Used Car Evaluation

PROJECT STAT378

Name: Boysen Mutembwa

01. INTRODUCTION

Fundamentals and story behind why we want to study this

02. METHODS

What techniques were used to find the results of study

03. RESULTS

The outcomes of the findings from the various techniques used

04. CONCLUSION

What we can draw from the results

05. FUTURE WORK

How else can we improve the study

06. REFERENCES

The sources used for the study

01

INRODUCTION

Fundamentals and story behind
why we want to study this

The Determinants of Price in Used Cars

A car's features might vary greatly dependent on the model and the manufacturer. The three most important considerations for buyers are price, safety, and luxury. All of these factors have an important role in lowering the frequency of accidents. There are several fundamental aspects that must be considered while acquiring a car. Cars include a variety of performance enhancers, amenities, and safety features. Safety should be a top priority while shopping for a vehicle, convenience elements including a door, a luggage compartment, and maintenance are also included. In the short term, assessing the impact of these policies is difficult because of the complexity of the automobile industry.

Facts About The **Used Car Industry**

Sales hit **over 15 million** vehicles in 2021, up 3.4% from 2020

40.42 million used cars were sold in 2018, compared to just **17.21 million new vehicle**

61.4% of car buyers prefer to buy from a dealership.

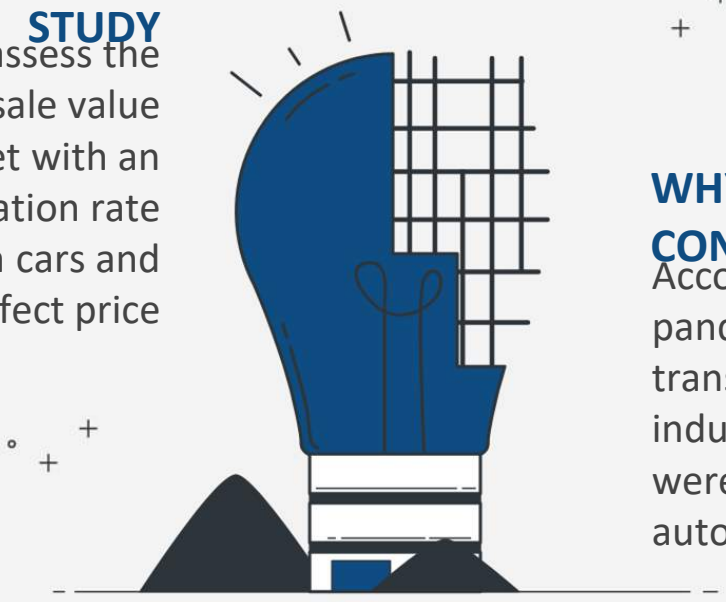
EDMUNDS CAR DEALERSHIP

Edmunds.com Inc. (stylized as edmunds) is an American online automotive resource that includes expert automobile reviews based on testing at the company's own facilities. Prices for new and used automobiles, dealer and inventory listings, a database of national and regional incentives and rebates, vehicle test drive evaluations, and tips and guidance on all areas of car purchases and ownership are all available on **Edmunds.com**. **Edmunds.com's** "True Market Value" pricing tools, which were established in 2000, give statistics. The **Edmunds.com** True Market Value New Vehicle Calculator shows the projected average price that people pay for new cars. **Edmunds.com's** True Market Value Used Vehicle Appraiser analyses actual transaction values for used automobiles purchased and sold by dealers and private individuals.

The Determinants of Price in Used Cars

OBJECTIVE OF THIS STUDY

This paper will assess the determinants of resale value in the used car market with an emphasis on depreciation rate variations between cars and what factors affect price



WHY STUDY WAS THE CONDUCTED?

According to DeLorenzo, the pandemic drove another transformation in the automobile industry: Last year, when automakers were obliged to reduce the number of automobiles for sale

02

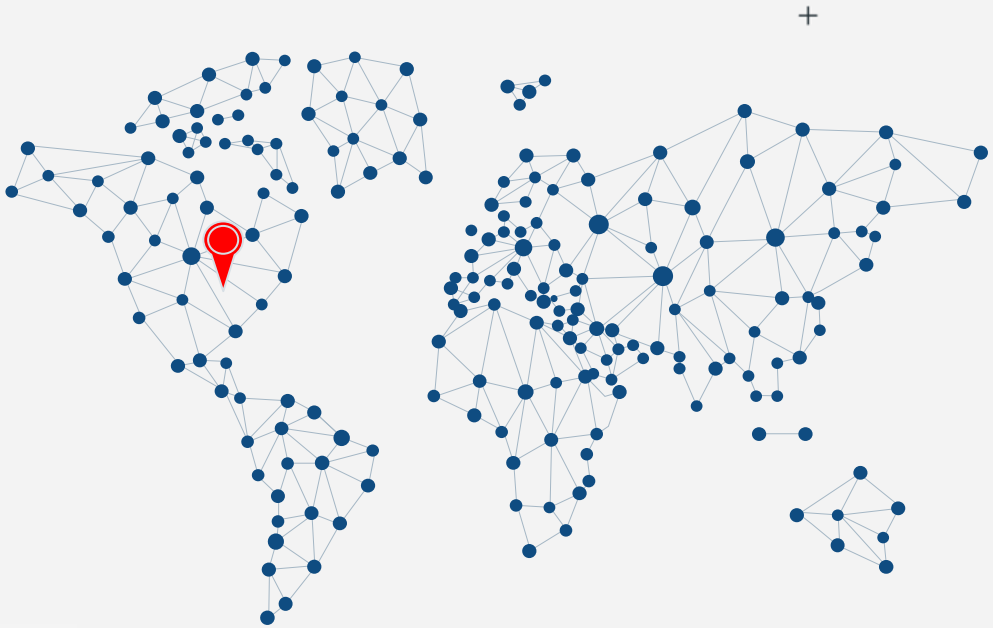
METHODS

What techniques were used to find the results of study

DATA COLLECTION

edmunds

This data was collect in the United States of America.



