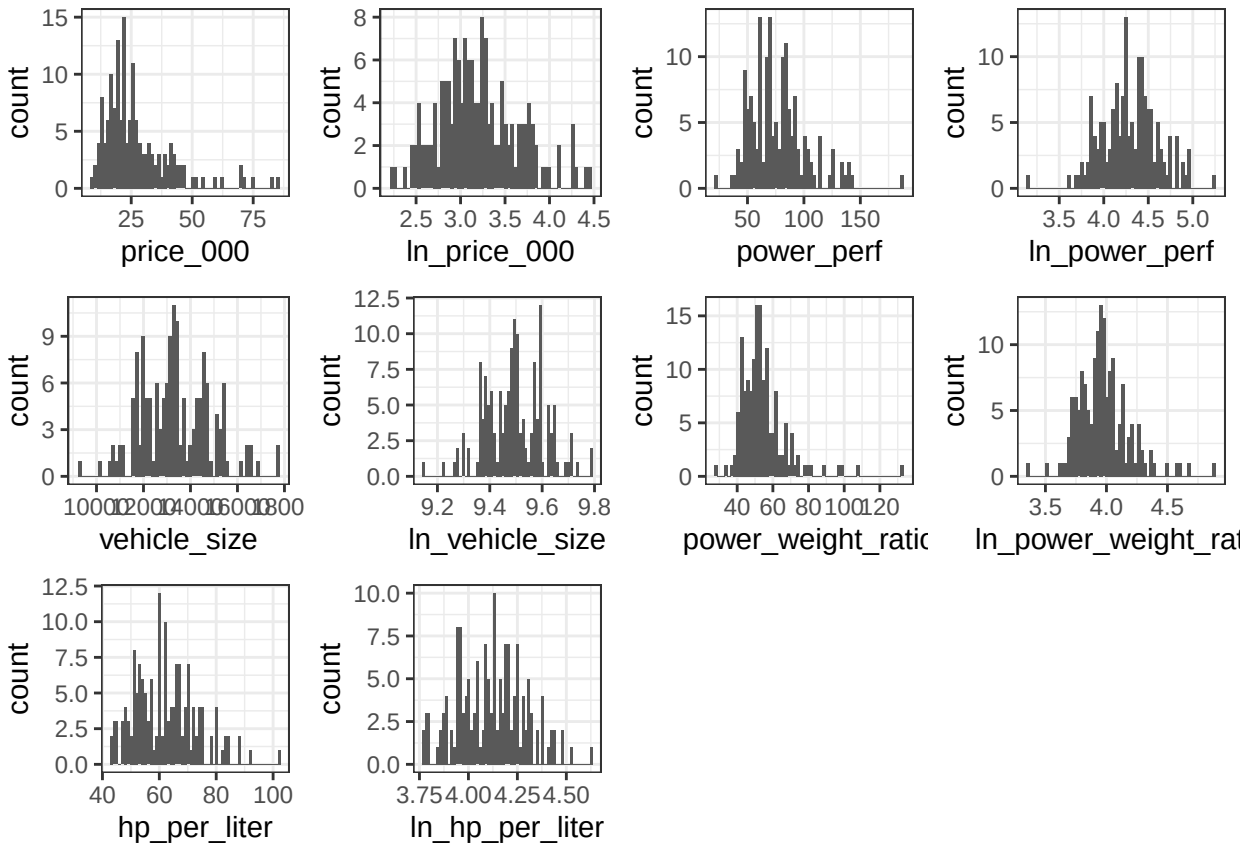
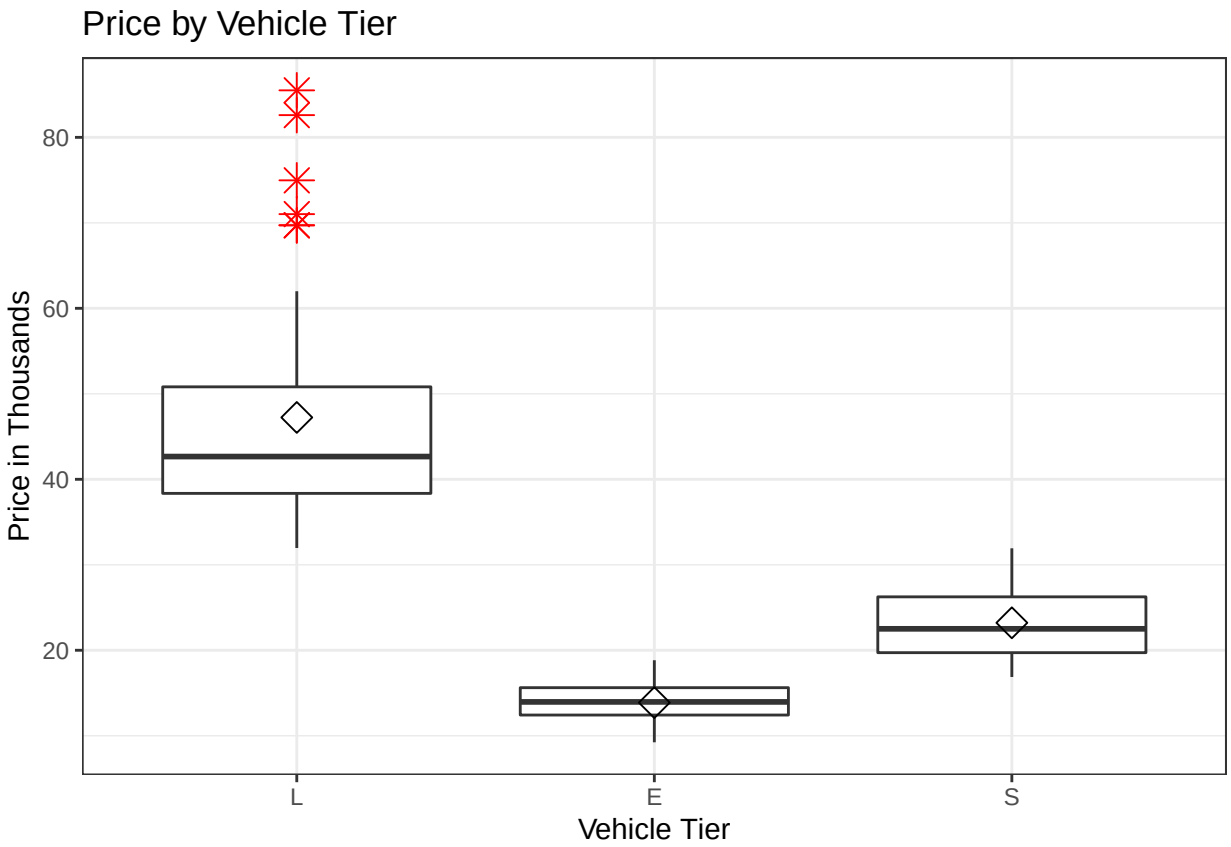
Car Price Estimating Model

The Principal Components Group - Ed Brown, Daphne Lin, Linh Tran, Lisa Wu

2022-07-24


count
price_000
ln_price_000
power_perf
ln_power_perf
vehicle_size
ln_vehicle_size
power_weight_ratio
ln_power_weight_rat
hp_per_liter
ln_hp_per_liter


Price by Vehicle Tier
Price in Thousands
80
60
40
20
L
E
S
Vehicle Tier

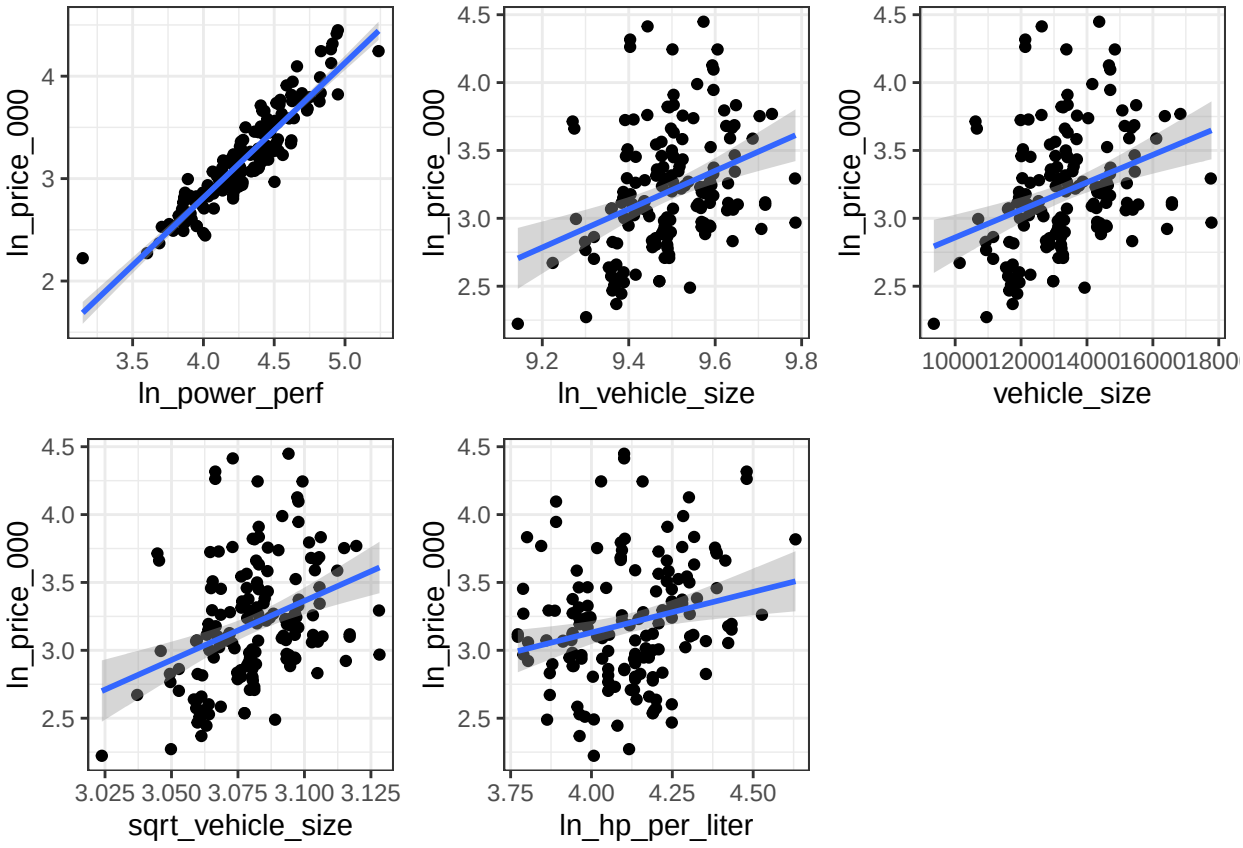

ln_price_000
ln_power_perf
ln_vehicle_size
vehicle_size
sqrt_vehicle_size
ln_hp_per_liter

