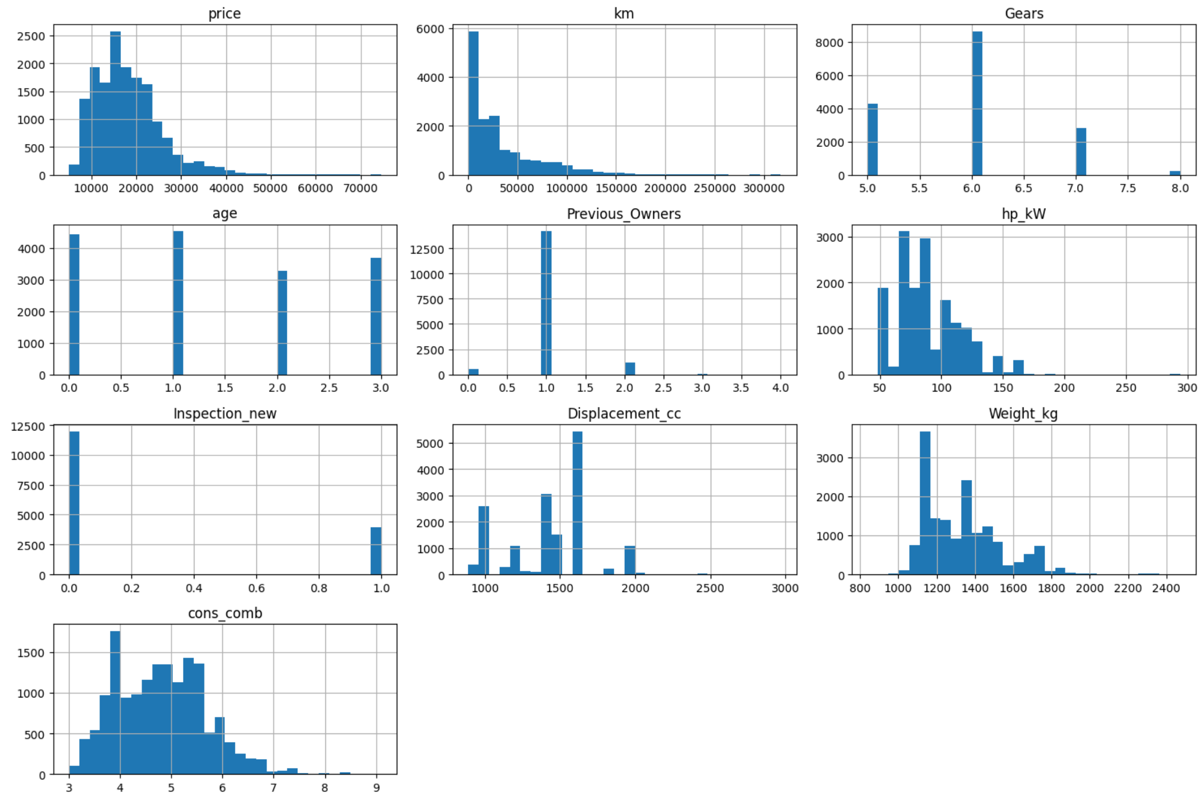
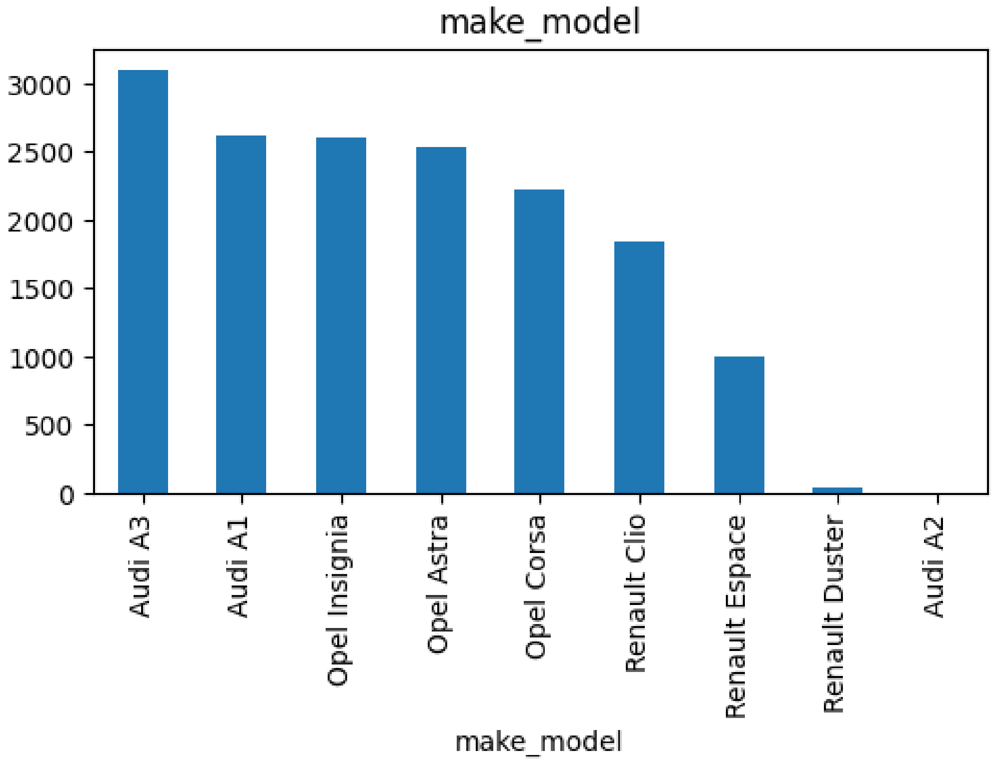
EDMLDS - C4 - Regularisation Assignment

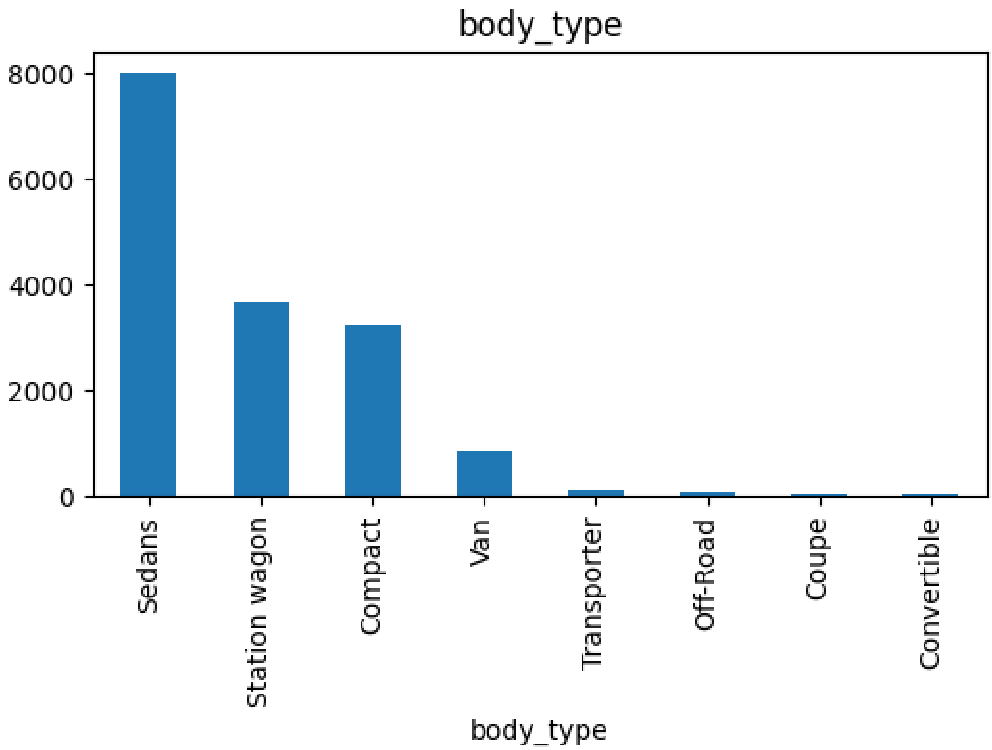
Rishabh Uniyal

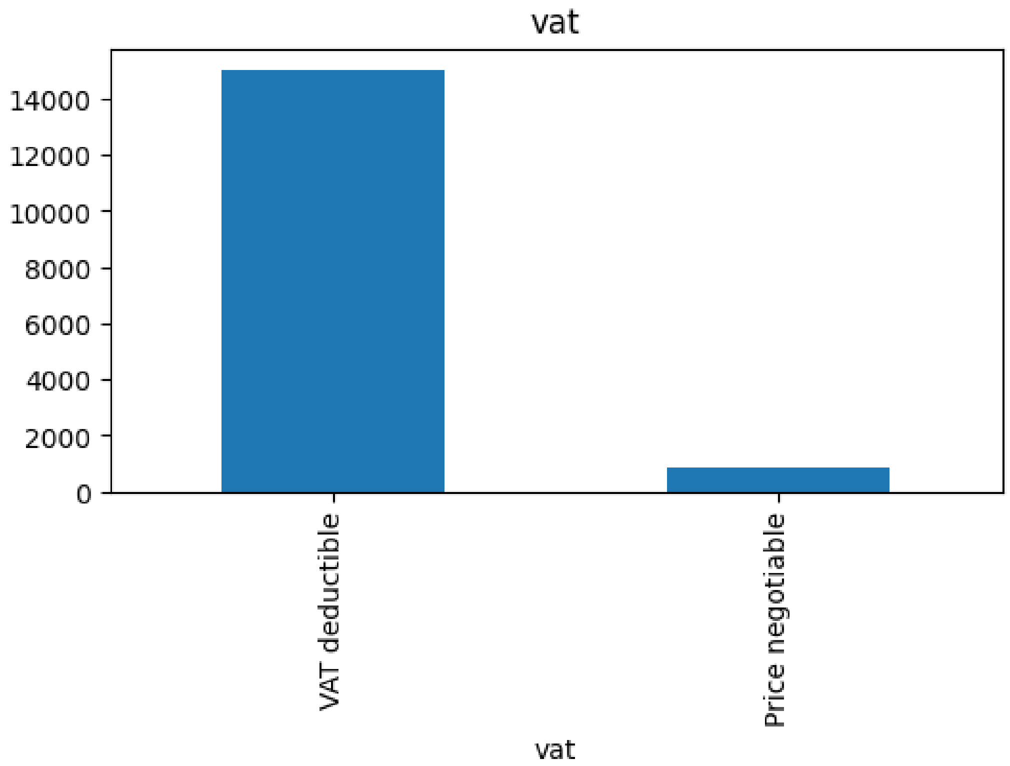
Identify numerical predictors and plot their frequency distributions.

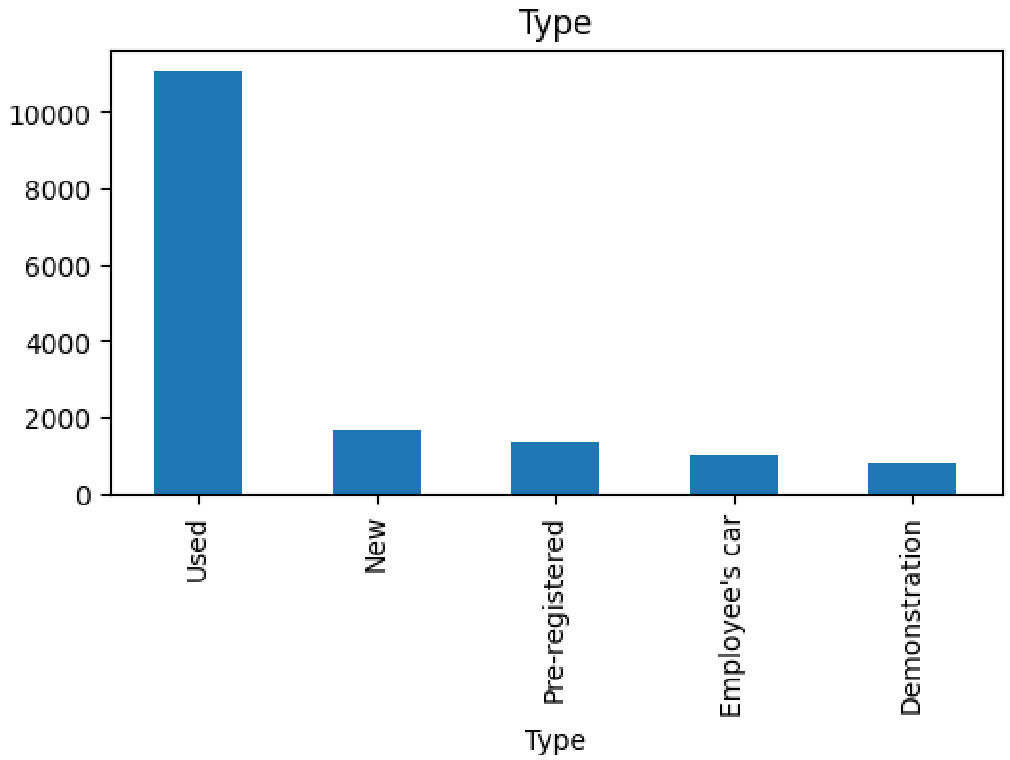


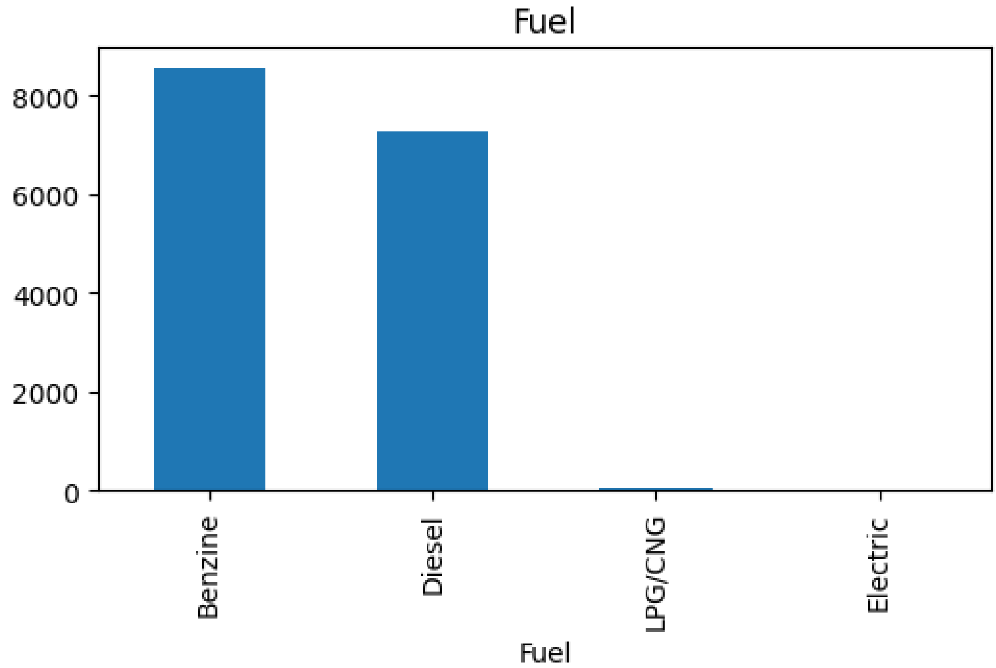
Identify categorical predictors and plot their frequency distributions.

