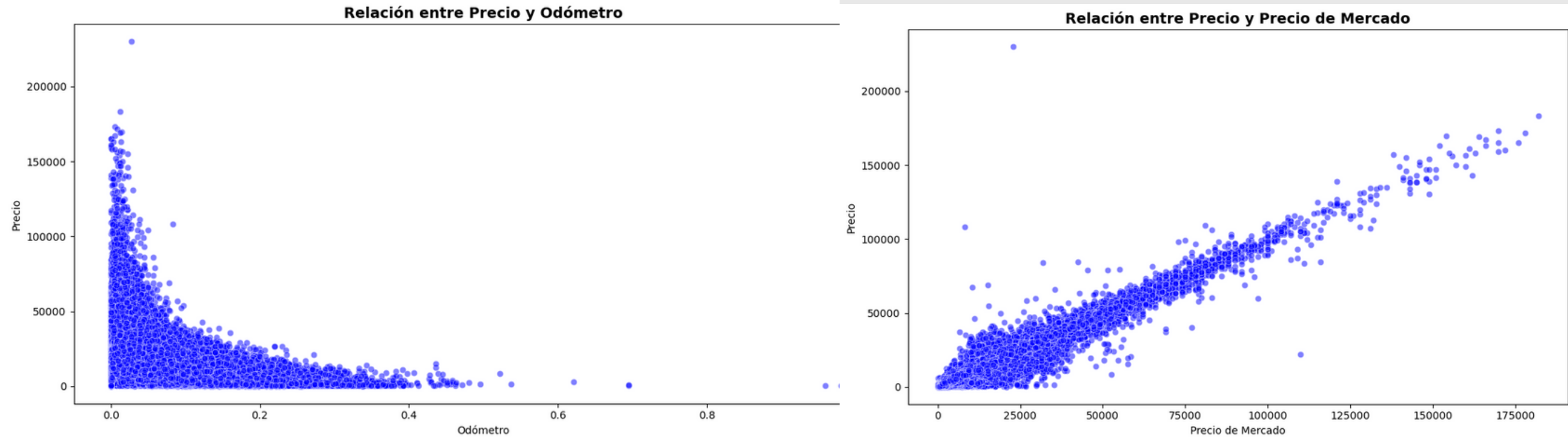
ODOMETER

Kilometraje del vehículo.