Figure 1: Car Price and Power Performance Plot by Vehicle Tier

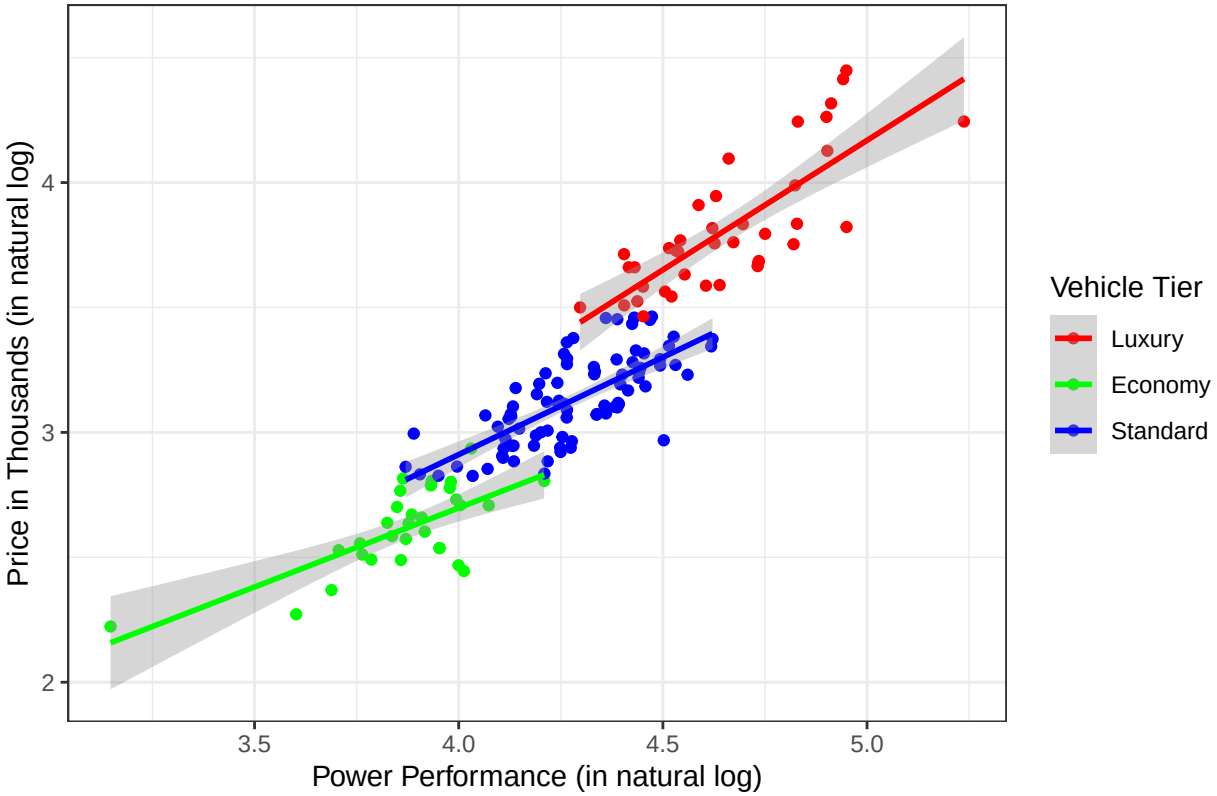
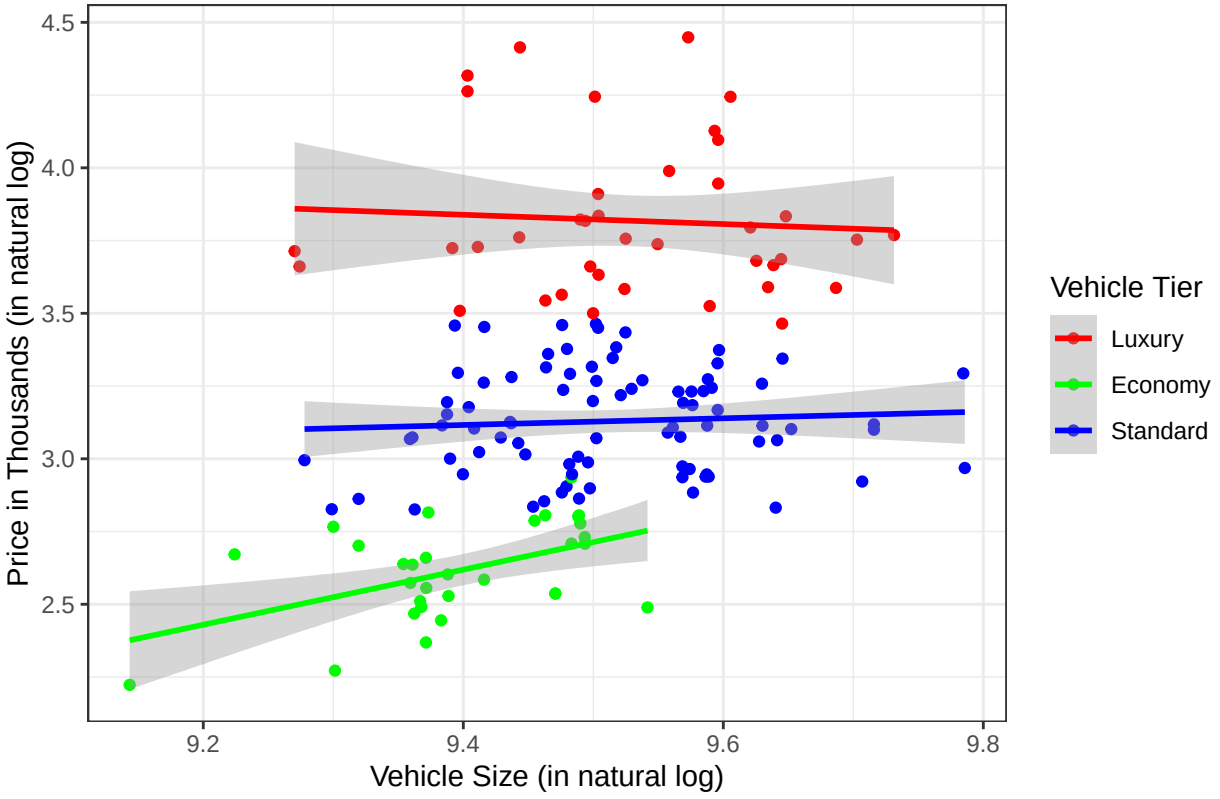
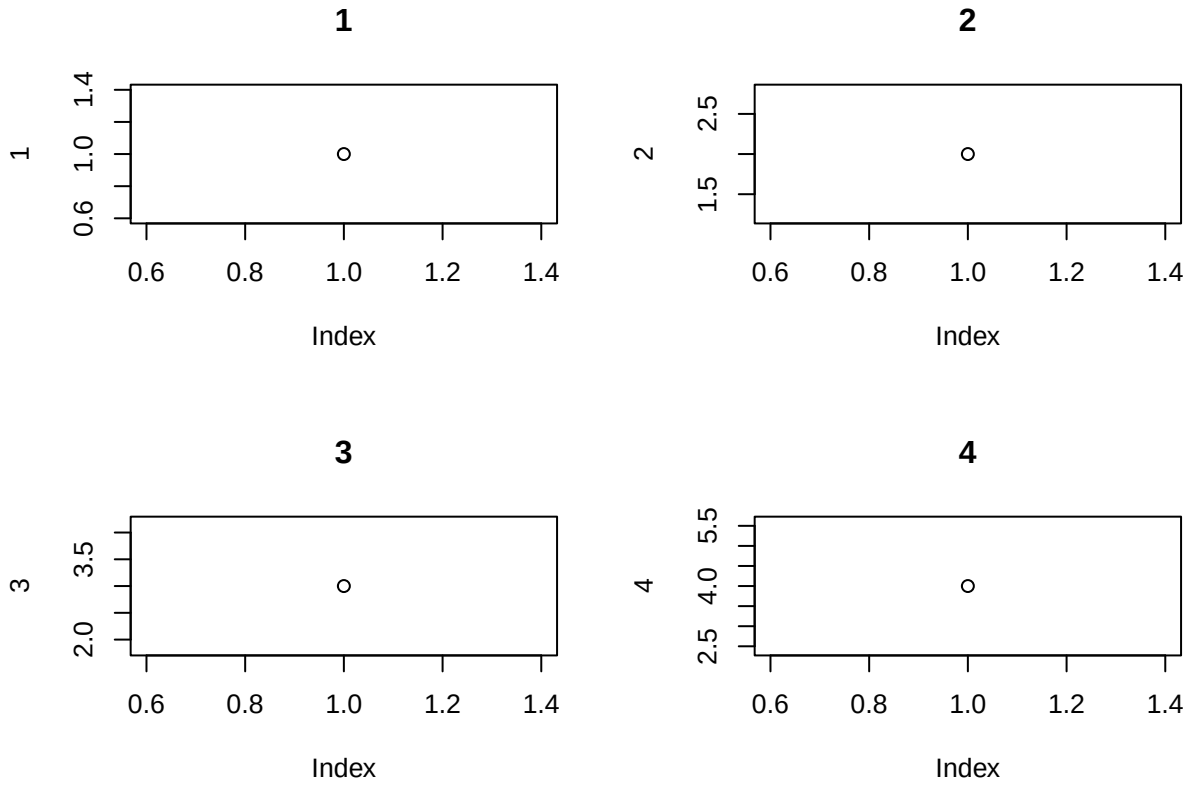


Figure 2: Car Price and Vehicle Size Plot by Vehicle Tier




```
## Proposed Second Model
layout(matrix(c(1, 2, 3, 4), nrow = 2, ncol = 2, byrow = TRUE))
plot(1, main = 1)
plot(2, main = 2)
plot(3, main = 3)
plot(4, main = 4)
```




```
car_base_size <- car_sales %>% lm(ln_price_000 ~ ln_power_perf + vehicle_tier, data = .)
car_base_size_se <- car_base_size %>%