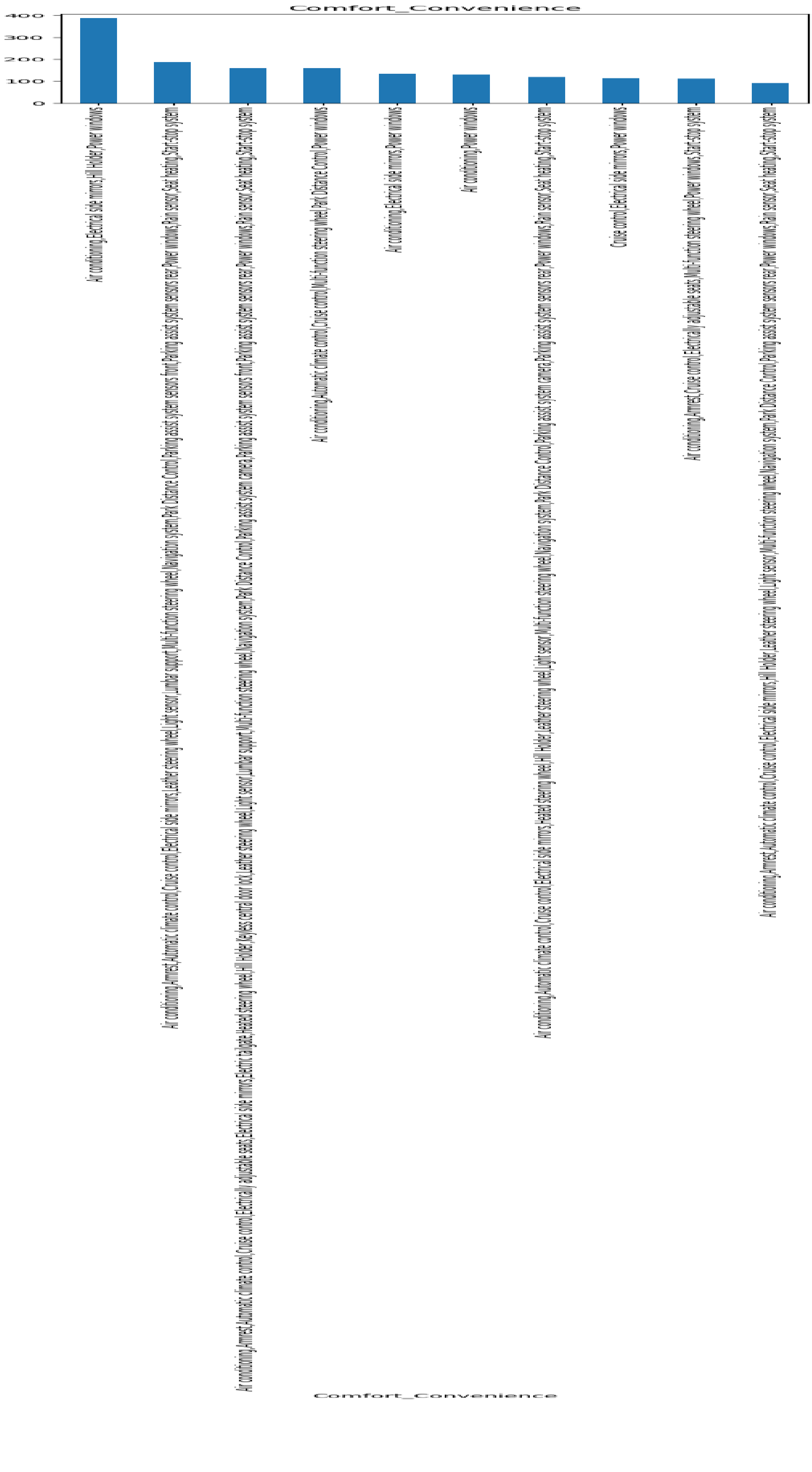

make_model
3000
2500
2000
1500
1000
500
0
Audi A3
Audi A1
Opel Insignia
Opel Astra
Opel Corsa
Renault Clio
Renault Espace
Renault Duster
Audi A2
make_model

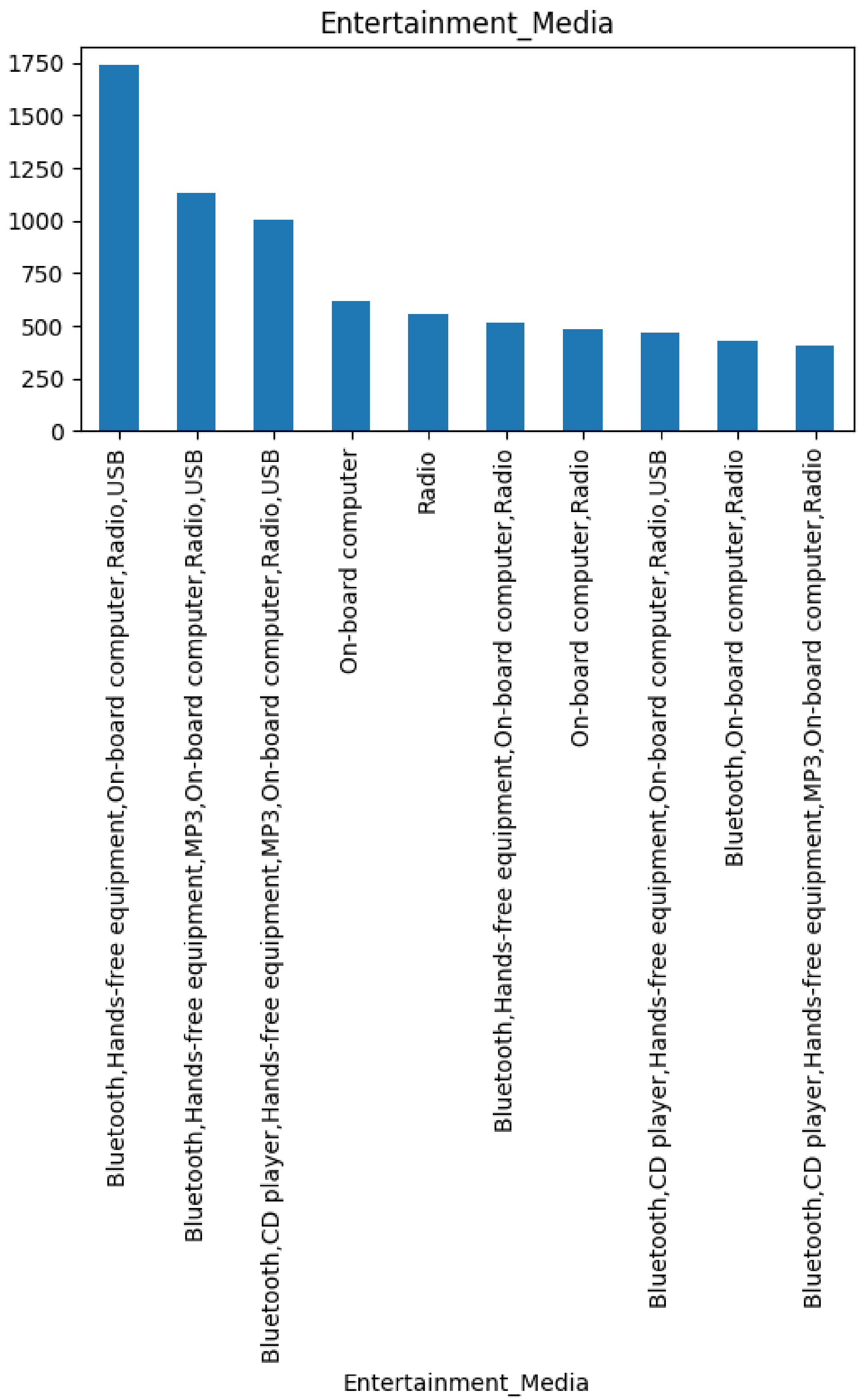

body_type
8000
6000
4000
2000
0
Sedans
Station wagon
Compact
Van
Transporter
Off-Road
Coupe
Convertible
body_type


vat
14000
12000
10000
8000
6000
4000
2000
0
VAT deductible
Price negotiable
vat

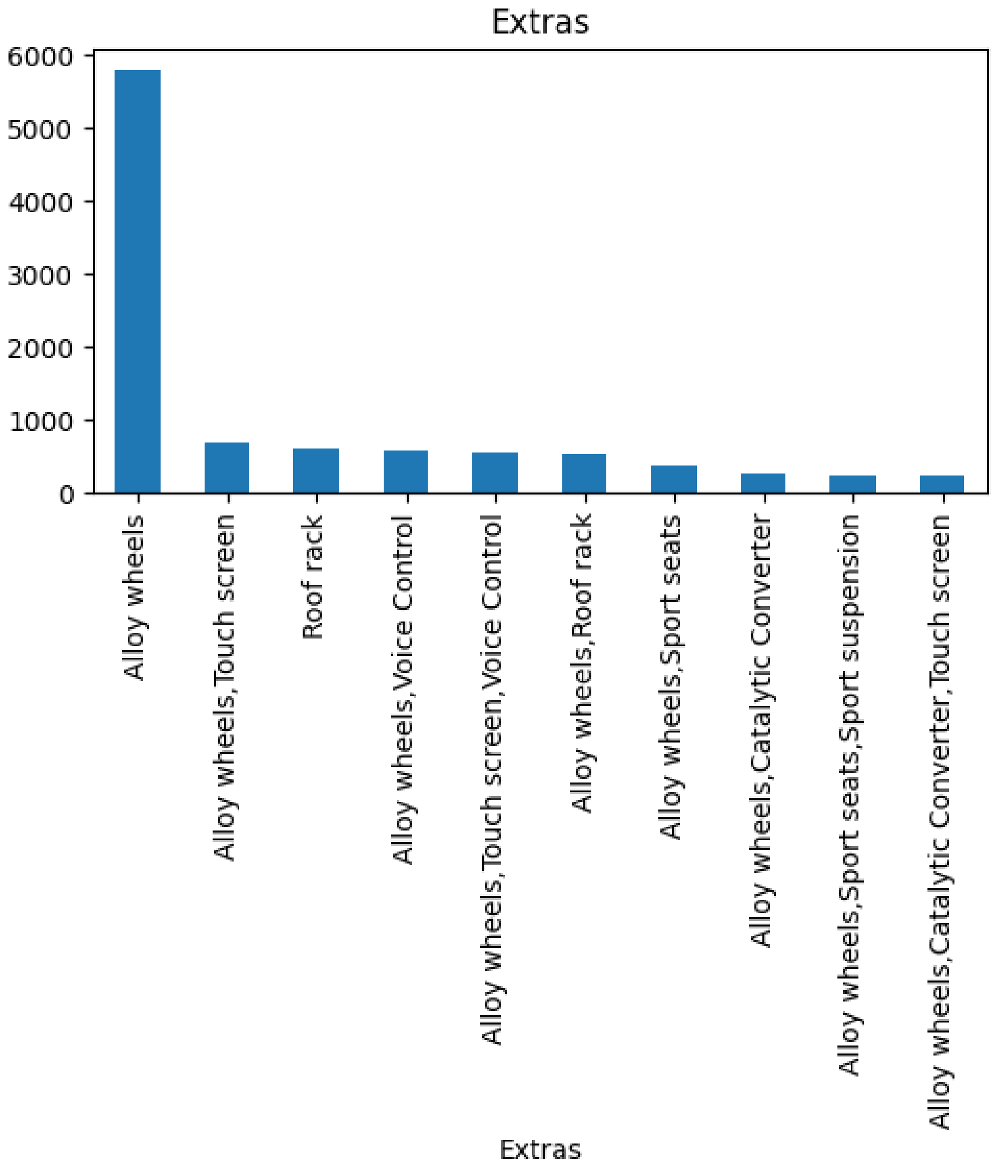

Type
10000
8000
6000
4000
2000
0
Used
New
Pre-registered
Employee's car
Demonstration
Type

