In [ ]:
```
import numpy as np
import pandas as pd
from matplotlib.pyplot import subplots
```

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

In [ ]:
```
from statsmodels.stats.outliers_influence \
  import variance_inflation_factor as VIF
from statsmodels.stats.anova import anova_lm
```

In [ ]:
```
#pip install ISLP
```

```
Collecting ISLP
  Downloading ISLP-0.4.0-py3-none-any.whl.metadata (7.0 kB)
Requirement already satisfied: numpy>=1.7.1 in /usr/local/lib/python3.11/dist-packa
ges (from ISLP) (1.26.4)
Requirement already satisfied: scipy>=0.9 in /usr/local/lib/python3.11/dist-package
s (from ISLP) (1.13.1)
Requirement already satisfied: pandas>=0.20 in /usr/local/lib/python3.11/dist-packa
ges (from ISLP) (2.2.2)
Requirement already satisfied: lxml in /usr/local/lib/python3.11/dist-packages (fro
m ISLP) (5.3.1)
Requirement already satisfied: scikit-learn>=1.2 in /usr/local/lib/python3.11/dist-
packages (from ISLP) (1.6.1)
Requirement already satisfied: joblib in /usr/local/lib/python3.11/dist-packages (f
rom ISLP) (1.4.2)
Requirement already satisfied: statsmodels>=0.13 in /usr/local/lib/python3.11/dist-
packages (from ISLP) (0.14.4)
Collecting lifelines (from ISLP)
  Downloading lifelines-0.30.0-py3-none-any.whl.metadata (3.2 kB)
Collecting pygam (from ISLP)
  Downloading pygam-0.9.1-py3-none-any.whl.metadata (7.1 kB)
Requirement already satisfied: torch in /usr/local/lib/python3.11/dist-packages (fr
om ISLP) (2.5.1+cu124)
Collecting pytorch-lightning (from ISLP)
  Downloading pytorch_lightning-2.5.0.post0-py3-none-any.whl.metadata (21 kB)
Collecting torchmetrics (from ISLP)
  Downloading torchmetrics-1.6.1-py3-none-any.whl.metadata (21 kB)
Requirement already satisfied: python-dateutil>=2.8.2 in /usr/local/lib/python3.11/
dist-packages (from pandas>=0.20->ISLP) (2.8.2)
Requirement already satisfied: pytz>=2020.1 in /usr/local/lib/python3.11/dist-packa
ges (from pandas>=0.20->ISLP) (2025.1)
Requirement already satisfied: tzdata>=2022.7 in /usr/local/lib/python3.11/dist-pac
kages (from pandas>=0.20->ISLP) (2025.1)
Requirement already satisfied: threadpoolctl>=3.1.0 in /usr/local/lib/python3.11/di
st-packages (from scikit-learn>=1.2->ISLP) (3.5.0)
Requirement already satisfied: patsy>=0.5.6 in /usr/local/lib/python3.11/dist-packa
ges (from statsmodels>=0.13->ISLP) (1.0.1)
Requirement already satisfied: packaging>=21.3 in /usr/local/lib/python3.11/dist-pa
ckages (from statsmodels>=0.13->ISLP) (24.2)
Requirement already satisfied: matplotlib>=3.0 in /usr/local/lib/python3.11/dist-pa
ckages (from lifelines->ISLP) (3.10.0)
Requirement already satisfied: autograd>=1.5 in /usr/local/lib/python3.11/dist-pack
ages (from lifelines->ISLP) (1.7.0)
Collecting autograd-gamma>=0.3 (from lifelines->ISLP)
  Downloading autograd-gamma-0.5.0.tar.gz (4.0 kB)
  Preparing metadata (setup.py) ... done
