Exercises Auto Multi

February 21, 2025

1 Multilinear Regression: Auto dataset

1.1 Import notebook functions

```
[1]: from notebookfuncs import *
```

1.2 Import standard libraries

```
import numpy as np
import pandas as pd

pd.set_option("display.max_rows", 1000)
pd.set_option("display.max_columns", 1000)
pd.set_option("display.width", 1000)
pd.set_option("display.max.colwidth", None)
import matplotlib.pyplot as plt
from matplotlib.pyplot import subplots
import seaborn as sns
import itertools
```

1.3 New imports

```
[3]: import statsmodels.api as sm
```

1.4 Import statsmodels.objects

1.5 Import ISLP objects

```
[5]: import ISLP from ISLP import models from ISLP import load_data
```

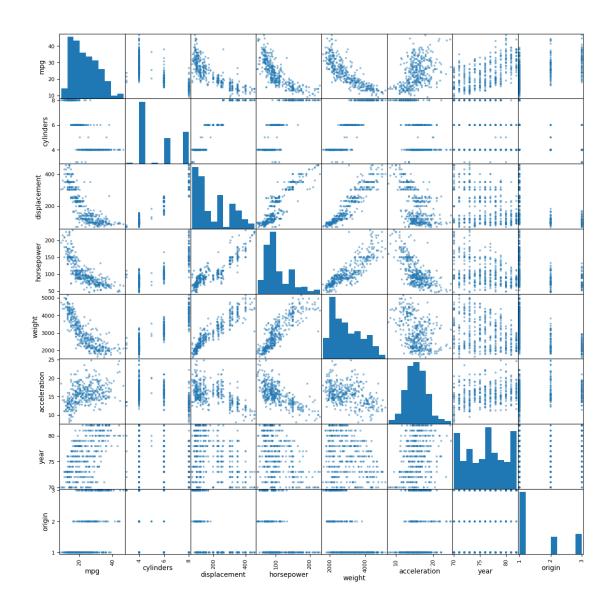
```
Import user functions
 [6]: from userfuncs import *
     Set level of significance (alpha)
 [7]: LOS Alpha = 0.01
 [7]: 0.01
 [8]: Auto = load_data("Auto")
      Auto = Auto.sort_values(by=["year"], ascending=True)
      Auto.head()
      Auto.columns
 [8]: Index(['mpg', 'cylinders', 'displacement', 'horsepower', 'weight',
      'acceleration', 'year', 'origin'], dtype='object')
 [9]: Auto.shape
 [9]: (392, 8)
[10]: Auto.describe()
[10]:
                           cylinders displacement horsepower
                                                                      weight
                    mpg
      acceleration
                          year
                                     origin
                         392.000000
      count 392.000000
                                        392.000000
                                                     392.000000
                                                                  392.000000
      392.000000 392.000000 392.000000
      mean
              23.445918
                            5.471939
                                        194.411990
                                                     104.469388
                                                                 2977.584184
      15.541327
                  75.979592
                                1.576531
      std
               7.805007
                            1.705783
                                        104.644004
                                                      38.491160
                                                                  849.402560
      2.758864
                  3.683737
                               0.805518
               9.000000
                            3.000000
                                         68.000000
                                                      46.000000
                                                                 1613.000000
      min
      8.000000
                 70.000000
                               1.000000
      25%
              17.000000
                            4.000000
                                        105.000000
                                                      75.000000
                                                                 2225.250000
      13.775000
                  73.000000
                                1.000000
      50%
              22.750000
                            4.000000
                                        151.000000
                                                      93.500000 2803.500000
      15.500000
                  76.000000
                                1.000000
      75%
                                                     126.000000 3614.750000
              29.000000
                            8.000000
                                        275.750000
      17.025000
                  79.000000
                                2.000000
                            8.000000
                                        455.000000
                                                     230.000000 5140.000000
      max
              46.600000
      24.800000
                  82.000000
                                3.000000
```

from ISLP.models import ModelSpec as MS, summarize, poly

- 1.6 9. This question involves the use of multiple linear regression on the Auto data set.
- 1.6.1 (a) Produce a scatterplot matrix which includes all of the variables in the data set.

```
[11]: pd.plotting.scatter_matrix(Auto, figsize=(14, 14))
[11]: array([[<Axes: xlabel='mpg', ylabel='mpg'>,
              <Axes: xlabel='cylinders', ylabel='mpg'>,
              <Axes: xlabel='displacement', ylabel='mpg'>,
              <Axes: xlabel='horsepower', ylabel='mpg'>,
              <Axes: xlabel='weight', ylabel='mpg'>,
              <Axes: xlabel='acceleration', ylabel='mpg'>,
              <Axes: xlabel='year', ylabel='mpg'>,
              <Axes: xlabel='origin', ylabel='mpg'>],
             [<Axes: xlabel='mpg', ylabel='cylinders'>,
              <Axes: xlabel='cylinders', ylabel='cylinders'>,
              <Axes: xlabel='displacement', ylabel='cylinders'>,
              <Axes: xlabel='horsepower', ylabel='cylinders'>,
              <Axes: xlabel='weight', ylabel='cylinders'>,
              <Axes: xlabel='acceleration', ylabel='cylinders'>,
              <Axes: xlabel='year', ylabel='cylinders'>,
              <Axes: xlabel='origin', ylabel='cylinders'>],
             [<Axes: xlabel='mpg', ylabel='displacement'>,
              <Axes: xlabel='cylinders', ylabel='displacement'>,
              <Axes: xlabel='displacement', ylabel='displacement'>,
              <Axes: xlabel='horsepower', ylabel='displacement'>,
              <Axes: xlabel='weight', ylabel='displacement'>,
              <Axes: xlabel='acceleration', ylabel='displacement'>,
              <Axes: xlabel='year', ylabel='displacement'>,
              <Axes: xlabel='origin', ylabel='displacement'>],
             [<Axes: xlabel='mpg', ylabel='horsepower'>,
              <Axes: xlabel='cylinders', ylabel='horsepower'>,
              <Axes: xlabel='displacement', ylabel='horsepower'>,
              <Axes: xlabel='horsepower', ylabel='horsepower'>,
              <Axes: xlabel='weight', ylabel='horsepower'>,
              <Axes: xlabel='acceleration', ylabel='horsepower'>,
              <Axes: xlabel='year', ylabel='horsepower'>,
              <Axes: xlabel='origin', ylabel='horsepower'>],
             [<Axes: xlabel='mpg', ylabel='weight'>,
              <Axes: xlabel='cylinders', ylabel='weight'>,
              <Axes: xlabel='displacement', ylabel='weight'>,
              <Axes: xlabel='horsepower', ylabel='weight'>,
              <Axes: xlabel='weight', ylabel='weight'>,
              <Axes: xlabel='acceleration', ylabel='weight'>,
              <Axes: xlabel='year', ylabel='weight'>,
              <Axes: xlabel='origin', ylabel='weight'>],
```

```
[<Axes: xlabel='mpg', ylabel='acceleration'>,
<Axes: xlabel='cylinders', ylabel='acceleration'>,
<Axes: xlabel='displacement', ylabel='acceleration'>,
<Axes: xlabel='horsepower', ylabel='acceleration'>,
<Axes: xlabel='weight', ylabel='acceleration'>,
<Axes: xlabel='acceleration', ylabel='acceleration'>,
<Axes: xlabel='year', ylabel='acceleration'>,
<Axes: xlabel='origin', ylabel='acceleration'>],
[<Axes: xlabel='mpg', ylabel='year'>,
<Axes: xlabel='cylinders', ylabel='year'>,
<Axes: xlabel='displacement', ylabel='year'>,
<Axes: xlabel='horsepower', ylabel='year'>,
<Axes: xlabel='weight', ylabel='year'>,
<Axes: xlabel='acceleration', ylabel='year'>,
<Axes: xlabel='year', ylabel='year'>,
<Axes: xlabel='origin', ylabel='year'>],
[<Axes: xlabel='mpg', ylabel='origin'>,
<Axes: xlabel='cylinders', ylabel='origin'>,
<Axes: xlabel='displacement', ylabel='origin'>,
<Axes: xlabel='horsepower', ylabel='origin'>,
<Axes: xlabel='weight', ylabel='origin'>,
<Axes: xlabel='acceleration', ylabel='origin'>,
<Axes: xlabel='year', ylabel='origin'>,
<Axes: xlabel='origin', ylabel='origin'>]], dtype=object)
```



1.6.2 (b) Compute the matrix of correlations between the variables using the DataFrame.corr() method.

```
[12]: Auto.corr()
[12]:
                              cylinders
                                         displacement
                                                       horsepower
                                                                     weight
                         mpg
      acceleration
                        year
                                origin
                              -0.777618
                    1.000000
                                            -0.805127
                                                        -0.778427 -0.832244
     mpg
      0.423329 0.580541 0.565209
      cylinders
                   -0.777618
                               1.000000
                                             0.950823
                                                         0.842983
                                                                   0.897527
      -0.504683 -0.345647 -0.568932
      displacement -0.805127
                               0.950823
                                             1.000000
                                                         0.897257 0.932994
      -0.543800 -0.369855 -0.614535
```

```
horsepower
             -0.778427
                         0.842983
                                       0.897257
                                                   1.000000 0.864538
-0.689196 -0.416361 -0.455171
weight
             -0.832244
                         0.897527
                                       0.932994
                                                  0.864538 1.000000
-0.416839 -0.309120 -0.585005
acceleration 0.423329 -0.504683
                                     -0.543800
                                                 -0.689196 -0.416839
1.000000 0.290316 0.212746
              0.580541 -0.345647
                                                 -0.416361 -0.309120
vear
                                     -0.369855
0.290316 1.000000 0.181528
              0.565209 -0.568932
                                                 -0.455171 -0.585005
origin
                                     -0.614535
0.212746 0.181528 1.000000
```

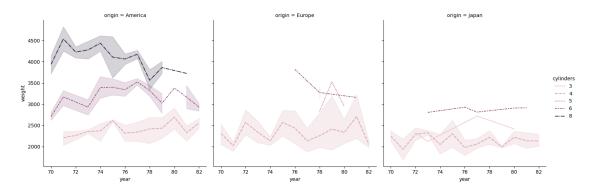
- 1.6.3 (c) Use the sm.OLS() function to perform a multiple linear regression with mpg as the response and all other variables except name as the predictors. Use the summarize() function to print the results. Comment on the output. For instance:
- 1.7 Convert year and origin columns to categorical types

```
[13]:
                          cylinders displacement horsepower
                                                                     weight
                    mpg
      acceleration
      count 392.000000 392.000000
                                       392.000000 392.000000
                                                                392.000000
      392.000000
              23.445918
                           5.471939
                                       194.411990 104.469388
     mean
                                                               2977.584184
      15.541327
               7.805007
                           1.705783
                                       104.644004
                                                    38.491160
                                                                849.402560
      std
      2.758864
     min
               9.000000
                           3.000000
                                        68.000000
                                                    46.000000 1613.000000
      8.000000
      25%
              17.000000
                           4.000000
                                       105.000000
                                                    75.000000 2225.250000
      13.775000
      50%
              22.750000
                           4.000000
                                       151.000000
                                                    93.500000 2803.500000
      15.500000
              29.000000
                           8.000000
                                       275.750000
                                                   126.000000 3614.750000
      75%
      17.025000
              46.600000
                           8.000000
                                       455.000000 230.000000 5140.000000
      max
      24.800000
```

```
[14]: sns.relplot(
    Auto,
    x="year",
    y="weight",
```

```
col="origin",
hue="cylinders",
style="cylinders",
estimator="mean",
kind="line",
)
```

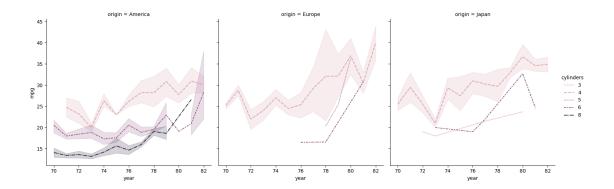
[14]: <seaborn.axisgrid.FacetGrid at 0x7cb747485d60>



The weight of the 8-cylinder American made models show a decline from the highs of 1972. It can also be seen that American made cars are heavier than their European and Japanese counterparts especially in the most common models with 4 cylinders.

```
[15]: sns.relplot(
    Auto,
    x="year",
    y="mpg",
    col="origin",
    hue="cylinders",
    style="cylinders",
    estimator="mean",
    kind="line",
)
```

[15]: <seaborn.axisgrid.FacetGrid at 0x7cb7473d12e0>



It can be seen that after the oil shock of 1973 and the regulations and actions taken by the US government, the mileage for American made cars rose across all models. This was, however, matched by the European and Japanese models which were already lighter and more fuel efficient.

1.7.1 Encode categorical variables as dummy variables dropping the first to remove multicollinearity.

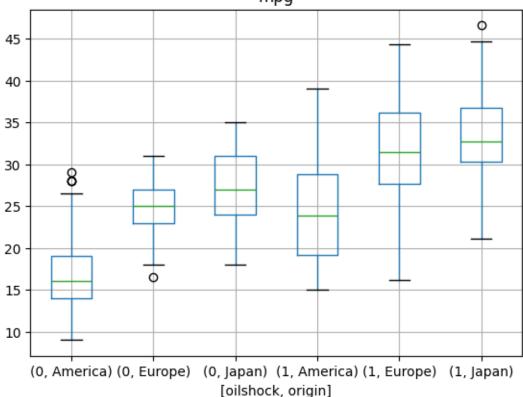
```
[16]: def categorize_for_oil_shock(row):
    # we add 3 years because it takes approximately that long for carumanufacturers to introduce a new model
    if row["year"] in (70, 71, 72, 73, 74, 75, 76):
        return 0
    return 1

Auto["oilshock"] = Auto.apply(categorize_for_oil_shock, axis=1)

[17]: Auto.boxplot(column="mpg", by=["oilshock", "origin"])
```

[17]: <Axes: title={'center': 'mpg'}, xlabel='[oilshock, origin]'>

Boxplot grouped by ['oilshock', 'origin'] mpg



```
[18]: Auto_os = Auto.drop(["year"], axis=1)
Auto_os.columns
```

```
[19]: # standardizing dataframes
Auto_os["oilshock"] = Auto_os["oilshock"].astype("category")
Auto_os = Auto_os.apply(standardize)
Auto_os.describe()
```

[19]: mpg cylinders displacement horsepower weight acceleration count 3.920000e+02 3.920000e+02 3.920000e+02 3.920000e+02 3.920000e+02 mean 1.812609e-16 -1.087565e-16 -7.250436e-17 -1.812609e-16 -3.625218e-17 -8.519262e-16 std 1.001278e+00 1.001278e+00 1.001278e+00 1.001278e+00 1.001278e+00

```
-1.853218e+00 -1.451004e+00 -1.209563e+00 -1.520975e+00 -1.608575e+00
     min
      -2.736983e+00
      25%
           -8.269250e-01 -8.640136e-01 -8.555316e-01 -7.665929e-01 -8.868535e-01
      -6.410551e-01
            -8.927701e-02 -8.640136e-01 -4.153842e-01 -2.853488e-01 -2.052109e-01
      -1.499869e-02
      75%
            7.125143e-01 1.483947e+00 7.782764e-01 5.600800e-01 7.510927e-01
      5.384714e-01
             2.970359e+00 1.483947e+00 2.493416e+00 3.265452e+00 2.549061e+00
     max
      3.360262e+00
[20]: Auto_os = pd.get_dummies(
          Auto_os, columns=list(["origin"]), drop_first=True, dtype=np.uint8
      Auto_os.columns
[20]: Index(['mpg', 'cylinders', 'displacement', 'horsepower', 'weight',
      'acceleration', 'oilshock', 'origin_Europe', 'origin_Japan'], dtype='object')
[21]: y = Auto_os["mpg"]
[21]: name
      chevrolet chevelle malibu
                                             -0.698638
     hi 1200d
                                             -1.853218
      dodge d200
                                             -1.596645
      chevy c20
                                             -1.724931
     ford f250
                                             -1.724931
     amc gremlin
                                             -0.313779
     bmw 2002
                                              0.327654
      saab 99e
                                              0.199368
                                              0.071081
      audi 100 ls
      volkswagen 1131 deluxe sedan
                                              0.327654
      datsun pl510
                                              0.455941
      ford maverick
                                             -0.313779
     amc hornet
                                             -0.698638
     plymouth duster
                                             -0.185492
     peugeot 504
                                              0.199368
     buick estate wagon (sw)
                                             -1.211785
      toyota corona mark ii
                                              0.071081
     plymouth satellite
                                             -0.698638
      amc rebel sst
                                             -0.955212
      ford torino
                                             -0.826925
     ford galaxie 500
                                             -1.083498
      chevrolet impala
                                             -1.211785
     buick skylark 320
                                             -1.083498
     pontiac catalina
                                             -1.211785
      amc ambassador dpl
                                             -1.083498
```

dodge challenger se	-1.083498
plymouth 'cuda 340	-1.211785
chevrolet monte carlo	-1.083498
plymouth fury iii	-1.211785
amc hornet sportabout (sw)	-0.698638
chevrolet vega (sw)	-0.185492
pontiac firebird	-0.570352
ford mustang	-0.698638
mercury capri 2000	-0.057205
toyota corolla 1200	0.969088
peugeot 304	0.840801
datsun 1200	1.482234
	0.455941
volkswagen model 111	
plymouth cricket	0.327654
pontiac safari (sw)	-1.340071
opel 1900	0.584228
ford country squire (sw)	-1.340071
fiat 124b	0.840801
plymouth fury iii	-1.211785
dodge monaco (sw)	-1.468358
chevrolet vega 2300	0.584228
3	0.199368
toyota corona	
amc gremlin	-0.570352
plymouth satellite custom	-0.955212
datsun pl510	0.455941
ford torino 500	-0.570352
amc matador	-0.698638
chevrolet impala	-1.211785
pontiac catalina brougham	-1.211785
ford galaxie 500	-1.211785
chevrolet chevelle malibu	-0.826925
chevrolet chevelle concours (sw)	
	-1.340071
plymouth satellite custom (sw)	-1.211785
volvo 145e (sw)	-0.698638
volkswagen 411 (sw)	-0.185492
peugeot 504 (sw)	-0.313779
ford pinto (sw)	-0.185492
datsun 510 (sw)	0.584228
toyouta corona mark ii (sw)	-0.057205
dodge colt (sw)	0.584228
amc matador (sw)	-1.083498
toyota corolla 1600 (sw)	0.455941
renault 12 (sw)	0.327654
mazda rx2 coupe	-0.570352
ford gran torino (sw)	-1.340071
oldsmobile delta 88 royale	-1.468358
chrysler newport royal	-1.340071
- •	

toyota corona hardtop	0.071081
volkswagen type 3	-0.057205
chevrolet vega	-0.442065
ford pinto runabout	-0.313779
chevrolet impala	-1.340071
dodge colt hardtop	0.199368
plymouth fury iii	-1.083498
ford galaxie 500	-1.211785
amc ambassador sst	-0.826925
mercury marquis	-1.596645
buick lesabre custom	-1.340071
pontiac catalina	-1.211785
-	0.327654
fiat 124 sport coupe	
amc gremlin	-0.698638
toyota carina	-0.442065
chevrolet vega	-0.313779
datsun 610	-0.185492
maxda rx3	-0.698638
ford pinto	-0.570352
mercury capri v6	-0.313779
chevrolet monte carlo s	-1.083498
saab 991e	0.071081
fiat 128	0.712514
opel manta	0.071081
audi 1001s	-0.442065
volvo 144ea	-0.570352
dodge dart custom	-1.083498
toyota mark ii	-0.442065
oldsmobile omega	-1.596645
oldsmobile vista cruiser	-1.468358
pontiac grand prix	-0.955212
plymouth custom suburb	-1.340071
	-0.698638
amc hornet	-1.596645
chevrolet impala	
buick century 350	-1.340071
amc matador	-1.211785
chevrolet malibu	-1.340071
dodge coronet custom	-1.083498
mercury marquis brougham	-1.468358
chevrolet caprice classic	-1.340071
ford 1td	-1.340071
plymouth fury gran sedan	-1.211785
ford gran torino	-1.211785
buick electra 225 custom	-1.468358
amc ambassador brougham	-1.340071
plymouth valiant	-0.698638
chevrolet nova custom	-0.955212

ford country	-1.468358
ford maverick	-0.698638
plymouth duster	-0.057205
volkswagen super beetle	0.327654
chrysler new yorker brougham	-1.340071
audi fox	0.712514
volkswagen dasher	0.327654
opel manta	0.327654
toyota corona	0.969088
datsun 710	1.097374
subaru	0.327654
fiat 128	0.071081
fiat 124 tc	0.327654
honda civic	0.071081
fiat x1.9	0.969088
amc matador (sw)	-1.211785
dodge colt	0.584228
ford gran torino (sw)	-1.211785
ford gran torino	-0.955212
buick century luxus (sw)	-1.340071
dodge coronet custom (sw)	-1.211785
plymouth duster	-0.442065
amc hornet	-0.570352
chevrolet nova	-1.083498
ford pinto	0.327654
datsun b210	0.969088
	0.199368
chevrolet vega chevrolet chevelle malibu classic	
	-0.955212
amc matador	-0.955212
plymouth satellite sebring	-0.698638
toyota corolla 1200	1.097374
datsun 710	0.071081
pontiac astro	-0.057205
amc gremlin	-0.442065
toyota corona	0.071081
volkswagen dasher	0.199368
ford pinto	-0.698638
saab 991e	0.199368
amc pacer	-0.570352
audi 100ls	-0.057205
peugeot 504	-0.057205
volvo 244dl	-0.185492
honda civic cvcc	1.225661
ford pinto	-0.057205
-	0.712514
volkswagen rabbit	
toyota corolla	0.712514
plymouth valiant custom	-0.570352

chevrolet monza 2+2	-0.442065
ford mustang ii	-1.340071
chevrolet nova	-0.698638
mercury monarch	-1.083498
pontiac catalina	-0.955212
chevrolet bel air	-1.083498
plymouth grand fury	-0.955212
ford maverick	-1.083498
buick century	-0.826925
chevroelt chevelle malibu	-0.955212
amc matador	-1.083498
plymouth fury	-0.698638
buick skyhawk	-0.313779
ford 1td	-1.211785
ford pinto	0.391798
pontiac ventura sj	-0.634495
amc pacer d/l	-0.762782
volkswagen rabbit	0.776658
datsun b-210	1.097374
toyota corolla	0.584228
volvo 245	-0.442065
ford f108	-1.340071
peugeot 504	-0.570352
toyota mark ii	-0.570352
mercedes-benz 280s	-0.891068
cadillac seville	-0.891068
chevy c10	-1.340071
dodge d100	-1.340071
ford granada ghia	-0.698638
plymouth volare premier v8	-1.340071
dodge aspen se	-0.442065
vw rabbit	0.712514
opel 1900	0.199368
honda civic	1.225661
fiat 131	0.584228
capri ii	0.199368
dodge colt	0.327654
renault 12tl	0.455941
dodge coronet brougham	-0.955212
amc matador	-1.019355
chevrolet chevelle malibu classic	-0.762782
plymouth valiant	-0.185492
chevrolet nova	-0.185492
ford maverick	0.071081
amc hornet	-0.121349
chevrolet chevette	0.712514
chevrolet woody	0.135225