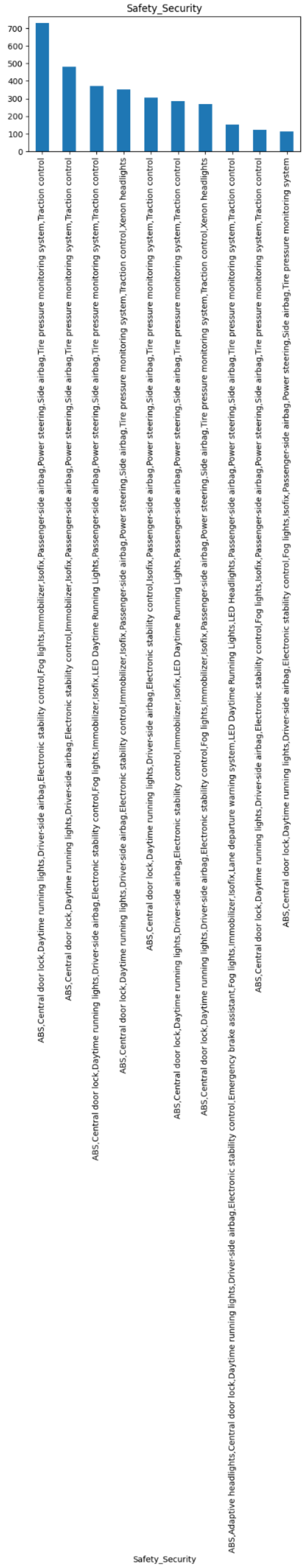

Fuel
8000
6000
4000
2000
0
Benzine
Diesel
LPG/CNG
Electric
Fuel

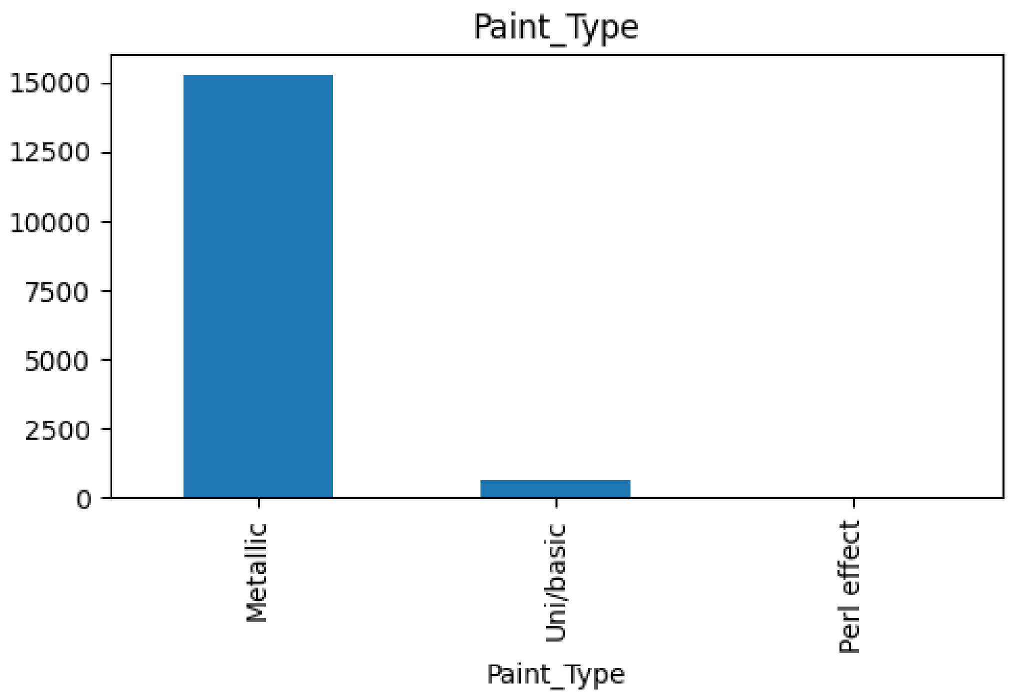

Comfort_Convenience
400
300
200
100
0
Air conditioning,Electrical side mirrors,Hill Holder,Power windows
Air conditioning,Armrest,Automatic climate control,Cruise control,Electrical side mirrors,Leather steering wheel,Light sensor,Lumbar support,Multi-function steering wheel,Navigation system,Park Distance Control,Parking assist system sensors front,Parking assist system sensors rear,Power windows,Rain sensor,Seat heating,Start-stop system
Air conditioning,Armrest,Automatic climate control,Cruise control,Electrically adjustable seats,Electrical side mirrors,Electric tailgate,Heated steering wheel,Hill Holder,Keyless central door lock,Leather steering wheel,Light sensor,Lumbar support,Multi-function steering wheel,Navigation system,Park Distance Control,Parking assist system camera,Parking assist system sensors front,Parking assist system sensors rear,Power windows,Rain sensor,Seat heating,Start-stop system
Air conditioning,Automatic climate control,Cruise control,Multi-function steering wheel,Park Distance Control,Power windows
Air conditioning,Electrical side mirrors,Power windows
Air conditioning,Power windows
Air conditioning,Automatic climate control,Cruise control,Electrical side mirrors,Heated steering wheel,Hill Holder,Leather steering wheel,Light sensor,Multi-function steering wheel,Navigation system,Park Distance Control,Parking assist system camera,Parking assist system sensors rear,Power windows,Rain sensor,Seat heating,Start-stop system
Cruise control,Electrical side mirrors,Power windows
Air conditioning,Armrest,Cruise control,Electrically adjustable seats,Multi-function steering wheel,Power windows,Start-stop system
Air conditioning,Armrest,Automatic climate control,Cruise control,Electrical side mirrors,Hill Holder,Leather steering wheel,Light sensor,Multi-function steering wheel,Navigation system,Park Distance Control,Parking assist system sensors rear,Power windows,Rain sensor,Seat heating,Start-stop system
Comfort_Convenience

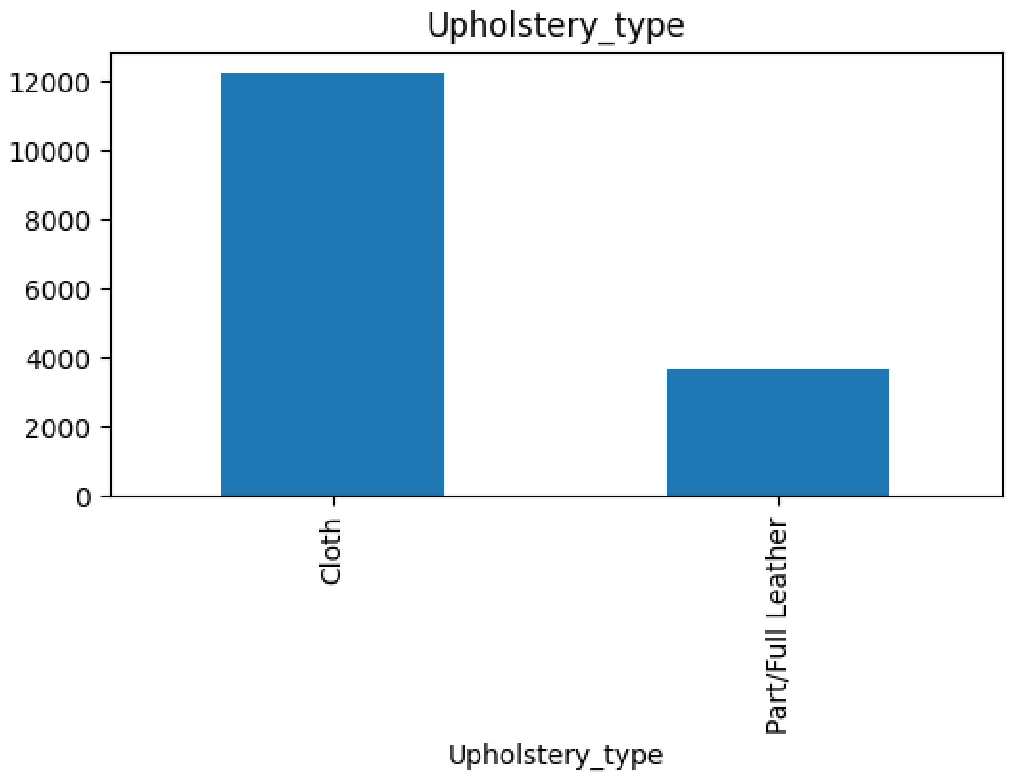


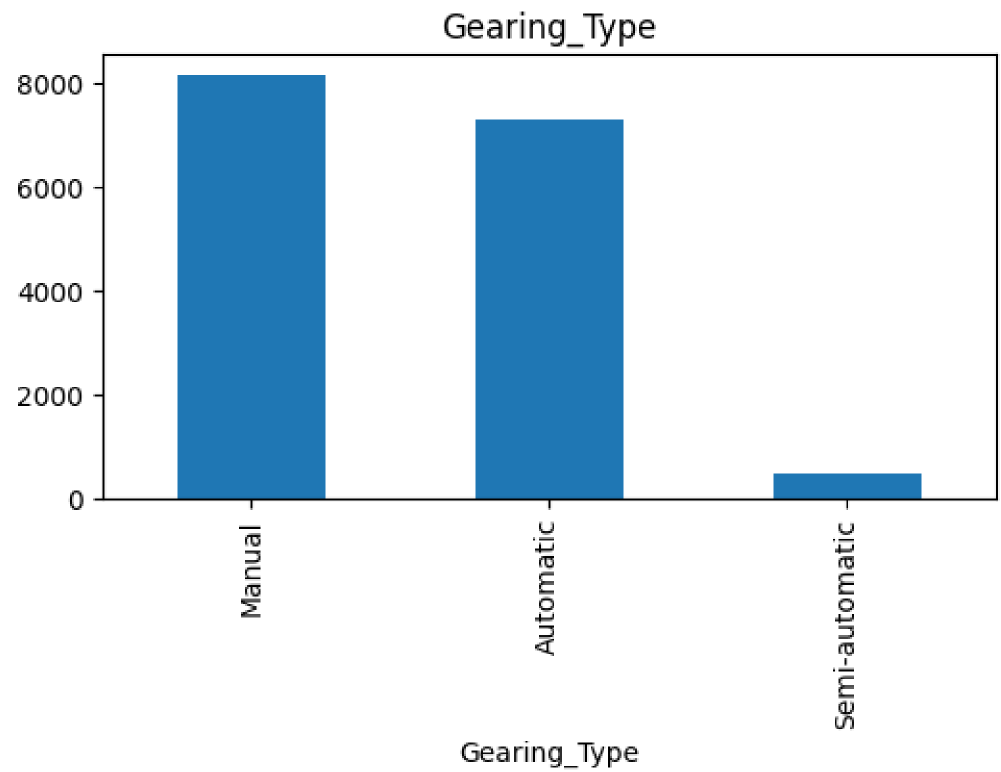
Entertainment_Media
1750
1500
1250
1000
750
500
250
0
Bluetooth,Hands-free equipment,On-board computer,Radio,USB
Bluetooth,Hands-free equipment,MP3,On-board computer,Radio,USB
Bluetooth,CD player,Hands-free equipment,MP3,On-board computer,Radio,USB
On-board computer
Radio
Bluetooth,Hands-free equipment,On-board computer,Radio
On-board computer,Radio
Bluetooth,CD player,Hands-free equipment,On-board computer,Radio,USB
Bluetooth,On-board computer,Radio
Bluetooth,CD player,Hands-free equipment,MP3,On-board computer,Radio
Entertainment_Media

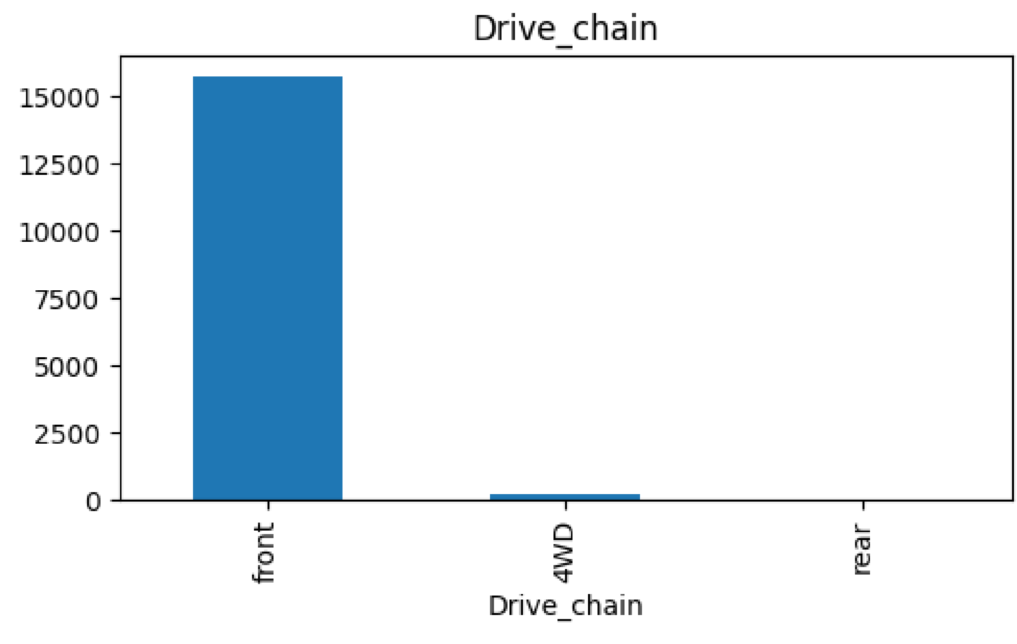

Extras
6000
5000
4000
3000
2000
1000
0
Alloy wheels
Alloy wheels,Touch screen
Roof rack
Alloy wheels,Voice Control
Alloy wheels,Touch screen,Voice Control
Alloy wheels,Roof rack
Alloy wheels,Sport seats
Alloy wheels,Catalytic Converter
Alloy wheels,Sport seats,Sport suspension
Alloy wheels,Catalytic Converter,Touch screen
Extras