DATA DESCRIPTION

01

50063 rows of Used car data

02

23 columns of Used car data

03

Used 8 columns for the 3 models

04

Response variable: Price

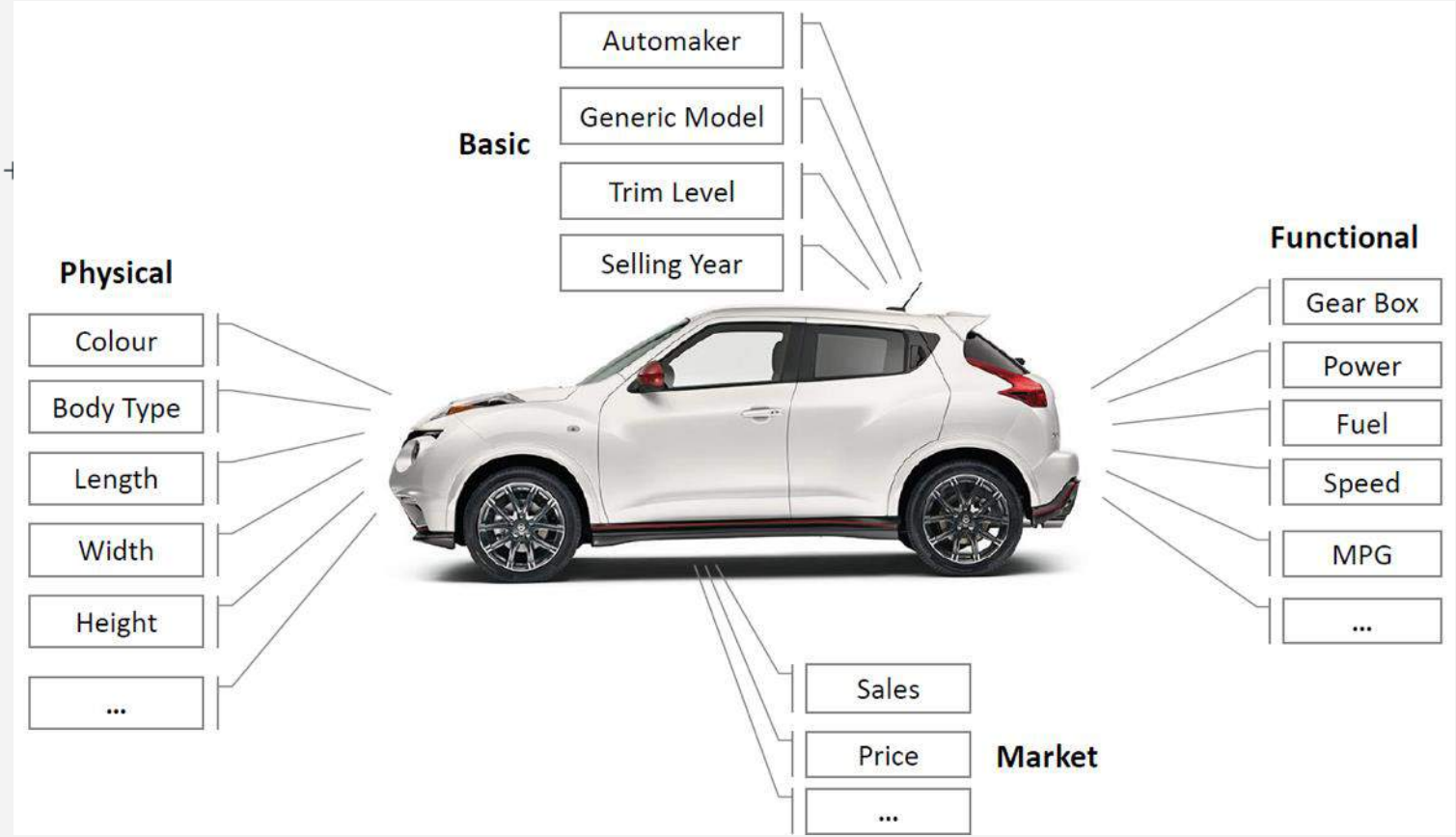
05