ford gran torino	-1.147642
ford mustang ii 2+2	0.263511
volkswagen rabbit custom	0.712514
pontiac sunbird coupe	0.135225
toyota corolla liftback	0.327654
chevrolet chevette	0.904944
bmw 320i	-0.249635
subaru dl	0.840801
volkswagen dasher	0.904944
datsun 810	-0.185492
mazda rx-4	-0.249635
ford thunderbird	-0.955212
dodge colt m/m	1.289804
chrysler cordoba	-1.019355
chevrolet monte carlo landau	-1.019355
plymouth arrow gs	0.263511
buick opel isuzu deluxe	0.840801
renault 5 gtl	1.610521
datsun f-10 hatchback	1.289804
pontiac grand prix lj	-0.955212
oldsmobile cutlass supreme	-0.826925
chevrolet caprice classic	-0.762782
mercury cougar brougham	-1.083498
chevrolet concours	-0.762782
buick skylark	-0.377922
plymouth volare custom	-0.570352
ford granada	-0.634495
dodge monaco brougham	-1.019355
honda accord cvcc	1.033231
datsun 510	0.481598
toyota corona	0.520084
chevrolet chevette	0.840801
buick regal sport coupe (turbo)	-0.737124
ford futura	-0.685810
dodge omni	0.956259
dodge magnum xe	-0.762782
toyota celica gt liftback	-0.300950
peugeot 604sl	-0.929554
oldsmobile starfire sx	0.929334
datsun 200-sx	0.043424
audi 5000	-0.403579
	-0.403379
volvo 264gl	
saab 99gle	-0.236807
volkswagen scirocco	1.033231
honda accord lx	0.776658
plymouth sapporo	-0.031548
chevrolet monte carlo landau	-0.544694

mazda glc deluxe	1.200003
dodge aspen	-0.621666
volkswagen rabbit custom diesel	2.521356
_	
ford fiesta	1.623349
datsun b210 gx	2.046695
honda civic cvcc	1.623349
amc concord d/l	-0.685810
dodge diplomat	-0.519037
-	
mercury monarch ghia	-0.416408
oldsmobile cutlass salon brougham	-0.454894
chevrolet malibu	-0.377922
ford fairmont (auto)	-0.416408
ford fairmont (man)	0.212197
plymouth volare	-0.377922
amc concord	-0.519037
buick century special	-0.365093
mercury zephyr	-0.339436
pontiac phoenix lj	-0.544694
plymouth horizon	1.379605
mercedes benz 300d	0.250683
cadillac eldorado	-0.057205
peugeot 504	0.481598
oldsmobile cutlass salon brougham	0.058253
plymouth horizon tc3	1.418091
amc spirit dl	0.507256
fiat strada custom	1.777293
buick skylark limited	0.635542
chevrolet citation	0.686857
oldsmobile omega brougham	0.430284
pontiac phoenix	1.289804
datsun 210	1.071717
dodge colt hatchback custom	1.572035
_	
dodge st. regis	-0.672981
vw rabbit custom	1.084545
mercury zephyr 6	-0.467723
ford fairmont 4	-0.147006
amc concord dl 6	-0.416408
dodge aspen 6	-0.365093
-	
chevrolet caprice classic	-0.826925
ford 1td landau	-0.749953
pontiac lemans v6	-0.249635
maxda glc deluxe	1.366776
buick estate wagon (sw)	-0.839754
ford country squire (sw)	-1.019355
chevrolet malibu classic (sw)	-0.544694
chrysler lebaron town @ country (sw)	-0.634495
mercury grand marquis	-0.891068

vw rabbit c (diesel)	2.675299
vw dasher (diesel)	2.559841
audi 5000s (diesel)	1.661835
mercedes-benz 240d	0.840801
honda civic 1500 gl	2.713785
datsun 280-zx	1.187175
	0.815144
vokswagen rabbit	
mazda rx-7 gs	0.032595
triumph tr7 coupe	1.482234
honda accord	1.148689
datsun 210	2.226296
subaru dl	1.328290
dodge colt	0.571399
mazda glc	2.970359
toyota corolla	1.123031
vw rabbit	2.316097
toyota corolla tercel	1.879922
chevrolet chevette	1.110203
chevrolet citation	0.584228
ford fairmont	0.378969
datsun 310	1.764465
dodge aspen	-0.557523
audi 4000	1.392433
toyota corona liftback	0.815144
mazda 626	1.007574
datsun 510 hatchback	1.738807
amc concord	0.109567
peugeot 505s turbo diesel	0.597056
honda prelude	1.315461
toyota corolla	1.148689
datsun 200sx	1.212832
mazda 626	1.046059
volvo diesel	0.930602
chrysler lebaron salon	-0.749953
datsun 810 maxima	0.096739
buick century	-0.134177
oldsmobile cutlass ls	0.404626
ford granada gl	-0.416408
volkswagen jetta	1.225661
toyota cressida	0.250683
ford escort 2h	0.827972
plymouth reliant	0.481598
plymouth horizon 4	1.443748
ford escort 4w	1.405262
buick skylark	0.404626
dodge aries wagon (sw)	0.301997
plymouth reliant	0.840801

```
0.006938
      chevrolet citation
      honda civic 1300
                                               1.495063
      subaru
                                               1.135860
      datsun 210 mpg
                                               1.738807
      toyota tercel
                                               1.828608
      mazda glc 4
                                               1.366776
      plymouth champ
                                               1.995380
      chrysler lebaron medallion
                                               0.327654
      honda civic (auto)
                                               1.097374
      datsun 310 gx
                                               1.867094
      buick century limited
                                               0.199368
      oldsmobile cutlass ciera (diesel)
                                               1.867094
      ford granada 1
                                              -0.185492
      dodge rampage
                                               1.097374
      dodge charger 2.2
                                               1.610521
      chevrolet camaro
                                               0.455941
      ford mustang gl
                                               0.455941
      vw pickup
                                               2.636813
      honda civic
                                               1.867094
      toyota celica gt
                                               1.097374
      toyota corolla
                                               1.353947
      ford ranger
                                               0.584228
     nissan stanza xe
                                               1.610521
      mercury lynx l
                                               1.610521
      plymouth horizon miser
                                               1.867094
     mazda glc custom
                                               0.969088
     mazda glc custom l
                                               1.738807
      volkswagen rabbit l
                                               1.610521
                                               0.071081
      ford fairmont futura
      pontiac phoenix
                                               0.455941
      dodge aries se
                                               0.712514
      pontiac j2000 se hatchback
                                               0.969088
      chevrolet cavalier 2-door
                                               1.353947
      chevrolet cavalier wagon
                                               0.455941
      chevrolet cavalier
                                               0.584228
      honda accord
                                               1.610521
      chevy s-10
                                               0.969088
      Name: mpg, dtype: float64
[22]: cols = list(Auto_os.columns)
      cols.remove("mpg")
      formula = " + ".join(cols)
      model = smf.ols(f"mpg ~ {formula}", data=Auto_os)
      results = model.fit()
      results.summary()
[22]:
```

2.008209

toyota starlet

				_		
Dep. Variable:		mpg		R-squared:		0.808
Model:		OLS		Adj. R-squared:		0.804
Method:	Lea	st Squares	$\mathbf{F} ext{-sta}$	itistic:		201.7
Date:	Fri, 2	21 Feb 202	5 Prob	Prob (F-statistic):		
Time:	1	5:48:44	Log-	Likelihoo	od:	-232.60
No. Observation	ıs:	392	AIC:			483.2
Df Residuals:		383	BIC:			518.9
Df Model:		8				
Covariance Type	e: no	onrobust				
	coef	std err	t	\mathbf{P} > $ \mathbf{t} $	[0.025	0.975]
Intercept	-0.4186	0.041	-10.263	0.000	-0.499	-0.338
${ m oilshock}[{ m T.1}]$	0.6363	0.048	13.204	0.000	0.542	0.731
cylinders	-0.1382	0.073	-1.885	0.060	-0.282	0.006
${f displacement}$	0.2845	0.107	2.659	0.008	0.074	0.495
horsepower	-0.2213	0.069	-3.192	0.002	-0.358	-0.085
\mathbf{weight}	-0.5923	0.073	-8.085	0.000	-0.736	-0.448
acceleration	0.0053	0.036	0.146	0.884	-0.066	0.076
origin_Europe	0.3038	0.076	4.015	0.000	0.155	0.453
origin_Japan	0.3819	0.074	5.156	0.000	0.236	0.528
Omnibus:	Omnibus: 20.03		9 Durbin-Watson:			31
Prob(Omni	ibus):	0.000 J	arque-Bera (JB):		27.583	
Skew:	(0.413 P	$\operatorname{rob}(\operatorname{JB})$	}	1.02e-06	
Kurtosis:		4.004 C	ond. No	•	11	.9

Notes:

- [1] Standard Errors assume that the covariance matrix of the errors is correctly specified.
- i. Is there a relationship between the predictors and the response? Use the anova $_{\rm lm}()$ function from statsmodels to answer this question.
- ii. Which predictors appear to have a statistically significant relationship to the response?

iii. What does the coefficient for the year variable suggest?

[23]:	anova_lm(resul	ts)				
[23]:		df	sum_sq	mean_sq	F	PR(>F)
	oilshock	1.0	106.483141	106.483141	542.347509	2.293307e-75
	cylinders	1.0	170.845795	170.845795	870.163964	1.267122e-100
	displacement	1.0	11.934469	11.934469	60.785485	6.078862e-14
	horsepower	1.0	3.951021	3.951021	20.123619	9.610639e-06
	weight	1.0	17.796189	17.796189	90.640818	1.988543e-19
	acceleration	1.0	0.009116	0.009116	0.046430	8.295108e-01
	origin_Europe	1.0	0.564108	0.564108	2.873155	9.088094e-02
	origin_Japan	1.0	5.218909	5.218909	26.581317	4.058867e-07
	Residual	383.0	75.197253	0.196337	NaN	NaN

There seems to be a statistical relationship between all of the predictors and the response variable, mpg, except for acceleration.

though some of \mathbf{the} categorical variables are insignificant, one of the levels is significant, \mathbf{it} isadvisable to retain them model. https://stats.stackexchange.com/questions/24298/ all inthe can-i-ignore-coefficients-for-non-significant-levels-of-factors-in-a-linear-mode

Note: Year has been converted to a categorical variable oilshock to better capture the effects of the oil shock of 1973 on the mileage.

1.7.2 (d) Produce some of diagnostic plots of the linear regression fit as described in the lab. Comment on any problems you see with the fit. Do the residual plots suggest any unusually large outliers? Does the leverage plot identify any observations with unusually high leverage?

Before producing the diagnostic plots, let's first test for collinearity using correlation matrix and variance inflation factors.

```
[24]: Auto_os.corr(numeric_only=True)
```

```
[24]:
                               cylinders displacement horsepower
                                                                        weight
                          mpg
                    origin_Europe origin_Japan
      acceleration
                                                           -0.778427 -0.832244
                     1.000000
                                -0.777618
                                              -0.805127
      mpg
      0.423329
                     0.244313
                                    0.451454
      cylinders
                    -0.777618
                                 1.000000
                                               0.950823
                                                            0.842983
                                                                     0.897527
      -0.504683
                     -0.352324
                                    -0.404209
      displacement
                    -0.805127
                                 0.950823
                                               1.000000
                                                            0.897257
                                                                      0.932994
      -0.543800
                     -0.371633
                                    -0.440825
     horsepower
                    -0.778427
                                 0.842983
                                               0.897257
                                                            1.000000
                                                                      0.864538
      -0.689196
                     -0.284948
                                    -0.321936
                    -0.832244
                                               0.932994
      weight
                                 0.897527
                                                            0.864538
                                                                      1.000000
      -0.416839
                     -0.293841
                                    -0.447929
      acceleration
                     0.423329
                                -0.504683
                                                           -0.689196 -0.416839
                                              -0.543800
      1.000000
                     0.208298
                                    0.115020
      origin_Europe
                     0.244313
                                -0.352324
                                              -0.371633
                                                           -0.284948 -0.293841
      0.208298
                     1.000000
                                   -0.230157
                                -0.404209
                                                           -0.321936 -0.447929
      origin_Japan
                     0.451454
                                              -0.440825
      0.115020
                    -0.230157
                                    1.000000
```

```
[25]: vifdf = calculate_VIFs("mpg ~ " + " + ".join(Auto_os.columns) + " - mpg", □

→Auto_os)

vifdf
```

```
[25]: VIF
Feature
oilshock[T.1] 1.149269
cylinders 10.737464
displacement 22.861475
```

```
horsepower 9.594564
weight 10.715246
acceleration 2.614133
origin_Europe 1.639338
origin_Japan 1.762590
```

```
[26]: identify_highest_VIF_feature(vifdf)
```

We find the highest VIF in this model is displacement with a VIF of 22.861474853464927

Hence, we drop displacement from the model to be fitted.

```
[26]: ('displacement', 22.861474853464927)
```

```
[27]: vifdf = calculate_VIFs(
        "mpg ~ " + " + ".join(Auto_os.columns) + " - mpg - displacement", Auto_os
)
vifdf
```

```
[27]: VIF
```

Feature
oilshock[T.1] 1.139339
cylinders 6.190903
horsepower 8.641303
weight 9.024884
acceleration 2.591157

origin_Europe 1.450726 origin_Japan 1.591434

[28]: identify_highest_VIF_feature(vifdf)

No variables are significantly collinear.

1.7.3 Linear Regression for mpg ~ cylinders + horsepower + weight + acceleration + oilshock + origin_Europe + origin_Japan

```
[29]: cols = list(Auto_os.columns)
    cols.remove("mpg")
    cols.remove("displacement")
    formula = " + ".join(cols)
    results = perform_analysis("mpg", formula, Auto_os)
```

OLS Regression Results

Dep. Variable: R-squared: 0.805 mpg Model: OLS Adj. R-squared: 0.801 F-statistic: Method: Least Squares 225.9 Prob (F-statistic): Date: Fri, 21 Feb 2025 6.41e-132

Time: 15:48:44 Log-Likelihood: -236.18

No. Observation Df Residuals: Df Model: Covariance Type		392 384 7 nonrobust	AIC: BIC:			488.4 520.1
0.975]	coef	std err	t	P> t	[0.025	
0.719 cylinders 0.099 horsepower -0.033 weight -0.382 acceleration 0.068	-0.3890 0.6243 -0.0113 -0.1632 -0.5149 -0.0038 0.2356 0.3205	0.040 0.048 0.056 0.066 0.068 0.036 0.072	-9.837 12.911 -0.202 -2.461 -7.599 -0.103 3.283 4.518	0.000 0.000 0.840 0.014 0.000 0.918 0.001	-0.467 0.529 -0.122 -0.294 -0.648 -0.075 0.095 0.181	
Omnibus: Prob(Omnibus): Skew: Kurtosis:		25.646 0.000	Durbin-Was Jarque-Borob (JB) Cond. No	atson: era (JB): :	1	1.305 40.287 .79e-09 7.67

Notes

[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.

	df	sum_sq	${\tt mean_sq}$	F	PR(>F)
oilshock	1.0	106.483141	106.483141	533.906720	1.149811e-74
cylinders	1.0	170.845795	170.845795	856.621225	7.985118e-100
horsepower	1.0	12.927972	12.927972	64.820882	1.039468e-14
weight	1.0	20.649905	20.649905	103.538670	1.085729e-21
acceleration	1.0	0.003626	0.003626	0.018183	8.928058e-01
origin_Europe	1.0	0.432514	0.432514	2.168627	1.416711e-01
origin_Japan	1.0	4.071523	4.071523	20.414626	8.312108e-06
Residual	384.0	76.585524	0.199441	NaN	NaN

[29]: <statsmodels.regression.linear_model.RegressionResultsWrapper at 0x7cb747038080>

[30]: identify_least_significant_feature(results, alpha=LOS_Alpha)

We find the least significant variable in this model is acceleration with a p-value of 0.9177499426057751 and a coefficient of -0.003752017674274216 Using the backward methodology, we suggest dropping acceleration from the new model

Linear Regression after dropping acceleration. The model now is mpg \sim cylinders + horsepower + weight + oilshock + origin_Europe + origin_Japan

OLS Regression Results

Dep. Variable: Model: Method: Date: Time: No. Observations Df Residuals: Df Model: Covariance Type:	Fri,	mpg OLS Least Squares Fri, 21 Feb 2025 15:48:44 392 385 6 nonrobust		uared: ic: tatistic):	0.805 0.802 264.3 3.80e-133 -236.19 486.4 514.2
0.975]	coef	std err	t	P> t	[0.025
	-0.3889	0.039	-9.849	0.000	-0.467
oilshock[T.1] 0.719 cylinders 0.099	0.6245	0.048	12.935 -0.189	0.000	0.530
horsepower -0.064 weight	-0.1585 -0.5182	0.048	-3.285 -8.704	0.001	-0.253 -0.635

```
-0.401
origin_Europe 0.2352 0.072 3.287 0.001
                                    0.095
0.376
origin_Japan 0.3202 0.071 4.524 0.000 0.181
0.459
______
Omnibus:
                 25.330 Durbin-Watson:
                                         1.305
Prob(Omnibus):
                  0.000 Jarque-Bera (JB):
                                        39.508
Skew:
                  0.454 Prob(JB):
                                       2.64e-09
Kurtosis:
                  4.263 Cond. No.
                                          6.84
______
```

Notes:

[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.

	df	sum_sq	mean_sq	F	PR(>F)
oilshock	1.0	106.483141	106.483141	535.282217	7.450502e-75
cylinders	1.0	170.845795	170.845795	858.828126	4.453195e-100
horsepower	1.0	12.927972	12.927972	64.987879	9.610211e-15
weight	1.0	20.649905	20.649905	103.805416	9.634452e-22
origin_Europe	1.0	0.434982	0.434982	2.186618	1.400323e-01
origin_Japan	1.0	4.070552	4.070552	20.462339	8.111817e-06
Residual	385.0	76.587654	0.198929	NaN	NaN

Linear Regression after dropping cylinders. The model now is mpg \sim horsepower + weight + oilshock + origin_Europe + origin_Japan

OLS Regression Results

```
_____
Dep. Variable:
                            R-squared:
                                                    0.805
                        mpg
Model:
                        OLS Adj. R-squared:
                                                    0.802
Method:
                Least Squares F-statistic:
                                                    317.9
             Fri, 21 Feb 2025 Prob (F-statistic): 2.06e-134
Date:
Time:
                    15:48:44 Log-Likelihood:
                                                  -236.21
No. Observations:
                        392 AIC:
                                                    484.4
Df Residuals:
                           BTC:
                                                    508.2
                        386
```

Df Model: Covariance Type	:	5 nonrobust				
=	coef	std err	 t	P> t	[0.025	
0.975]	COGI	Stu ell	Ü	17 0	[0.020	
_						
Intercept -0.314	-0.3901	0.039	-10.030	0.000	-0.467	
oilshock[T.1] 0.720	0.6250	0.048	12.983	0.000	0.530	
horsepower	-0.1613	0.046	-3.510	0.001	-0.252	
weight -0.427	-0.5245	0.050	-10.576	0.000	-0.622	
origin_Europe 0.375	0.2386	0.069	3.448	0.001	0.103	
origin_Japan 0.460	0.3222		4.611	0.000	0.185	
Omnibus:	=======	24.971				1.304
Prob(Omnibus):		0.000	Jarque-Be	era (JB):		38.456
Skew:		0.453	Prob(JB):		4	.46e-09
Kurtosis:		4.239				5.60
=========		========	=======			=====
Notes: [1] Standard Errspecified.	rors assume	that the co	variance ma	atrix of the	e errors is	correctly
1	df	sum_sq m	ean_sq	F	PR(>F)	

	df	sum_sq	${\tt mean_sq}$	F	PR(>F)
oilshock	1.0	106.483141	106.483141	536.622900	4.863116e-75
horsepower	1.0	165.048555	165.048555	831.763917	2.445119e-98
weight	1.0	39.079210	39.079210	196.940090	1.939884e-36
origin_Europe	1.0	0.574647	0.574647	2.895939	8.960825e-02
origin_Japan	1.0	4.219706	4.219706	21.265252	5.446537e-06
Residual	386.0	76.594741	0.198432	NaN	NaN

We can now try and plot the diagnostics for the model.

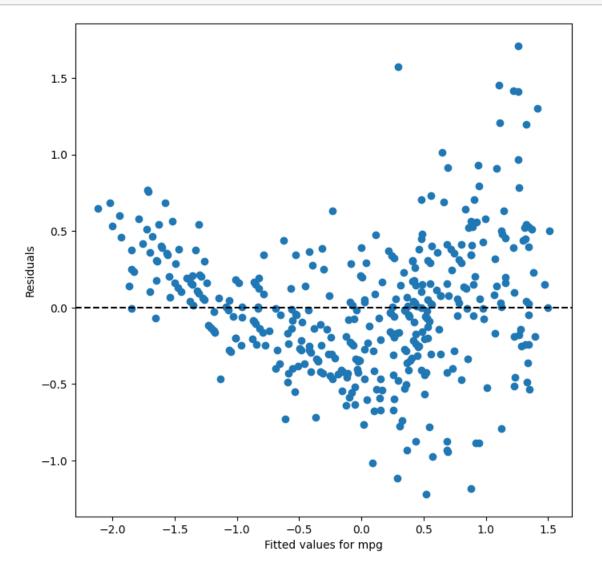
```
[33]: TSS = np.sum((y - np.mean(y)) ** 2)
TSS
RSS = np.sum((y - results.fittedvalues) ** 2)
RSS
RSE = np.sqrt(RSS / results.df_model)
display("RSE " + str(RSE))
display("R-squared adjusted : " + str(results.rsquared_adj))
display("F-statistic : " + str(results.fvalue))
```

'RSE 3.9139428061794668'

'R-squared adjusted : 0.8020742313429469'

'F-statistic : 317.8976193276657'

[34]: display_residuals_plot(results)



There is some evidence of non-linearity and heteroskedasticity from the residuals plot above.

1.7.4 (e) Fit some models with interactions as described in the lab. Do any interactions appear to be statistically significant?