Safety_Security
0
100
200
300
400
500
600
700
ABS,Central door lock,Daytime running lights,Driver-side airbag,Electronic stability control,Fog lights,Immobilizer,Isofix,Passenger-side airbag,Power steering,Side airbag,Tire pressure monitoring system,Traction control
ABS,Central door lock,Daytime running lights,Driver-side airbag,Electronic stability control,Immobilizer,Isofix,Passenger-side airbag,Power steering,Side airbag,Tire pressure monitoring system,Traction control
ABS,Central door lock,Daytime running lights,Driver-side airbag,Electronic stability control,Fog lights,Immobilizer,Isofix,LED Daytime Running Lights,Passenger-side airbag,Power steering,Side airbag,Tire pressure monitoring system,Traction control
ABS,Central door lock,Daytime running lights,Driver-side airbag,Electronic stability control,Immobilizer,Isofix,Passenger-side airbag,Power steering,Side airbag,Tire pressure monitoring system,Traction control,Xenon headlights
ABS,Central door lock,Daytime running lights,Driver-side airbag,Electronic stability control,Isofix,Passenger-side airbag,Power steering,Side airbag,Tire pressure monitoring system,Traction control
ABS,Central door lock,Daytime running lights,Driver-side airbag,Electronic stability control,Immobilizer,Isofix,LED Daytime Running Lights,Passenger-side airbag,Power steering,Side airbag,Tire pressure monitoring system,Traction control
ABS,Central door lock,Daytime running lights,Driver-side airbag,Electronic stability control,Fog lights,Immobilizer,Isofix,Passenger-side airbag,Power steering,Side airbag,Tire pressure monitoring system,Traction control,Xenon headlights
ABS,Adaptive headlights,Central door lock,Daytime running lights,Driver-side airbag,Electronic stability control,Emergency brake assistant,Fog lights,Immobilizer,Isofix,Lane departure warning system,LED Daytime Running Lights,LED Headlights,Passenger-side airbag,Power steering,Side airbag,Tire pressure monitoring system,Traction control
ABS,Central door lock,Daytime running lights,Driver-side airbag,Electronic stability control,Fog lights,Isofix,Passenger-side airbag,Power steering,Side airbag,Tire pressure monitoring system,Traction control
ABS,Central door lock,Daytime running lights,Driver-side airbag,Electronic stability control,Fog lights,Isofix,Passenger-side airbag,Power steering,Side airbag,Tire pressure monitoring system
Safety_Security

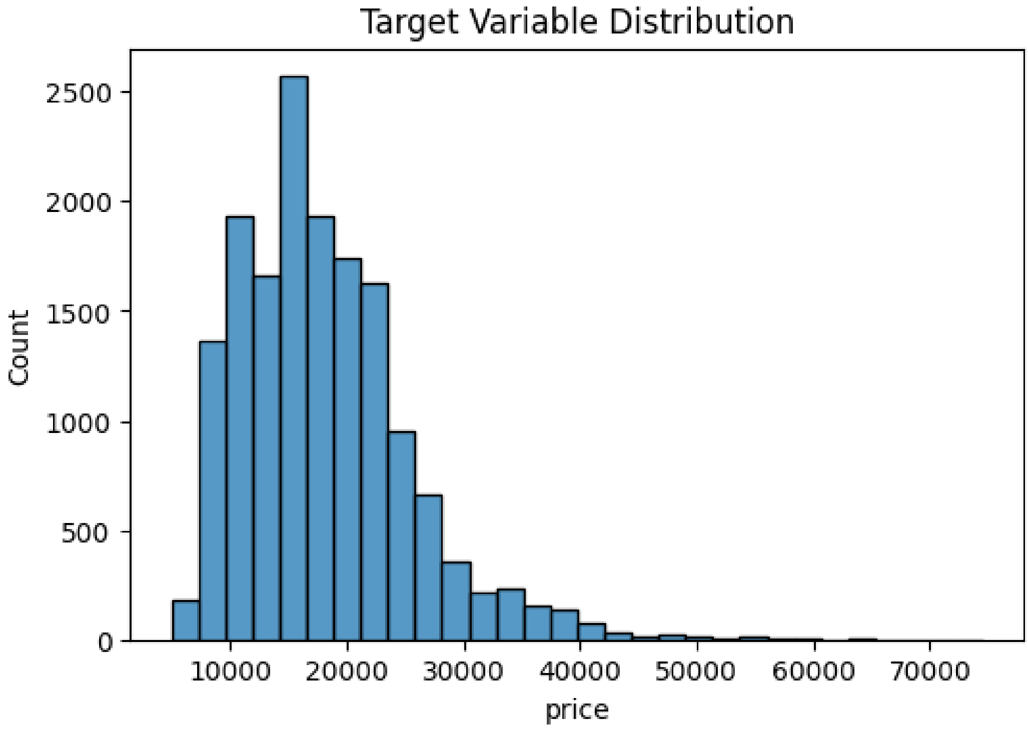

Paint_Type
15000
12500
10000
7500
5000
2500
0
Metallic
Uni/basic
Perl effect
Paint_Type


Upholstery_type
12000
10000
8000
6000
4000
2000
0
Cloth
Part/Full Leather
Upholstery_type

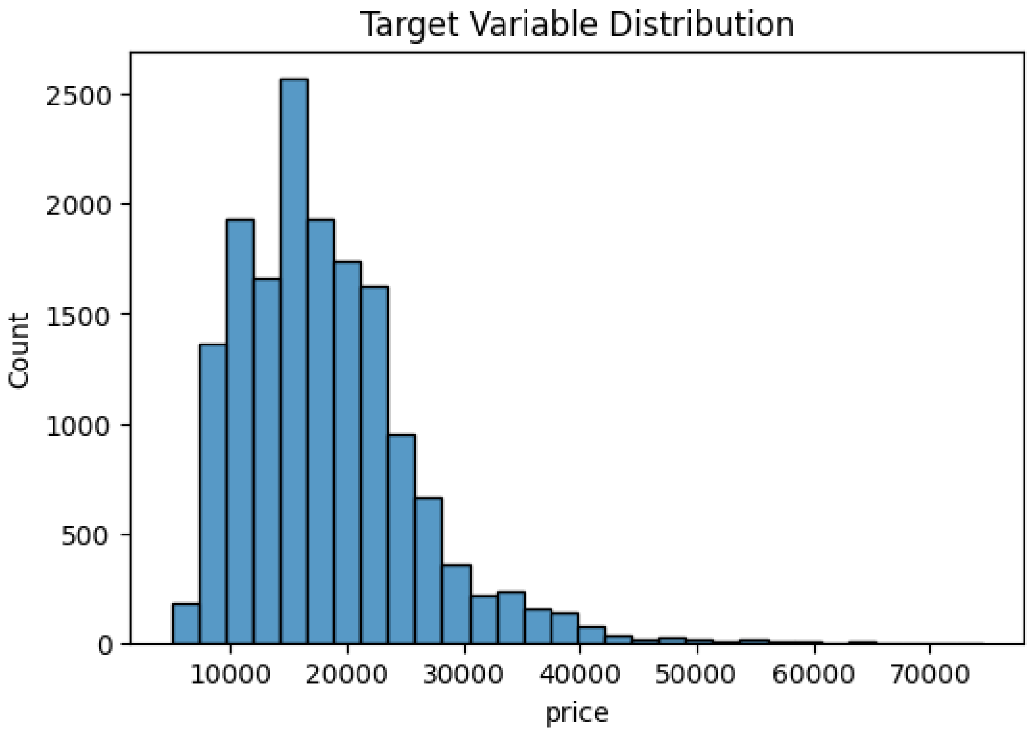




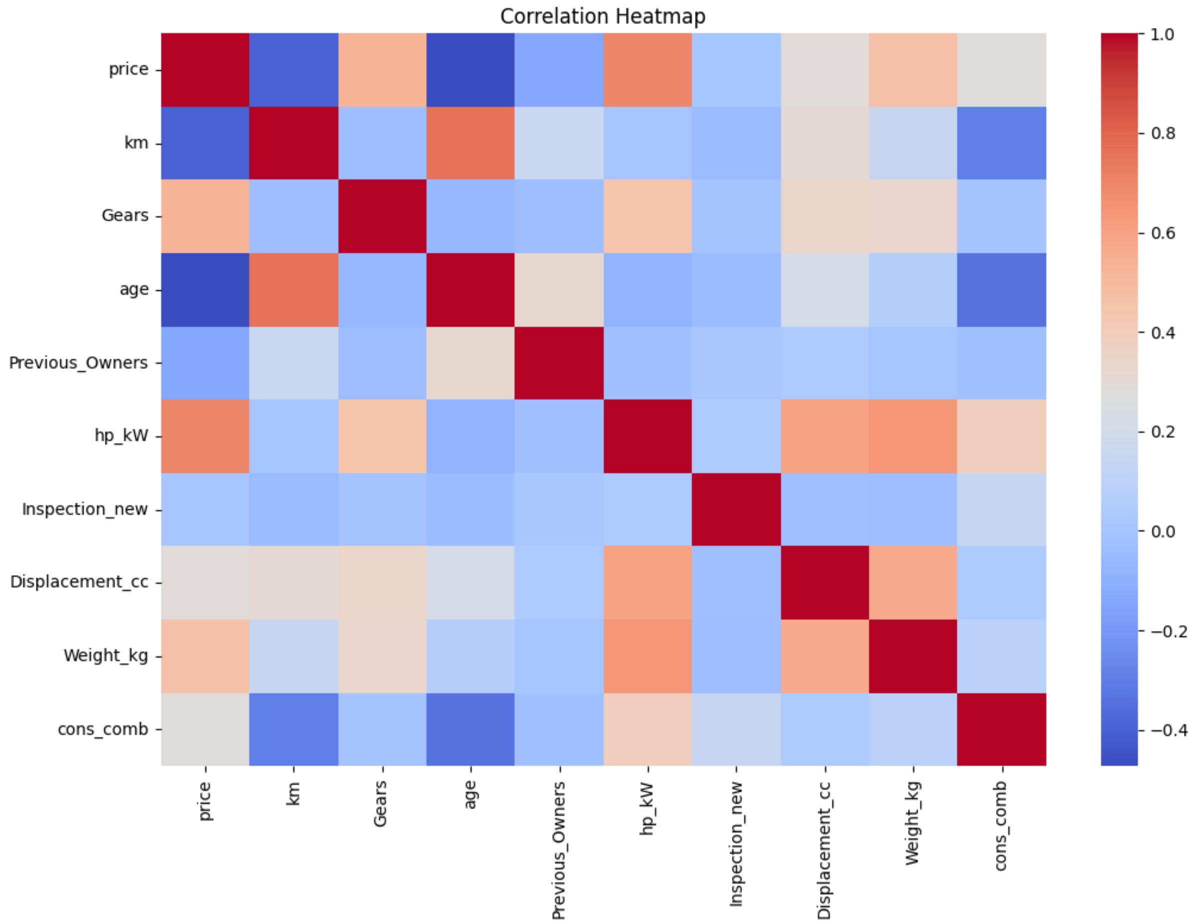
Identify target variable and plot the frequency distributions



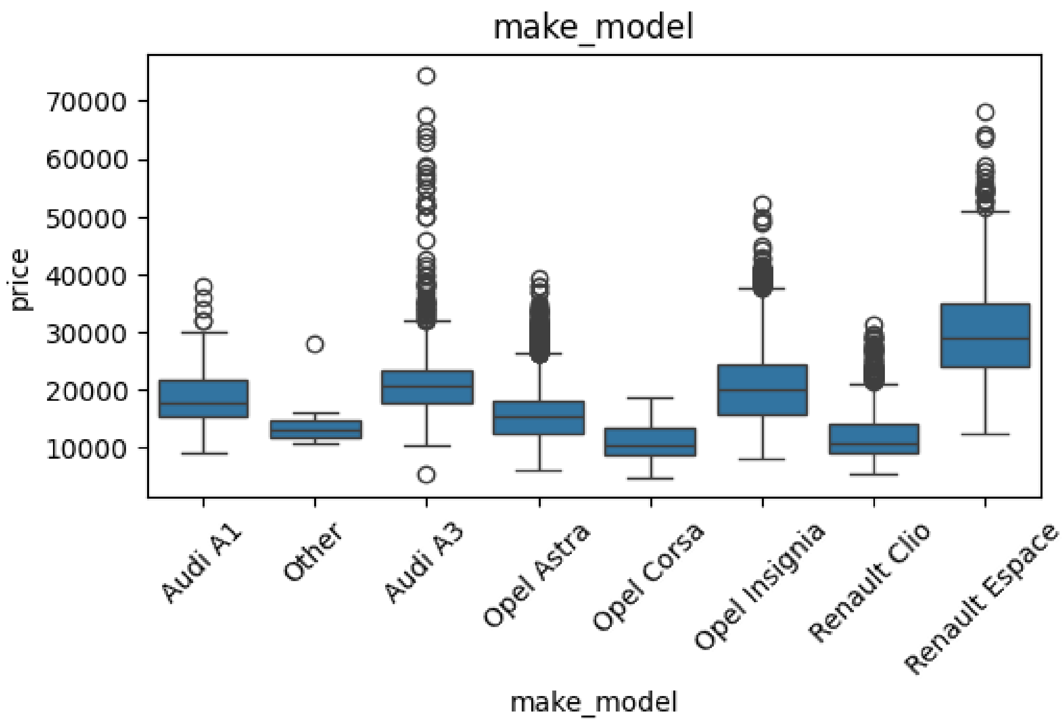
Transform the target feature

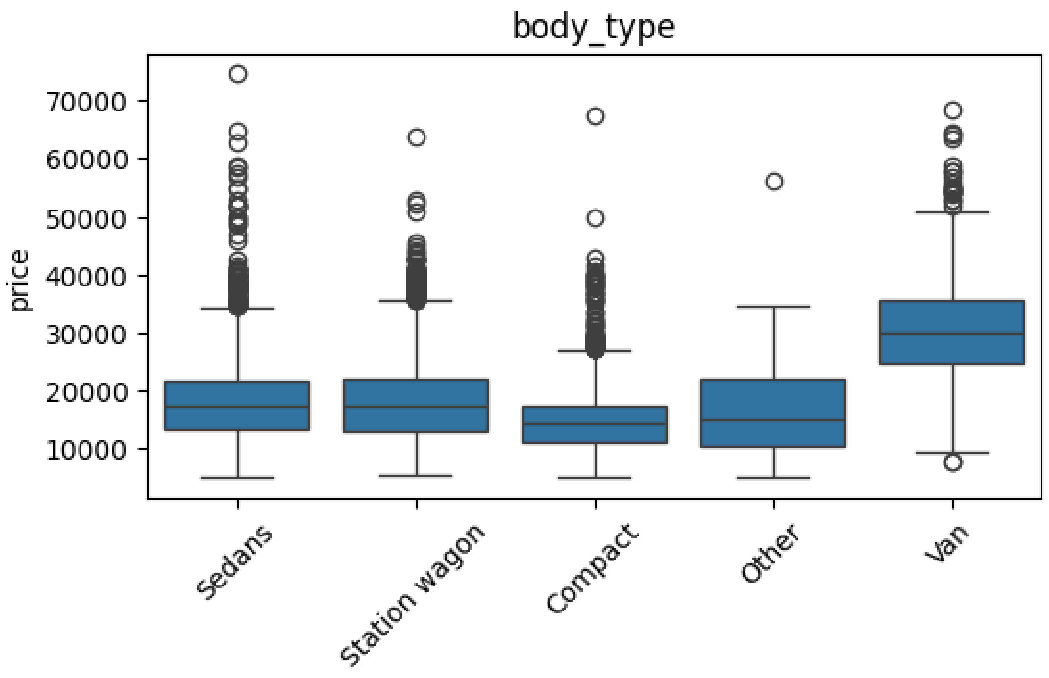


Plot the correlation map between features and target variable.