Collecting formulaic>=0.2.2 (from lifelines->ISLP)
  Downloading formulaic-1.1.1-py3-none-any.whl.metadata (6.9 kB)
Requirement already satisfied: progressbar2<5.0.0,>=4.2.0 in /usr/local/lib/python
3.11/dist-packages (from pygam->ISLP) (4.5.0)
Collecting scipy>=0.9 (from ISLP)
  Downloading scipy-1.11.4-cp311-cp311-manylinux_2_17_x86_64.manylinux2014_x86_64.w
hl.metadata (60 kB)
     ━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━ 60.4/60.4 kB 2.4 MB/s eta 0:00:00
Requirement already satisfied: tqdm>=4.57.0 in /usr/local/lib/python3.11/dist-packa
ges (from pytorch-lightning->ISLP) (4.67.1)
Requirement already satisfied: PyYAML>=5.4 in /usr/local/lib/python3.11/dist-packag
es (from pytorch-lightning->ISLP) (6.0.2)
Requirement already satisfied: fsspec>=2022.5.0 in /usr/local/lib/python3.11/dist-p
ackages (from fsspec[http]>=2022.5.0->pytorch-lightning->ISLP) (2024.10.0)
```

```
Requirement already satisfied: typing-extensions>=4.4.0 in /usr/local/lib/python3.1
1/dist-packages (from pytorch-lightning->ISLP) (4.12.2)
Collecting lightning-utilities>=0.10.0 (from pytorch-lightning->ISLP)
  Downloading lightning_utilities-0.12.0-py3-none-any.whl.metadata (5.6 kB)
Requirement already satisfied: filelock in /usr/local/lib/python3.11/dist-packages
(from torch->ISLP) (3.17.0)
Requirement already satisfied: networkx in /usr/local/lib/python3.11/dist-packages
(from torch->ISLP) (3.4.2)
Requirement already satisfied: jinja2 in /usr/local/lib/python3.11/dist-packages (f
rom torch->ISLP) (3.1.5)
Collecting nvidia-cuda-nvrtc-cu12==12.4.127 (from torch->ISLP)
  Downloading nvidia_cuda_nvrtc_cu12-12.4.127-py3-none-manylinux2014_x86_64.whl.met
adata (1.5 kB)
Collecting nvidia-cuda-runtime-cu12==12.4.127 (from torch->ISLP)
  Downloading nvidia_cuda_runtime_cu12-12.4.127-py3-none-manylinux2014_x86_64.whl.m
etadata (1.5 kB)
Collecting nvidia-cuda-cupti-cu12==12.4.127 (from torch->ISLP)
  Downloading nvidia_cuda_cupti_cu12-12.4.127-py3-none-manylinux2014_x86_64.whl.met
adata (1.6 kB)
Collecting nvidia-cudnn-cu12==9.1.0.70 (from torch->ISLP)
  Downloading nvidia_cudnn_cu12-9.1.0.70-py3-none-manylinux2014_x86_64.whl.metadata
(1.6 kB)
Collecting nvidia-cublas-cu12==12.4.5.8 (from torch->ISLP)
  Downloading nvidia_cublas_cu12-12.4.5.8-py3-none-manylinux2014_x86_64.whl.metadat
a (1.5 kB)
Collecting nvidia-cufft-cu12==11.2.1.3 (from torch->ISLP)
  Downloading nvidia_cufft_cu12-11.2.1.3-py3-none-manylinux2014_x86_64.whl.metadata
(1.5 kB)
Collecting nvidia-curand-cu12==10.3.5.147 (from torch->ISLP)
  Downloading nvidia_curand_cu12-10.3.5.147-py3-none-manylinux2014_x86_64.whl.metad
ata (1.5 kB)
Collecting nvidia-cusolver-cu12==11.6.1.9 (from torch->ISLP)
  Downloading nvidia_cusolver_cu12-11.6.1.9-py3-none-manylinux2014_x86_64.whl.metad
ata (1.6 kB)
Collecting nvidia-cusparse-cu12==12.3.1.170 (from torch->ISLP)