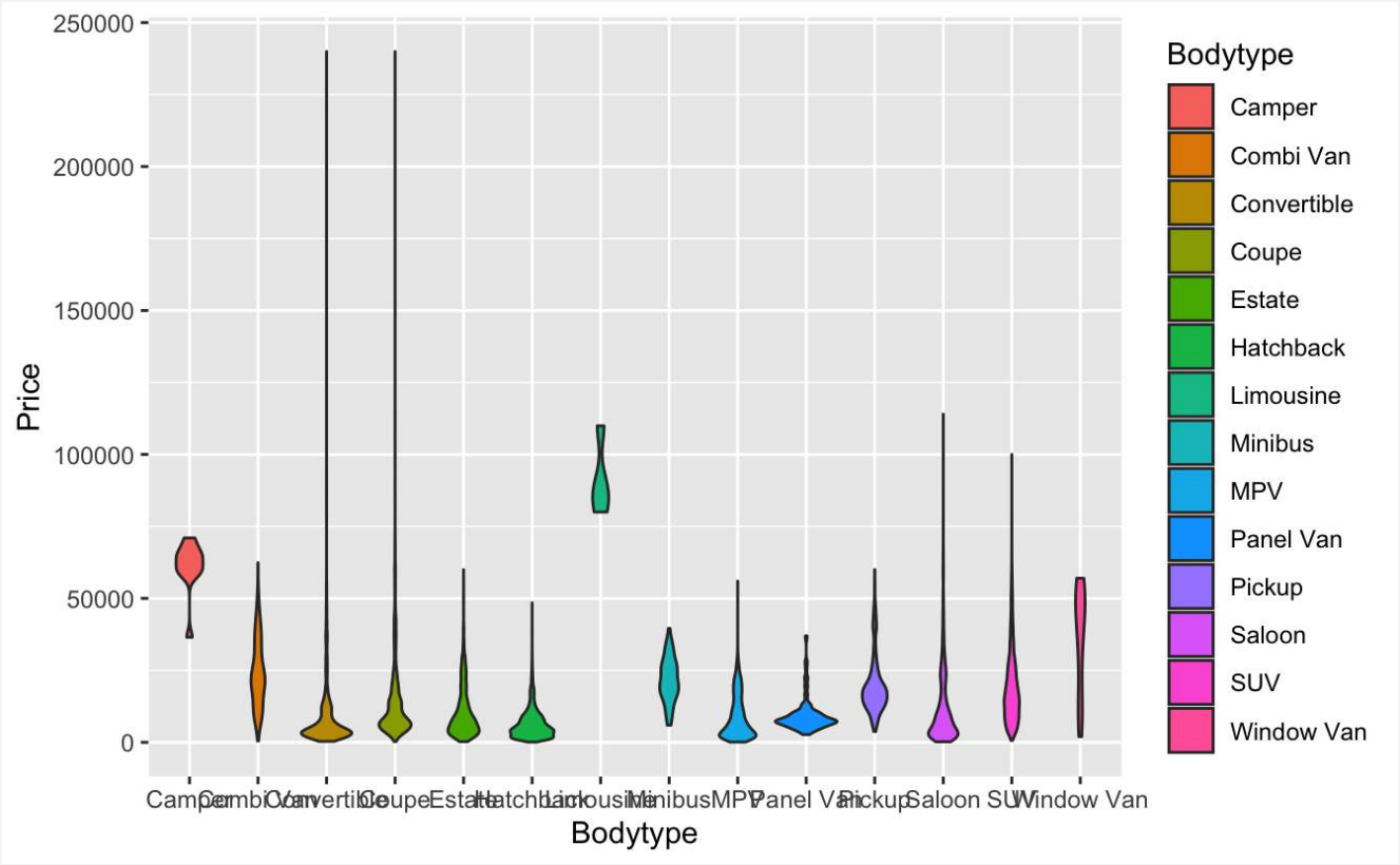
Used Cars ranging from 1968-2021

06

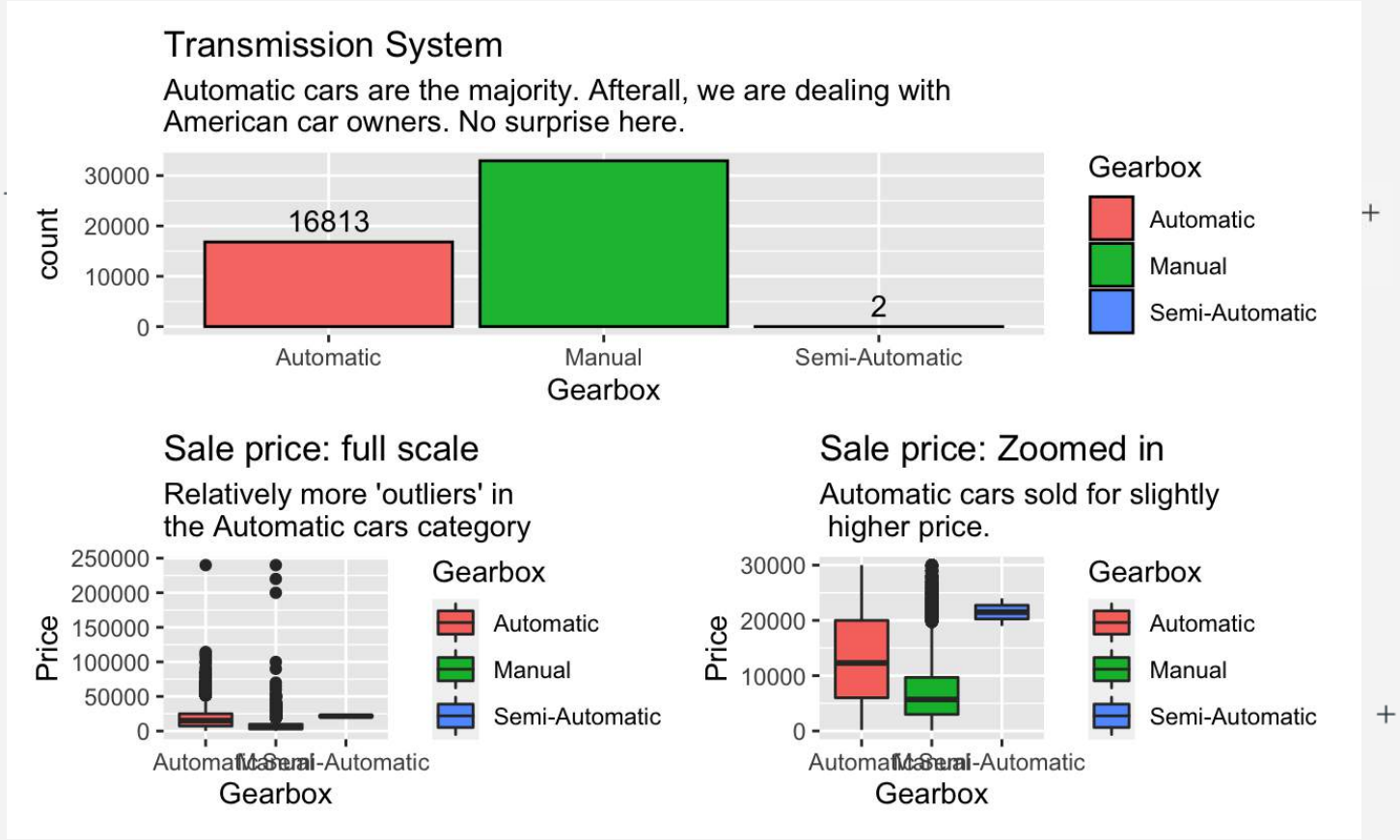
The data used for this regression will be quantitative and categorical in nature.