OLS Regression Results

=======================================	UI ==========	S Regress	ion Results ========	========	.========
Dep. Variable: Model: Method: Date: Time: No. Observations: Df Residuals: Df Model: Covariance Type:	0.847 0.845 355.8 1.15e-153 -187.98 390.0 417.8				
0.975]	coef	std err	t	======= P> t	[0.025
Intercept -0.488	-0.5631	0.038	-14.715	0.000	-0.638
oilshock[T.1]	0.6508	0.043	15.243	0.000	0.567
horsepower -0.283	-0.3723	0.045	-8.185	0.000	-0.462
weight -0.406	-0.4926	0.044	-11.192	0.000	-0.579
origin_Europe 0.278	0.1565	0.062	2.535	0.012	0.035
origin_Japan 0.330	0.2061	0.063	3.278	0.001	0.082
horsepower:weight	0.2300	0.022	10.364	0.000	0.186
Omnibus:	:=======	27.116	 Durbin-Watson	======= :	 1.364

```
      Prob(Omnibus):
      0.000 Jarque-Bera (JB):
      45.242

      Skew:
      0.457 Prob(JB):
      1.50e-10

      Kurtosis:
      4.390 Cond. No.
      7.09
```

Notes

[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.

```
sum_sq
                                                     F
                                                              PR(>F)
                   df
                                    mean_sq
oilshock
                  1.0 106.483141 106.483141
                                             684.553356
                                                         1.927144e-87
                 1.0 165.048555 165.048555 1061.055689 1.099814e-112
horsepower
weight
                 1.0 39.079210 39.079210 251.230425 6.522402e-44
                 1.0
origin_Europe
                        0.574647
                                  0.574647
                                              3.694260
                                                         5.533771e-02
                 1.0 4.219706 4.219706 27.127429
                                                         3.109409e-07
origin_Japan
                1.0 16.707504 16.707504
                                             107.408348
                                                         2.313061e-22
horsepower:weight
Residual
                385.0 59.887237 0.155551
                                                   \mathtt{NaN}
                                                                 NaN
```

OLS Regression Results

_____ Dep. Variable: R-squared: mpg 0.861 Model: OLS Adj. R-squared: 0.858 Method: Least Squares F-statistic: 297.5 Date: Fri, 21 Feb 2025 Prob (F-statistic): 3.50e-159 Time: 15:48:45 Log-Likelihood: -168.92No. Observations: ATC: 392 355.8 Df Residuals: 383 BIC: 391.6 Df Model: 8 Covariance Type: nonrobust ______ ========= coef std err t P>|t| Γ0.025 0.975]

<pre>Omnibus: Prob(Omnibus): Skew: Kurtosis:</pre>	23.934 0.000 0.377 4.450	Durbin-W Jarque-B Prob(JB) Cond. No	era (JB):		1.456 43.610 3.39e-10 11.3
horsepower:weight 0.217	0.1715	0.023	7.403	0.000	0.126
origin_Japan 0.311	0.1929	0.060	3.209	0.001	0.075
origin_Europe 0.297	0.1804	0.059	3.039	0.003	0.064
oilshock[T.1]:weight	-0.0963	0.082	-1.181	0.238	-0.257
-0.055 weight -0.361	-0.4591	0.050	-9.226	0.000	-0.557
-0.179 oilshock[T.1]:horsepower	-0.2276	0.088	-2.598	0.010	-0.400
0.674 horsepower	-0.2801	0.051	-5.456	0.000	-0.381
-0.463 oilshock[T.1]	0.5913	0.042	14.071	0.000	0.509
Intercept	-0.5350	0.037	-14.531	0.000	-0.607

Notes:

[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.

	df	sum_sq	${\tt mean_sq}$	F	PR(>F)
oilshock	1.0	106.483141	106.483141	750.563317	2.852400e-92
horsepower	1.0	165.048555	165.048555	1163.370934	3.985476e-118
oilshock:horsepower	1.0	19.155472	19.155472	135.020381	6.081794e-27
weight	1.0	35.071635	35.071635	247.207986	2.471356e-43
oilshock:weight	1.0	0.468549	0.468549	3.302640	6.994992e-02
origin_Europe	1.0	0.971997	0.971997	6.851278	9.208817e-03
origin_Japan	1.0	2.687942	2.687942	18.946386	1.727234e-05
horsepower:weight	1.0	7.776131	7.776131	54.811293	8.507915e-13
Residual	383.0	54.336579	0.141871	NaN	NaN

```
[37]: formula = " + ".join(cols)
  formula += " + " + "oilshock: horsepower"
  formula += " + " + "origin_Europe: horsepower"
  formula += " + " + "origin_Japan: horsepower"
  formula += " + " + "origin_Europe: weight"
  formula += " + " + "origin_Japan: weight"
  formula += " + " + "oilshock: weight"
  formula += " + " + "oilshock: horsepower"
```

results = perform_analysis("mpg", formula, Auto_os)
origin_interactions = results

OLS	Reg	ress	sion	Re	sul	ts
====	====	====		===	===	===

=======================================		:======			=======	
Dep. Variable:	mpg	R-squar	ed:	0.855		
Model:	OLS	Adj. R-	squared:		0.851	
Method:	Least Squares	F-stati:	stic:		204.1	
Date: Fri	, 21 Feb 2025	Prob (F	-statistic):	(6.22e-152	
Time:	15:48:45	Log-Like	elihood:		-177.44	
No. Observations:	392	AIC:			378.9	
Df Residuals:	380	BIC:			426.5	
Df Model:	11					
Covariance Type:	nonrobust					
		.=======		.======		
========						
	coef	std err	t	P> t	[0.025	
0.975]						
Intercept	-0.4245	0.035	-12.121	0.000	-0.493	
-0.356						
oilshock[T.1]	0.5690	0.044	13.054	0.000	0.483	
0.655						
horsepower	-0.0708	0.048	-1.470	0.142	-0.166	
0.024						
oilshock[T.1]:horsepower	-0.1615	0.096	-1.687	0.092	-0.350	
0.027						
weight	-0.4713	0.054	-8.712	0.000	-0.578	
-0.365						
oilshock[T.1]:weight	-0.2333	0.087	-2.694	0.007	-0.404	
-0.063						
origin_Europe	0.0297	0.082	0.363	0.717	-0.131	
0.191						
origin_Japan	-0.0010	0.131	-0.007	0.994	-0.259	
0.257						
origin_Europe:horsepower	-0.5852	0.130	-4.515	0.000	-0.840	
-0.330						
origin_Japan:horsepower	-0.2801	0.204	-1.370	0.172	-0.682	
0.122						
origin_Europe:weight	0.1640	0.120	1.370	0.171	-0.071	
0.399						
origin_Japan:weight	-0.1326	0.245	-0.541	0.589	-0.614	
0.349	0.1010	0.2.0	0.011	0.000	0.021	
		=======		.======	=======	
Omnibus:	20.717	Durbin-	Watson:		1.585	
Prob(Omnibus):	0.000		Bera (JB):		32.838	
Skew:	0.373	Prob(JB)			7.40e-08	
~	0.0.0	1 1 3 5 (3 D	, -		100 00	

Notes:

[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.

•	df	sum_sq	${\tt mean_sq}$	F
PR(>F)				
oilshock	1.0	106.483141	106.483141	712.972849
3.365342e-89				
horsepower	1.0	165.048555	165.048555	1105.105820
1.586816e-114				
oilshock:horsepower	1.0	19.155472	19.155472	128.258155
8.134152e-26				
weight	1.0	35.071635	35.071635	234.827067
1.302452e-41				
oilshock:weight	1.0	0.468549	0.468549	3.137234
7.732495e-02				
origin_Europe	1.0	0.971997	0.971997	6.508146
1.112920e-02				
origin_Japan	1.0	2.687942	2.687942	17.997494
2.781984e-05				
origin_Europe:horsepower	1.0	3.024522	3.024522	20.251113
9.040705e-06				
origin_Japan:horsepower	1.0	1.977640	1.977640	13.241566
3.116551e-04				
origin_Europe:weight	1.0	0.313437	0.313437	2.098659
1.482531e-01				
origin_Japan:weight	1.0	0.043767	0.043767	0.293050
5.885897e-01				
Residual	380.0	56.753344	0.149351	NaN
NaN				

[37]: <statsmodels.regression.linear_model.RegressionResultsWrapper at 0x7cb746f2bb60>

- From the above analysis, we can see that there is no significant interaction between origin and weight.
- So we can omit them from the model.

```
[38]: formula = " + ".join(cols)
  formula += " + " + "oilshock: horsepower"
  formula += " + " + "origin_Europe: horsepower"
  formula += " + " + "origin_Japan: horsepower"
  formula += " + " + "oilshock: weight"
  formula += " + " + "oilshock: horsepower"
  results = perform_analysis("mpg", formula, Auto_os)
  origin_interactions = results
```

OLS Regression Results

	mpg R-squared: OLS Adj. R-squared: Least Squares F-statistic: ri, 21 Feb 2025 Prob (F-statistic): 15:48:45 Log-Likelihood: 392 AIC: 382 BIC:				0.854 0.851 248.9 8.18e-154 -178.67 377.3 417.1	
Df Model:	9	DIC.			417.1	
Covariance Type:	nonrobust					
=========	========	=======	========	======		
0.975]	coef	std err	t	P> t	[0.025	
Intercept -0.363	-0.4296	0.034	-12.650	0.000	-0.496	
oilshock[T.1] 0.654	0.5693	0.043	13.280	0.000	0.485	
horsepower 0.013	-0.0797	0.047	-1.689	0.092	-0.172	
oilshock[T.1]:horsepower -0.015	-0.1958	0.092	-2.133	0.034	-0.376	
weight -0.357	-0.4571	0.051	-8.947	0.000	-0.558	
oilshock[T.1]:weight	-0.2019	0.084	-2.408	0.016	-0.367	
origin_Europe 0.162	0.0044	0.080	0.055	0.956	-0.153	
origin_Japan 0.240	0.0750	0.084	0.892	0.373	-0.090	
origin_Europe:horsepower	-0.4667	0.096	-4.884	0.000	-0.655	
origin_Japan:horsepower -0.169	-0.3682	0.101	-3.637	0.000	-0.567	
Omnibus: Prob(Omnibus): Skew: Kurtosis:	20.114 0.000 0.366 4.183	Durbin-W	era (JB):	== ===	1.577 31.613 1.37e-07 10.2	

Notes:

[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.

df sum_sq mean_sq F

PR(>F)

oilshock	1.0	106.483141	106.483141	712.242504
2.550768e-89				
horsepower	1.0	165.048555	165.048555	1103.973788
9.879451e-115				
oilshock:horsepower	1.0	19.155472	19.155472	128.126771
8.220841e-26				
weight	1.0	35.071635	35.071635	234.586518
1.270254e-41				
oilshock:weight	1.0	0.468549	0.468549	3.134021
7.747196e-02				
origin_Europe	1.0	0.971997	0.971997	6.501479
1.116822e-02				
origin_Japan	1.0	2.687942	2.687942	17.979058
2.804561e-05				
origin_Europe:horsepower	1.0	3.024522	3.024522	20.230368
9.121050e-06				
origin_Japan:horsepower	1.0	1.977640	1.977640	13.228002
3.136339e-04				
Residual	382.0	57.110548	0.149504	NaN
NaN				

[38]: <statsmodels.regression.linear_model.RegressionResultsWrapper at 0x7cb746ff5d60>

• From the above analysis, it is evident that with the interaction between origin and horsepower, the interaction between oilshock and weight and horsepower is insignificant. We can drop these from the model as well.

OLS Regression Results

```
Dep. Variable:
                                       R-squared:
                                                                       0.852
                                 mpg
Model:
                                 OLS
                                       Adj. R-squared:
                                                                       0.849
Method:
                       Least Squares
                                      F-statistic:
                                                                       275.8
Date:
                   Fri, 21 Feb 2025
                                      Prob (F-statistic):
                                                                 8.41e-154
                            15:48:45
                                       Log-Likelihood:
Time:
                                                                     -181.63
No. Observations:
                                 392
                                       AIC:
                                                                       381.3
```

Df Residuals: Df Model: Covariance Type:	nonrol	383 8 oust	BIC:					417.0
=========	======		=====	====	======	====	======	
0.975]	CO6	ef 	std e	rr 		t 	P> t	[0.025
Intercept	-0.426	67	0.0	34	-12.49	5	0.000	-0.494
-0.360	0.12			-				0 7 10 1
oilshock[T.1] 0.647	0.562	28	0.0	43	13.07	3	0.000	0.478
horsepower 0.056	-0.026	35	0.0	42	-0.63	3	0.527	-0.109
oilshock[T.1]:horsepower	-0.380)4	0.0	51	-7.49	7	0.000	-0.480
weight -0.438	-0.523	31	0.0	43	-12.05	1	0.000	-0.608
origin_Europe 0.163	0.004	11	0.0	81	0.05	1	0.959	-0.155
origin_Japan 0.254	0.08	77	0.0	84	1.03	9	0.299	-0.078
origin_Europe:horsepower	-0.442	23	0.0	96	-4.62	6	0.000	-0.630
origin_Japan:horsepower	-0.358	39	0.1	02	-3.52	6	0.000	-0.559
Omnibus:	19	-=== . 159	===== Durb	==== in-W	atson:	====		1.576
Prob(Omnibus):		.000	-		Bera (JB)	:		29.046
Skew: Kurtosis:		.362 .119	Prob Cond					4.93e-07 9.30
	======	====	=====	. NC	·	====		9.50
Notes: [1] Standard Errors assume specified.	that tl	ne co	varian	ce m	natrix of	the	errors :	is correctly
apoolina.	df	s	um_sq		mean_sq		F	
PR(>F) oilshock	1.0	106.4	83141	106	3.483141	703	3.425918	
9.826044e-89 horsepower	1.0	165 0	48555	165	5.048555	1090	0.308104	
4.259700e-114	1.0		10000	100		100		
oilshock:horsepower 1.468876e-25	1.0	19.1	55472	19	.155472	126	6.540737	
weight 2.992243e-41	1.0	35.0	71635	35	5.071635	23:	1.682658	
and adm. Process	1 0	0.7	00007	_	700007		- 027001	

0.792887

5.237801

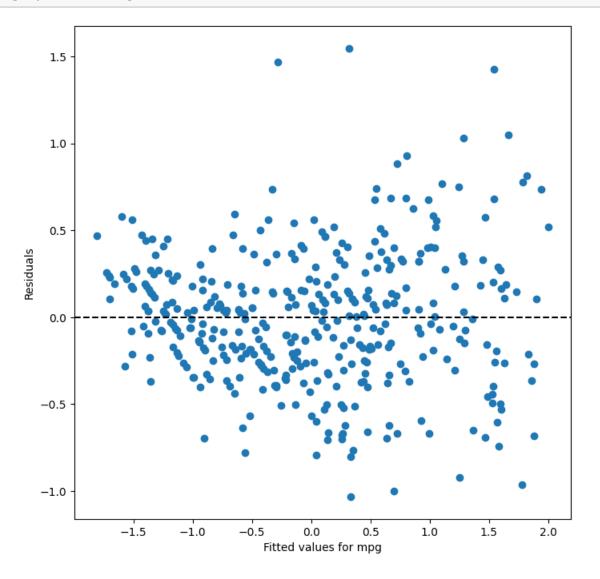
0.792887

1.0

origin_Europe

1.0	2.840217	2.840217	18.762427
1.0	2.748574	2.748574	18.157033
1.0	1.881783	1.881783	12.431031
383.0	57.977737	0.151378	NaN
	1.0	1.0 2.748574 1.0 1.881783	1.0 2.748574 2.748574 1.0 1.881783 1.881783

[40]: display_residuals_plot(results)



[41]: anova_lm(simple_model, numeric_interactions, oilshock_interactions, origin_interactions)

```
[41]:
        df_resid
                        ssr df_diff
                                        ss_diff
                                                          F
                                                                   Pr(>F)
     0
           386.0 76.594741
                                 0.0
                                            {\tt NaN}
                                                        NaN
                                                                      NaN
      1
           385.0 59.887237
                                 1.0 16.707504 110.369506 7.218441e-23
      2
           383.0 54.336579
                                 2.0
                                      5.550658
                                                  18.333780 2.489897e-08
      3
           383.0 57.977737
                                -0.0 -3.641158
                                                        inf
                                                                      NaN
[42]: pd.DataFrame(models)
[42]:
                         name
     model R-squared adjusted
                  simple model
      mpg ~ horsepower + weight + oilshock + origin_Europe + origin_Japan
      0.802074
         numeric_interactions
     mpg ~ horsepower + weight + oilshock + origin_Europe + origin_Japan +
                                   0.844846
     horsepower: weight
                                              mpg ~ horsepower + weight + oilshock +
      2 oilshock_interactions
      origin_Europe + origin_Japan + horsepower: weight + oilshock: weight + oilshock:
     horsepower
                           0.858491
          origin_interactions mpg ~ horsepower + weight + oilshock + origin_Europe +
      origin_Japan + oilshock: horsepower + origin_Europe: horsepower + origin_Japan:
     horsepower
                           0.849008
```

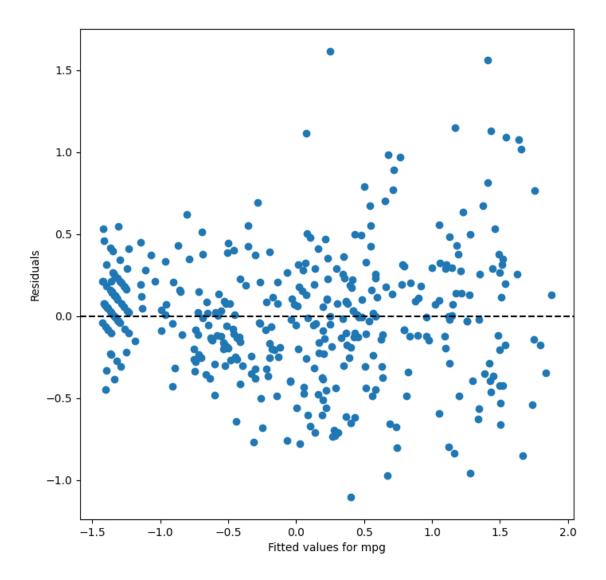
1.7.5 (f) Try a few different transformations of the variables, such as $\log(X)$, \sqrt{X} , X2 . Comment on your findings.

OLS Regression Results

Dep. Variable: R-squared: 0.848 mpg Model: OLS Adj. R-squared: 0.846 Method: F-statistic: Least Squares 306.8 Prob (F-statistic): Date: Fri, 21 Feb 2025 5.75e-153 Log-Likelihood: Time: 15:48:46 -186.58

No. Observations: Df Residuals: Df Model: Covariance Type:	1	392 384 7 nonrobust	AIC: BIC:		389.2 420.9		
0.975]	coe	======= f std e	======= rr	t P> t	[0.025		
Intercept -0.524	-0.6020	0.0	40 -15.07	72 0.000	-0.681		
oilshock[T.1] 0.742	0.6580	0.0	43 15.33	0.000	0.574		
horsepower -0.274	-0.3798	3 0.0	54 -7.05	0.000	-0.486		
weight -0.408	-0.5069	9 0.0	50 -10.04	0.000	-0.606		
origin_Europe 0.265	0.1436	6 0.0	62 2.32	0.021	0.022		
origin_Japan 0.322	0.1959	9 0.0	64 3.04	0.002	0.069		
<pre>I(horsepower ** 2) 0.139</pre>	0.0976	6 0.0	21 4.66	0.000	0.056		
I(weight ** 2) 0.193	0.1413	3 0.0	26 5.38	0.000	0.090		
Omnibus: Prob(Omnibus): Skew: Kurtosis:		26.672 0.000	Durbin-Wat Jarque-Ber Prob(JB): Cond. No.	cson:	1.394 46.435 8.26e-11 10.5		
Notes: [1] Standard Errors assume that the covariance matrix of the errors is correctly specified.							
	df	sum_sq	mean_sq	F	PR(>F)		
oilshock		06.483141	106.483141	687.677177	1.332036e-87		
horsepower		65.048555	165.048555	1065.897602	7.775452e-113		
weight		39.079210	39.079210	252.376864	4.864770e-44		
origin_Europe	1.0	0.574647	0.574647	3.711118	5.478894e-02		
origin_Japan	1.0	4.219706	4.219706	27.251220	2.932534e-07		
I(horsepower ** 2)	1.0	12.648520	12.648520	81.685220	7.946432e-18		
I(weight ** 2) Residual		4.485871 59.460351	4.485871 0.154845	28.970134 NaN	1.281966e-07 NaN		

[44]: display_residuals_plot(results)



```
[45]: anova_lm(simple_model, squared_transformations)
[45]:
         df_resid
                              df_diff
                                          ss_diff
                                                           F
                                                                     Pr(>F)
                         ssr
      0
            386.0
                   76.594741
                                   0.0
                                              NaN
                                                                        NaN
                                                         {\tt NaN}
      1
            384.0 59.460351
                                   2.0
                                        17.134391
                                                   55.327677 7.681995e-22
[46]: pd.DataFrame(models)
[46]:
                           name
      model R-squared adjusted
                   simple_model
      mpg ~ horsepower + weight + oilshock + origin_Europe + origin_Japan
      0.802074
      1
           numeric_interactions
```

```
mpg ~ horsepower + weight + oilshock + origin_Europe + origin_Japan +
horsepower: weight
                              0.844846
    oilshock_interactions
                                          mpg ~ horsepower + weight + oilshock +
origin Europe + origin Japan + horsepower: weight + oilshock: weight + oilshock:
horsepower
                      0.858491
      origin_interactions mpg ~ horsepower + weight + oilshock + origin_Europe
+ origin_Japan + oilshock: horsepower + origin_Europe: horsepower +
origin_Japan: horsepower
                                    0.849008
4 squared transformation
                                                                      mpg ~
horsepower + weight + oilshock + origin_Europe + origin_Japan + I(horsepower**2)
+ I(weight**2)
                          0.845550
```

- Since we've standardized the variables, we cannot run log or square root transformations on the negative valued columns.
- We can reload the data and run the log and sqrt transformations on the original unstandardized data.

```
[47]: Auto = load_data("Auto")
Auto = Auto.sort_values(by=["year"], ascending=True)
Auto.columns
```

[47]: Index(['mpg', 'cylinders', 'displacement', 'horsepower', 'weight', 'acceleration', 'year', 'origin'], dtype='object')

```
[48]: print("Minimums:")
  print(Auto.min())
  print("Maximums:")
  print(Auto.max())
```

Minimums:

mpg

cylinders 3.0 displacement 68.0 horsepower 46.0 weight 1613.0 acceleration 8.0 year 70.0 origin 1.0 dtype: float64 Maximums: 46.6 mpg 8.0 cylinders displacement 455.0 horsepower 230.0 weight 5140.0 acceleration 24.8 year 82.0 3.0 origin

9.0

dtype: float64

- From the above, we can see that the values for displacement, horsepower and weight are quite large.
- Hence, we log or square root transform only these variables.