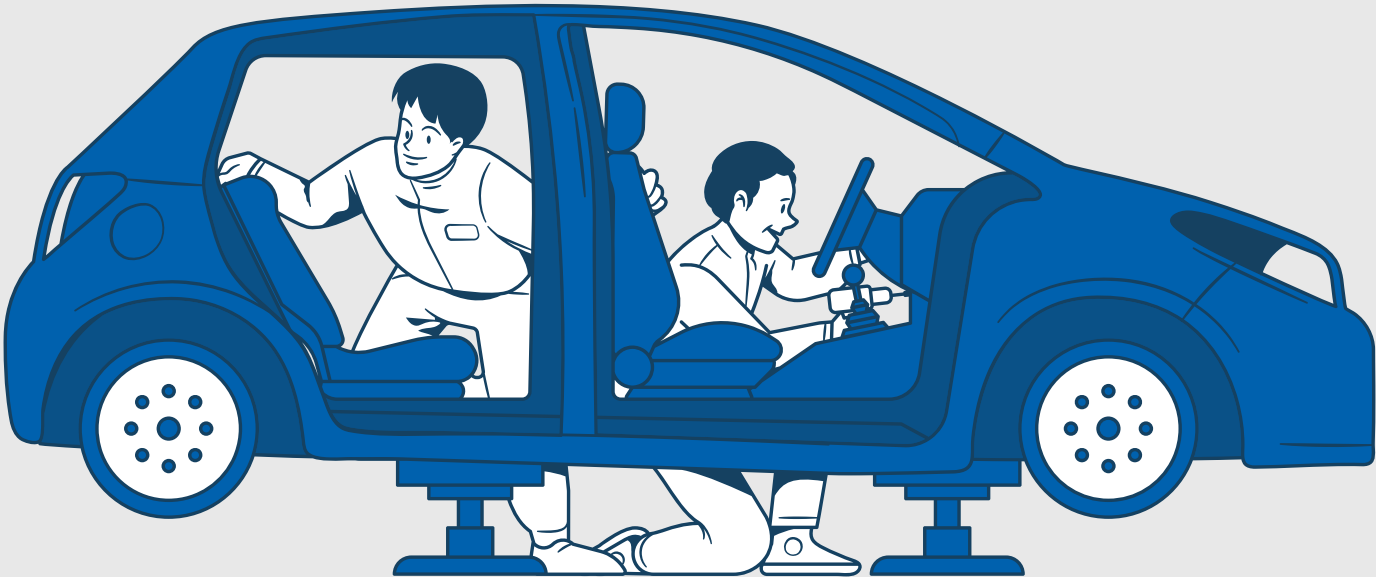
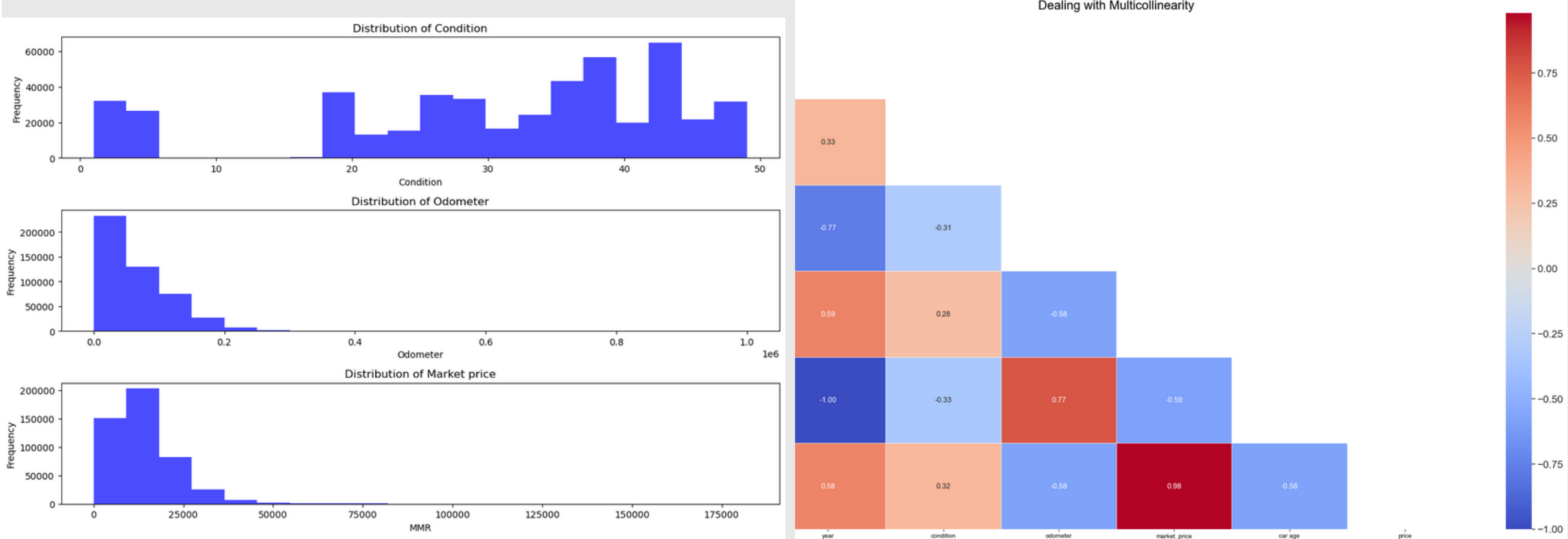
CONDITION