DATA **DESCRIPTION**





Violine Plot: Bodytype



Box Plots & Bar graph: Transmission

GLM MODELS

Inverse Gaussian

Tweedie

Gamma

GLM: Inverse gaussian vs. Gamma vs. Tweedie

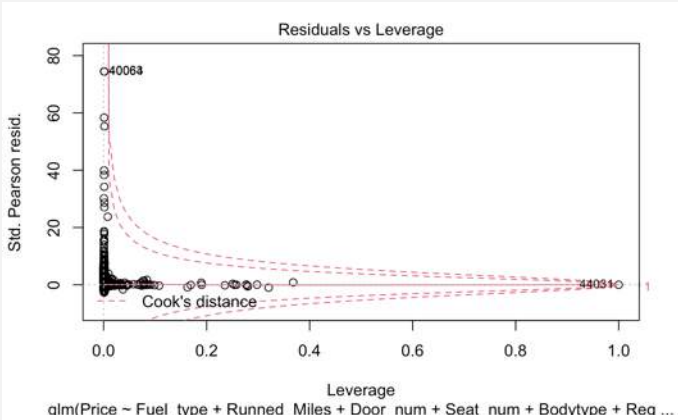
Variables that take positive and continuous values often measure the amount of some physical quantity that is always present like the used car data. The two most common GLMs for this type of data are based on the gamma and inverse Gaussian distributions. Tweedie EDMs are distributions that generalize many of the EDMs already seen (the normal, Poisson, gamma and inverse Gaussian distributions are special cases) and include other distributions also

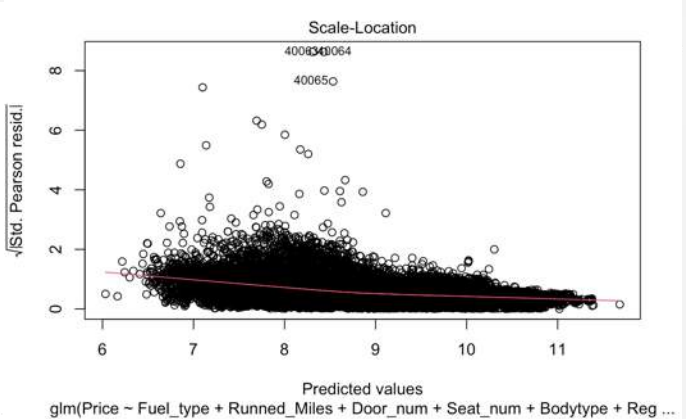
03

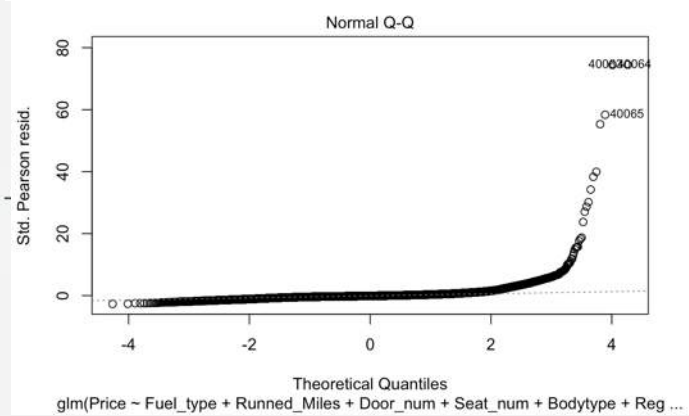
RESULTS

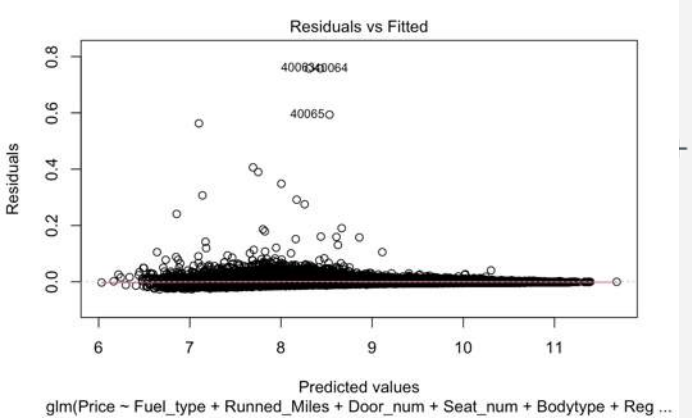
The outcomes of the findings from the various techniques used