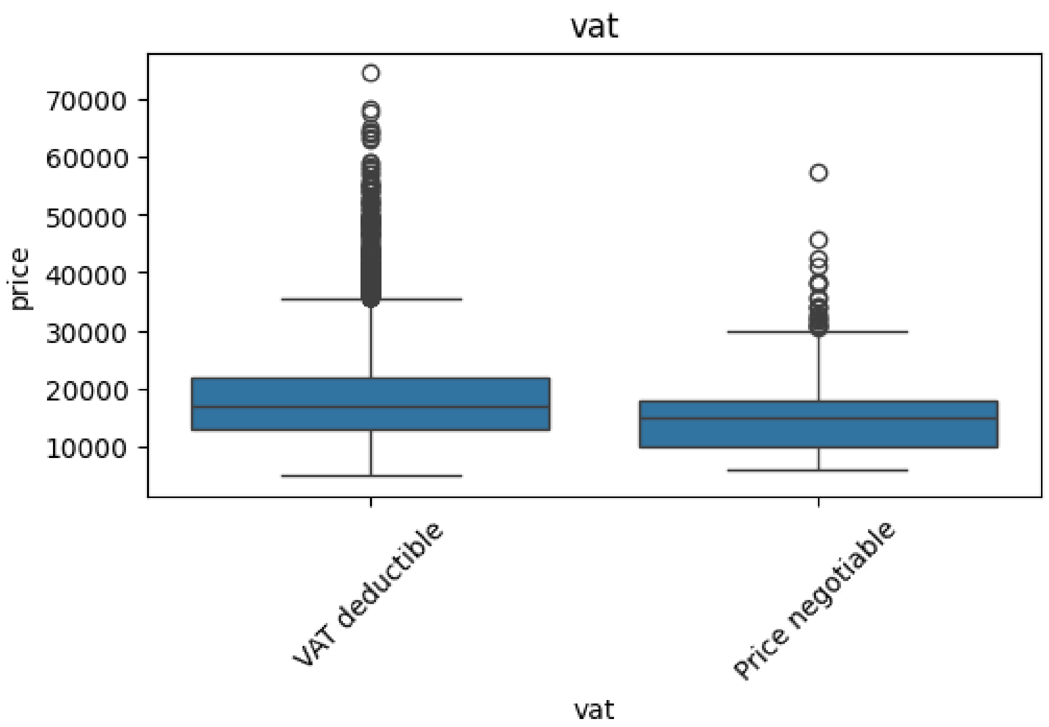


Analyse correlation between categorical features and target variable.

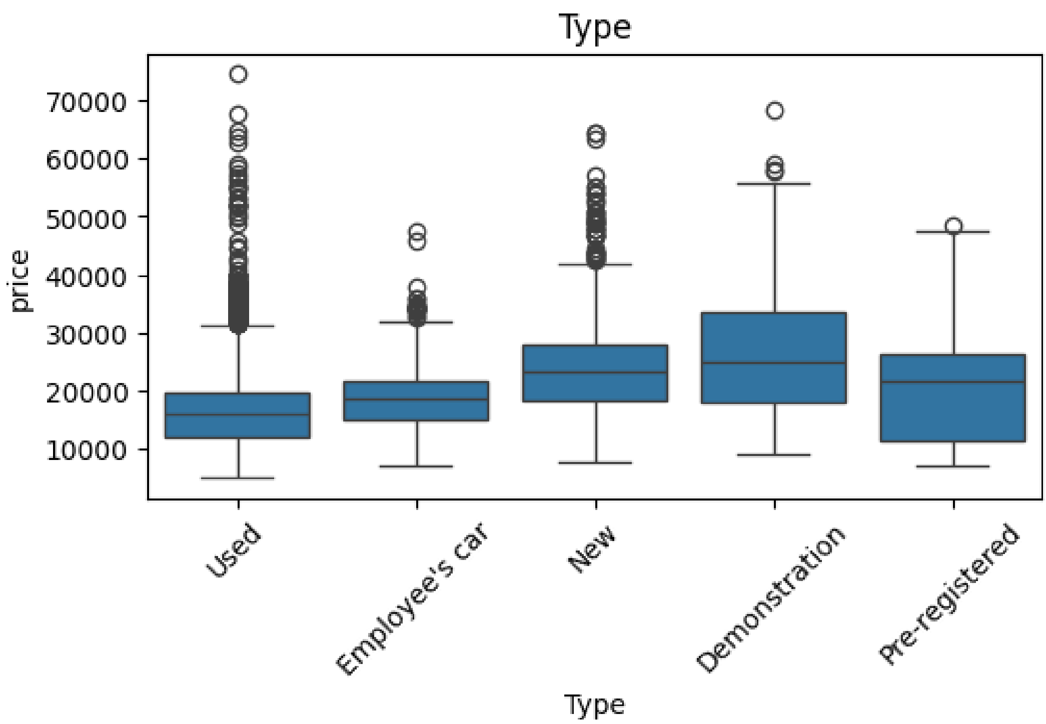


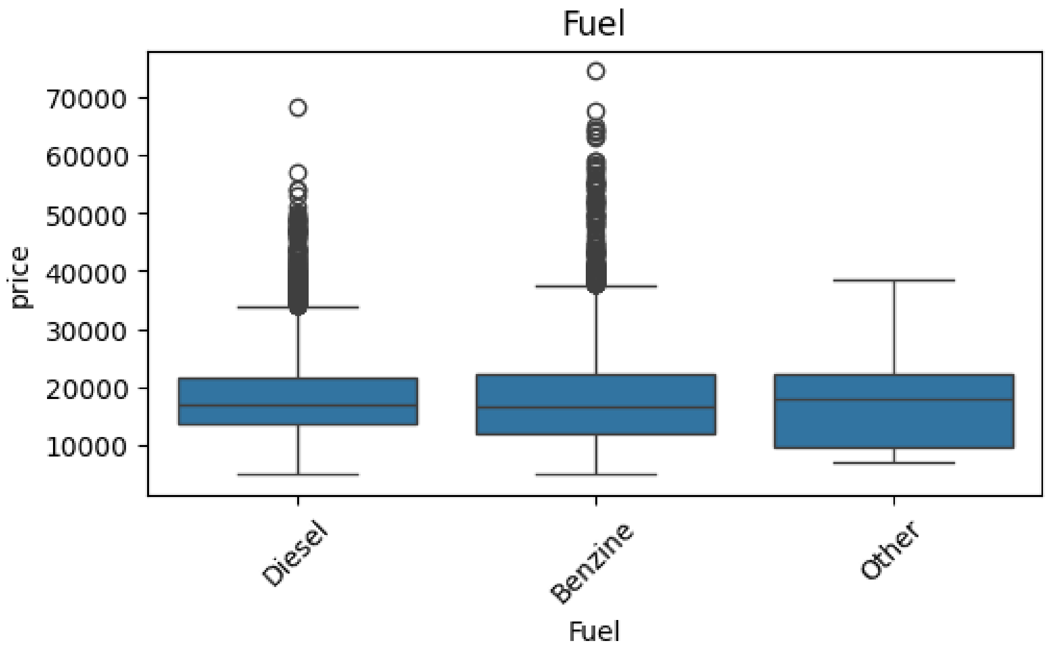


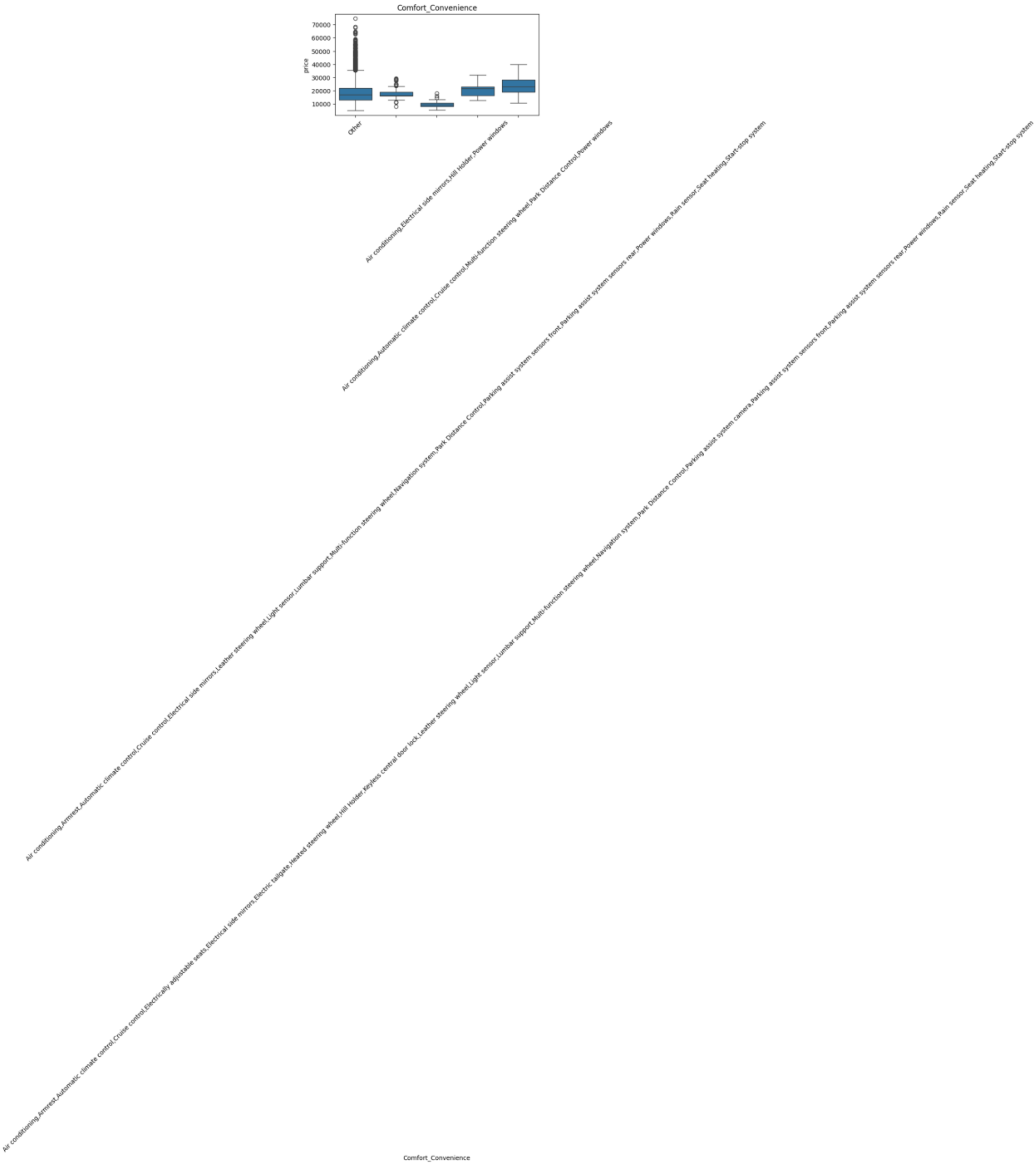
body_type
70000
60000
50000
40000
30000
20000
10000
price
Sedans
Station wagon
Compact
Other
Van
body_type