1.7.6 Now let's categorize the variables

1.8 Log Transformed Model

```
[50]: Auto_log = Auto.copy(deep=True)
```

[50]:						linders	displacement	horsepower
	weight accelo	eration	yea	r origin	oilshock			
	chevrolet che	velle m	alih	11	18.0	8	307.0	130
	3504	12.0	70	America	0	Ü	001.0	100
	hi 1200d	12.0	. 0	1111101100	9.0	8	304.0	193
	4732	18.5	70	America	0		30210	200
	dodge d200				11.0	8	318.0	210
	4382	13.5	70	America	0			
	chevy c20				10.0	8	307.0	200
	4376	15.0	70	America	0			
	ford f250				10.0	8	360.0	215
	4615	14.0	70	America	0			
	amc gremlin				21.0	6	199.0	90
	2648	15.0	70	America	0			
	bmw 2002				26.0	4	121.0	113
	2234	12.5	70	Europe	0			
	saab 99e				25.0	4	104.0	95
	2375	17.5	70	Europe	0			
	audi 100 ls				24.0	4	107.0	90
	2430	14.5	70	Europe	0			
	volkswagen 113				26.0	4	97.0	46
	1835	20.5	70	Europe	0	_		
	datsun pl510			_	27.0	4	97.0	88
	2130	14.5	70	Japan	0			
	ford maverick				21.0	6	200.0	85
	2587	16.0	70	America	0	0	400.0	07
	amc hornet	45 5	70		18.0	6	199.0	97
	2774	15.5	70	America	0			

plymouth duster	A	22.0	6	198.0	95
2833 15.5 70 peugeot 504	America	0 25.0	4	110.0	87
2672 17.5 70 buick estate wagon (sw)	Europe	0 14.0	8	455.0	225
3086 10.0 70	America	0	O	400.0	220
toyota corona mark ii 2372 15.0 70	Japan	24.0 0	4	113.0	95
plymouth satellite	Japan	18.0	8	318.0	150
3436 11.0 70 amc rebel sst	America	0 16.0	8	304.0	150
3433 12.0 70	America	0	Ü	001.0	100
ford torino 3449 10.5 70	America	17.0 0	8	302.0	140
ford galaxie 500	America	15.0	8	429.0	198
4341 10.0 70 chevrolet impala	America	0 14.0	8	454.0	220
4354 9.0 70	America	0	O	101.0	220
buick skylark 320 3693 11.5 70	America	15.0 0	8	350.0	165
pontiac catalina	America	14.0	8	455.0	225
4425 10.0 70 amc ambassador dpl	America	0 15.0	8	390.0	190
3850 8.5 70	America	0	O	330.0	190
dodge challenger se 3563 10.0 70	America	15.0 0	8	383.0	170
plymouth 'cuda 340	America	14.0	8	340.0	160
3609 8.0 70 chevrolet monte carlo	America	0 15.0	8	400.0	150
3761 9.5 70	America	0	O	400.0	130
plymouth fury iii 4312 8.5 70	Amorica	14.0 0	8	440.0	215
amc hornet sportabout (s	America sw)	18.0	6	258.0	110
2962 13.5 71 chevrolet vega (sw)	America	0 22.0	4	140.0	72
2408 19.0 71	America	0	4	140.0	12
pontiac firebird	A	19.0	6	250.0	100
3282 15.0 71 ford mustang	America	0 18.0	6	250.0	88
3139 14.5 71	America	0	4	100.0	0.0
mercury capri 2000 2220 14.0 71	America	23.0 0	4	122.0	86
toyota corolla 1200	-	31.0	4	71.0	65
1773 19.0 71 peugeot 304	Japan	0 30.0	4	79.0	70
2074 19.5 71	Europe	0			
datsun 1200		35.0	4	72.0	69

1613 18.0 71 Japan	0			
volkswagen model 111	27.0	4	97.0	60
1834 19.0 71 Europe	0			
plymouth cricket	26.0	4	91.0	70
1955 20.5 71 America	0			
pontiac safari (sw)	13.0	8	400.0	175
5140 12.0 71 America	0	•	440.0	
opel 1900	28.0	4	116.0	90
2123 14.0 71 Europe	0 13.0	8	400.0	170
ford country squire (sw) 4746 12.0 71 America	13.0	0	400.0	170
fiat 124b	30.0	4	88.0	76
2065 14.5 71 Europe	0	1	00.0	70
plymouth fury iii	14.0	8	318.0	150
4096 13.0 71 America	0			
dodge monaco (sw)	12.0	8	383.0	180
4955 11.5 71 America	0			
chevrolet vega 2300	28.0	4	140.0	90
2264 15.5 71 America	0			
toyota corona	25.0	4	113.0	95
2228 14.0 71 Japan	0			
amc gremlin	19.0	6	232.0	100
2634 13.0 71 America	0			
plymouth satellite custom	16.0	6	225.0	105
3439 15.5 71 America	0			
datsun pl510	27.0	4	97.0	88
2130 14.5 71 Japan	0			
ford torino 500	19.0	6	250.0	88
3302 15.5 71 America	0			
amc matador	18.0	6	232.0	100
3288 15.5 71 America	0	_		
chevrolet impala	14.0	8	350.0	165
4209 12.0 71 America	0			
pontiac catalina brougham	14.0	8	400.0	175
4464 11.5 71 America	0	0	054.0	150
ford galaxie 500	14.0	8	351.0	153
4154 13.5 71 America	0		050.0	100
chevrolet chevelle malibu	17.0	6	250.0	100
3329 15.5 71 America	12.0	0	207.0	120
chevrolet chevelle concours (sw) 4098 14.0 72 America	13.0 0	8	307.0	130
plymouth satellite custom (sw)	14.0	8	318.0	150
4077 14.0 72 America	0	0	310.0	150
volvo 145e (sw)	18.0	4	121.0	112
2933 14.5 72 Europe	0	4	121.0	112
volkswagen 411 (sw)	22.0	4	121.0	76
2511 18.0 72 Europe	0	4	121.0	70
2011 10.0 12 Lutope	J			

peugeot 504 (sw)	21.0	4	120.0	87
2979 19.5 72 Europe ford pinto (sw)	0 22.0	4	122.0	86
2395 16.0 72 America	0	I	122.0	00
datsun 510 (sw)	28.0	4	97.0	92
2288 17.0 72 Japan toyouta corona mark ii (sw)	0 23.0	4	120.0	97
2506 14.5 72 Japan	0			
dodge colt (sw)	28.0	4	98.0	80
2164 15.0 72 America amc matador (sw)	0 15.0	8	304.0	150
3892 12.5 72 America	0			
toyota corolla 1600 (sw)	27.0	4	97.0	88
2100 16.5 72 Japan	0			
renault 12 (sw)	26.0	4	96.0	69
2189 18.0 72 Europe	0			
mazda rx2 coupe	19.0	3	70.0	97
2330 13.5 72 Japan	0			
ford gran torino (sw)	13.0	8	302.0	140
4294 16.0 72 America	0			
oldsmobile delta 88 royale	12.0	8	350.0	160
4456 13.5 72 America	0			
chrysler newport royal	13.0	8	400.0	190
4422 12.5 72 America	0			
toyota corona hardtop	24.0	4	113.0	95
2278 15.5 72 Japan	0			
volkswagen type 3	23.0	4	97.0	54
2254 23.5 72 Europe	0			
chevrolet vega	20.0	4	140.0	90
2408 19.5 72 America	0			
ford pinto runabout	21.0	4	122.0	86
2226 16.5 72 America	0			
chevrolet impala	13.0	8	350.0	165
4274 12.0 72 America	0			
dodge colt hardtop	25.0	4	97.5	80
2126 17.0 72 America	0			
plymouth fury iii	15.0	8	318.0	150
4135 13.5 72 America	0			
ford galaxie 500	14.0	8	351.0	153
4129 13.0 72 America	0			
amc ambassador sst	17.0	8	304.0	150
3672 11.5 72 America	0			
mercury marquis	11.0	8	429.0	208
4633 11.0 72 America	0			
buick lesabre custom	13.0	8	350.0	155
4502 13.5 72 America	0			
pontiac catalina	14.0	8	400.0	175

4385	12.0	72	America	0			
fiat 124 sport			ımoı ioa	26.0	4	98.0	90
2265	15.5	73	Europe	0			
amc gremlin			•	18.0	6	232.0	100
2789	15.0	73	America	0			
toyota carina				20.0	4	97.0	88
2279	19.0	73	Japan	0			
chevrolet vega	ì.			21.0	4	140.0	72
2401	19.5	73	America	0			
datsun 610				22.0	4	108.0	94
2379	16.5	73	Japan	0			
maxda rx3			_	18.0	3	70.0	90
2124	13.5	73	Japan	0	_		
ford pinto				19.0	4	122.0	85
2310	18.5	73	America	0		455.0	407
mercury capri		72	۸	21.0	6	155.0	107
2472 chevrolet mont	14.0	73	America	0 15.0	8	350.0	145
4082	13.0	73	America	0	0	350.0	145
saab 991e	13.0	13	America	24.0	4	121.0	110
2660	14.0	73	Europe	24.0	-	121.0	110
fiat 128	11.0	, 0	Багоро	29.0	4	68.0	49
1867	19.5	73	Europe	0	-	00.0	10
opel manta			r	24.0	4	116.0	75
2158	15.5	73	Europe	0			
audi 1001s			•	20.0	4	114.0	91
2582	14.0	73	Europe	0			
volvo 144ea				19.0	4	121.0	112
2868	15.5	73	Europe	0			
dodge dart cus	stom			15.0	8	318.0	150
3399	11.0	73	America	0			
toyota mark ii				20.0	6	156.0	122
2807		73	Japan	0			
oldsmobile ome	_			11.0	8	350.0	180
3664	11.0		America	0			
oldsmobile vis				12.0	8	350.0	180
4499	12.5	73	America	0	0	400.0	000
pontiac grand	-	72	۸	16.0	8	400.0	230
4278	9.5	73	America	13.0	0	260.0	170
plymouth custo 4654	m subu 13.0	73	America	13.0 0	8	360.0	170
amc hornet	13.0	13	America	18.0	6	232.0	100
2945	16.0	73	America	0	O	202.0	100
chevrolet impa		. 0		11.0	8	400.0	150
4997	14.0	73	America	0	J	200.0	
buick century				13.0	8	350.0	175
4100	13.0	73	America	0			

amc matador	14.0	8	304.0	150
3672 11.5 73 America	0			
chevrolet malibu	13.0	8	350.0	145
3988 13.0 73 America	0	•	040.0	450
dodge coronet custom	15.0	8	318.0	150
3777 12.5 73 America	0 12.0	0	420.0	100
mercury marquis brougham 4952 11.5 73 America	0	8	429.0	198
chevrolet caprice classic	13.0	8	400.0	150
4464 12.0 73 America	0	G	100.0	100
ford 1td	13.0	8	351.0	158
4363 13.0 73 America	0			
plymouth fury gran sedan	14.0	8	318.0	150
4237 14.5 73 America	0			
ford gran torino	14.0	8	302.0	137
4042 14.5 73 America	0			
buick electra 225 custom	12.0	8	455.0	225
4951 11.0 73 America	0			
amc ambassador brougham	13.0	8	360.0	175
3821 11.0 73 America	0			
plymouth valiant	18.0	6	225.0	105
3121 16.5 73 America	0	_		
chevrolet nova custom	16.0	6	250.0	100
3278 18.0 73 America	0		100.0	4.05
ford country	12.0	8	400.0	167
4906 12.5 73 America	0	6	050.0	00
ford maverick	18.0	6	250.0	88
3021 16.5 73 America	0	6	100.0	٥٦
plymouth duster	23.0	6	198.0	95
2904 16.0 73 America	0 26.0	4	07.0	16
volkswagen super beetle 1950 21.0 73 Europe	0	4	97.0	46
-	13.0	8	440.0	215
chrysler new yorker brougham 4735 11.0 73 America	0	0	440.0	215
audi fox	29.0	4	98.0	83
2219 16.5 74 Europe	0	-	00.0	00
volkswagen dasher	26.0	4	79.0	67
1963 15.5 74 Europe	0	-	10.0	01
opel manta	26.0	4	97.0	78
2300 14.5 74 Europe	0			
toyota corona	31.0	4	76.0	52
1649 16.5 74 Japan	0			
datsun 710	32.0	4	83.0	61
2003 19.0 74 Japan	0			
subaru	26.0	4	108.0	93
2391 15.5 74 Japan	0			
fiat 128	24.0	4	90.0	75

2108	15.5	74	Europe	0			
fiat 124 tc			1	26.0	4	116.0	75
2246	14.0	74	Europe	0			
honda civic				24.0	4	120.0	97
2489	15.0	74	Japan	0			
fiat x1.9				31.0	4	79.0	67
2000	16.0	74	Europe	0			
amc matador (14.0	8	304.0	150
4257	15.5	74	America	0			
dodge colt				28.0	4	90.0	75
2125	14.5	74	America	0			
ford gran tor				14.0	8	302.0	140
4638	16.0	74	America	0			
ford gran tor				16.0	8	302.0	140
4141	14.0	74		0			
buick century	luxus	(wa)		13.0	8	350.0	150
4699	14.5	74	America	0			
dodge coronet	custom	(sw	·)	14.0	8	318.0	150
4457	13.5	74	America	0			
plymouth dust	er			20.0	6	198.0	95
3102	16.5	74	America	0			
amc hornet				19.0	6	232.0	100
2901	16.0	74	America	0			
chevrolet nov				15.0	6	250.0	100
3336	17.0	74	America	0	-		
ford pinto				26.0	4	122.0	80
2451	16.5	74	America	0	-	122.0	00
datsun b210	10.0		1111101 100	31.0	4	79.0	67
1950	19.0	74	Japan	0	-	70.0	01
chevrolet veg		1-1	Japan	25.0	4	140.0	75
2542	a 17.0	74	America	0	4	140.0	13
chevrolet che				16.0	6	250.0	100
					6	250.0	100
3781	17.0	74	America	0		050.0	440
amc matador	40.0	- 4		16.0	6	258.0	110
3632			America	0			
plymouth sate			J	18.0	6	225.0	105
3613	16.5	74	America	0			
toyota coroll				32.0	4	71.0	65
1836	21.0	74	Japan	0			
datsun 710				24.0	4	119.0	97
2545	17.0	75	Japan	0			
pontiac astro				23.0	4	140.0	78
2592	18.5	75	America	0			
amc gremlin				20.0	6	232.0	100
2914	16.0	75	America	0			
toyota corona				24.0	4	134.0	96
2702	13.5	75	Japan	0			
			-				

volkswagen dasher			25.0	4	90.0	71
2223 16.5	75	Europe	0	_		
ford pinto		•	18.0	6	171.0	97
2984 14.5	75	America	0			
saab 991e			25.0	4	121.0	115
2671 13.5	75	Europe	0			
amc pacer			19.0	6	232.0	90
3211 17.0	75	America	0	4	115.0	0.5
audi 1001s 2694 15.0	75	Furana	23.0 0	4	115.0	95
peugeot 504	13	Europe	23.0	4	120.0	88
2957 17.0	75	Europe	0	-	120.0	00
volvo 244dl	. •		22.0	4	121.0	98
2945 14.5	75	Europe	0			
honda civic cvcc		_	33.0	4	91.0	53
1795 17.5	75	Japan	0			
ford pinto			23.0	4	140.0	83
2639 17.0	75	America	0			
volkswagen rabbit		_	29.0	4	90.0	70
1937 14.0	75	Europe	0	4	07.0	70
toyota corolla 2171 16.0	75	Innon	29.0 0	4	97.0	75
plymouth valiant cus		Japan	19.0	6	225.0	95
3264 16.0	75	America	0	Ü	220.0	50
chevrolet monza 2+2			20.0	8	262.0	110
3221 13.5	75	America	0			
ford mustang ii			13.0	8	302.0	129
3169 12.0	75	America	0			
chevrolet nova			18.0	6	250.0	105
3459 16.0	75	America	0	_		
mercury monarch	75		15.0	6	250.0	72
3432 21.0	75	America	0 16.0	8	400.0	170
pontiac catalina 4668 11.5	75	America	0	0	400.0	170
chevrolet bel air	10	America	15.0	8	350.0	145
4440 14.0	75	America	0	· ·		-10
plymouth grand fury			16.0	8	318.0	150
4498 14.5	75	America	0			
ford maverick			15.0	6	250.0	72
3158 19.5	75	America	0			
buick century			17.0	6	231.0	110
3907 21.0	75	America	0		050.0	105
chevroelt chevelle			16.0	6	250.0	105
3897 18.5 amc matador	75	America	0 15.0	6	258.0	110
3730 19.0	75	America	0	O	200.0	110
plymouth fury	. 0	101 100	18.0	6	225.0	95
. J			- -	-	- · · -	

3785	19.0	75	America	0			
buick skyhaw				21.0	6	231.0	110
3039	15.0	75	America	0			
ford ltd				14.0	8	351.0	148
4657	13.5	75	America	0			
ford pinto				26.5	4	140.0	72
2565	13.6	76	America	0			
pontiac vent	-			18.5	6	250.0	110
3645	16.2	76	America	0			
amc pacer d/				17.5	6	258.0	95
3193	17.8	76	America	0	_		
volkswagen r		7.0	-	29.5	4	97.0	71
1825	12.2	76	Europe	0	4	05.0	70
datsun b-210	17 0	76	T	32.0	4	85.0	70
1990	17.0	76	Japan	0 28.0	4	97.0	75
toyota coroli 2155	16.4	76	Ianan	20.0	4	91.0	75
volvo 245	10.4	70	Japan	20.0	4	130.0	102
3150	15.7	76	Europe	0	T	100.0	102
ford f108	10.1	, 0	дагоро	13.0	8	302.0	130
3870	15.0	76	America	0	· ·	002.0	100
peugeot 504				19.0	4	120.0	88
3270	21.9	76	Europe	0			
toyota mark	ii		•	19.0	6	156.0	108
2930	15.5	76	Japan	0			
mercedes-ben	z 280s			16.5	6	168.0	120
3820	16.7	76	Europe	0			
cadillac sev	ille			16.5	8	350.0	180
4380	12.1	76	America	0			
chevy c10				13.0	8	350.0	145
4055	12.0	76	America	0			
dodge d100	4.4.0			13.0	8	318.0	150
3755	14.0	76	America	0	6	050.0	70
ford granada	•	76	Amorrico	18.0 0	6	250.0	78
3574 plymouth vol	21.0		America	13.0	8	318.0	150
3940	13.2	76	America	0	0	310.0	130
dodge aspen		70	America	20.0	6	225.0	100
3651	17.7	76	America	0	· ·	220.0	100
vw rabbit				29.0	4	90.0	70
1937	14.2	76	Europe	0			
opel 1900			_	25.0	4	116.0	81
2220	16.9	76	Europe	0			
honda civic				33.0	4	91.0	53
1795	17.4	76	Japan	0			
fiat 131				28.0	4	107.0	86
2464	15.5	76	Europe	0			

capri ii		25.0	4	140.0	92
	merica	0			7.0
dodge colt 2255 17.7 76 A	morico	26.0	4	98.0	79
2255 17.7 76 A renault 12tl	merica	27.0	4	101.0	83
	Europe	0	•	101.0	00
dodge coronet brougham	1	16.0	8	318.0	150
4190 13.0 76 A	merica	0			
amc matador		15.5	8	304.0	120
	merica	0	_		
chevrolet chevelle malibu		17.5	8	305.0	140
	merica	0		005.0	400
plymouth valiant		22.0	6	225.0	100
	merica	0	6	050.0	405
chevrolet nova		22.0	6	250.0	105
	merica	0 24.0	6	200 0	01
ford maverick 3012 17.6 76 A	merica	0	6	200.0	81
amc hornet	merica	22.5	6	232.0	90
	merica	0	0	202.0	30
chevrolet chevette	merica	29.0	4	85.0	52
	merica	0	-	33.3	02
chevrolet woody		24.5	4	98.0	60
· ·	merica	0			
ford gran torino		14.5	8	351.0	152
_	merica	0			
ford mustang ii 2+2		25.5	4	140.0	89
2755 15.8 77 A	merica	1			
volkswagen rabbit custom		29.0	4	97.0	78
1940 14.5 77	Europe	1			
pontiac sunbird coupe		24.5	4	151.0	88
	merica	1			
toyota corolla liftback		26.0	4	97.0	75
2265 18.2 77	Japan	1	_		
chevrolet chevette		30.5	4	98.0	63
	merica	1		101 0	
bmw 320i	-	21.5	4	121.0	110
	Europe	1	4	07.0	67
subaru dl 1985 16.4 77	Ionon	30.0	4	97.0	67
1985 16.4 77 volkswagen dasher	Japan	1 30.5	4	97.0	78
_	Europe	1	4	91.0	70
datsun 810	Lurope	22.0	6	146.0	97
2815 14.5 77	Japan	1	•	_ 10.0	01
mazda rx-4		21.5	3	80.0	110
2720 13.5 77	Japan	1	-	· *	
ford thunderbird	1	16.0	8	351.0	149

4335	14.5	77	America	1			
dodge colt m/		• •	ımor roa	33.5	4	98.0	83
2075	 15.9	77	America	1	_	00.0	
chrysler cord	oba			15.5	8	400.0	190
4325	12.2	77	America	1			
chevrolet mon				15.5	8	350.0	170
4165	11.4	77	America	1			
plymouth arro				25.5	4	122.0	96
2300	15.5	77	America	1			
buick opel is				30.0	4	111.0	80
2155	14.8	77	America	1			
renault 5 gtl				36.0	4	79.0	58
1825	18.6	77	Europe	1			
datsun f-10 h	atchba	ck	-	33.5	4	85.0	70
1945	16.8	77	Japan	1			
pontiac grand	prix	1j	-	16.0	8	400.0	180
4220	11.1	77	America	1			
oldsmobile cu	tlass	supre	eme	17.0	8	260.0	110
4060	19.0	77	America	1			
chevrolet cap	rice c	lassi	.C	17.5	8	305.0	145
3880	12.5	77	America	1			
mercury couga	r brou	gham		15.0	8	302.0	130
4295	14.9	77	America	1			
chevrolet con	cours			17.5	6	250.0	110
3520	16.4	77	America	1			
buick skylark				20.5	6	231.0	105
3425	16.9	77	America	1			
plymouth vola	re cus	tom		19.0	6	225.0	100
3630	17.7	77	America	1			
ford granada				18.5	6	250.0	98
3525	19.0	77	America	1			
dodge monaco	brough	am		15.5	8	318.0	145
4140	13.7	77	America	1			
honda accord				31.5	4	98.0	68
2045	18.5	77	Japan	1			
datsun 510				27.2	4	119.0	97
2300	14.7	78	Japan	1			
toyota corona				27.5	4	134.0	95
2560	14.2	78	Japan	1			
chevrolet che				30.0	4	98.0	68
2155	16.5	78		1			
buick regal s	_	_		17.7	6	231.0	165
3445	13.4	78	America	1			
ford futura		- -		18.1	8	302.0	139
3205	11.2	78	America	1	A	105.0	==
dodge omni	44 =	70		30.9	4	105.0	75
2230	14.5	78	America	1			

dodge magnum xe			17.5	8	318.0	140
•	7 78	America	1			
toyota celica gt	Liftbac	k	21.1	4	134.0	95
2515 14.8	3 78	Japan	1	_		
peugeot 604sl	. 70	-	16.2	6	163.0	133
3410 15.8		Europe	1	4	151 0	85
oldsmobile starfing 2855 17.6		America	23.8 1	4	151.0	00
datsun 200-sx	, 10	America	23.9	4	119.0	97
2405 14.9	78	Japan	1	-	110.0	01
audi 5000		1	20.3	5	131.0	103
2830 15.9	78	Europe	1			
volvo 264gl			17.0	6	163.0	125
3140 13.6	5 78	Europe	1			
saab 99gle			21.6	4	121.0	115
2795 15.7		Europe	1	_		
volkswagen sciroco			31.5	4	89.0	71
1990 14.9	9 78	Europe	1 29.5	4	00.0	60
2135 16.6	5 78	Japan	29.5 1	4	98.0	68
plymouth sapporo	, 10	Japan	23.2	4	156.0	105
2745 16.7	7 78	America	1	-	100.0	100
chevrolet monte ca			19.2	8	305.0	145
3425 13.2	2 78	America	1			
mazda glc deluxe			32.8	4	78.0	52
1985 19.4	1 78	Japan	1			
dodge aspen			18.6	6	225.0	110
3620 18.7		America	1	_		
volkswagen rabbit			43.1	4	90.0	48
1985 21.5	5 78	Europe	1 36.1	4	00.0	66
ford fiesta 1800 14.4	1 78	America	1	4	98.0	66
datsun b210 gx	10	America	39.4	4	85.0	70
2070 18.6	5 78	Japan	1	-		
honda civic cvcc		1	36.1	4	91.0	60
1800 16.4	1 78	Japan	1			
amc concord d/l			18.1	6	258.0	120
3410 15.3	L 78	America	1			
dodge diplomat			19.4	8	318.0	140
3735 13.2		America	1	0	200	400
mercury monarch gl 3570 12.8		۸	20.2	8	302.0	139
3570 12.8 oldsmobile cutlass		America	1 19.9	8	260.0	110
3365 15.5		America	19.9	J	200.0	110
chevrolet malibu			20.5	6	200.0	95
3155 18.2	2 78	America	1	•	-	
ford fairmont (aut	50)		20.2	6	200.0	85