  Downloading nvidia_cusparse_cu12-12.3.1.170-py3-none-manylinux2014_x86_64.whl.met
adata (1.6 kB)
Requirement already satisfied: nvidia-nccl-cu12==2.21.5 in /usr/local/lib/python3.1
1/dist-packages (from torch->ISLP) (2.21.5)
Requirement already satisfied: nvidia-nvtx-cu12==12.4.127 in /usr/local/lib/python
3.11/dist-packages (from torch->ISLP) (12.4.127)
Collecting nvidia-nvjitlink-cu12==12.4.127 (from torch->ISLP)
  Downloading nvidia_nvjitlink_cu12-12.4.127-py3-none-manylinux2014_x86_64.whl.meta
data (1.5 kB)
Requirement already satisfied: triton==3.1.0 in /usr/local/lib/python3.11/dist-pack
ages (from torch->ISLP) (3.1.0)
Requirement already satisfied: sympy==1.13.1 in /usr/local/lib/python3.11/dist-pack
ages (from torch->ISLP) (1.13.1)
Requirement already satisfied: mpmath<1.4,>=1.1.0 in /usr/local/lib/python3.11/dist
-packages (from sympy==1.13.1->torch->ISLP) (1.3.0)
Collecting interface-meta>=1.2.0 (from formulaic>=0.2.2->lifelines->ISLP)
  Downloading interface_meta-1.3.0-py3-none-any.whl.metadata (6.7 kB)
Requirement already satisfied: wrapt>=1.0 in /usr/local/lib/python3.11/dist-package
s (from formulaic>=0.2.2->lifelines->ISLP) (1.17.2)
Requirement already satisfied: aiohttp!=4.0.0a0,!=4.0.0a1 in /usr/local/lib/python
3.11/dist-packages (from fsspec[http]>=2022.5.0->pytorch-lightning->ISLP) (3.11.12)
Requirement already satisfied: setuptools in /usr/local/lib/python3.11/dist-package
s (from lightning-utilities>=0.10.0->pytorch-lightning->ISLP) (75.1.0)
Requirement already satisfied: contourpy>=1.0.1 in /usr/local/lib/python3.11/dist-p
```

```
ackages (from matplotlib>=3.0->lifelines->ISLP) (1.3.1)
Requirement already satisfied: cycler>=0.10 in /usr/local/lib/python3.11/dist-packa
ges (from matplotlib>=3.0->lifelines->ISLP) (0.12.1)
Requirement already satisfied: fonttools>=4.22.0 in /usr/local/lib/python3.11/dist-
packages (from matplotlib>=3.0->lifelines->ISLP) (4.56.0)
Requirement already satisfied: kiwisolver>=1.3.1 in /usr/local/lib/python3.11/dist-
packages (from matplotlib>=3.0->lifelines->ISLP) (1.4.8)
Requirement already satisfied: pillow>=8 in /usr/local/lib/python3.11/dist-packages
(from matplotlib>=3.0->lifelines->ISLP) (11.1.0)
Requirement already satisfied: pyparsing>=2.3.1 in /usr/local/lib/python3.11/dist-p
ackages (from matplotlib>=3.0->lifelines->ISLP) (3.2.1)
Requirement already satisfied: python-utils>=3.8.1 in /usr/local/lib/python3.11/dis
t-packages (from progressbar2<5.0.0,>=4.2.0->pygam->ISLP) (3.9.1)
Requirement already satisfied: six>=1.5 in /usr/local/lib/python3.11/dist-packages
(from python-dateutil>=2.8.2->pandas>=0.20->ISLP) (1.17.0)
Requirement already satisfied: MarkupSafe>=2.0 in /usr/local/lib/python3.11/dist-pa
ckages (from jinja2->torch->ISLP) (3.0.2)
Requirement already satisfied: aiohappyeyeballs>=2.3.0 in /usr/local/lib/python3.1