Condición o estado del vehículo

Estadísticas

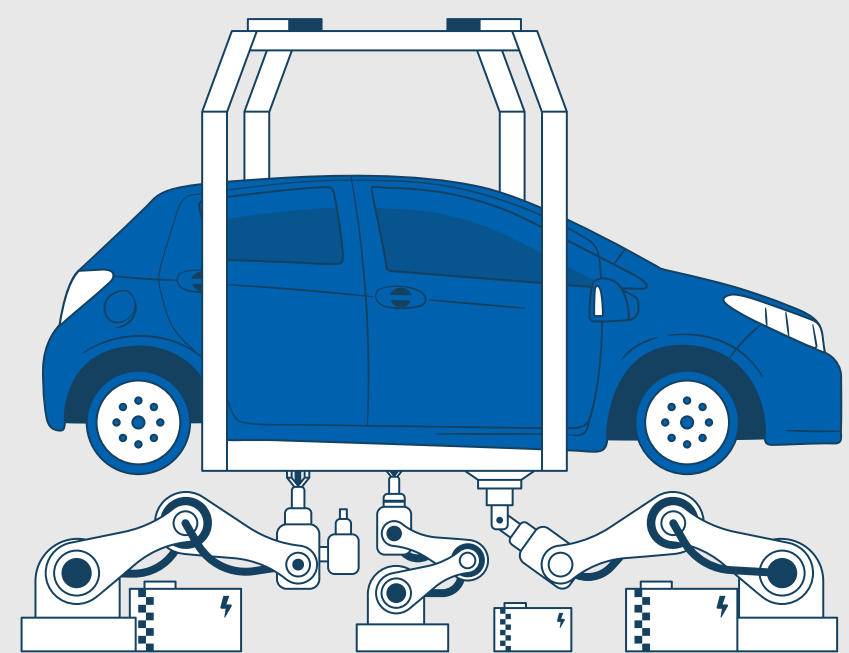


Estadísticas



Estadísticas

	year	condition	odometer	market_price	price	car age
count	472325.0	472325.0	472325.0	472325.0	472325.0	472325.0
mean	2010.0	31.0	66702.0	13837.0	13691.0	6.0
std	4.0	13.0	51940.0	9532.0	9613.0	4.0
min	1990.0	1.0	1.0	25.0	1.0	0.0
25%	2008.0	24.0	28137.0	7425.0	7200.0	3.0
50%	2012.0	35.0	51085.0	12300.0	12200.0	4.0
75%	2013.0	41.0	96590.0	18300.0	18200.0	8.0
max	2015.0	49.0	999999.0	182000.0	230000.0	26.0



INFORMACIÓN TÉCNICA

1

	R ²	RMSE	MSE	MAE
Model				
Linear Regression	0.97	1622.76	2633354.10	1040.47
Ridge	0.97	1622.76	2633354.09	1040.47
Lasso	0.97	1622.76	2633335.89	1040.43
Decision Tree	0.95	2108.40	4445356.53	1368.50
KNN	0.96	1831.98	3356145.16	1212.73
XGBoost	0.96	1985.20	3941017.85	1070.18

Eligo el Modelo Ridge

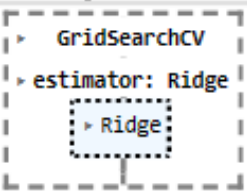
2

Primera mejora

```
# Normalizar características
scaler = MinMaxScaler()
X_train_scaled = scaler.fit_transform(X_train)
X_test_scaled = scaler.transform(X_test)
```

Segunda mejora

```
# Entrenar el modelo Ridge y seleccionar los mejores hiperparámetros
param_grid = {'alpha': [0.001, 0.01, 0.1, 1, 10, 100]}
grid_search = GridSearchCV(Ridge(), param_grid, cv=5)
grid_search.fit(X_train_scaled, y_train)
```



Tercera mejora

```
# Funciones polinómicas
poly = PolynomialFeatures(degree=2)
X_train_poly = poly.fit_transform(X_train_scaled)
X_test_poly = poly.transform(X_test_scaled)
```