GLM: Inverse Gaussian





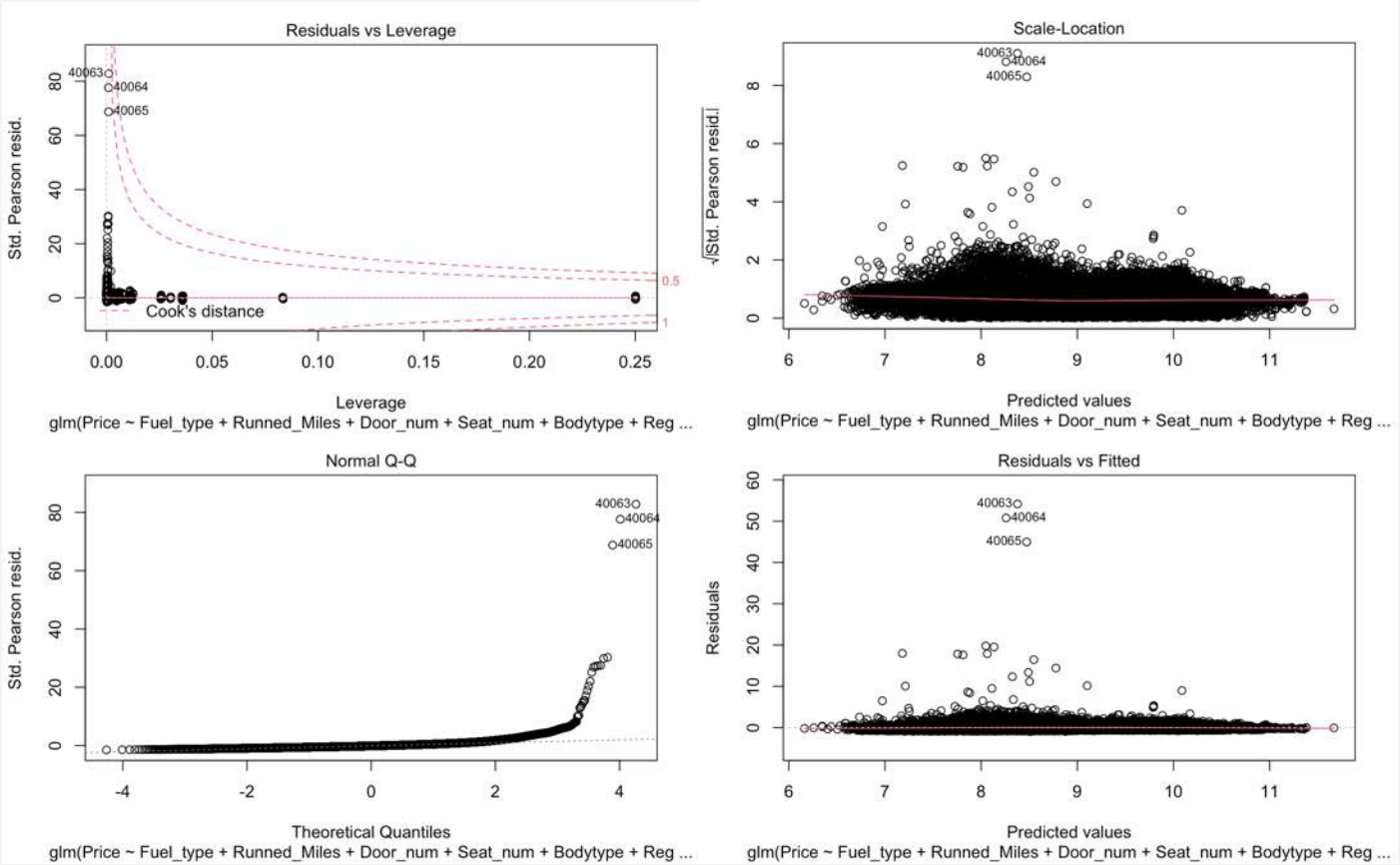




GLM: Inverse Gaussian

```
glm(formula = Price ~ Fuel_type + Runned_Miles + Door_num + Seat_num +
    Bodytype + Reg_year, family = inverse.gaussian(link = "log"),
    data = d, maxit = 1000)

Deviance Residuals:
      Min         1Q     Median         3Q        Max
-0.097958  -0.003726  -0.001212   0.001488   0.124007

Coefficients:
                                    Estimate Std. Error t value Pr(>|t|)
(Intercept)                       -2.111e+02  2.207e+00 -95.659  < 2e-16 ***
Fuel_typeDiesel                    9.107e-01  1.606e-01   5.672 1.42e-08 ***
Fuel_typeDiesel Hybrid             1.025e+00  3.724e-01   2.751  0.00594 **
Fuel_typeElectric                  1.800e+00  1.883e+00   0.956  0.33915
Fuel_typeHybrid  Diesel/Electric  1.169e+00  1.356e+00   0.862  0.38891
Fuel_typeHybrid  Diesel/Electric Plug-in  1.263e+00  2.968e-01   4.256 2.08e-05 ***
Fuel_typeHybrid  Petrol/Electric  1.422e+00  1.643e-01   8.654  < 2e-16 ***
Fuel_typeHybrid  Petrol/Electric Plug-in  1.472e+00  2.097e-01   7.018 2.29e-12 ***
Fuel_typePetrol                    7.247e-01  1.606e-01   4.514 6.38e-06 ***
Fuel_typePetrol Ethanol            6.783e-01  3.381e-01   2.006  0.04484 *
Fuel_typePetrol Hybrid             1.013e+00  1.051e+00   0.964  0.33505
Fuel_typePetrol Plug-in Hybrid     1.068e+00  2.686e-01   3.975 7.04e-05 ***
Runned_Miles                      -7.118e-06  1.088e-07 -65.410  < 2e-16 ***
Door_num                           3.099e-02  4.765e-03   6.503 7.92e-11 ***
Seat_num                           9.304e-02  5.905e-03  15.756  < 2e-16 ***
BodytypeCombi Van                 -4.384e-01  7.312e-01  -0.600  0.54881
BodytypeConvertible               -2.086e-01  7.250e-01  -0.288  0.77357
BodytypeCoupe                     -3.068e-01  7.250e-01  -0.423  0.67216
BodytypeEstate                    -9.548e-01  7.249e-01  -1.317  0.18777
BodytypeHatchback                 -1.345e+00  7.248e-01  -1.856  0.06344 .
BodytypeLimousine                  2.206e+00  1.692e+00   1.304  0.19228
BodytypeMinibus                   -2.575e-01  7.434e-01  -0.346  0.72907
BodytypeMPV                       -1.410e+00  7.249e-01  -1.946  0.05169 .
BodytypePanel Van                 -6.120e-01  7.282e-01  -0.840  0.40064
BodytypePickup                    -3.059e-01  7.270e-01  -0.421  0.67389
BodytypeSaloon                    -7.525e-01  7.249e-01  -1.038  0.29919
BodytypeSUV                       -4.759e-01  7.249e-01  -0.657  0.51144
BodytypeWindow Van                -3.901e-01  7.761e-01  -0.503  0.61520
Reg_year                           1.093e-01  1.032e-03 105.957  < 2e-16 ***
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for inverse.gaussian family taken to be 0.0001036823)

    Null deviance: 7.0777  on 49756  degrees of freedom
Residual deviance: 2.3133  on 49728  degrees of freedom
AIC: 968467

Number of Fisher Scoring iterations: 15
```