2965	15.8	78	America	1			
ford fairmont				25.1	4	140.0	88
2720	15.4	78	America	1			
plymouth volar	re			20.5	6	225.0	100
3430	17.2	78	America	1			
amc concord				19.4	6	232.0	90
3210	17.2	78	America	1		204.2	405
buick century	specia. 15.8		A	20.6	6	231.0	105
3380		78	America	1 20.8	6	200.0	O.E.
mercury zephym 3070	16.7	78	America	20.8	0	200.0	85
pontiac phoeni		10	America	19.2	6	231.0	105
3535	19.2	78	America	1	Ü	201.0	100
plymouth horiz				34.2	4	105.0	70
2200	13.2	79	America	1			
mercedes benz	300d			25.4	5	183.0	77
3530	20.1	79	Europe	1			
cadillac eldon	rado		-	23.0	8	350.0	125
3900	17.4	79	America	1			
peugeot 504				27.2	4	141.0	71
3190	24.8	79	Europe	1			
oldsmobile cut	class s	alon	brougham	23.9	8	260.0	90
3420	22.2	79	America	1			
plymouth horiz	zon tc3			34.5	4	105.0	70
2150	14.9	79	America	1			
amc spirit dl				27.4	4	121.0	80
2670	15.0	79	America	1	_		
fiat strada cu			_	37.3	4	91.0	69
2130	14.7	79	Europe	1	4	454.0	0.0
buick skylark				28.4	4	151.0	90
2670	16.0	79	America	1	6	172 0	115
chevrolet cita		70	Amonias	28.8	6	173.0	115
2595 oldsmobile ome	11.3	79 uæba	America m	1 26.8	6	173.0	115
2700	12.9	_	m America	20.0	O	173.0	113
pontiac phoeni		13	America	33.5	4	151.0	90
2556	13.2	79	America	1	1	101.0	30
datsun 210	10.2	10	nmor roa	31.8	4	85.0	65
2020	19.2	79	Japan	1	-	33.1	
dodge colt hat			_	35.7	4	98.0	80
1915	14.4	79		1			
dodge st. regi				18.2	8	318.0	135
3830	15.2	79	America	1			
vw rabbit cust	om			31.9	4	89.0	71
1925	14.0	79	Europe	1			
mercury zephyr	6			19.8	6	200.0	85
2990	18.2	79	America	1			

2890	ford fairmont 4			22.3	4	140.0	88
3265 18.2 79 America 1		79	America				
dodge aspen 6 20.6 6 225.0 110 3360 16.6 79 America 1 170 8 305.0 130 3340 15.4 79 America 1 170 8 305.0 130 3340 15.4 79 America 1 170 8 302.0 129 3725 13.4 79 America 1 150 15.2 15.4 79 America 1 150 15.4 15.4 79 America 1 150 15.4 15.4 15.4 15.4 15.4 15.4 15.5 6 231.0 115 15.5 15.4 15.5 15.				20.2	6	232.0	90
3360 16.6 79 America 1 Chevrolet caprice classic 17.0 8 305.0 130 3840 15.4 79 America 1 17.6 8 302.0 129 3725 13.4 79 America 1 17.6 8 302.0 129 3725 13.4 79 America 1 17.5 6 231.0 115 3245 15.4 79 America 1 3245 15.2 79 America 1 3245 15.2 79 America 1 3245 35.0 355.0		79	America	_	_		
Chevrolet caprice class		70	A		6	225.0	110
3840 15.4 79				_	8	305.0	130
ford ltd landau	_				O	505.0	100
pontiac lemans v6 21.5 6 231.0 115 3245 15.4 79 America 1 amaxda glc deluxe 34.1 4 86.0 65 1975 15.2 79 Japan 1 amaxda glc deluxe 16.9 8 350.0 155 4360 14.9 79 America 1 amaxda glc deluxe 16.9 8 350.0 155 4360 14.9 79 America 1 amaxda glc deluxe 1 1 2 351.0 142 4 351.0 142 4 360.0 125 360.6 150 125 360.6 150 125 360.0 150 125 360.0 150 <td></td> <td></td> <td></td> <td>17.6</td> <td>8</td> <td>302.0</td> <td>129</td>				17.6	8	302.0	129
3245 15.4 79 America 1 maxda glc deluxe 34.1 4 86.0 65 1975 15.2 79 Japan 1 buick estate wagon (sw) 16.9 8 350.0 155 4360 14.9 79 America 1 ford country squire (sw) 15.5 8 351.0 142 4054 14.3 79 America 1 chevrolet malibu classic (sw) 19.2 8 267.0 125 3605 15.0 79 America 1 chrysler lebaron town	3725 13.4	79	America	1			
maxda glc deluxe 34.1 4 86.0 65 1975 15.2 79 Japan 1	_				6	231.0	115
1975		79	America		4	00.0	C.F.
buick estate wagon (sw)		70	Janan		4	86.0	65
4360 14.9 79 America 1 ford country squire (sw) 15.5 8 351.0 142 4054 14.3 79 America 1 1 chevrolet malibu classic (sw) 19.2 8 267.0 125 3605 15.0 79 America 1 1 chrysler lebaron town country (sw) 18.5 8 360.0 150 3940 13.0 79 America 1 360.5 351.0 138 3940 13.0 79 America 1 360.5 351.0 138 3940 13.2 79 America 1 360.5 351.0 138 361.0 138 3955 13.2 79 America 1 44.3 4 90.0 48 2085 21.7 80 Europe 1 1 43.4 4 90.0 48 2335 23.7 80 Europe 1 46.0 67 2250 12.1 67 67 225			Japan		8	350 0	155
A054	•		America		Ü	200.0	100
chevrolet malibu classic (sw) 19.2 8 267.0 125 3605 15.0 79 America 1 chrysler lebaron town @ country (sw) 18.5 8 360.0 150 3940 13.0 79 America 1 mercury grand marquis 16.5 8 351.0 138 3955 13.2 79 America 1 vw rabbit c (diesel) 44.3 4 90.0 48 2085 21.7 80 Europe 1 1 1 1 1 1 1 1 1 1 1 2 1 2 1 4 90.0 48 2 2085 21.7 80 Europe 1 2 1 4 90.0 48 2 2085 23.7 80 Europe 1 6 7 2 1 6 7 2 1 6 7 2 1 6 7 2 1 6 7 3 1 6 7 1 1 <				15.5	8	351.0	142
3605 15.0 79 America 1 chrysler lebaron town 0 country (sw) 18.5 8 360.0 150 3940 13.0 79 America 1 mercury grand marquis 16.5 8 351.0 138 3955 13.2 79 America 1 vw rabbit c (diesel) 40 44.3 4 90.0 48 2085 21.7 80 Europe 1 1 2085 23.7 80 Europe 1 20 48 2335 23.7 80 Europe 1 43.4 4 90.0 48 2335 23.7 80 Europe 1 42 90.0 48 2335 23.7 80 Europe 1 44 90.0 67 67 2950 19.9 80 Europe 1 44 91.0 67 67 3250 21.8 80 Japan 1 44 89.0 62 62 188.0 132 29	4054 14.3	79	America	1			
chrysler lebaron town € country (sw) 18.5 8 360.0 150 3940 13.0 79 America 1 mercury grand marquis 16.5 8 351.0 138 3955 13.2 79 America 1 vw rabbit c (diesel) 44.3 4 90.0 48 2085 21.7 80 Europe 1 ww dasher (diesel) 43.4 4 90.0 48 2335 23.7 80 Europe 1 audi 5000s (diesel) 5 121.0 67 2950 19.9 80 Europe 1 mercedes-benz 240d 30.0 4 146.0 67 3250 21.8 80 Europe 1 honda civic 1500 gl 44.6 4 91.0 67 1850 13.8 80 Japan 1 vokswagen rabbit 29.8 4 89.0 62 1845 15.3 80 Europe 1 mazda rx-7 gs 23.7					8	267.0	125
3940 13.0 79 America 1 mercury grand marquis 16.5 8 351.0 138 3955 13.2 79 America 1 vw rabbit c (diesel) 44.3 4 90.0 48 2085 21.7 80 Europe 1 vw dasher (diesel) 43.4 4 90.0 67 2335 23.7 80 Europe 1 audi 5000s (diesel) 36.4 5 121.0 67 2950 19.9 80 Europe 1 mercedes-benz 240d 30.0 4 146.0 67 3250 21.8 80 Europe 1 honda civic 1500 gl 44.6 4 91.0 67 1850 13.8 80 Japan 1 datsun 280-zx 32.7 32.7 6 168.0 132 2910 11.4 80 Japan 1 vokswagen rabbit 29.8 4 89.0 62 1845 15.3 80 Europe 1 mazda rx-7 gs 23.7 3 70.0 100 2420 12.5 80 Japan 1 triumph tr7 coupe 35.0 4 122.0 88 2500 15.1 80 Europe 1 honda accord 32.4 4 107.0 72 2290 17.0 80 Japan 1 datsun 210 40.8 4 85.0 65 2110 19.2 80 Japan 1				_	0	200 0	450
mercury grand marquis 16.5 8 351.0 138 3955 13.2 79 America 1 vw rabbit c (diesel) 44.3 4 90.0 48 2085 21.7 80 Europe 1 20.0 48 2335 23.7 80 Europe 1 20.0 121.0 67 2950 19.9 80 Europe 1 121.0 67 2950 19.9 80 Europe 1 146.0 67 2950 21.8 80 Europe 1 146.0 67 3250 21.8 80 Europe 1 91.0 67 1850 13.8 80 Japan 1	•		•		ð	360.0	150
3955 13.2 79 America 1 vw rabbit c (diesel) 44.3 4 90.0 48 2085 21.7 80 Europe 1 vw dasher (diesel) 43.4 4 90.0 48 2335 23.7 80 Europe 1 audi 5000s (diesel) 36.4 5 121.0 67 2950 19.9 80 Europe 1 67 2950 19.9 80 Europe 1 67 3250 21.8 80 Europe 1 6 44.6 4 91.0 67 1850 13.8 80 Japan 1 6 168.0 132 2910 11.4 80 Japan 1 89.0 62 1845 15.3 80 Europe 1 89.0 62 1845 15.3 80 Europe 1 89.0 100 2420 12.5 80 Japan 1 100 100 2420			America		8	351.0	138
2085 21.7 80 Europe 1 vw dasher (diesel) 43.4 4 90.0 48 2335 23.7 80 Europe 1 audi 5000s (diesel) 36.4 5 121.0 67 2950 19.9 80 Europe 1 146.0 67 2950 21.8 80 Europe 1 146.0 67 3250 21.8 80 Europe 1 91.0 67 1850 13.8 80 Japan 1 168.0 132 2910 11.4 80 Japan 1 168.0 132 2910 11.4 80 Japan 1 1 100 62 1845 15.3 80 Europe 1			America		•	332.0	
vw dasher (diesel) 43.4 4 90.0 48 2335 23.7 80 Europe 1 1 audi 5000s (diesel) 36.4 5 121.0 67 2950 19.9 80 Europe 1 1 mercedes-benz 240d 30.0 4 146.0 67 3250 21.8 80 Europe 1 44.6 4 91.0 67 1850 13.8 80 Japan 1 5 168.0 132 2910 11.4 80 Japan 1 5 168.0 132 2910 11.4 80 Japan 1 5 168.0 132 2910 11.4 80 Japan 1 5 168.0 132 1845 15.3 80 Europe 1 5 70.0 100 2420 12.5 80 Japan 1 1 1 107.0 72 290 17.0 80 Japan 1 1 107.0	vw rabbit c (diese	L)		44.3	4	90.0	48
2335 23.7 80 Europe 1 audi 5000s (diesel) 36.4 5 121.0 67 2950 19.9 80 Europe 1 146.0 67 mercedes-benz 240d 30.0 4 146.0 67 3250 21.8 80 Europe 1 146.0 67 1850 13.8 80 Japan 1 91.0 67 1850 13.8 80 Japan 1 6 168.0 132 2910 11.4 80 Japan 1 89.0 62 1845 15.3 80 Europe 1 89.0 62 1845 15.3 80 Europe 1 70.0 100 2420 12.5 80 Japan 1 122.0 88 2500 15.1 80 Europe 1 107.0 72 2290 17.0 80 Japan 1 4 107.0 72 2290 17.0 80 Japan		80	Europe				
audi 5000s (diesel) 36.4 5 121.0 67 2950 19.9 80 Europe 1 mercedes-benz 240d 30.0 4 146.0 67 3250 21.8 80 Europe 1 honda civic 1500 gl 44.6 4 91.0 67 1850 13.8 80 Japan 1 datsun 280-zx 32.7 6 168.0 132 2910 11.4 80 Japan 1 vokswagen rabbit 29.8 4 89.0 62 1845 15.3 80 Europe 1 mazda rx-7 gs 23.7 3 70.0 100 2420 12.5 80 Japan 1 triumph tr7 coupe 35.0 4 122.0 88 2500 15.1 80 Europe 1 honda accord 32.4 4 107.0 72 2290 17.0 80 Japan 1 datsun 210 40.8 4 85.0			_		4	90.0	48
2950 19.9 80 Europe 1 mercedes-benz 240d 30.0 4 146.0 67 3250 21.8 80 Europe 1 honda civic 1500 gl 44.6 4 91.0 67 1850 13.8 80 Japan 1 6 168.0 132 2910 11.4 80 Japan 1 89.0 62 1845 15.3 80 Europe 1 89.0 62 1845 15.3 80 Europe 1 9 100 100 2420 12.5 80 Japan 1 122.0 88 2500 15.1 80 Europe 1 1 107.0 72 2290 17.0 80 Japan 1 1 107.0 72 2290 17.0 80 Japan 1 4 85.0 65 2110 19.2 80 Japan 1 1 1 1 1 1 <			Europe	_	_	101.0	67
mercedes-benz 240d 30.0 4 146.0 67 3250 21.8 80 Europe 1 honda civic 1500 gl 44.6 4 91.0 67 1850 13.8 80 Japan 1 6 168.0 132 2910 11.4 80 Japan 1 89.0 62 1845 15.3 80 Europe 1 89.0 62 1845 15.3 80 Europe 1 70.0 100 2420 12.5 80 Japan 1 122.0 88 2500 15.1 80 Europe 1 107.0 72 2290 17.0 80 Japan 1 107.0 72 2290 17.0 80 Japan 1 4 85.0 65 2110 19.2 80 Japan 1 4 85.0 65			Furone		5	121.0	67
3250 21.8 80 Europe 1 honda civic 1500 gl 44.6 4 91.0 67 1850 13.8 80 Japan 1 datsun 280-zx 32.7 6 168.0 132 2910 11.4 80 Japan 1 vokswagen rabbit 29.8 4 89.0 62 1845 15.3 80 Europe 1 mazda rx-7 gs 23.7 3 70.0 100 2420 12.5 80 Japan 1 triumph tr7 coupe 35.0 4 122.0 88 2500 15.1 80 Europe 1 honda accord 32.4 4 107.0 72 2290 17.0 80 Japan 1 datsun 210 40.8 4 85.0 65 2110 19.2 80 Japan 1		00	Lurope		4	146.0	67
1850 13.8 80 Japan 1 datsun 280-zx 32.7 6 168.0 132 2910 11.4 80 Japan 1 vokswagen rabbit 29.8 4 89.0 62 1845 15.3 80 Europe 1 mazda rx-7 gs 23.7 3 70.0 100 2420 12.5 80 Japan 1 triumph tr7 coupe 35.0 4 122.0 88 2500 15.1 80 Europe 1 honda accord 32.4 4 107.0 72 2290 17.0 80 Japan 1 datsun 210 40.8 4 85.0 65 2110 19.2 80 Japan 1		80	Europe				
datsun 280-zx 32.7 6 168.0 132 2910 11.4 80 Japan 1 vokswagen rabbit 29.8 4 89.0 62 1845 15.3 80 Europe 1 mazda rx-7 gs 23.7 3 70.0 100 2420 12.5 80 Japan 1 triumph tr7 coupe 35.0 4 122.0 88 2500 15.1 80 Europe 1 honda accord 32.4 4 107.0 72 2290 17.0 80 Japan 1 datsun 210 40.8 4 85.0 65 2110 19.2 80 Japan 1	honda civic 1500 g	L	-	44.6	4	91.0	67
2910 11.4 80 Japan 1 vokswagen rabbit 29.8 4 89.0 62 1845 15.3 80 Europe 1 mazda rx-7 gs 23.7 3 70.0 100 2420 12.5 80 Japan 1 triumph tr7 coupe 35.0 4 122.0 88 2500 15.1 80 Europe 1 honda accord 32.4 4 107.0 72 2290 17.0 80 Japan 1 datsun 210 40.8 4 85.0 65 2110 19.2 80 Japan 1		80	Japan				
vokswagen rabbit 29.8 4 89.0 62 1845 15.3 80 Europe 1 mazda rx-7 gs 23.7 3 70.0 100 2420 12.5 80 Japan 1 triumph tr7 coupe 35.0 4 122.0 88 2500 15.1 80 Europe 1 honda accord 32.4 4 107.0 72 2290 17.0 80 Japan 1 datsun 210 40.8 4 85.0 65 2110 19.2 80 Japan 1		00	.		6	168.0	132
1845 15.3 80 Europe 1 mazda rx-7 gs 23.7 3 70.0 100 2420 12.5 80 Japan 1 triumph tr7 coupe 35.0 4 122.0 88 2500 15.1 80 Europe 1 honda accord 32.4 4 107.0 72 2290 17.0 80 Japan 1 datsun 210 40.8 4 85.0 65 2110 19.2 80 Japan 1		80	Japan		1	90.0	60
mazda rx-7 gs 23.7 3 70.0 100 2420 12.5 80 Japan 1 triumph tr7 coupe 35.0 4 122.0 88 2500 15.1 80 Europe 1 honda accord 32.4 4 107.0 72 2290 17.0 80 Japan 1 datsun 210 40.8 4 85.0 65 2110 19.2 80 Japan 1	_	80	Furone		4	09.0	02
2420 12.5 80 Japan 1 triumph tr7 coupe 35.0 4 122.0 88 2500 15.1 80 Europe 1 honda accord 32.4 4 107.0 72 2290 17.0 80 Japan 1 datsun 210 40.8 4 85.0 65 2110 19.2 80 Japan 1		00	Багоро		3	70.0	100
2500 15.1 80 Europe 1 honda accord 32.4 4 107.0 72 2290 17.0 80 Japan 1 datsun 210 40.8 4 85.0 65 2110 19.2 80 Japan 1	•	80	Japan				
honda accord 32.4 4 107.0 72 2290 17.0 80 Japan 1 datsun 210 40.8 4 85.0 65 2110 19.2 80 Japan 1	_			35.0	4	122.0	88
2290 17.0 80 Japan 1 datsun 210 40.8 4 85.0 65 2110 19.2 80 Japan 1		80	Europe				
datsun 210 40.8 4 85.0 65 2110 19.2 80 Japan 1		00	T		4	107.0	72
2110 19.2 80 Japan 1		80	Japan		Λ	85 A	65
•		80	Japan		4	00.0	05
			r		4	97.0	67