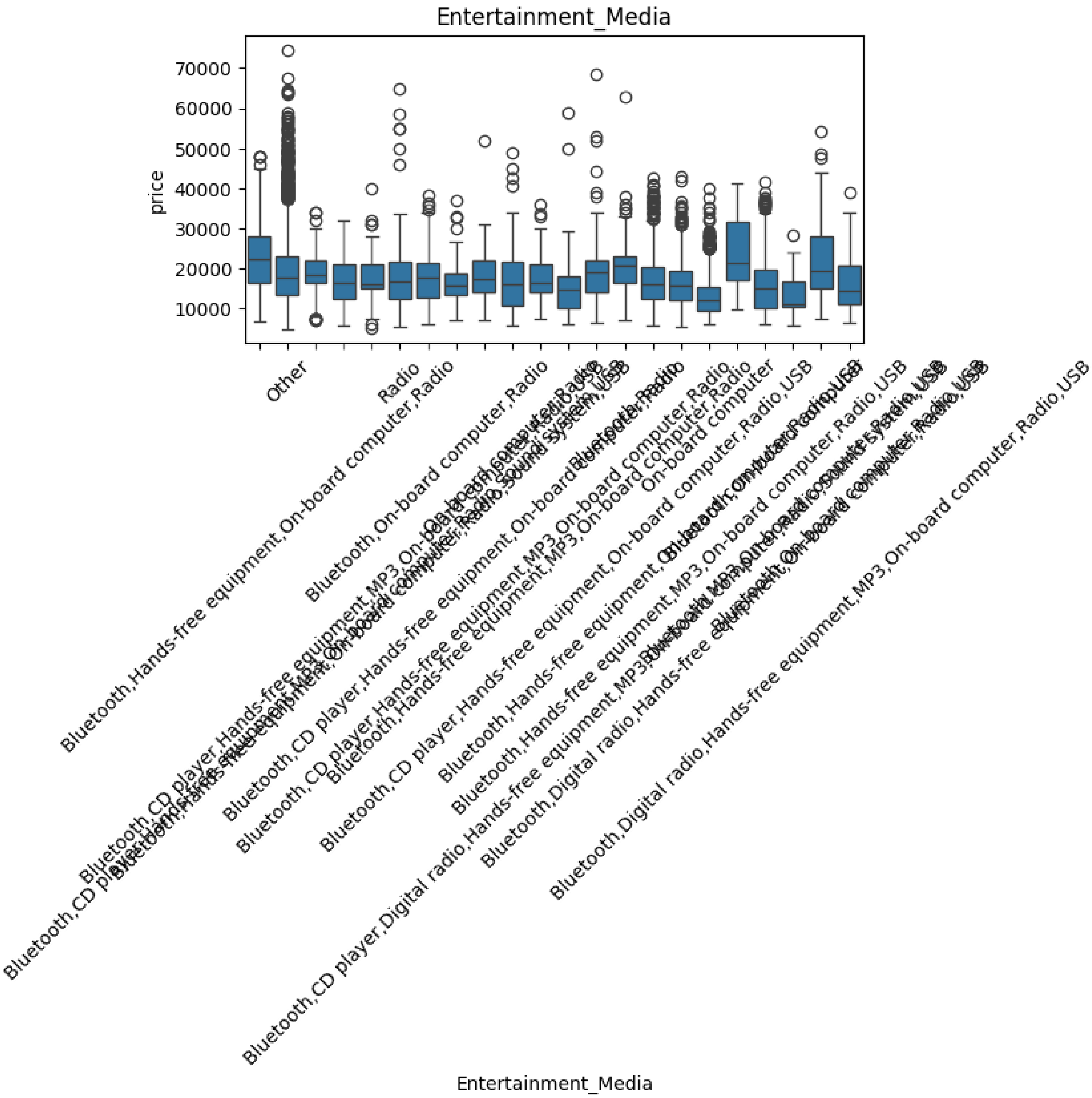
vat
70000
60000
50000
40000
30000
20000
10000
price
VAT deductible
Price negotiable
vat

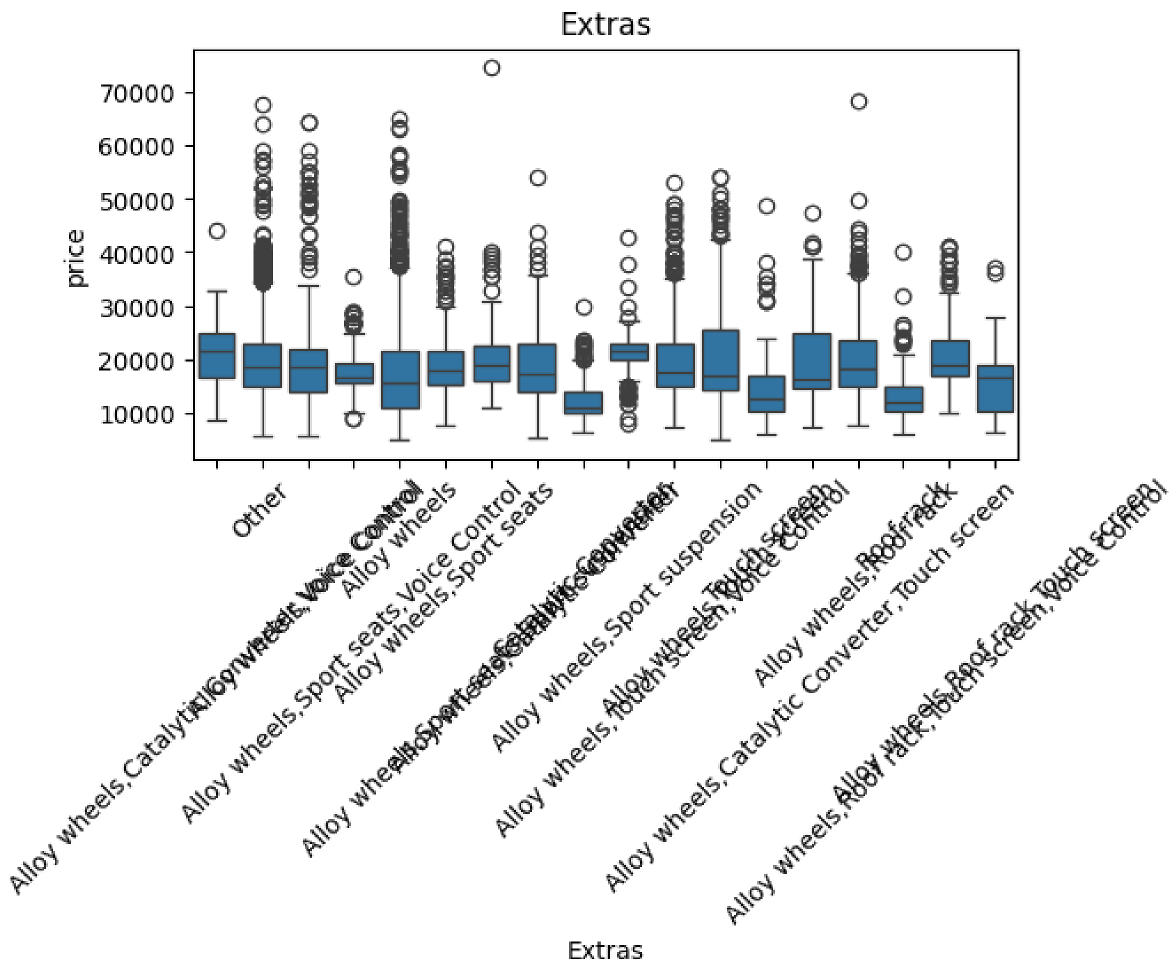

Type
price
70000
60000
50000
40000
30000
20000
10000
Used
Employee's car
New
Demonstration
Pre-registered
Type


Fuel
price
70000
60000
50000
40000
30000
20000
10000
Diesel
Benzine
Other
Fuel

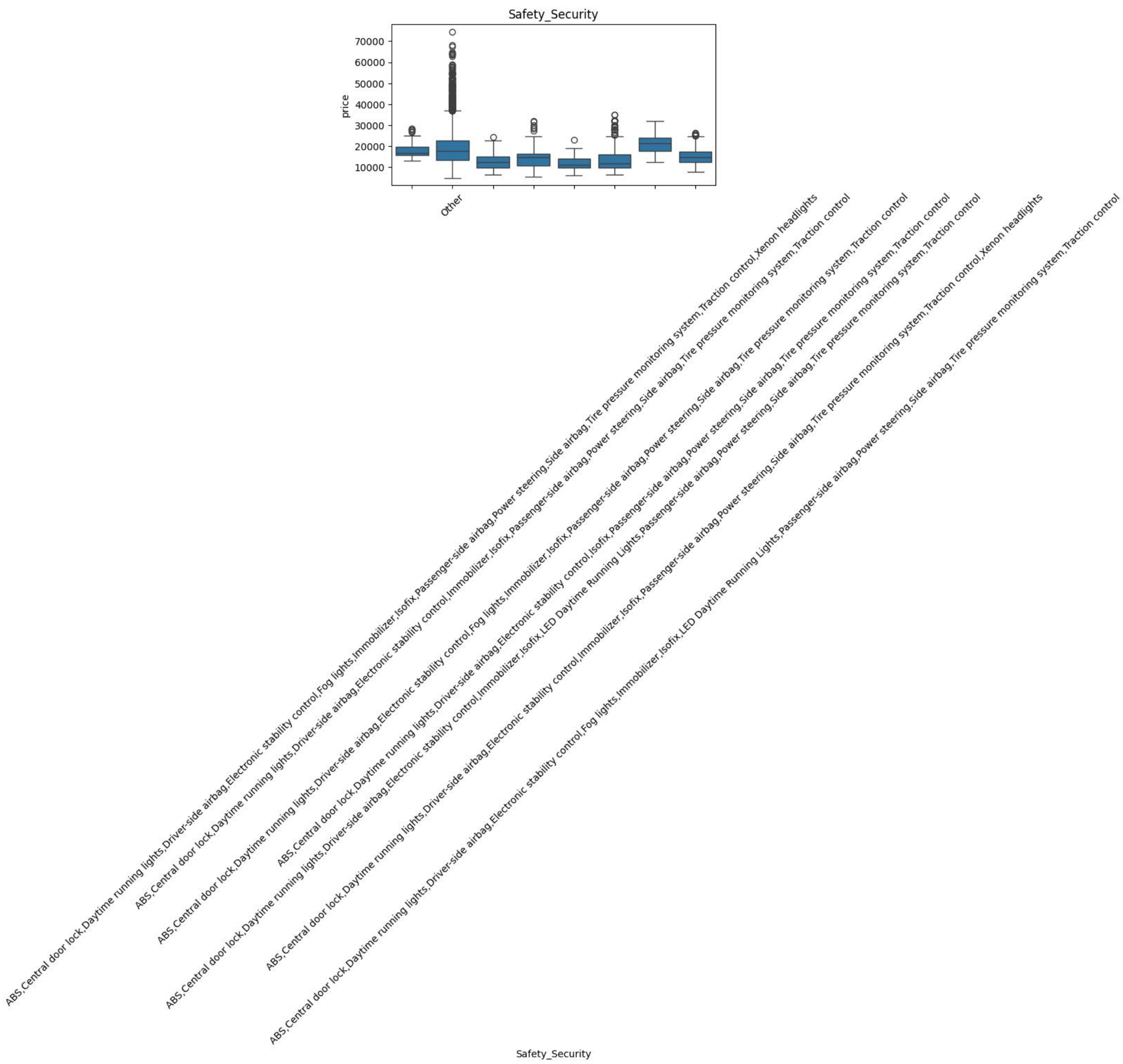


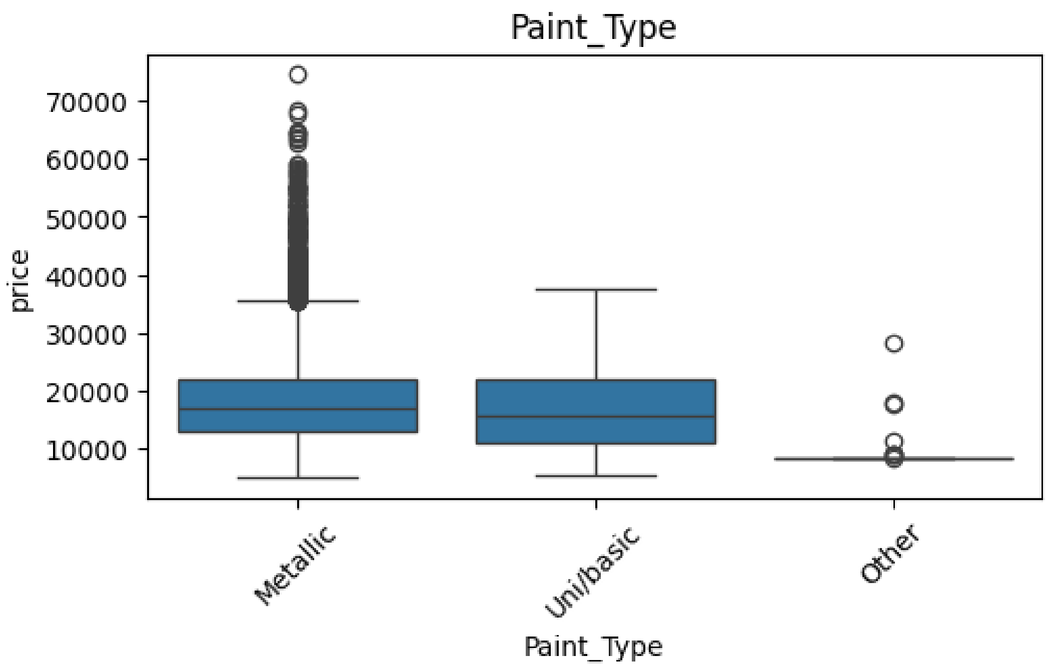
Comfort_Convenience
price
70000
60000
50000
40000
30000
20000
10000
Other
Air conditioning,Electrical side mirrors,Hill Holder,Power windows
Air conditioning,Automatic climate control,Cruise control,Multi-function steering wheel,Park Distance Control,Power windows
Air conditioning,Armrest,Automatic climate control,Cruise control,Electrical side mirrors,Leather steering wheel,Light sensor,Lumbar support,Multi-function steering wheel,Navigation system,Park Distance Control,Parking assist system sensors front,Parking assist system sensors rear,Power windows,Rain sensor,Seat heating,Start-stop system
Air conditioning,Armrest,Automatic climate control,Cruise control,Electrically adjustable seats,Electrical side mirrors,Electric tailgate,Heated steering wheel,Hill Holder,Keyless central door lock,Leather steering wheel,Light sensor,Lumbar support,Multi-function steering wheel,Navigation system,Park Distance Control,Parking assist system camera,Parking assist system sensors front,Parking assist system sensors rear,Power windows,Rain sensor,Seat heating,Start-stop system
Comfort_Convenience


Entertainment_Media
price
70000
60000
50000
40000
30000
20000
10000
Other
Radio
Bluetooth,Hands-free equipment,On-board computer,Radio
Bluetooth,On-board computer,Radio
Bluetooth,Digital radio,Hands-free equipment,MP3,On-board computer,Radio,USB
Entertainment_Media