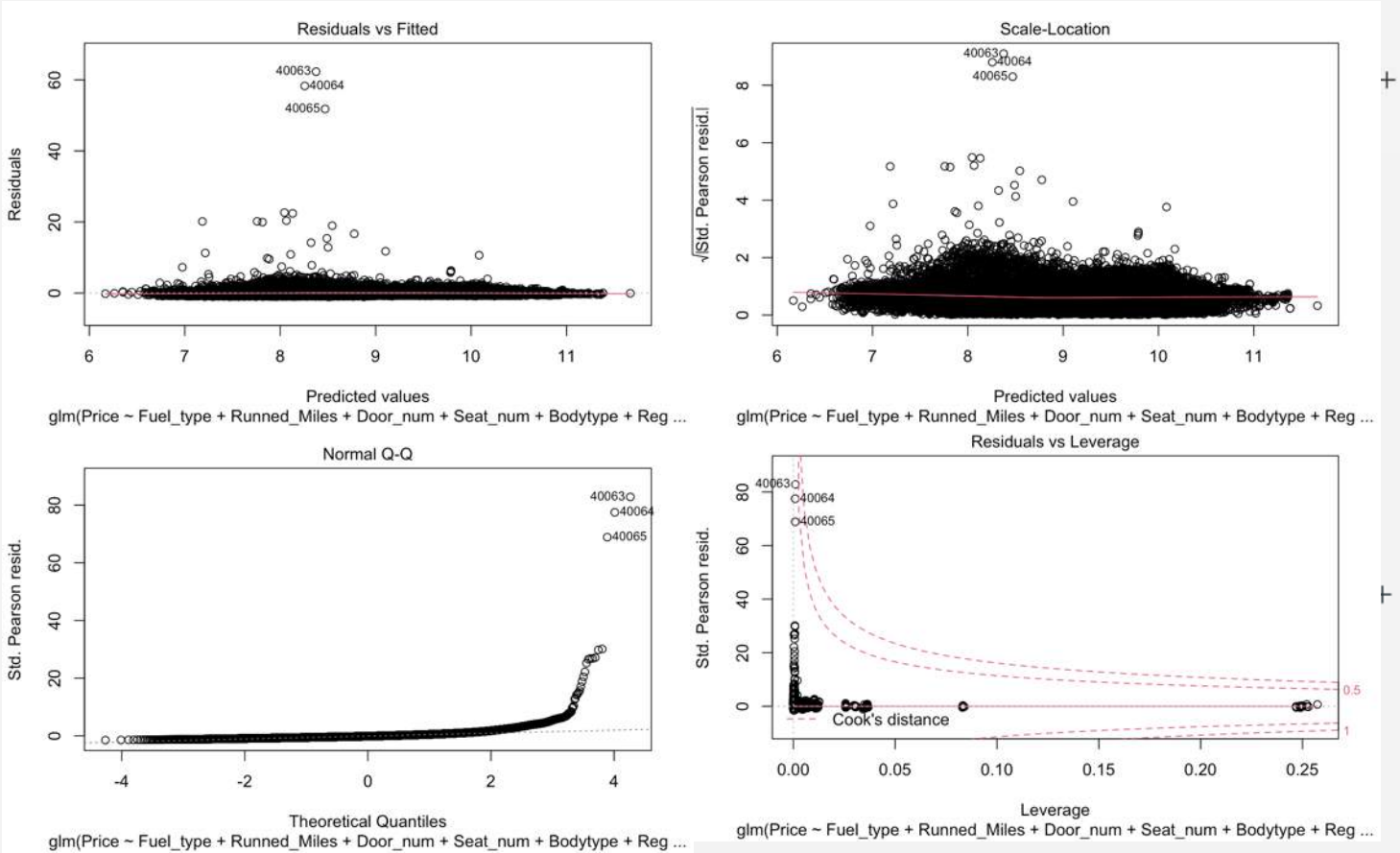
GLM: Gamma


Residuals vs Leverage
Std. Pearson resid.
40063
40064
40065
Cook's distance
0.5
1
Leverage
glm(Price ~ Fuel_type + Runned_Miles + Door_num + Seat_num + Bodytype + Reg ...
Scale-Location
√|Std. Pearson resid.|
Predicted values
glm(Price ~ Fuel_type + Runned_Miles + Door_num + Seat_num + Bodytype + Reg ...
Normal Q-Q
Std. Pearson resid.
Theoretical Quantiles
glm(Price ~ Fuel_type + Runned_Miles + Door_num + Seat_num + Bodytype + Reg ...
Residuals vs Fitted
Residuals
Predicted values
glm(Price ~ Fuel_type + Runned_Miles + Door_num + Seat_num + Bodytype + Reg ...