2145	18.0	80	Japan	1			
dodge colt			1	27.9	4	156.0	105
2800	14.4	80	America	1			
mazda glc				46.6	4	86.0	65
2110	17.9	80	Japan	1			
toyota coroll	a			32.2	4	108.0	75
2265	15.2	80	Japan	1			
vw rabbit			_	41.5	4	98.0	76
2144	14.7	80	Europe	1			
toyota coroll	a terc	el	_	38.1	4	89.0	60
1968	18.8	80	Japan	1			
chevrolet che	vette		_	32.1	4	98.0	70
2120	15.5	80	America	1			
chevrolet cit	ation			28.0	4	151.0	90
2678	16.5	80	America	1			
ford fairmont				26.4	4	140.0	88
2870	18.1	80	America	1			
datsun 310				37.2	4	86.0	65
2019	16.4	80	Japan	1			
dodge aspen			•	19.1	6	225.0	90
3381	18.7	80	America	1			
audi 4000				34.3	4	97.0	78
2188	15.8	80	Europe	1			
toyota corona			1	29.8	4	134.0	90
2711	15.5	80	Japan	1			
mazda 626			1	31.3	4	120.0	75
2542	17.5	80	Japan	1			
datsun 510 ha		k	•	37.0	4	119.0	92
2434	15.0	80	Japan	1			
amc concord			1	24.3	4	151.0	90
3003	20.1	80	America	1			
peugeot 505s				28.1	4	141.0	80
3230	20.4	81	Europe	1	_		
honda prelude			r	33.7	4	107.0	75
2210	14.4	81	Japan	1	_		
toyota coroll				32.4	4	108.0	75
2350	16.8	81	Japan	1	-		
datsun 200sx	10.0	01	oupun	32.9	4	119.0	100
2615	14.8	81	Japan	1	-	110.0	100
mazda 626	11.0	01	oupun	31.6	4	120.0	74
2635	18.3	81	Japan	1	-	120.0	, -
volvo diesel	10.0	01	oupun	30.7	6	145.0	76
3160	19.6	81	Europe	1	J	110.0	
chrysler leba				17.6	6	225.0	85
3465	16.6	81	America	1	Ü	220.0	30
datsun 810 ma		91		24.2	6	146.0	120
2930	13.8	81	Japan	1	J	110.0	120
	10.0	91	Japan	_			

buick century			22.4	6	231.0	110
3415 15.8		America	1			
oldsmobile cutlass 1		Α .	26.6	8	350.0	105
3725 19.0 ford granada gl	81	America	1 20.2	6	200.0	88
3060 17.1	81	America	20.2	O	200.0	00
volkswagen jetta	01	nmor roa	33.0	4	105.0	74
2190 14.2	81	Europe	1			
toyota cressida			25.4	6	168.0	116
2900 12.6	81	Japan	1			
ford escort 2h			29.9	4	98.0	65
2380 20.7	81	America	1	4	125.0	0.4
plymouth reliant 2490 15.7	81	America	27.2 1	4	135.0	84
plymouth horizon 4	01	America	34.7	4	105.0	63
2215 14.9	81	America	1	-	100.0	
ford escort 4w			34.4	4	98.0	65
2045 16.2	81	America	1			
buick skylark			26.6	4	151.0	84
2635 16.4		America	1	_		
dodge aries wagon (s			25.8	4	156.0	92
2620 14.4	81	America	1	4	125.0	0.4
plymouth reliant 2385 12.9	81	America	30.0 1	4	135.0	84
toyota starlet	01	America	39.1	4	79.0	58
1755 16.9	81	Japan	1	-	70.0	
chevrolet citation		•	23.5	6	173.0	110
2725 12.6	81	America	1			
honda civic 1300			35.1	4	81.0	60
1760 16.1	81	Japan	1			
subaru	- 4	_	32.3	4	97.0	67
2065 17.8	81	Japan	1	4	05.0	C.F.
datsun 210 mpg 1975 19.4	81	Japan	37.0 1	4	85.0	65
toyota tercel	01	Japan	37.7	4	89.0	62
2050 17.3	81	Japan	1	-	33.3	02
mazda glc 4			34.1	4	91.0	68
1985 16.0	81	Japan	1			
plymouth champ			39.0	4	86.0	64
1875 16.4			1			
chrysler lebaron med			26.0	4	156.0	92
2585 14.5	82	America	1	4	0.4.0	07
honda civic (auto)	99	Tonon	32.0	4	91.0	67
1965 15.7 datsun 310 gx	02	Japan	1 38.0	4	91.0	67
1995 16.2	82	Japan	30.0	4	31.0	01
buick century limite		Japan	25.0	6	181.0	110
J			-	•	•	-

2945 16.4 82	America	1			
oldsmobile cutlass ciera	(diesel)	38.0	6	262.0	85
3015 17.0 82	America	1			
ford granada l		22.0	6	232.0	112
2835 14.7 82	America	1			
dodge rampage		32.0	4	135.0	84
2295 11.6 82	America	1			
dodge charger 2.2		36.0	4	135.0	84
2370 13.0 82	America	1			
chevrolet camaro		27.0	4	151.0	90
2950 17.3 82	America	1			
ford mustang gl		27.0	4	140.0	86
2790 15.6 82	America	1			
vw pickup		44.0	4	97.0	52
2130 24.6 82	Europe	1			
honda civic		38.0	4	91.0	67
1965 15.0 82	Japan	1			
toyota celica gt		32.0	4	144.0	96
2665 13.9 82	Japan	1			
toyota corolla		34.0	4	108.0	70
2245 16.9 82	Japan	1			
ford ranger		28.0	4	120.0	79
2625 18.6 82	America	1			
nissan stanza xe		36.0	4	120.0	88
2160 14.5 82	Japan	1			
mercury lynx l		36.0	4	98.0	70
2125 17.3 82	America	1			
plymouth horizon miser		38.0	4	105.0	63
2125 14.7 82	America	1			
mazda glc custom		31.0	4	91.0	68
1970 17.6 82	Japan	1			
mazda glc custom l		37.0	4	91.0	68
2025 18.2 82	Japan	1			
volkswagen rabbit l		36.0	4	105.0	74
1980 15.3 82	Europe	1			
ford fairmont futura		24.0	4	140.0	92
2865 16.4 82	America	1			
pontiac phoenix		27.0	4	151.0	90
2735 18.0 82	America	1			
dodge aries se		29.0	4	135.0	84
2525 16.0 82	America	1			
pontiac j2000 se hatchba	ck	31.0	4	112.0	85
2575 16.2 82	America	1			
chevrolet cavalier 2-doo		34.0	4	112.0	88
2395 18.0 82	America	1			
chevrolet cavalier wagon		27.0	4	112.0	88
2640 18.6 82	America	1			

```
28.0
      chevrolet cavalier
                                                          4
                                                                    112.0
                                                                                    88
      2605
                    19.6
                           82 America
      honda accord
                                            36.0
                                                          4
                                                                    107.0
                                                                                    75
      2205
                    14.5
                           82
                                 Japan
      chevy s-10
                                            31.0
                                                          4
                                                                     119.0
                                                                                    82
      2720
                    19.4
                           82 America
                                               1
[51]: Auto_log["log_displacement"] = np.log(Auto_log["displacement"])
      Auto_log["log_horsepower"] = np.log(Auto_log["horsepower"])
      Auto_log["log_weight"] = np.log(Auto_log["weight"])
      Auto_log = Auto_log.drop(
          columns=[
              "displacement",
              "weight",
              "horsepower",
              "year",
          ]
      Auto_log.columns
[51]: Index(['mpg', 'cylinders', 'acceleration', 'origin', 'oilshock',
      'log_displacement', 'log_horsepower', 'log_weight'], dtype='object')
[52]: Auto_log.corr(numeric_only=True)
[52]:
                                  cylinders acceleration oilshock log_displacement
                             mpg
      log_horsepower log_weight
                        1.000000 -0.777618
                                                 0.423329
                                                           0.521192
                                                                            -0.828453
     mpg
      -0.817517
                  -0.844194
      cylinders
                       -0.777618
                                   1.000000
                                                -0.504683 -0.273703
                                                                             0.942814
      0.843204
                  0.884303
      acceleration
                        0.423329
                                  -0.504683
                                                 1.000000 0.195892
                                                                            -0.497107
      -0.698328
                 -0.401563
      oilshock
                        0.521192 -0.273703
                                                 0.195892 1.000000
                                                                            -0.268161
      -0.299037
                  -0.250520
                                                -0.497107 -0.268161
                                                                              1.000000
      log_displacement -0.828453
                                   0.942814
      0.872149
                  0.942850
      log_horsepower
                       -0.817517
                                   0.843204
                                                -0.698328 -0.299037
                                                                             0.872149
      1.000000
                  0.873956
                       -0.844194
                                   0.884303
                                                -0.401563 -0.250520
                                                                              0.942850
      log_weight
      0.873956
                  1,000000
[53]: Auto_log = pd.get_dummies(
          Auto_log, columns=list(["origin"]), drop_first=True, dtype=np.uint8
      Auto_log.columns
```

```
[53]: Index(['mpg', 'cylinders', 'acceleration', 'oilshock', 'log_displacement',
      'log_horsepower', 'log_weight', 'origin_Europe', 'origin_Japan'],
      dtype='object')
[54]: cols = list(Auto_log.columns)
      cols.remove("mpg")
[55]: vifdf = calculate_VIFs("mpg ~ " + " + ".join(cols), Auto_log)
      vifdf
[55]:
                              VIF
      Feature
      cylinders
                         9.828626
      acceleration
                         3.304749
      oilshock
                         1.147770
      log_displacement
                        25.969595
      log_horsepower
                        11.414709
      log_weight
                        16.146573
      origin_Europe
                         1.876698
      origin_Japan
                         2.097688
[56]: identify_highest_VIF_feature(vifdf)
     We find the highest VIF in this model is log_displacement with a VIF of
     25.96959512578754
     Hence, we drop log_displacement from the model to be fitted.
[56]: ('log_displacement', 25.96959512578754)
[57]: cols.remove("log_displacement")
      vifdf = calculate_VIFs("mpg ~ " + " + ".join(cols), Auto_log)
      vifdf
[57]:
                            VIF
      Feature
      cylinders
                       5.535070
      acceleration
                       3.179336
      oilshock
                       1.142791
      log_horsepower 11.411764
      log_weight
                      10.608718
      origin_Europe
                       1.451961
      origin_Japan
                       1.652749
[58]: identify_highest_VIF_feature(vifdf)
```

We find the highest VIF in this model is log_horsepower with a VIF of 11.411764499222897

Hence, we drop log_horsepower from the model to be fitted.

```
[58]: ('log_horsepower', 11.411764499222897)
[59]: cols.remove("log_horsepower")
     vifdf = calculate_VIFs("mpg ~ " + " + ".join(cols), Auto_log)
     vifdf
[59]:
                     VIF
    Feature
    cylinders
                 5.517868
    acceleration
                 1.377517
     oilshock
                 1.118666
    log_weight
                 5.014899
     origin_Europe 1.451265
    origin_Japan
                 1.608682
[60]: identify_highest_VIF_feature(vifdf)
    No variables are significantly collinear.
[61]: formula = " + ".join(cols)
     results = perform_analysis("mpg", formula, Auto_log)
                           OLS Regression Results
    ______
    Dep. Variable:
                                     R-squared:
                                                                 0.823
                                mpg
    Model:
                                OLS Adj. R-squared:
                                                                 0.821
    Method:
                        Least Squares F-statistic:
                                                                 299.3
    Date:
                    Fri, 21 Feb 2025 Prob (F-statistic):
                                                            1.36e-141
    Time:
                            15:48:47 Log-Likelihood:
                                                               -1021.3
    No. Observations:
                                    AIC:
                                                                 2057.
                                392
    Df Residuals:
                                385
                                    BIC:
                                                                 2084.
    Df Model:
                                  6
    Covariance Type:
                           nonrobust
    ______
                           std err
                                      t P>|t|
                                                         [0.025
                     coef
    0.975]
    Intercept
                 168.2416
                             9.696
                                     17.351
                                               0.000
                                                        149.177
    187.306
                             0.230
                                     0.056
                                               0.955
    cylinders
                   0.0129
                                                        -0.440
    0.465
    acceleration 0.1805
                             0.071
                                      2.538
                                               0.012
                                                         0.041
    0.320
    oilshock
                             0.355
                                     14.470
                                               0.000
                                                         4.434
                  5.1312
    5.828
    log_weight
                 -18.9156
                             1.331
                                    -14.211
                                               0.000
                                                        -21.533
    -16.299
```

origin_Europe 2.413	1.3692	0.531	2.578	0.010	0.325	
origin_Japan 2.598	1.5602	0.528	2.956	0.003	0.522	
Omnibus:		30.158	Durbin-W	atson:	1.25	- - 53
<pre>Prob(Omnibus):</pre>		0.000	Jarque-B	era (JB):	51.81	l1
Skew:		0.493	Prob(JB)	:	5.62e-1	12
Kurtosis:		4.484	Cond. No		1.08e+0)3
===========		========		========		==

Notes:

- [1] Standard Errors assume that the covariance matrix of the errors is correctly specified.
- [2] The condition number is large, 1.08e+03. This might indicate that there are strong multicollinearity or other numerical problems.

	df	sum_sq	mean_sq	F	PR(>F)
cylinders	1.0	14403.083079	14403.083079	1318.559127	2.133814e-126
acceleration	1.0	30.471304	30.471304	2.789557	9.569322e-02
oilshock	1.0	2422.051542	2422.051542	221.731566	6.308582e-40
log_weight	1.0	2639.878573	2639.878573	241.672978	1.215817e-42
origin_Europe	1.0	22.569818	22.569818	2.066199	1.514086e-01
origin_Japan	1.0	95.449335	95.449335	8.738101	3.308129e-03
Residual	385.0	4205.489819	10.923350	NaN	NaN

[61]: <statsmodels.regression.linear_model.RegressionResultsWrapper at 0x7cb746f5e8d0>

[62]: identify_least_significant_feature(results, alpha=LOS_Alpha)

We find the least significant variable in this model is cylinders with a p-value of 0.9554479997905154 and a coefficient of 0.012867295201156259 Using the backward methodology, we suggest dropping cylinders from the new model

```
[63]: cols.remove("cylinders")
formula = " + ".join(cols)
results = perform_analysis("mpg", formula, Auto_log)
```

OLS Regression Results

			=========
Dep. Variable:	mpg	R-squared:	0.823
Model:	OLS	Adj. R-squared:	0.821
Method:	Least Squares	F-statistic:	360.0
Date:	Fri, 21 Feb 2025	Prob (F-statistic):	6.83e-143
Time:	15:48:47	Log-Likelihood:	-1021.3
No. Observations:	392	AIC:	2055.
Df Residuals:	386	BIC:	2078.
Df Model:	5		
Covariance Type:	nonrobust		

0.975]	coef	std err	t	P> t	[0.025	
- Intercept 181.696	167.8689	7.032	23.870	0.000	154.042	
acceleration 0.311	0.1792	0.067	2.673	0.008	0.047	
oilshock 5.821	5.1289	0.352	14.574	0.000	4.437	
log_weight -17.245	-18.8571	0.820	-22.992	0.000	-20.470	
origin_Europe 2.381	1.3628	0.518	2.631	0.009	0.344	
origin_Japan 2.590	1.5576	0.525	2.967	0.003	0.525	
Omnibus: Prob(Omnibus): Skew:		30.308 0.000 0.493 4.492	Prob(JB)	era (JB): :	4	1.253 52.282 4.44e-12 751.
Kurtosis:		4.492	Cond. No	•		731.

Notes:

[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.

	df	sum_sq	mean_sq	F	PR(>F)
acceleration	1.0	4268.531557	4268.531557	391.783092	1.076057e-60
oilshock	1.0	4757.627552	4757.627552	436.674301	2.076230e-65
log_weight	1.0	10466.602734	10466.602734	960.667136	8.818191e-107
origin_Europe	1.0	24.823504	24.823504	2.278402	1.320051e-01
origin_Japan	1.0	95.884166	95.884166	8.800637	3.198611e-03
Residual	386.0	4205.523957	10.895140	NaN	NaN

[63]: <statsmodels.regression.linear_model.RegressionResultsWrapper at 0x7cb746ff62a0>

[64]: | identify_least_significant_feature(results, alpha=LOS_Alpha)

No variables are statistically insignificant.

The model mpg ~ acceleration + oilshock + log_weight + origin_Europe + origin_Japan cannot be pruned further.

```
}
[66]: pd.DataFrame(models)
[66]:
                           name
     model R-squared adjusted
                   simple_model
     mpg ~ horsepower + weight + oilshock + origin_Europe + origin_Japan
      0.802074
          numeric_interactions
     mpg ~ horsepower + weight + oilshock + origin_Europe + origin_Japan +
     horsepower: weight
                                    0.844846
          oilshock_interactions
                                                mpg ~ horsepower + weight + oilshock +
      origin_Europe + origin_Japan + horsepower: weight + oilshock: weight + oilshock:
     horsepower
                            0.858491
            origin_interactions mpg ~ horsepower + weight + oilshock + origin_Europe
      + origin_Japan + oilshock: horsepower + origin_Europe: horsepower +
      origin_Japan: horsepower
                                          0.849008
      4 squared transformation
     horsepower + weight + oilshock + origin_Europe + origin_Japan + I(horsepower**2)
      + I(weight**2)
                                0.845550
             log_transformation
     mpg ~ acceleration + oilshock + log weight + origin Europe + origin Japan
      0.821151
```

1.9 Square Root Transformed Model

12.5

70

Europe

2234

[67]: Auto_sqrt = Auto.copy(deep=True) mpg cylinders displacement horsepower [67]: origin oilshock weight acceleration year name chevrolet chevelle malibu 18.0 8 307.0 130 3504 12.0 70 America 0 hi 1200d 9.0 8 304.0 193 4732 18.5 70 America 0 dodge d200 11.0 8 318.0 210 4382 13.5 70 America 0 chevy c20 10.0 8 307.0 200 4376 15.0 70 America 0 ford f250 10.0 8 360.0 215 4615 14.0 70 America 0 199.0 amc gremlin 21.0 6 90 2648 15.0 70 America 0 bmw 2002 26.0 4 121.0 113

0

saab 99e			25.0	4	104.0	95
2375 17.5	70	Europe	0			
audi 100 ls			24.0	4	107.0	90
2430 14.5	70	Europe	0		0.77	10
volkswagen 1131 del			26.0	4	97.0	46
1835 20.5	70	Europe	0	4	07.0	00
datsun pl510 2130 14.5	70	Tonon	27.0 0	4	97.0	88
2130 14.5 ford maverick	70	Japan	21.0	6	200.0	85
2587 16.0	70	America	0	0	200.0	00
amc hornet	, ,	mioi i ca	18.0	6	199.0	97
2774 15.5	70	America	0	-		
plymouth duster			22.0	6	198.0	95
2833 15.5	70	America	0			
peugeot 504			25.0	4	110.0	87
2672 17.5	70	Europe	0			
buick estate wagon	(wa)		14.0	8	455.0	225
3086 10.0	70	America	0			
toyota corona mark			24.0	4	113.0	95
2372 15.0	70	Japan	0			
plymouth satellite			18.0	8	318.0	150
3436 11.0	70	America	0	•	004.0	450
amc rebel sst	70		16.0	8	304.0	150
3433 12.0	70	America	0	8	200.0	1.10
ford torino 3449 10.5	70	America	17.0 0	0	302.0	140
ford galaxie 500	70	America	15.0	8	429.0	198
4341 10.0	70	America	0	O	423.0	130
chevrolet impala	, ,	mioi i ca	14.0	8	454.0	220
4354 9.0	70	America	0	G	101.0	220
buick skylark 320			15.0	8	350.0	165
3693 11.5	70	America	0			
pontiac catalina			14.0	8	455.0	225
4425 10.0	70	America	0			
amc ambassador dpl			15.0	8	390.0	190
3850 8.5	70	America	0			
dodge challenger se			15.0	8	383.0	170
3563 10.0	70	America	0			
plymouth 'cuda 340			14.0	8	340.0	160
3609 8.0	70	America	0	0	400.0	450
chevrolet monte car 3761 9.5		Amonios	15.0 0	8	400.0	150
plymouth fury iii	70	America	14.0	8	440.0	215
4312 8.5	70	America	0	0	110.0	210
amc hornet sportabo			18.0	6	258.0	110
2962 13.5	71	America	0	-		
chevrolet vega (sw)			22.0	4	140.0	72
5						

2408 19.0	71	Amorica	0			
pontiac firebird	11	America	19.0	6	250.0	100
3282 15.0	71	America	0	0	250.0	100
ford mustang	1 1	America	18.0	6	250.0	88
3139 14.5	71	America	0	· ·	200.0	00
mercury capri 2000	'-	1111101 100	23.0	4	122.0	86
2220 14.0	71	America	0	1	122.0	00
toyota corolla 1200		America	31.0	4	71.0	65
1773 19.0	71	Japan	0	-	71.0	00
peugeot 304	' -	Japan	30.0	4	79.0	70
2074 19.5	71	Europe	0	1	73.0	10
datsun 1200	' -	Багоро	35.0	4	72.0	69
1613 18.0	71	Japan	0	1	12.0	03
volkswagen model 11		oupun	27.0	4	97.0	60
1834 19.0	71	Europe	0	1	37.0	00
plymouth cricket	' -	Багорс	26.0	4	91.0	70
1955 20.5	71	America	0	-	31.0	10
pontiac safari (sw)		America	13.0	8	400.0	175
5140 12.0	71	America	0	O	400.0	110
opel 1900	11	America	28.0	4	116.0	90
2123 14.0	71	Europe	0	-	110.0	50
ford country squire		Багорс	13.0	8	400.0	170
4746 12.0	71	America	0	0	400.0	170
fiat 124b	1 1	America	30.0	4	88.0	76
2065 14.5	71	Europe	0	4	00.0	70
plymouth fury iii	11	Lurope	14.0	8	318.0	150
4096 13.0	71	America	0	O	310.0	100
dodge monaco (sw)	11	America	12.0	8	383.0	180
4955 11.5	71	America	0	O	303.0	100
chevrolet vega 2300		America	28.0	4	140.0	90
2264 15.5	71	America	0	-	140.0	50
toyota corona	11	America	25.0	4	113.0	95
2228 14.0	71	Japan	0	-	110.0	30
amc gremlin	11	Japan	19.0	6	232.0	100
	71	America	0	Ü	202.0	100
plymouth satellite			16.0	6	225.0	105
3439 15.5		 America	0	Ü	220.0	100
datsun pl510	11	America	27.0	4	97.0	88
2130 14.5	71	Japan	0	-	37.0	00
ford torino 500	' -	Japan	19.0	6	250.0	88
3302 15.5	71	America	0	Ü	200.0	00
amc matador	' -	America	18.0	6	232.0	100
3288 15.5	71	America	0	Ü	202.0	100
chevrolet impala	1 1	ımorroa	14.0	8	350.0	165
4209 12.0	71	America	0	Ü	550.0	100
pontiac catalina br			14.0	8	400.0	175
4464 11.5	71		0	3	400.0	170
11.0	1 1	Amorica	U			

ford galaxie 500	14.0	8	351.0	153
4154 13.5 71 America chevrolet chevelle malibu	0 17.0	6	250.0	100
3329 15.5 71 America chevrolet chevelle concours (sw)	0 13.0	8	307.0	130
4098 14.0 72 America	0			
plymouth satellite custom (sw) 4077 14.0 72 America	14.0 0	8	318.0	150
volvo 145e (sw)	18.0	4	121.0	112
2933 14.5 72 Europe volkswagen 411 (sw)	0 22.0	4	121.0	76
2511 18.0 72 Europe peugeot 504 (sw)	0 21.0	4	120.0	87
2979 19.5 72 Europe	0	Ŧ	120.0	01
ford pinto (sw) 2395 16.0 72 America	22.0 0	4	122.0	86
datsun 510 (sw)	28.0	4	97.0	92
2288 17.0 72 Japan toyouta corona mark ii (sw)	0 23.0	4	120.0	97
2506 14.5 72 Japan	0	4	00.0	00
dodge colt (sw) 2164 15.0 72 America	28.0 0	4	98.0	80
amc matador (sw) 3892 12.5 72 America	15.0 0	8	304.0	150
toyota corolla 1600 (sw)	27.0	4	97.0	88
2100 16.5 72 Japan renault 12 (sw)	0 26.0	4	96.0	69
2189 18.0 72 Europe	0			
mazda rx2 coupe 2330 13.5 72 Japan	19.0 0	3	70.0	97
ford gran torino (sw)	13.0	8	302.0	140
4294 16.0 72 America oldsmobile delta 88 royale	0 12.0	8	350.0	160
4456 13.5 72 America	0	0	400.0	100
chrysler newport royal 4422 12.5 72 America	13.0 0	8	400.0	190
toyota corona hardtop 2278 15.5 72 Japan	24.0 0	4	113.0	95
volkswagen type 3	23.0	4	97.0	54
2254 23.5 72 Europe chevrolet vega	0 20.0	4	140.0	90
2408 19.5 72 America	0			
ford pinto runabout 2226 16.5 72 America	21.0 0	4	122.0	86
chevrolet impala	13.0	8	350.0	165
4274 12.0 72 America dodge colt hardtop	0 25.0	4	97.5	80