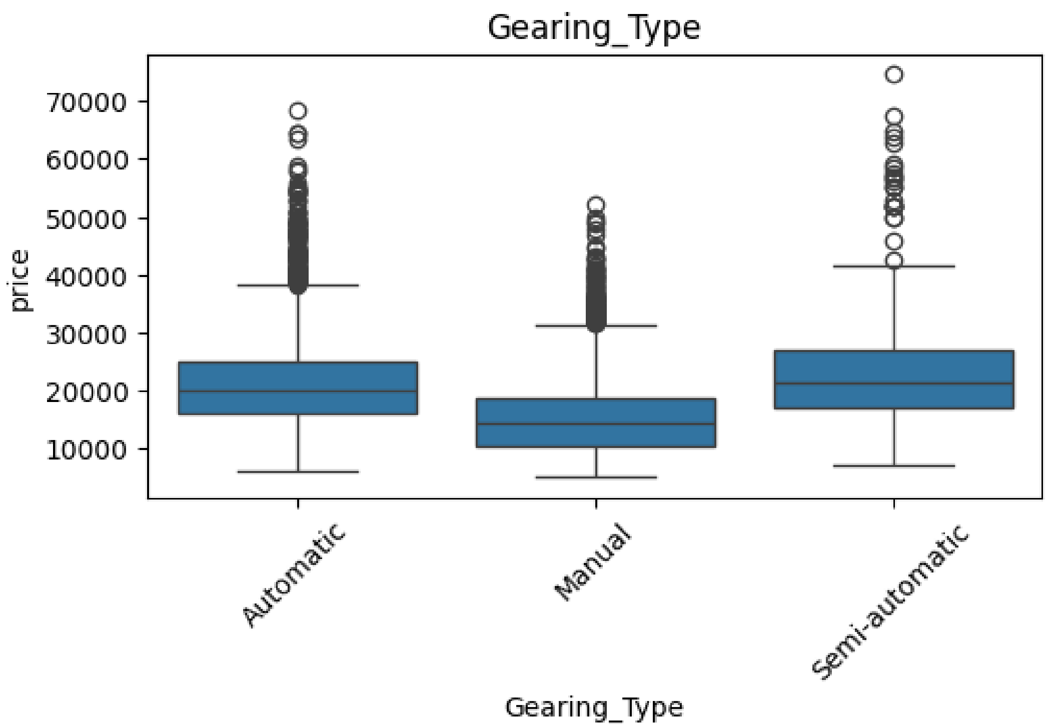


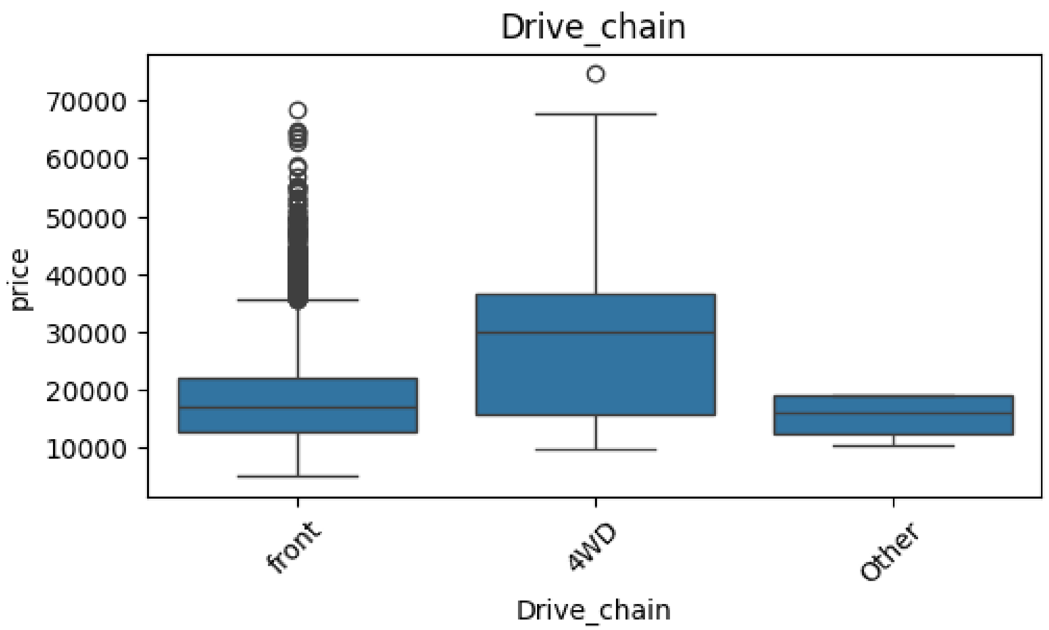
Extras
price
70000
60000
50000
40000
30000
20000
10000
Other
Alloy wheels
Alloy wheels,Sport seats,Voice Control
Alloy wheels,Sport seats
Alloy wheels,Sport suspension
Alloy wheels,Catalytic Converter,Touch screen
Extras


Safety_Security
price
70000
60000
50000
40000
30000
20000
10000
ABS,Central door lock,Daytime running lights,Driver-side airbag,Electronic stability control,Fog lights,Immobilizer,Isofix,Passenger-side airbag,Power steering,Side airbag,Tire pressure monitoring system,Traction control,Xenon headlights
Other
ABS,Central door lock,Daytime running lights,Driver-side airbag,Electronic stability control,Immobilizer,Isofix,Passenger-side airbag,Power steering,Side airbag,Tire pressure monitoring system,Traction control
ABS,Central door lock,Daytime running lights,Driver-side airbag,Electronic stability control,Fog lights,Immobilizer,Isofix,Passenger-side airbag,Power steering,Side airbag,Tire pressure monitoring system,Traction control
ABS,Central door lock,Daytime running lights,Driver-side airbag,Electronic stability control,Isofix,Passenger-side airbag,Power steering,Side airbag,Tire pressure monitoring system,Traction control
ABS,Central door lock,Daytime running lights,Driver-side airbag,Electronic stability control,Immobilizer,Isofix,LED Daytime Running Lights,Passenger-side airbag,Power steering,Side airbag,Tire pressure monitoring system,Traction control
ABS,Central door lock,Daytime running lights,Driver-side airbag,Electronic stability control,Immobilizer,Isofix,Passenger-side airbag,Power steering,Side airbag,Tire pressure monitoring system,Traction control,Xenon headlights
ABS,Central door lock,Daytime running lights,Driver-side airbag,Electronic stability control,Fog lights,Immobilizer,Isofix,LED Daytime Running Lights,Passenger-side airbag,Power steering,Side airbag,Tire pressure monitoring system,Traction control
Safety_Security


Paint_Type
price
70000
60000
50000
40000
30000
20000
10000
Metallic
Uni/basic
Other
Paint_Type

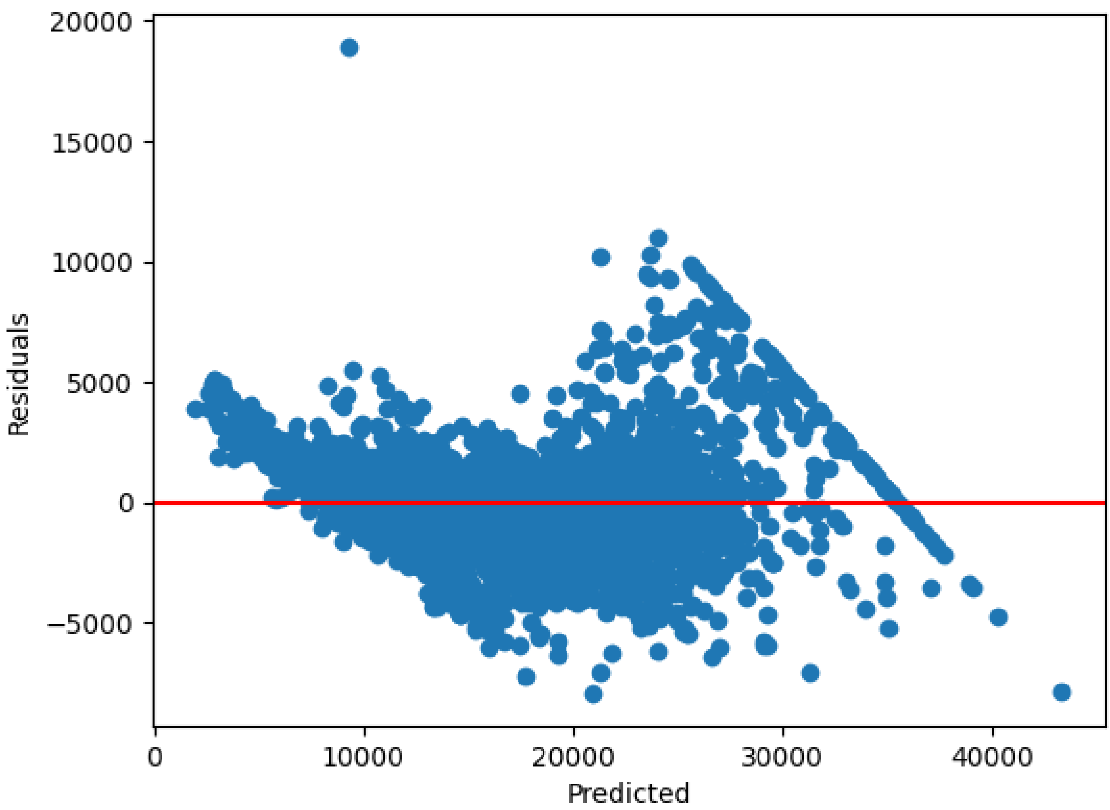

Upholstery_type
price
70000
60000
50000
40000
30000
20000
10000
Cloth
Part/Full Leather
Upholstery_type


Gearing_Type
price
70000
60000
50000
40000
30000
20000
10000
Automatic
Manual
Semi-automatic
Gearing_Type

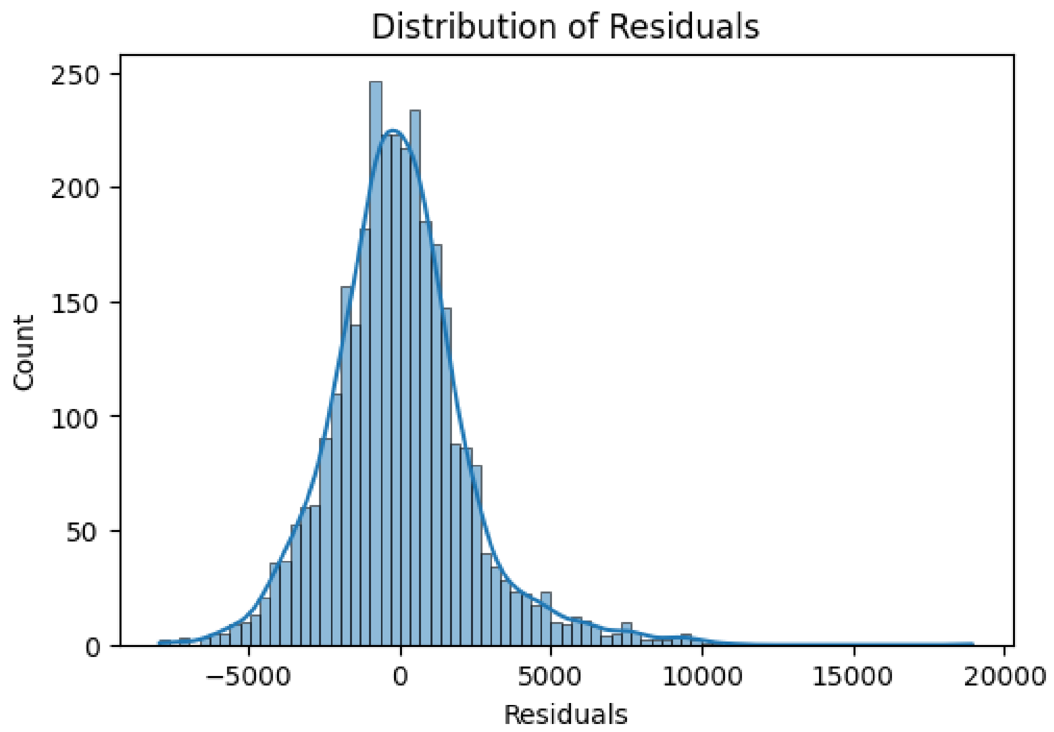


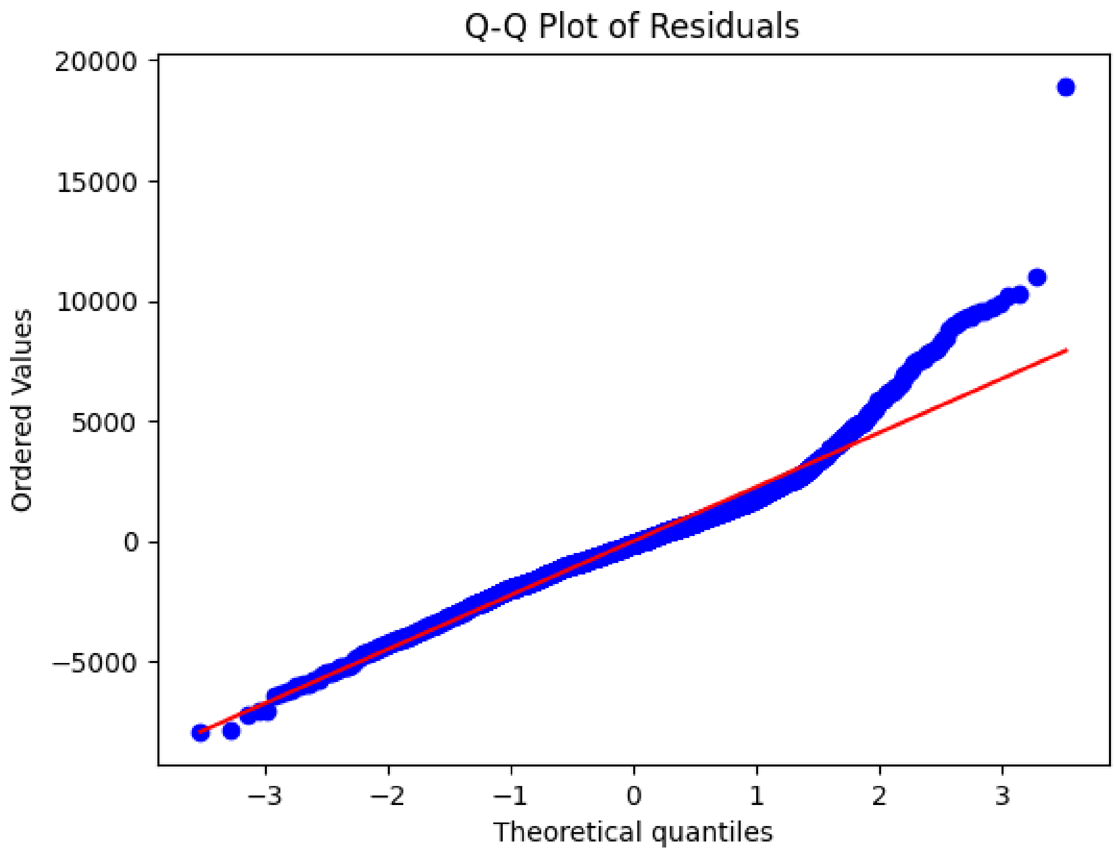
Linear Regression Models

Analyse residuals and check other assumptions of linear regression.