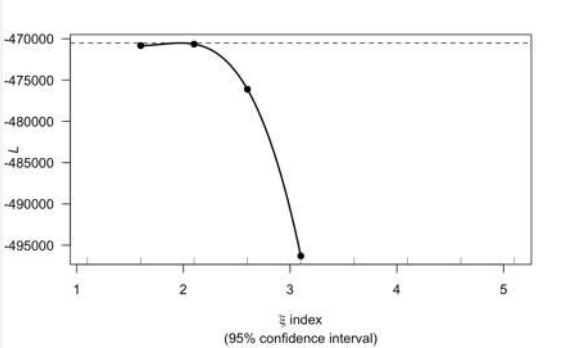
GLM: Gamma

```
Call:
glm(formula = Price ~ Fuel_type + Runned_Miles + Door_num + Seat_num +
    Bodytype + Reg_year, family = Gamma(link = log), data = d)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-2.4186  -0.3124  -0.0922   0.1573  10.0154

Coefficients:
                                      Estimate Std. Error t value Pr(>|t|)
(Intercept)                         -1.973e+02  2.170e+00 -90.929  < 2e-16 ***
Fuel_typeDiesel                      9.200e-01  3.273e-01   2.811 0.004937 **
Fuel_typeDiesel Hybrid               1.118e+00  3.437e-01   3.252 0.001148 **
Fuel_typeElectric                    1.857e+00  7.316e-01   2.538 0.011155 *
Fuel_typeHybrid  Diesel/Electric     1.134e+00  7.316e-01   1.550 0.121188
Fuel_typeHybrid  Diesel/Electric Plug-in  1.254e+00  3.467e-01   3.617 0.000298 ***
Fuel_typeHybrid  Petrol/Electric     1.309e+00  3.281e-01   3.989 6.63e-05 ***
Fuel_typeHybrid  Petrol/Electric Plug-in  1.594e+00  3.304e-01   4.824 1.41e-06 ***
Fuel_typePetrol                      7.331e-01  3.273e-01   2.240 0.025095 *
Fuel_typePetrol Ethanol              6.723e-01  4.628e-01   1.453 0.146334
Fuel_typePetrol Hybrid               1.095e+00  4.628e-01   2.365 0.018014 *
Fuel_typePetrol Plug-in Hybrid       1.169e+00  3.349e-01   3.491 0.000481 ***
Runned_Miles                        -6.948e-06  1.130e-07 -61.466  < 2e-16 ***
Door_num                             3.082e-02  4.909e-03   6.278 3.45e-10 ***
Seat_num                             1.016e-01  5.033e-03  20.195  < 2e-16 ***
BodytypeCombi Van                   -5.316e-01  1.942e-01  -2.738 0.006191 **
BodytypeConvertible                 -3.680e-01  1.899e-01  -1.938 0.052613 .
BodytypeCoupe                       -5.372e-01  1.894e-01  -2.837 0.004561 **
BodytypeEstate                      -1.005e+00  1.893e-01  -5.307 1.12e-07 ***
BodytypeHatchback                   -1.407e+00  1.891e-01  -7.438 1.04e-13 ***
BodytypeLimousine                    2.092e+00  3.780e-01   5.535 3.12e-08 ***
BodytypeMinibus                     -7.841e-01  2.019e-01  -3.884 0.000103 ***
BodytypeMPV                         -1.391e+00  1.896e-01  -7.335 2.24e-13 ***
BodytypePanel Van                   -6.802e-01  1.958e-01  -3.474 0.000514 ***
BodytypePickup                      -5.391e-01  1.910e-01  -2.822 0.004771 **
BodytypeSaloon                      -8.345e-01  1.893e-01  -4.408 1.05e-05 ***
BodytypeSUV                         -6.400e-01  1.892e-01  -3.383 0.000718 ***
BodytypeWindow Van                  -2.732e-01  2.258e-01  -1.210 0.226294
Reg_year                             1.025e-01  1.061e-03  96.602  < 2e-16 ***
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for Gamma family taken to be 0.4280841)

    Null deviance: 42441.9  on 49756  degrees of freedom
Residual deviance:  9083.3  on 49728  degrees of freedom
AIC: 940547

Number of Fisher Scoring iterations: 8
```

GLM: Tweedie


Residuals vs Fitted
Residuals
40063
40064
40065
Predicted values
glm(Price ~ Fuel_type + Runned_Miles + Door_num + Seat_num + Bodytype + Reg ...
Scale-Location
√|Std. Pearson resid.|
40063
40064
40065
Predicted values
glm(Price ~ Fuel_type + Runned_Miles + Door_num + Seat_num + Bodytype + Reg ...
Normal Q-Q
Std. Pearson resid.
40063
40064
40065
Theoretical Quantiles
glm(Price ~ Fuel_type + Runned_Miles + Door_num + Seat_num + Bodytype + Reg ...
Residuals vs Leverage
Std. Pearson resid.
40063
40064
40065
Cook's distance
0.5
1
Leverage
glm(Price ~ Fuel_type + Runned_Miles + Door_num + Seat_num + Bodytype + Reg ...