3

```
# Función para preprocesar la entrada del usuario
def preprocess_user_input(year, condition, odometer, car_age, market_price):
    user_input_df = pd.DataFrame({
        'year': [year],
        'condition': [condition],
        'odometer': [odometer],
        'market_price': [market_price],
        'car_age': [car_age]
    })

    # Reorganizar las columnas en el orden correcto
    user_input_df = user_input_df[['year', 'condition', 'odometer', 'market_price', 'car_age']]

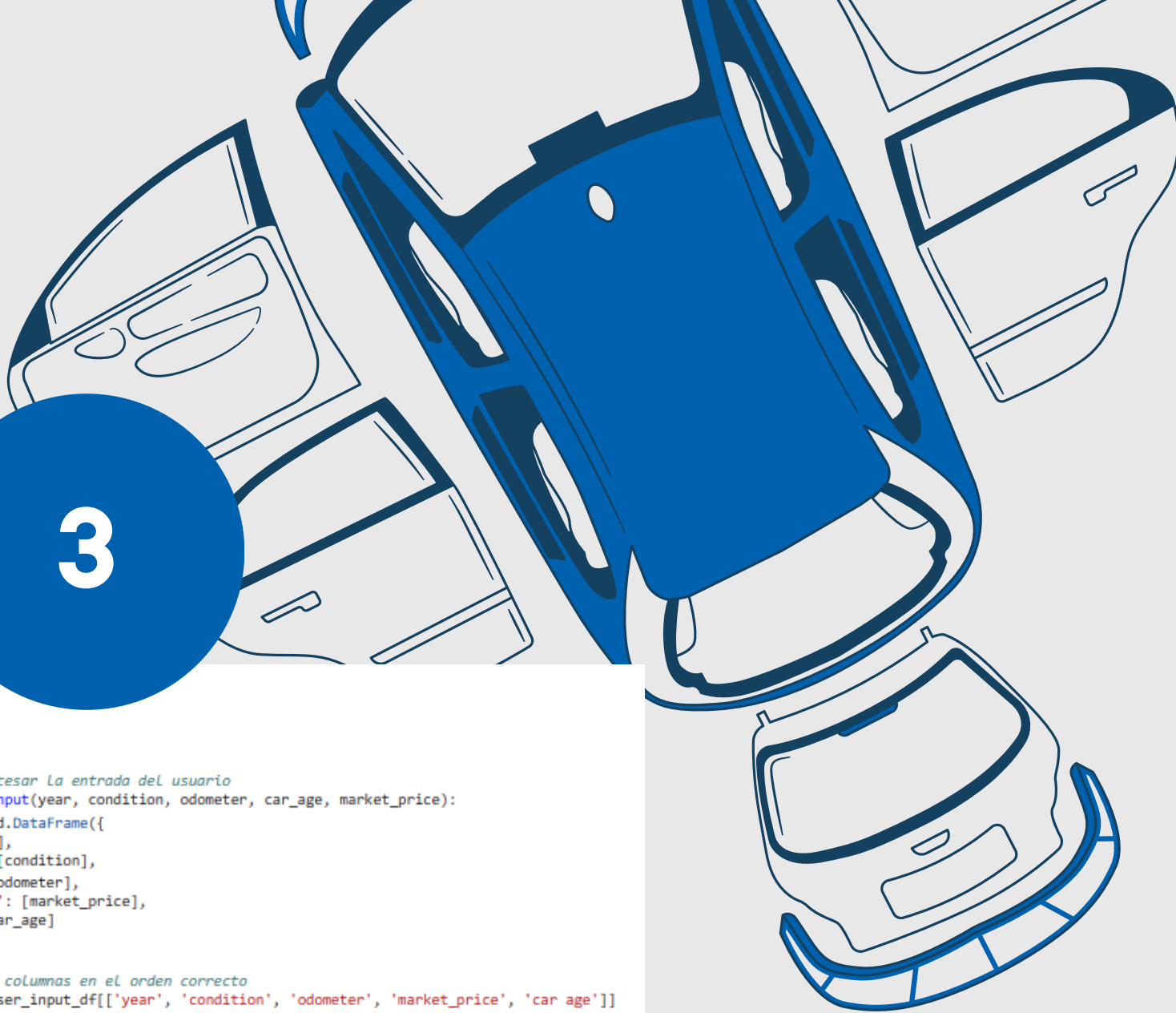
    # Normalizar características
    user_input_scaled = scaler.transform(user_input_df)

    # Aplicar funciones polinómicas
    user_input_poly = poly.transform(user_input_scaled)

    return user_input_poly
```

Es importante escalar y aplicar las mismas funciones que cuando entrenas el modelo

El orden de las columnas tiene que ser el mismo que en el input



ASPECTOS A MEJORAR



Realizar una aplicación web real

Uso de bibliotecas como steamlit

Uso de librerías mas avanzadas para las visualizaciones

Uso de bibliotecas como Altair

Uso de bibliotecas como Holoviews

Uso de bibliotecas como Geopandas

Mejorar la robustez de los inputs

Uso de mas funciones para simplificar código

Muchas Gracias

Podéis encontrar mis datos de contacto a continuación para resolver cualquier duda o consulta, así como si queréis ver el código detrás del trabajo en GitHub.

GitHub: carloscuerda9

EMAIL: carloscuerdarenedo@gmail.com

Bibliografía: —

- seaborn
- Ridge vs Lasso
- SteamLit
- Data set Kaggle
- Página ejemplo