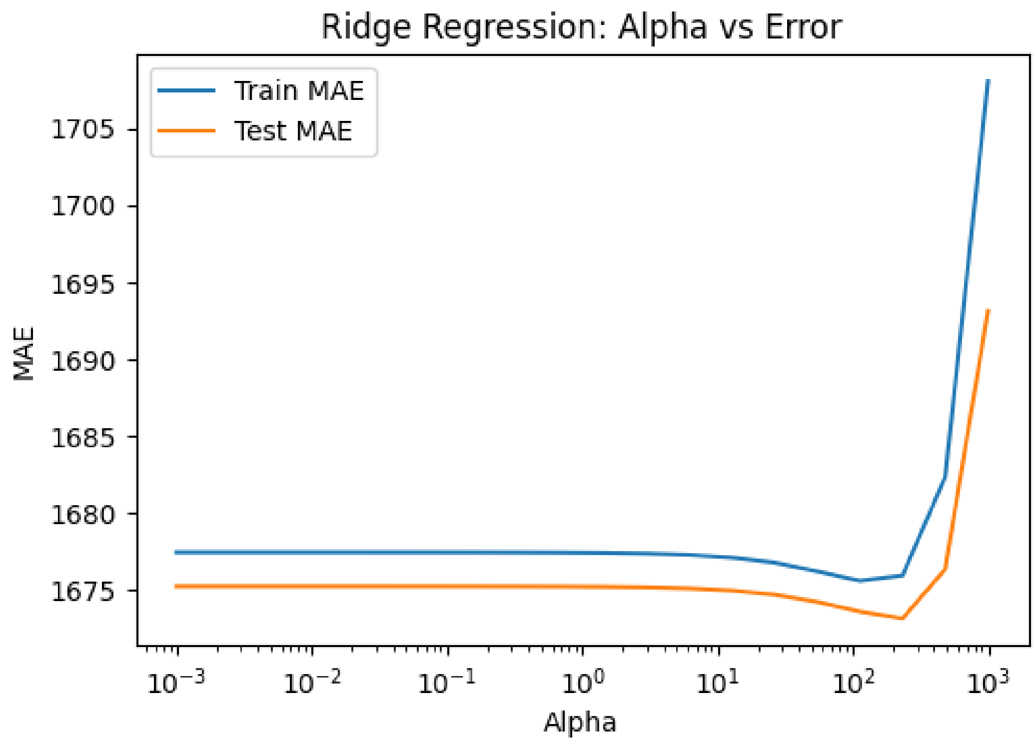
normality in residual distribution



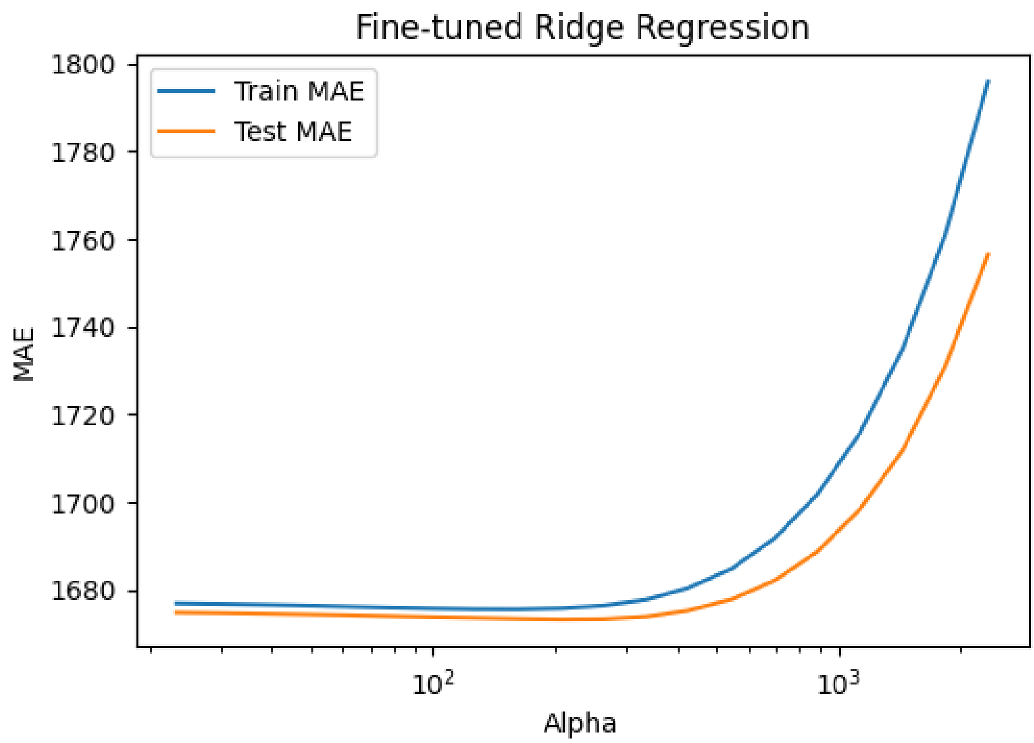


Apply Ridge Regularisation and find the best value of alpha from the list

Plot train and test scores against alpha



Plot the error-alpha graph again and find the actual optimal value for alpha.

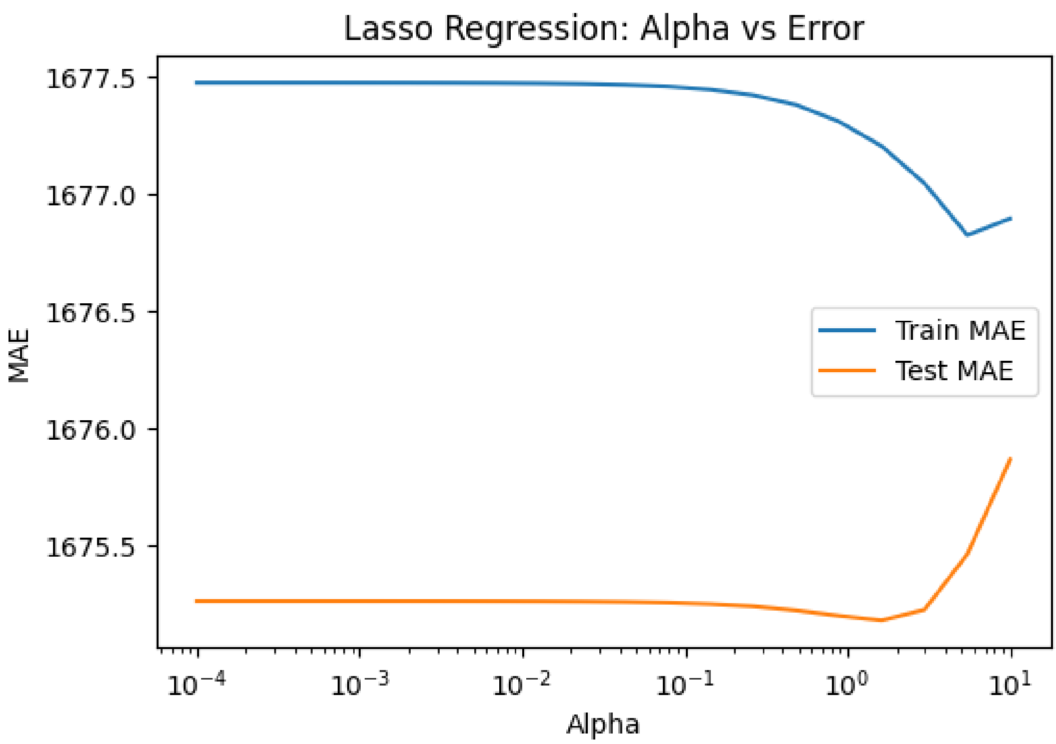


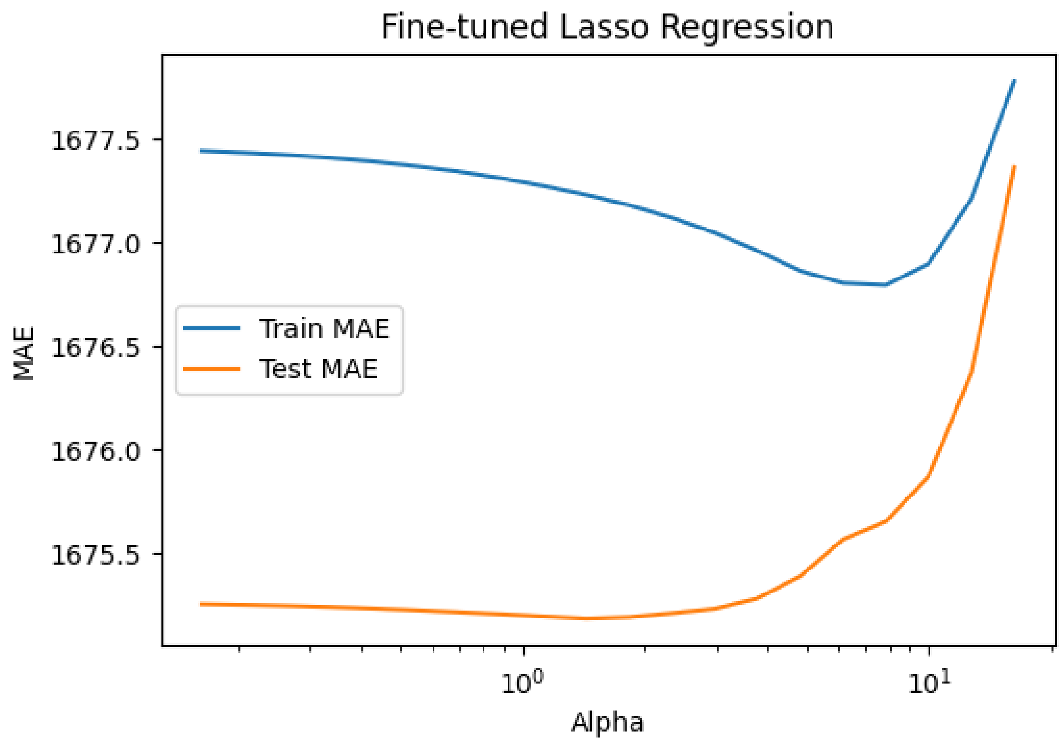
Ridge(alpha=np.float64(206.9138081114788))

	0
hp_kW	2042.446649

	0
age	-1907.786564
make_model_Opel Corsa	-1895.323680
make_model_Renault Clio	-1713.646690
km	-1290.357447
make_model_Opel Astra	-1228.392091
Type_Used	-1189.136552
Gearing_Type_Manual	-905.134462
make_model_Renault Espace	891.978514
Type_Employee's car	-690.264104

Lasso Regression Implementation





Lasso(alpha=np.float64(1.438449888287662), max_iter=5000)

In a Jupyter environment, please rerun this cell to show the HTML representation or trust the notebook.
On GitHub, the HTML representation is unable to render, please try loading this page with nbviewer.org.

Lasso

?Documentation for LassoiFitted

Lasso(alpha=np.float64(1.438449888287662), max_iter=5000)

	0
hp_kW	2126.655867
age	-1955.670793
make_model_Opel Corsa	-1951.300206
make_model_Renault Clio	-1789.776225
make_model_Opel Astra	-1299.732279
Type_Used	-1299.472968
km	-1284.116944
Gearing_Type_Manual	-930.379001

	0
make_model_Renault Espace	904.889208
Type_Employee's car	-763.572395

dtype: float64

Regularisation Comparison & Analysis

Compare metrics for each model

	Model	R2	MAE
0	Linear	0.886127	1675.265187
1	Ridge	0.886029	1673.195330
2	Lasso	0.886146	1675.181613

Compare the coefficients for the three models.

	Linear	Ridge	Lasso
hp_kW	2138.384463	2042.446649	2126.655867
age	1956.363814	1907.786564	1955.670793
make_model_Opel Corsa	1946.802631	1895.323680	1951.300206
make_model_Renault Clio	1793.202014	1713.646690	1789.776225
make_model_Opel Astra	1304.113830	1228.392091	1299.732279
Type_Used	1316.331277	1189.136552	1299.472968
km	1285.918999	1290.357447	1284.116944
Gearing_Type_Manual	930.690630	905.134462	930.379001
make_model_Renault Espace	904.776010	891.978514	904.889208
Type_Employee's car	774.466049	690.264104	763.572395

Conclusion & Key Takeaways

This analysis aimed to build a reliable pricing model for used cars while understanding how different vehicle characteristics influence market value. The baseline linear regression model already performed strongly, explaining nearly 89% of the variation in prices. This indicates that the dataset was rich and informative, and that many pricing drivers follow a largely linear relationship.

Regularisation helped refine the model rather than dramatically improve performance. Ridge regression stabilised coefficient estimates by reducing the impact of multicollinearity, making the model more robust and dependable for real-world use. Lasso regression offered an additional benefit by automatically removing less important features, resulting in a simpler and more interpretable model without sacrificing accuracy. The similar performance across all three models suggests that overfitting was minimal and the data size was sufficient for this task.

From a business perspective, vehicle age, mileage, engine power, and safety and comfort features consistently emerged as the most influential price drivers. This aligns well with buyer expectations and real-world pricing behaviour. Overall, linear models with regularisation are well-suited for used car price prediction. However, to capture more complex market dynamics, future work could explore non-linear models such as tree-based or ensemble methods.