GLM: Tweedie




```
MLE of xi CI for xi1 CI for xi2
 1.967347   1.998084   1.998084
```

```
glm(formula = Price ~ Fuel_type + Runned_Miles + Door_num + Seat_num +
    Bodytype + Reg_year, family = tweedie(var.power = xi.est,
    link.power = 0), data = d)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-2.7083  -0.3613  -0.1066   0.1818  11.6623

Coefficients:
                                        Estimate Std. Error t value Pr(>|t|)
(Intercept)                           -1.972e+02  2.173e+00 -90.762  < 2e-16 ***
Fuel_typeDiesel                        9.179e-01  3.351e-01   2.739 0.006167 **
Fuel_typeDiesel Hybrid                 1.117e+00  3.502e-01   3.190 0.001425 **
Fuel_typeElectric                      1.856e+00  7.177e-01   2.586 0.009710 **
Fuel_typeHybrid  Diesel/Electric       1.130e+00  7.239e-01   1.561 0.118478
Fuel_typeHybrid  Diesel/Electric Plug-in 1.250e+00  3.533e-01   3.539 0.000402 ***
Fuel_typeHybrid  Petrol/Electric       1.302e+00  3.359e-01   3.876 0.000106 ***
Fuel_typeHybrid  Petrol/Electric Plug-in 1.593e+00  3.380e-01   4.714 2.44e-06 ***
Fuel_typePetrol                        7.312e-01  3.352e-01   2.182 0.029131 *
Fuel_typePetrol Ethanol                6.698e-01  4.698e-01   1.426 0.153988
Fuel_typePetrol Hybrid                 1.093e+00  4.608e-01   2.372 0.017700 *
Fuel_typePetrol Plug-in Hybrid         1.169e+00  3.422e-01   3.416 0.000635 ***
Runned_Miles                          -6.932e-06  1.131e-07 -61.293  < 2e-16 ***
Door_num                               3.037e-02  4.908e-03   6.189 6.11e-10 ***
Seat_num                               1.024e-01  4.988e-03  20.523  < 2e-16 ***
BodytypeCombi Van                     -5.350e-01  1.867e-01  -2.865 0.004165 **
BodytypeConvertible                   -3.717e-01  1.825e-01  -2.037 0.041658 *
BodytypeCoupe                         -5.420e-01  1.820e-01  -2.978 0.002900 **
BodytypeEstate                        -1.005e+00  1.819e-01  -5.525 3.32e-08 ***
BodytypeHatchback                     -1.408e+00  1.817e-01  -7.748 9.52e-15 ***
BodytypeLimousine                      2.090e+00  3.614e-01   5.782 7.41e-09 ***
BodytypeMinibus                       -7.960e-01  1.944e-01  -4.095 4.23e-05 ***
BodytypeMPV                           -1.389e+00  1.822e-01  -7.627 2.45e-14 ***
BodytypePanel Van                     -6.808e-01  1.886e-01  -3.610 0.000306 ***
BodytypePickup                        -5.440e-01  1.836e-01  -2.964 0.003042 **
BodytypeSaloon                        -8.355e-01  1.819e-01  -4.593 4.38e-06 ***
BodytypeSUV                           -6.443e-01  1.818e-01  -3.545 0.000394 ***
BodytypeWindow Van                    -2.712e-01  2.176e-01  -1.246 0.212825
Reg_year                               1.025e-01  1.062e-03  96.450  < 2e-16 ***
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for Tweedie family taken to be 0.5661851)

    Null deviance: 56895  on 49756  degrees of freedom
Residual deviance: 12086  on 49728  degrees of freedom
AIC: NA

Number of Fisher Scoring iterations: 8
```

AIC COMPARISON

GLM	Pseudo-R^2	AIC	Dispersion
Inverse Gaussian	0.6731551	968467	0.0001037
Gamma	0.785983	940547	0.4280841
Tweedie	0.7875824	940372.9	0.5661851

04

CONCLUSION

What we can draw from the results

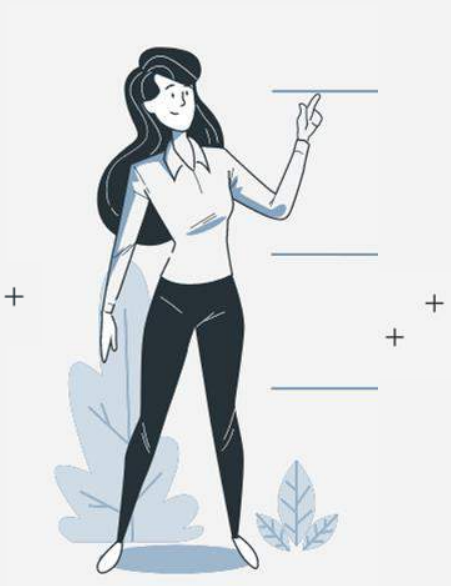
CONCLUSION

Findings 1

Tweedie model was the best based on the AIC and Pseudo-R^2

Findings 2

Body Type of car and year were a big contributor to price



Findings 3

Age and mileage are negative contributors to the resale value

Findings 4

Electric cars tend to retain their price

05

FUTURE WORK

How else can we improve the study

FUTURE WORK

- More explorative study on comparing prices of different car dealerships
- Study on how price of the same model of car decreases or increases price based on mileage

06

REFERENCES

How else can we improve the study

Bajari, P., & Hortacsu, A. (2004). Economic insights from internet auctions. Journal of Economic Literature, 42(2), 457–486 (June)

Cho, Sung Jin. "The Determinants of Used Rental Car Prices." Journal of Economic Research 10, no. 2 (2005): 277-304.

Dunn, P.K., Smyth, G.K.: Randomized quantile residuals. Journal of Computational and Graphical Statistics 5(3), 236–244 (1996) [5]

Feigl, P., Zelen, M.: Estimation of exponential survival probabilities with concomitant information. Biometrics 21, 826–838 (1965) [6] Fox, R.,

Zhu, Rui (Juliet), Xinlei (Jack) Chen, and Srabana Dasgupta. "Can Trade-Ins Hurt You? Exploring the Effect of a Trade-In on Consumers' Willingness to Pay for a New Product." Journal of Marketing Research (JMR) 45, no. 2 (2008): 159-170.

THANKS!

Do you have any questions about the project?