0100	17.0	70	A	0			
2126	17.0	72	America	0	0	210.0	150
plymouth fury		70	A	15.0	8	318.0	150
4135	13.5	72	America	0	0	254 0	150
ford galaxie 5		70	A	14.0	8	351.0	153
4129	13.0	72	America	0	0	204.0	450
amc ambassador		70		17.0	8	304.0	150
3672	11.5	72	America	0	0	400.0	000
mercury marqui		70	A	11.0	8	429.0	208
4633	11.0	72	America	0	0	250.0	455
buick lesabre			A	13.0	8	350.0	155
4502	13.5	72	America	0	0	400.0	175
pontiac catali		70	A	14.0 0	8	400.0	175
4385	12.0	72	America	-	4	00.0	00
fiat 124 sport	_		P	26.0	4	98.0	90
2265	15.5	73	Europe	0	C	020.0	100
amc gremlin	1	70	A	18.0	6	232.0	100
2789	15.0	73	America	0	4	07.0	00
toyota carina	40.0	70	.	20.0	4	97.0	88
2279	19.0	73	Japan	0	4	4.40.0	70
chevrolet vega		70		21.0	4	140.0	72
2401	19.5	73	America	0	4	400.0	0.4
datsun 610	40 5	70	-	22.0	4	108.0	94
2379	16.5	73	Japan	0		70.0	00
maxda rx3	40 5	70	.	18.0	3	70.0	90
2124	13.5	73	Japan	0	4	400.0	0.5
ford pinto	40 5	70		19.0	4	122.0	85
2310	18.5	73	America	0	0	455.0	407
mercury capri		70		21.0	6	155.0	107
2472	14.0	73	America	0	0	250.0	1.4.5
chevrolet mont			A	15.0	8	350.0	145
4082	13.0	73	America	0	4	101.0	110
saab 991e	11 0	70	P	24.0	4	121.0	110
2660	14.0	73	Europe	0	4	60.0	40
fiat 128	10 F	70	P	29.0	4	68.0	49
1867	19.5	73	Europe	0	4	110.0	7.5
opel manta	45 5	70	П	24.0	4	116.0	75
2158	15.5	73	Europe	0	4	111 0	0.1
audi 1001s	11 0	70	P	20.0	4	114.0	91
2582	14.0	73	Europe	0	4	101.0	110
volvo 144ea	4 F F	70	P	19.0	4	121.0	112
2868	15.5	73	Europe	0	0	210.0	150
dodge dart cus		70	A	15.0	8	318.0	150
3399	11.0	73	America	0	C	150.0	100
toyota mark ii		70	T	20.0	6	156.0	122
2807	13.5	73	Japan	0	0	250.0	100
oldsmobile ome	_	70	۸	11.0	8	350.0	180
3664	11.0	73	America	0			

oldsmobile vista cruiser	12.0	8	350.0	180
4499 12.5 73 Am	erica 0			
pontiac grand prix	16.0	8	400.0	230
4278 9.5 73 Am	erica 0			
plymouth custom suburb	13.0	8	360.0	170
	erica 0			
amc hornet	18.0	6	232.0	100
	erica 0	_		
chevrolet impala	11.0	8	400.0	150
	erica 0		252.2	486
buick century 350	13.0	8	350.0	175
4100 13.0 73 Am		0	204.0	450
amc matador	14.0	8	304.0	150
	erica 0	0	250.0	4.4.5
chevrolet malibu	13.0	8	350.0	145
	erica 0	0	210 0	150
dodge coronet custom 3777 12.5 73 Am	15.0 erica 0	8	318.0	150
	erica 0 12.0	8	429.0	198
mercury marquis brougham 4952 11.5 73 Am		0	429.0	190
chevrolet caprice classic	13.0	8	400.0	150
4464 12.0 73 Am		O	400.0	130
ford ltd	13.0	8	351.0	158
	erica 0	O	331.0	130
plymouth fury gran sedan	14.0	8	318.0	150
	erica 0	Ŭ	010.0	100
ford gran torino	14.0	8	302.0	137
	erica 0	•	002.0	
buick electra 225 custom	12.0	8	455.0	225
	erica 0	-		
amc ambassador brougham	13.0	8	360.0	175
G	erica 0			
plymouth valiant	18.0	6	225.0	105
	erica 0			
chevrolet nova custom	16.0	6	250.0	100
3278 18.0 73 Am	erica 0			
ford country	12.0	8	400.0	167
4906 12.5 73 Am	erica 0			
ford maverick	18.0	6	250.0	88
3021 16.5 73 Am	erica 0			
plymouth duster	23.0	6	198.0	95
2904 16.0 73 Am	erica 0			
volkswagen super beetle	26.0	4	97.0	46
1950 21.0 73 E	urope 0			
chrysler new yorker brougha		8	440.0	215
4735 11.0 73 Am				
audi fox	29.0	4	98.0	83

2219	16.5	74	Europe	0			
volkswagen das		1.1	паторс	26.0	4	79.0	67
1963	15.5	74	Europe	0	-	, , , ,	01
opel manta			r	26.0	4	97.0	78
2300	14.5	74	Europe	0			
toyota corona			•	31.0	4	76.0	52
1649	16.5	74	Japan	0			
datsun 710			-	32.0	4	83.0	61
2003	19.0	74	Japan	0			
subaru			_	26.0	4	108.0	93
2391	15.5	74	Japan	0			
fiat 128			_	24.0	4	90.0	75
2108	15.5	74	Europe	0			
fiat 124 tc			_	26.0	4	116.0	75
2246	14.0	74	Europe	0			
honda civic			_	24.0	4	120.0	97
2489	15.0	74	Japan	0			
fiat x1.9			-	31.0	4	79.0	67
2000	16.0	74	Europe	0			
amc matador (sw)		-	14.0	8	304.0	150
4257	15.5	74	America	0			
dodge colt				28.0	4	90.0	75
2125	14.5	74	America	0			
ford gran tor:	ino (sw	·)		14.0	8	302.0	140
4638	16.0	74	America	0			
ford gran tor:	ino			16.0	8	302.0	140
4141	14.0	74	America	0			
buick century	luxus	(wa)		13.0	8	350.0	150
4699	14.5	74	America	0			
dodge coronet	custom	(sw)	14.0	8	318.0	150
4457	13.5		America	0			
plymouth dust	er			20.0	6	198.0	95
3102		74	America	0			
amc hornet				19.0	6	232.0	100
2901	16.0	74	America	0			
chevrolet nova	a			15.0	6	250.0	100
3336	17.0	74	America	0			
ford pinto				26.0	4	122.0	80
2451	16.5	74	America	0			
datsun b210				31.0	4	79.0	67
1950	19.0	74	Japan	0			
chevrolet veg	a		-	25.0	4	140.0	75
2542	17.0	74	America	0			
chevrolet che	velle m	alib	u classic	16.0	6	250.0	100
3781	17.0	74	America	0			
amc matador				16.0	6	258.0	110
3632	18.0	74	America	0			

plymouth satelli	te sebri	nσ	18.0	6	225.0	105
3613 16		America	0	J	220.0	100
toyota corolla 1			32.0	4	71.0	65
1836 21		Japan	0			
datsun 710			24.0	4	119.0	97
2545 17	.0 75	Japan	0			
pontiac astro			23.0	4	140.0	78
2592 18	.5 75	America	0			
amc gremlin	0 75		20.0	6	232.0	100
2914 16	.0 75	America	0	4	124 0	06
toyota corona 2702 13	.5 75	Ianan	24.0 0	4	134.0	96
volkswagen dashe		Japan	25.0	4	90.0	71
2223 16		Europe	0	-	30.0	71
ford pinto	.0 10	Багоро	18.0	6	171.0	97
2984 14	.5 75	America	0	•	2.2.0	0.
saab 991e			25.0	4	121.0	115
2671 13	.5 75	Europe	0			
amc pacer		_	19.0	6	232.0	90
3211 17	.0 75	America	0			
audi 1001s			23.0	4	115.0	95
2694 15	.0 75	Europe	0			
peugeot 504			23.0	4	120.0	88
2957 17	.0 75	Europe	0			
volvo 244dl			22.0	4	121.0	98
2945 14	.5 75	Europe	0	_		
honda civic cvcc		-	33.0	4	91.0	53
1795 17	.5 75	Japan	0	4	440.0	00
ford pinto	0 75	A	23.0	4	140.0	83
2639 17		America	0 29.0	4	90.0	70
volkswagen rabbi 1937 14		Europe	29.0	4	90.0	70
toyota corolla	.0 73	Europe	29.0	4	97.0	75
2171 16	.0 75	Japan	0	1	31.0	70
plymouth valiant		0 ap a	19.0	6	225.0	95
	.0 75	America	0			
chevrolet monza			20.0	8	262.0	110
3221 13	.5 75	America	0			
ford mustang ii			13.0	8	302.0	129
3169 12	.0 75	America	0			
chevrolet nova			18.0	6	250.0	105
3459 16	.0 75	America	0			
mercury monarch			15.0	6	250.0	72
3432 21		America	0			
pontiac catalina			16.0	8	400.0	170
4668 11		America	0	^	252.2	4.45
chevrolet bel ai:	r		15.0	8	350.0	145

4440	14.0	75	Amorica	0			
plymouth grand		13	America	16.0	8	318.0	150
	•	75	America	0	0	310.0	130
ford maverick	14.0	10	America	15.0	6	250.0	72
	19.5	75	America	0	O	200.0	12
buick century	10.0	10	America	17.0	6	231.0	110
•	21.0	75	America	0	O	201.0	110
chevroelt chev				16.0	6	250.0	105
		75	a America	0	O	250.0	103
amc matador	10.0	10	America	15.0	6	258.0	110
	19.0	75	America	0	O	200.0	110
plymouth fury	13.0	10	America	18.0	6	225.0	95
	19.0	75	America	0	O	220.0	33
buick skyhawk	13.0	10	America	21.0	6	231.0	110
•	15.0	75	America	0	O	231.0	110
ford ltd	15.0	13	America	14.0	8	351.0	148
	13.5	75	America	0	0	331.0	140
	13.5	15	America	26.5	4	140.0	70
ford pinto 2565	12 6	76	Amonico	20.5	4	140.0	72
		10	America		6	050.0	110
pontiac ventur	_	76	A	18.5	6	250.0	110
	16.2	76	America	0		050.0	٥٦
amc pacer d/l	17.0	70		17.5	6	258.0	95
		76	America	0	4	07.0	7.1
volkswagen rab		7.0	П	29.5	4	97.0	71
	12.2	76	Europe	0	4	05.0	70
datsun b-210			-	32.0	4	85.0	70
		76	Japan	0			
toyota corolla			-	28.0	4	97.0	75
	16.4	76	Japan	0	_		
volvo 245			_	20.0	4	130.0	102
	15.7	76	Europe	0	_		
ford f108				13.0	8	302.0	130
	15.0	76	America	0			
peugeot 504				19.0	4	120.0	88
		76	Europe	0			
toyota mark ii				19.0	6	156.0	108
		76	Japan	0			
mercedes-benz				16.5	6	168.0	120
		76	Europe	0			
cadillac sevil	le			16.5	8	350.0	180
4380	12.1	76	America	0			
chevy c10				13.0	8	350.0	145
	12.0	76	America	0			
dodge d100				13.0	8	318.0	150
		76	America	0			
ford granada g				18.0	6	250.0	78
3574	21.0	76	America	0			

plymouth volare premier	v8	13.0	8	318.0	150
3940 13.2 76	America	0			
dodge aspen se		20.0	6	225.0	100
3651 17.7 76	America	0			
vw rabbit		29.0	4	90.0	70
1937 14.2 76	Europe	0			
opel 1900	_	25.0	4	116.0	81
2220 16.9 76	Europe	0	4	04.0	F.0
honda civic	7	33.0	4	91.0	53
1795 17.4 76	Japan	0	4	107.0	0.0
fiat 131	F	28.0	4	107.0	86
2464 15.5 76	Europe	0	4	140.0	00
capri ii	A	25.0 0	4	140.0	92
2572 14.9 76	America	26.0	4	00.0	70
dodge colt 2255 17.7 76	America	26.0	4	98.0	79
renault 12tl	America	27.0	4	101.0	83
2202 15.3 76	Europe	0	4	101.0	03
dodge coronet brougham	Lurope	16.0	8	318.0	150
4190 13.0 76	America	0	O	310.0	100
amc matador	America	15.5	8	304.0	120
3962 13.9 76	America	0	Ü	001.0	120
chevrolet chevelle malik		17.5	8	305.0	140
4215 13.0 76	America	0	· ·	333.3	110
plymouth valiant		22.0	6	225.0	100
3233 15.4 76	America	0			
chevrolet nova		22.0	6	250.0	105
3353 14.5 76	America	0			
ford maverick		24.0	6	200.0	81
3012 17.6 76	America	0			
amc hornet		22.5	6	232.0	90
3085 17.6 76	America	0			
chevrolet chevette		29.0	4	85.0	52
2035 22.2 76	America	0			
chevrolet woody		24.5	4	98.0	60
2164 22.1 76	America	0			
ford gran torino		14.5	8	351.0	152
4215 12.8 76	America	0			
ford mustang ii 2+2		25.5	4	140.0	89
2755 15.8 77	America	1			
volkswagen rabbit custor	n	29.0	4	97.0	78
1940 14.5 77	Europe	1			
pontiac sunbird coupe		24.5	4	151.0	88
2740 16.0 77	America	1			
toyota corolla liftback		26.0	4	97.0	75
2265 18.2 77	Japan	1			
chevrolet chevette		30.5	4	98.0	63

2051	17.0	77	America	1			
bmw 320i	11.0	' '	America	21.5	4	121.0	110
2600	12.8	77	Europe	1	_	121.0	110
subaru dl				30.0	4	97.0	67
1985	16.4	77	Japan	1			
volkswagen d			•	30.5	4	97.0	78
2190	14.1	77	Europe	1			
datsun 810			•	22.0	6	146.0	97
2815	14.5	77	Japan	1			
mazda rx-4			_	21.5	3	80.0	110
2720	13.5	77	Japan	1			
ford thunder	bird		_	16.0	8	351.0	149
4335	14.5	77	America	1			
dodge colt m	/m			33.5	4	98.0	83
2075	15.9	77	America	1			
chrysler cor	doba			15.5	8	400.0	190
4325	12.2	77	America	1			
chevrolet mo	nte car	lo la	ndau	15.5	8	350.0	170
4165	11.4	77	America	1			
plymouth arr	ow gs			25.5	4	122.0	96
2300	15.5	77	America	1			
buick opel i	suzu de	luxe		30.0	4	111.0	80
2155	14.8		America	1			
renault 5 gt	1			36.0	4	79.0	58
1825	18.6	77	Europe	1			
datsun f-10	hatchba	ck	-	33.5	4	85.0	70
1945	16.8	77	Japan	1			
pontiac gran	d prix	1j	-	16.0	8	400.0	180
4220	11.1		America	1			
oldsmobile c	utlass	supre	me	17.0	8	260.0	110
4060		_	America	1			
chevrolet ca	price c	lassi	.c	17.5	8	305.0	145
3880	-			1			
mercury coug	ar brou	gham		15.0	8	302.0	130
4295	14.9	77	America	1			
chevrolet co	ncours			17.5	6	250.0	110
3520	16.4	77	America	1			
buick skylar	k			20.5	6	231.0	105
3425	16.9	77	America	1			
plymouth vol	are cus	tom		19.0	6	225.0	100
3630	17.7	77	America	1			
ford granada				18.5	6	250.0	98
3525	19.0	77	America	1			
dodge monaco	brough	am		15.5	8	318.0	145
4140	13.7	77	America	1			
honda accord	cvcc			31.5	4	98.0	68
2045	18.5	77	Japan	1			
			-				

dotann E10	07.0	4	110 0	07
datsun 510	27.2	4	119.0	97
2300 14.7 78 Japan		4	124 0	O.F.
toyota corona	27.5	4	134.0	95
2560 14.2 78 Japan				20
chevrolet chevette	30.0	4	98.0	68
2155 16.5 78 America				
buick regal sport coupe (turbo)	17.7	6	231.0	165
3445 13.4 78 America	. 1			
ford futura	18.1	8	302.0	139
3205 11.2 78 America	. 1			
dodge omni	30.9	4	105.0	75
2230 14.5 78 America	. 1			
dodge magnum xe	17.5	8	318.0	140
4080 13.7 78 America	. 1			
toyota celica gt liftback	21.1	4	134.0	95
2515 14.8 78 Japan				
peugeot 604sl	16.2	6	163.0	133
3410 15.8 78 Europe		Ü	100.0	100
oldsmobile starfire sx	23.8	4	151.0	85
		4	131.0	03
2855 17.6 78 America		4	110 0	07
datsun 200-sx	23.9	4	119.0	97
2405 14.9 78 Japan		_		
audi 5000	20.3	5	131.0	103
2830 15.9 78 Europe				
volvo 264gl	17.0	6	163.0	125
3140 13.6 78 Europe	. 1			
saab 99gle	21.6	4	121.0	115
2795 15.7 78 Europe	: 1			
volkswagen scirocco	31.5	4	89.0	71
1990 14.9 78 Europe	. 1			
honda accord lx	29.5	4	98.0	68
2135 16.6 78 Japan	. 1			
plymouth sapporo	23.2	4	156.0	105
2745 16.7 78 America				
chevrolet monte carlo landau	19.2	8	305.0	145
3425 13.2 78 America		· ·	000.0	110
	32.8	4	78.0	52
mazda glc deluxe 1985 19.4 78 Japan		4	10.0	52
•		C	005 0	110
dodge aspen	18.6	6	225.0	110
3620 18.7 78 America				
volkswagen rabbit custom diesel		4	90.0	48
1985 21.5 78 Europe				
ford fiesta	36.1	4	98.0	66
1800 14.4 78 America	. 1			
datsun b210 gx	39.4	4	85.0	70
2070 18.6 78 Japan	. 1			
honda civic cvcc	36.1	4	91.0	60

1800	16.4	78	Japan	1			
amc concord d/		70	Japan	18.1	6	258.0	120
		78	America	10.1	O	200.0	120
dodge diplomat		10	America	19.4	8	318.0	140
		78	America	1	· ·	010.0	110
mercury monarc			IMIOI 10a	20.2	8	302.0	139
	12.8	78	America	1	Ü	002.0	100
oldsmobile cut				19.9	8	260.0	110
	15.5	78	America	1	· ·	200.0	110
chevrolet mali			11111011100	20.5	6	200.0	95
	18.2	78	America	1	· ·	200.0	
ford fairmont		. •		20.2	6	200.0	85
	15.8	78	America	1			
ford fairmont				25.1	4	140.0	88
	15.4	78	America	1	-		
plymouth volar				20.5	6	225.0	100
		78	America	1			
amc concord				19.4	6	232.0	90
	17.2	78	America	1			
buick century				20.6	6	231.0	105
•	15.8	78	America	1			
mercury zephyr				20.8	6	200.0	85
	16.7	78	America	1			
pontiac phoeni				19.2	6	231.0	105
-	19.2	78	America	1			
plymouth horiz	on			34.2	4	105.0	70
	13.2	79	America	1			
mercedes benz	300d			25.4	5	183.0	77
3530	20.1	79	Europe	1			
cadillac eldor	ado		•	23.0	8	350.0	125
3900	17.4	79	America	1			
peugeot 504				27.2	4	141.0	71
3190	24.8	79	Europe	1			
oldsmobile cut	lass sa	alon	brougham	23.9	8	260.0	90
3420	22.2	79	America	1			
plymouth horiz	on tc3			34.5	4	105.0	70
2150	14.9	79	America	1			
amc spirit dl				27.4	4	121.0	80
2670	15.0	79	America	1			
fiat strada cu	stom			37.3	4	91.0	69
2130	14.7	79	Europe	1			
buick skylark	limited	i		28.4	4	151.0	90
2670	16.0	79	America	1			
chevrolet cita	tion			28.8	6	173.0	115
2595	11.3	79	America	1			
oldsmobile ome	ga brou	ıghaı	m	26.8	6	173.0	115
2700	12.9	79	America	1			

pontiac phoenix	33.5	4	151.0	90
2556 13.2 79 America datsun 210	1 31.8	4	85.0	65
2020 19.2 79 Japan	1	4	05.0	03
dodge colt hatchback custom	35.7	4	98.0	80
1915 14.4 79 America	1			
dodge st. regis	18.2	8	318.0	135
3830 15.2 79 America	1	4	00.0	7.4
vw rabbit custom 1925 14.0 79 Europe	31.9 1	4	89.0	71
mercury zephyr 6	19.8	6	200.0	85
2990 18.2 79 America	1			
ford fairmont 4	22.3	4	140.0	88
2890 17.3 79 America	1			
amc concord dl 6	20.2	6	232.0	90
3265 18.2 79 America	1 20.6	6	225 0	110
dodge aspen 6 3360 16.6 79 America	20.6	0	225.0	110
chevrolet caprice classic	17.0	8	305.0	130
3840 15.4 79 America	1			
ford 1td landau	17.6	8	302.0	129
3725 13.4 79 America	1			
pontiac lemans v6	21.5	6	231.0	115
3245 15.4 79 America	1			25
maxda glc deluxe 1975 15.2 79 Japan	34.1 1	4	86.0	65
1975 15.2 79 Japan buick estate wagon (sw)	16.9	8	350.0	155
4360 14.9 79 America	10.3	O	000.0	100
ford country squire (sw)	15.5	8	351.0	142
4054 14.3 79 America	1			
chevrolet malibu classic (sw)	19.2	8	267.0	125
3605 15.0 79 America	1			
chrysler lebaron town @ country (sw)	18.5	8	360.0	150
3940 13.0 79 America	16 5	0	251 0	120
mercury grand marquis 3955 13.2 79 America	16.5 1	8	351.0	138
vw rabbit c (diesel)	44.3	4	90.0	48
2085 21.7 80 Europe	1	-		
vw dasher (diesel)	43.4	4	90.0	48
2335 23.7 80 Europe	1			
audi 5000s (diesel)	36.4	5	121.0	67
2950 19.9 80 Europe	1	4	4.4.0.0	07
mercedes-benz 240d	30.0	4	146.0	67
3250 21.8 80 Europe honda civic 1500 gl	1 44.6	4	91.0	67
1850 13.8 80 Japan	1	-1	51.0	01
datsun 280-zx	32.7	6	168.0	132