  vcovHC(type = "HC1") %>%
  diag() %>%
  sqrt()
stargazer(car_base_size, type = "text", se = list(car_base_size_se))
```

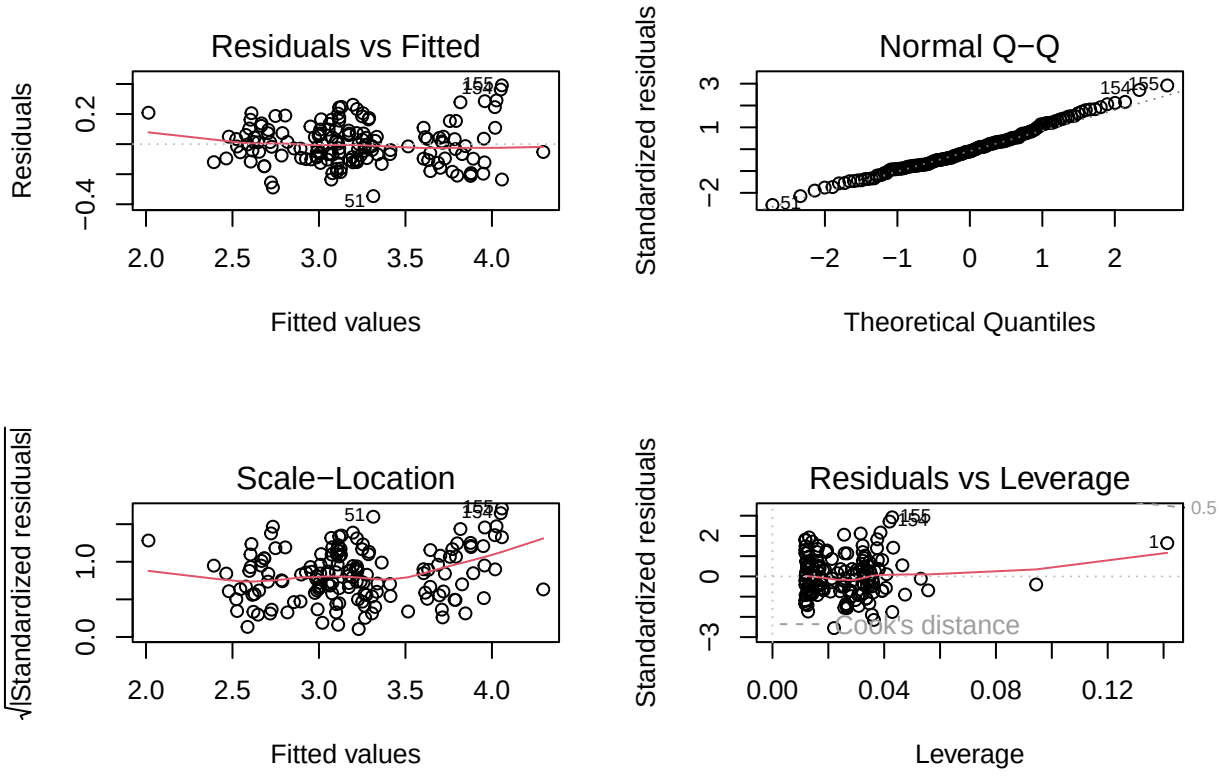
```
##
## ================================================
##                         Dependent variable:
##                     ---------------------------
##                              ln_price_000
## ------------------------------------------------
## ln_power_perf                 0.831***
##                                (0.068)
##
## vehicle_tierE                 -0.545***
##                                (0.053)
##
## vehicle_tierS                 -0.371***
##                                (0.033)
##
## Constant                       -0.057
##                                (0.305)
##
## ------------------------------------------------
```

```
## Observations                    155
## R2                             0.912
## Adjusted R2                    0.910
## Residual Std. Error      0.137 (df = 151)
## F Statistic          519.058*** (df = 3; 151)
## ================================================
## Note:                *p<0.1; **p<0.05; ***p<0.01
```

```
coeftest(car_base_size, vconv = vcovHC(type = "HC1"))
```

```
##
## t test of coefficients:
##
##                Estimate Std. Error  t value  Pr(>|t|)
## (Intercept)   -0.056931   0.291350  -0.1954    0.8453
## ln_power_perf  0.831288   0.062327  13.3375 < 2.2e-16 ***
## vehicle_tierE -0.545152   0.059163  -9.2144 2.511e-16 ***
## vehicle_tierS -0.371459   0.035598 -10.4348 < 2.2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
plot(car_base_size)
```




```
lmtest::bptest(car_base_size)
```

```
##
##  studentized Breusch-Pagan test
##
```

```
## data:  car_base_size
## BP = 16.807, df = 3, p-value = 0.0007742
```

```
ols_vif_tol(car_base_size)
```

```
##       Variables Tolerance      VIF
## 1 ln_power_perf 0.3091439 3.234740
## 2 vehicle_tierE 0.2158614 4.632602
## 3 vehicle_tierS 0.3852018 2.596042
```

Table 1: Estimated Car Price Linear Regression Models

	Output Variable: Price in Thousands of Dollars (in natural log)		
	(1)	(2)	(3)
Power Performance Ratio (in natural log)	1.318***	0.831***	0.831***
	(0.061)	(0.068)	(0.068)
Vehicle Tier-Economy		−0.545***	−0.545***
		(0.053)	(0.053)
Vehicle Tier-Mid		−0.371***	−0.371***
		(0.033)	(0.033)
Constant	−2.458***	−0.057	−0.057
	(0.259)	(0.305)	(0.305)
Observations	155	155	155
R^2	0.846	0.912	0.912
Adjusted R^2	0.845	0.910	0.910
Residual Std. Error	0.179 (df = 153)	0.137 (df = 151)	0.137 (df = 151)
F Statistic	841.934*** (df = 1; 153)	519.058*** (df = 3; 151)	519.058*** (df = 3; 15
Note:	*p<0.05; **p<0.01; ***p<0.001 HC_1 robust standard errors in parentheses. Luxury Vehicles are the base Tier		