2910	11.4	80	Japan	1			
vokswagen rabb				29.8	4	89.0	62
1845	15.3	80	Europe	1			
mazda rx-7 gs				23.7	3	70.0	100
2420	12.5	80	Japan	1			
triumph tr7 co	oupe			35.0	4	122.0	88
2500	15.1	80	Europe	1			
honda accord				32.4	4	107.0	72
2290	17.0	80	Japan	1			
datsun 210				40.8	4	85.0	65
2110	19.2	80	Japan	1			
subaru dl				33.8	4	97.0	67
2145	18.0	80	Japan	1			
dodge colt				27.9	4	156.0	105
2800	14.4	80	America	1			
mazda glc				46.6	4	86.0	65
2110	17.9	80	Japan	1			
toyota corolla	ì.		•	32.2	4	108.0	75
2265	15.2	80	Japan	1			
vw rabbit			•	41.5	4	98.0	76
2144	14.7	80	Europe	1			
toyota corolla			Ī	38.1	4	89.0	60
1968	18.8	80	Japan	1	_	00.10	
chevrolet chev			oupun	32.1	4	98.0	70
2120	15.5	80	America	1	-	00.0	, 0
chevrolet cita			1111101 100	28.0	4	151.0	90
2678	16.5	80	America	1	-	101.0	00
ford fairmont	10.0	00	America	26.4	4	140.0	88
2870	18.1	80	America	1	1	140.0	00
datsun 310	10.1	80	America	37.2	4	86.0	65
2019	16.4	80	Innon	1	4	80.0	05
	10.4	80	Japan	19.1	6	225.0	00
dodge aspen	10 7	90	Amonios		6	225.0	90
3381	18.7	80	America	1	4	07.0	70
audi 4000	15 0	00	F	34.3	4	97.0	78
2188	15.8		Europe	1	4	124 0	00
toyota corona			.	29.8	4	134.0	90
2711	15.5	80	Japan	1		400.0	
mazda 626			-	31.3	4	120.0	75
2542	17.5	80	Japan	1	_		
datsun 510 hat			_	37.0	4	119.0	92
2434	15.0	80	Japan	1	_		
amc concord				24.3	4	151.0	90
3003	20.1		America	1			
peugeot 505s t				28.1	4	141.0	80
3230	20.4	81	Europe	1			
honda prelude				33.7	4	107.0	75
2210	14.4	81	Japan	1			

toyota corolla		32.4	4	108.0	75
2350 16.8 81 datsun 200sx	Japan	1 32.9	4	119.0	100
2615 14.8 81	Japan	32.9 1	4	119.0	100
mazda 626	•	31.6	4	120.0	74
2635 18.3 81	Japan	1	2	4.45.0	7.0
volvo diesel 3160 19.6 81	Europe	30.7 1	6	145.0	76
chrysler lebaron salon	Zuropo	17.6	6	225.0	85
3465 16.6 81	America	1			
datsun 810 maxima 2930 13.8 81	Ionon	24.2 1	6	146.0	120
buick century	Japan	22.4	6	231.0	110
3415 15.8 81	America	1			
oldsmobile cutlass ls		26.6	8	350.0	105
3725 19.0 81 ford granada gl	America	1 20.2	6	200.0	88
3060 17.1 81	America	1	O	200.0	00
volkswagen jetta		33.0	4	105.0	74
2190 14.2 81	Europe	1	•	4.00	
toyota cressida 2900 12.6 81	Japan	25.4 1	6	168.0	116
ford escort 2h	Japan	29.9	4	98.0	65
2380 20.7 81	America	1			
plymouth reliant		27.2	4	135.0	84
2490 15.7 81 plymouth horizon 4	America	1 34.7	4	105.0	63
2215 14.9 81	America	1	1	100.0	00
ford escort 4w		34.4	4	98.0	65
2045 16.2 81	America	1		454.0	0.4
buick skylark 2635 16.4 81	America	26.6 1	4	151.0	84
dodge aries wagon (sw)	America	25.8	4	156.0	92
2620 14.4 81	America	1			
plymouth reliant		30.0	4	135.0	84
2385 12.9 81 toyota starlet	America	1 39.1	4	79.0	58
1755 16.9 81	Japan	1	1	10.0	00
chevrolet citation	•	23.5	6	173.0	110
2725 12.6 81	America	1	4	04.0	20
honda civic 1300 1760 16.1 81	Japan	35.1 1	4	81.0	60
subaru	Japan	32.3	4	97.0	67
2065 17.8 81	Japan	1			
datsun 210 mpg	-	37.0	4	85.0	65
1975 19.4 81 toyota tercel	Japan	1 37.7	4	89.0	62
coyota tercer		51.1	7	03.0	02

2050	17.3	81	Japan	1			
mazda glc 4			•	34.1	4	91.0	68
1985	16.0	81	Japan	1			
plymouth cham	-			39.0	4	86.0	64
1875	16.4		America	1			
chrysler leba				26.0	4	156.0	92
2585	14.5	82	America	1			
honda civic (00	-	32.0	4	91.0	67
1965	15.7	82	Japan	1	4	01.0	67
datsun 310 gx 1995	16.2	82	Japan	38.0 1	4	91.0	67
buick century			Japan	25.0	6	181.0	110
2945		82	America	1	O	101.0	110
oldsmobile cu				38.0	6	262.0	85
3015	17.0	82	America	1	Ü	202.0	00
ford granada		02	ımor roa	22.0	6	232.0	112
2835	_ 14.7	82	America	1			
dodge rampage				32.0	4	135.0	84
2295	11.6	82	America	1			
dodge charger	2.2			36.0	4	135.0	84
2370	13.0	82	America	1			
chevrolet cam	aro			27.0	4	151.0	90
2950	17.3	82	America	1			
ford mustang	_			27.0	4	140.0	86
2790	15.6	82	America	1			
vw pickup				44.0	4	97.0	52
2130	24.6	82	Europe	1			
honda civic	45.0		_	38.0	4	91.0	67
1965	15.0	82	Japan	1	4	444.0	0.0
toyota celica	_	00	T	32.0	4	144.0	96
2665	13.9	82	Japan	1 34.0	4	108.0	70
toyota corolla 2245	a 16.9	82	Japan	34.0	4	100.0	70
ford ranger	10.9	02	Japan	28.0	4	120.0	79
2625	18.6	82	America	1	-	120.0	13
nissan stanza		02	ımor roa	36.0	4	120.0	88
2160	14.5	82	Japan	1	_		
mercury lynx			•	36.0	4	98.0	70
2125	17.3	82	America	1			
plymouth horis	zon mis	er		38.0	4	105.0	63
2125	14.7	82	America	1			
mazda glc cus	tom			31.0	4	91.0	68
1970	17.6	82	Japan	1			
mazda glc cus				37.0	4	91.0	68
2025	18.2	82	Japan	1			
volkswagen ra			_	36.0	4	105.0	74
1980	15.3	82	Europe	1			

```
24.0
                                                                     140.0
      ford fairmont futura
                                                           4
                                                                                     92
      2865
                    16.4
                           82 America
                                                1
      pontiac phoenix
                                                                     151.0
                                             27.0
                                                           4
                                                                                     90
      2735
                    18.0
                               America
                                                1
      dodge aries se
                                             29.0
                                                           4
                                                                      135.0
                                                                                     84
      2525
                    16.0
                           82 America
                                                1
      pontiac j2000 se hatchback
                                             31.0
                                                           4
                                                                      112.0
                                                                                     85
      2575
                    16.2
                           82 America
                                                1
      chevrolet cavalier 2-door
                                             34.0
                                                           4
                                                                      112.0
                                                                                     88
      2395
                    18.0
                           82 America
                                                1
      chevrolet cavalier wagon
                                             27.0
                                                                      112.0
                                                                                     88
                                                           4
                    18.6
                           82
                               America
                                                1
      chevrolet cavalier
                                             28.0
                                                           4
                                                                      112.0
                                                                                     88
      2605
                    19.6
                           82 America
                                                1
      honda accord
                                             36.0
                                                           4
                                                                     107.0
                                                                                     75
      2205
                    14.5
                           82
                                  Japan
                                                1
      chevy s-10
                                             31.0
                                                           4
                                                                     119.0
                                                                                     82
      2720
                    19.4
                           82
                               America
                                                1
[68]: Auto_sqrt["sqrt_displacement"] = np.sqrt(Auto_sqrt["displacement"])
      Auto_sqrt["sqrt_horsepower"] = np.sqrt(Auto_sqrt["horsepower"])
      Auto sqrt["sqrt weight"] = np.sqrt(Auto sqrt["weight"])
      Auto_sqrt = Auto_sqrt.drop(
          columns=[
              "displacement",
              "weight",
              "horsepower",
              "year",
      )
      Auto_sqrt.columns
[68]: Index(['mpg', 'cylinders', 'acceleration', 'origin', 'oilshock',
      'sqrt displacement', 'sqrt horsepower', 'sqrt weight'], dtype='object')
[69]: Auto sqrt.corr(numeric only=True)
[69]:
                                   cylinders acceleration oilshock
                              mpg
      sqrt_displacement
                         sqrt_horsepower sqrt_weight
                         1.000000 -0.777618
                                                   0.423329
                                                             0.521192
      mpg
      -0.821331
                       -0.802311
                                     -0.840095
      cylinders
                        -0.777618
                                    1.000000
                                                  -0.504683 -0.273703
      0.953208
                       0.849266
                                    0.893465
      acceleration
                         0.423329 -0.504683
                                                   1.000000 0.195892
      -0.521812
                       -0.696702
                                    -0.409829
      oilshock
                         0.521192 -0.273703
                                                   0.195892 1.000000
      -0.284587
                       -0.306247
                                    -0.260664
```

```
-0.521812 -0.284587
      sqrt_displacement -0.821331
                                    0.953208
      1.000000
                       0.886470
                                    0.939868
      sqrt_horsepower -0.802311
                                    0.849266
                                                 -0.696702 -0.306247
      0.886470
                       1.000000
                                    0.872045
      sqrt_weight
                       -0.840095
                                    0.893465
                                                 -0.409829 -0.260664
      0.939868
                                    1.000000
                       0.872045
[70]: Auto_sqrt = pd.get_dummies(
          Auto_sqrt, columns=list(["origin"]), drop_first=True, dtype=np.uint8
      Auto_sqrt.columns
[70]: Index(['mpg', 'cylinders', 'acceleration', 'oilshock', 'sqrt_displacement',
      'sqrt_horsepower', 'sqrt_weight', 'origin_Europe', 'origin_Japan'],
      dtype='object')
[71]: cols = list(Auto_sqrt.columns)
      cols.remove("mpg")
[72]: vifdf = calculate_VIFs("mpg ~ " + " + ".join(cols), Auto_sqrt)
      vifdf
[72]:
                               VIF
     Feature
      cylinders
                         11.465746
      acceleration
                          3.010771
      oilshock
                         1.151324
     sqrt_displacement 27.042946
      sqrt_horsepower
                         10.615281
      sqrt weight
                         13.450552
      origin_Europe
                         1.774827
      origin_Japan
                          1.944729
[73]: identify_highest_VIF_feature(vifdf)
     We find the highest VIF in this model is sqrt_displacement with a VIF of
     27.042946454149405
     Hence, we drop sqrt_displacement from the model to be fitted.
[73]: ('sqrt_displacement', 27.042946454149405)
[74]: cols.remove("sqrt_displacement")
      vifdf = calculate_VIFs("mpg ~ " + " + ".join(cols), Auto_sqrt)
      vifdf
[74]:
                             VIF
      Feature
      cylinders
                        5.974510
```

```
sqrt_horsepower 10.446261
      sqrt_weight
                        9.963350
      origin_Europe
                        1.450840
      origin_Japan
                        1.623907
[75]: identify_highest_VIF_feature(vifdf)
     We find the highest VIF in this model is sqrt_horsepower with a VIF of
     10.446261176837464
     Hence, we drop sqrt_horsepower from the model to be fitted.
[75]: ('sqrt_horsepower', 10.446261176837464)
[76]: cols.remove("sqrt_horsepower")
      vifdf = calculate_VIFs("mpg ~ " + " + ".join(cols), Auto_sqrt)
      vifdf
[76]:
                          VIF
     Feature
      cylinders
                     5.907717
      acceleration
                     1.377206
      oilshock
                     1.119726
      sqrt_weight
                     5.331435
      origin_Europe 1.446456
      origin_Japan
                     1.581460
[77]: identify_highest_VIF_feature(vifdf)
     No variables are significantly collinear.
[78]: formula = " + ".join(cols)
      results = perform_analysis("mpg", formula, Auto_sqrt)
                                 OLS Regression Results
     Dep. Variable:
                                             R-squared:
                                                                               0.814
                                       mpg
     Model:
                                       OLS
                                             Adj. R-squared:
                                                                               0.811
     Method:
                             Least Squares
                                            F-statistic:
                                                                               281.0
     Date:
                          Fri, 21 Feb 2025 Prob (F-statistic):
                                                                           2.76e-137
     Time:
                                  15:48:49
                                            Log-Likelihood:
                                                                             -1031.4
     No. Observations:
                                             AIC:
                                                                               2077.
                                       392
     Df Residuals:
                                       385
                                             BIC:
                                                                               2105.
     Df Model:
                                         6
     Covariance Type:
                                 nonrobust
                         coef
                                 std err
                                                  t
                                                         P>|t|
                                                                     Γ0.025
```

acceleration

oilshock

2.934605

1.141428

0.975

Intercept 54.3622 2.396 22.687 0.000 49.651 59.073 cylinders 0.0148 0.244 0.061 0.952 -0.466 0.495 acceleration 0.1748 0.073 2.395 0.017 0.031 0.318 oilshock 5.0506 0.364 13.873 0.000 4.335 5.766 sqrt_weight -0.6785 0.052 -13.130 0.000 -0.780 -0.577 origin_Europe 1.5511 0.544 2.851 0.005 0.481 2.621 origin_Japan 1.9033 0.537 3.544 0.000 0.847 2.959							
cylinders 0.0148 0.244 0.061 0.952 -0.466 0.495 acceleration 0.1748 0.073 2.395 0.017 0.031 0.318 oilshock 5.0506 0.364 13.873 0.000 4.335 5.766 sqrt_weight -0.6785 0.052 -13.130 0.000 -0.780 -0.577 origin_Europe 1.5511 0.544 2.851 0.005 0.481 2.621 origin_Japan 1.9033 0.537 3.544 0.000 0.847 2.959	-	54.3622	2.396	22.687	0.000	49.651	
acceleration 0.1748 0.073 2.395 0.017 0.031 0.318 0ilshock 5.0506 0.364 13.873 0.000 4.335 5.766 sqrt_weight -0.6785 0.052 -13.130 0.000 -0.780 -0.577 origin_Europe 1.5511 0.544 2.851 0.005 0.481 2.621 origin_Japan 1.9033 0.537 3.544 0.000 0.847 2.959	cylinders	0.0148	0.244	0.061	0.952	-0.466	
oilshock 5.0506 0.364 13.873 0.000 4.335 5.766 sqrt_weight -0.6785 0.052 -13.130 0.000 -0.780 -0.577 origin_Europe 1.5511 0.544 2.851 0.005 0.481 2.621 origin_Japan 1.9033 0.537 3.544 0.000 0.847 2.959	acceleration	0.1748	0.073	2.395	0.017	0.031	
sqrt_weight -0.6785 0.052 -13.130 0.000 -0.780 -0.577 origin_Europe 1.5511 0.544 2.851 0.005 0.481 2.621 origin_Japan 1.9033 0.537 3.544 0.000 0.847 2.959	oilshock	5.0506	0.364	13.873	0.000	4.335	
origin_Europe 1.5511 0.544 2.851 0.005 0.481 2.621 origin_Japan 1.9033 0.537 3.544 0.000 0.847 2.959	sqrt_weight	-0.6785	0.052	-13.130	0.000	-0.780	
origin_Japan 1.9033 0.537 3.544 0.000 0.847 2.959		1.5511	0.544	2.851	0.005	0.481	
Omnibus: 25.773 Durbin-Watson: 1.269 Prob(Omnibus): 0.000 Jarque-Bera (JB): 41.514		1.9033	0.537	3.544	0.000	0.847	
Prob(Omnibus): 0.000 Jarque-Bera (JB): 41.514	2.959			=======		.=========	==
•	Omnibus:		25.773	Durbin-Wa	atson:	1.26	69
Skew: 0.449 Prob(JB): 9.67e-10	Prob(Omnibus):		0.000	Jarque-B	era (JB):	41.51	14
	Skew:		0.449	Prob(JB)	:	9.67e-1	10
Kurtosis: 4.317 Cond. No. 803.	Kurtosis:		4.317	Cond. No		803	3.

Notes:

[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.

	df	sum_sq	mean_sq	F	PR(>F)
cylinders	1.0	14403.083079	14403.083079	1252.209748	4.496819e-123
acceleration	1.0	30.471304	30.471304	2.649187	1.044209e-01
oilshock	1.0	2422.051542	2422.051542	210.574120	2.284075e-38
sqrt_weight	1.0	2365.801178	2365.801178	205.683691	1.125346e-37
origin_Europe	1.0	24.780703	24.780703	2.154444	1.429743e-01
origin_Japan	1.0	144.484455	144.484455	12.561536	4.421781e-04
Residual	385.0	4428.321210	11.502133	NaN	NaN

[78]: <statsmodels.regression.linear_model.RegressionResultsWrapper at 0x7cb746e94110>

[79]: identify_least_significant_feature(results, alpha=LOS_Alpha)

We find the least significant variable in this model is cylinders with a p-value of 0.9516256870315071 and a coefficient of 0.014835798307379022
Using the backward methodology, we suggest dropping cylinders from the new model

```
[80]: cols.remove("cylinders")
  formula = " + ".join(cols)
  results = perform_analysis("mpg", formula, Auto_sqrt)
```

OLS Regression Results

Dep. Variable: Model: Method: Date: Time: No. Observations Df Residuals: Df Model: Covariance Type:	Fri,	ast Squares 21 Feb 2025 15:48:49 392 386 5 nonrobust	Adj. R-sq F-statist: Prob (F-s Log-Likel: AIC: BIC:	uared: ic: tatistic): ihood:	-1031.4 2075. 2099.
0.975]			t		
- Intercept	54.3324	2.342	23.198	0.000	49.728
58.937					
acceleration 0.309	0.1733	0.069	2.512	0.012	0.038
oilshock 5.760	5.0486	0.362	13.946	0.000	4.337
sqrt_weight -0.616	-0.6760	0.031	-21.968	0.000	-0.737
origin_Europe 2.586	1.5438	0.530	2.913	0.004	0.502
origin_Japan 2.949	1.8999	0.534	3.561	0.000	0.851
Omnibus:		======================================	====== Durbin-Wa	======== tson·	 1.269
Prob(Omnibus):		0.000			41.952
Skew:			Prob(JB):	(02).	7.77e-10
Kurtosis:		4.326	Cond. No.		783.
=======================================	=======	=======	=======	=======	
Notes: [1] Standard Errospecified.	ors assume	that the co	variance ma	trix of the	errors is correct
	df	sum_sq	mean_sq	F	PR(>F)

	df	sum_sq	mean_sq	F	PR(>F)
acceleration	1.0	4268.531557	4268.531557	372.068179	1.546709e-58
oilshock	1.0	4757.627552	4757.627552	414.700418	3.906846e-63
sqrt_weight	1.0	10191.062423	10191.062423	888.307839	3.798699e-102
origin_Europe	1.0	27.910362	27.910362	2.432817	1.196386e-01
origin_Japan	1.0	145.497978	145.497978	12.682387	4.152294e-04
Residual	386.0	4428.363596	11.472445	NaN	NaN

[80]: <statsmodels.regression.linear_model.RegressionResultsWrapper at 0x7cb7470272f0>

[81]: identify_least_significant_feature(results, alpha=LOS_Alpha)

We find the least significant variable in this model is acceleration with a p-value of 0.01239680442008935 and a coefficient of 0.17332821877689886 Using the backward methodology, we suggest dropping acceleration from the new model

```
[82]: cols.remove("acceleration")
formula = " + ".join(cols)
results = perform_analysis("mpg", formula, Auto_sqrt)
```

OLS Regression Results

============	========	uLs kegres =========	sion kesuit	.ຮ -=======	========	======
Dep. Variable: Model: Method: Date: Time: No. Observations Df Residuals: Df Model: Covariance Type:	Fri, 2	392 387 4 nonrobust	Adj. R-so F-statist	quared: cic: statistic):	1	0.811 0.809 415.3 .50e-138 -1034.6 2079. 2099.
0.975]	coef	std err	t	P> t	[0.025	
- Intercept 61.714 oilshock 5.867 sqrt_weight	58.2668 5.1556 -0.6999	1.753 0.362 0.029	33.230 14.244 -23.759	0.000 0.000 0.000	54.819 4.444 -0.758	
-0.642	1.6530 1.8263	0.532 0.536	3.108 3.405	0.002	0.607	
Omnibus: Prob(Omnibus): Skew: Kurtosis:		31.883 0.000 0.483 4.664	Jarque-Be Prob(JB): Cond. No.	era (JB):		1.241 60.472 7.39e-14 574.

Notes:

^[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.

```
df
                                              mean_sq
                                                                            PR(>F)
                                 sum_sq
     oilshock
                      1.0
                            6470.207217
                                          6470.207217
                                                         556.341155
                                                                      7.004177e-77
     sqrt_weight
                      1.0 12674.256790 12674.256790 1089.796728 1.354903e-114
     origin_Europe
                      1.0
                              38.907051
                                            38.907051
                                                           3.345425
                                                                      6.816142e-02
     origin Japan
                                                          11.594270
                                                                      7.307851e-04
                      1.0
                             134.840521
                                           134.840521
     Residual
                    387.0
                           4500.781890
                                            11.629927
                                                                NaN
                                                                               NaN
[82]: <statsmodels.regression.linear_model.RegressionResultsWrapper at 0x7cb746fdda90>
[83]: models.append(
          {
              "name": "sqrt_transformation",
              "model": results.model.formula,
              "R-squared adjusted": results.rsquared adj,
          }
      )
[84]: pd.DataFrame(models)
[84]:
                           name
     model R-squared adjusted
                   simple model
     mpg ~ horsepower + weight + oilshock + origin_Europe + origin_Japan
      0.802074
           numeric_interactions
     mpg ~ horsepower + weight + oilshock + origin_Europe + origin_Japan +
     horsepower: weight
                                    0.844846
          oilshock interactions
                                                mpg ~ horsepower + weight + oilshock +
      origin_Europe + origin_Japan + horsepower: weight + oilshock: weight + oilshock:
      horsepower
                            0.858491
            origin_interactions mpg ~ horsepower + weight + oilshock + origin_Europe
      + origin Japan + oilshock: horsepower + origin Europe: horsepower +
      origin_Japan: horsepower
                                          0.849008
      4 squared_transformation
                                                                            mpg ~
     horsepower + weight + oilshock + origin_Europe + origin_Japan + I(horsepower**2)
      + I(weight**2)
                                0.845550
             log transformation
      mpg ~ acceleration + oilshock + log_weight + origin_Europe + origin_Japan
      0.821151
            sqrt_transformation
      mpg ~ oilshock + sqrt_weight + origin_Europe + origin_Japan
                                                                             0.809089
[85]: allDone()
```

<IPython.lib.display.Audio object>