1/dist-packages (from aiohttp!=4.0.0a0,!=4.0.0a1->fsspec[http]>=2022.5.0->pytorch-l
ightning->ISLP) (2.4.6)
Requirement already satisfied: aiosignal>=1.1.2 in /usr/local/lib/python3.11/dist-p
ackages (from aiohttp!=4.0.0a0,!=4.0.0a1->fsspec[http]>=2022.5.0->pytorch-lightning
->ISLP) (1.3.2)
Requirement already satisfied: attrs>=17.3.0 in /usr/local/lib/python3.11/dist-pack
ages (from aiohttp!=4.0.0a0,!=4.0.0a1->fsspec[http]>=2022.5.0->pytorch-lightning->I
SLP) (25.1.0)
Requirement already satisfied: frozenlist>=1.1.1 in /usr/local/lib/python3.11/dist-
packages (from aiohttp!=4.0.0a0,!=4.0.0a1->fsspec[http]>=2022.5.0->pytorch-lightnin
g->ISLP) (1.5.0)
Requirement already satisfied: multidict<7.0,>=4.5 in /usr/local/lib/python3.11/dis
t-packages (from aiohttp!=4.0.0a0,!=4.0.0a1->fsspec[http]>=2022.5.0->pytorch-lightn
ing->ISLP) (6.1.0)
Requirement already satisfied: propcache>=0.2.0 in /usr/local/lib/python3.11/dist-p
ackages (from aiohttp!=4.0.0a0,!=4.0.0a1->fsspec[http]>=2022.5.0->pytorch-lightning
->ISLP) (0.2.1)
Requirement already satisfied: yarl<2.0,>=1.17.0 in /usr/local/lib/python3.11/dist-
packages (from aiohttp!=4.0.0a0,!=4.0.0a1->fsspec[http]>=2022.5.0->pytorch-lightnin
g->ISLP) (1.18.3)
Requirement already satisfied: idna>=2.0 in /usr/local/lib/python3.11/dist-packages
(from yarl<2.0,>=1.17.0->aiohttp!=4.0.0a0,!=4.0.0a1->fsspec[http]>=2022.5.0->pytorc
h-lightning->ISLP) (3.10)
Downloading ISLP-0.4.0-py3-none-any.whl (3.6 MB)
   ━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━ 3.6/3.6 MB 33.9 MB/s eta 0:00:00
Downloading lifelines-0.30.0-py3-none-any.whl (349 kB)
   ━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━ 349.3/349.3 kB 29.4 MB/s eta 0:00:00
Downloading pygam-0.9.1-py3-none-any.whl (522 kB)
   ━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━ 522.0/522.0 kB 35.2 MB/s eta 0:00:00
Downloading scipy-1.11.4-cp311-cp311-manylinux_2_17_x86_64.manylinux2014_x86_64.whl
(36.4 MB)
   ━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━ 36.4/36.4 MB 20.8 MB/s eta 0:00:00
Downloading pytorch_lightning-2.5.0.post0-py3-none-any.whl (819 kB)
   ━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━ 819.3/819.3 kB 33.2 MB/s eta 0:00:00
Downloading nvidia_cublas_cu12-12.4.5.8-py3-none-manylinux2014_x86_64.whl (363.4 M
B)
   ━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━ 363.4/363.4 MB 4.7 MB/s eta 0:00:00
Downloading nvidia_cuda_cupti_cu12-12.4.127-py3-none-manylinux2014_x86_64.whl (13.8
MB)
   ━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━ 13.8/13.8 MB 61.1 MB/s eta 0:00:00
Downloading nvidia_cuda_nvrtc_cu12-12.4.127-py3-none-manylinux2014_x86_64.whl (24.6
```

```
MB)
   ━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━ 24.6/24.6 MB 34.5 MB/s eta 0:00:00
Downloading nvidia_cuda_runtime_cu12-12.4.127-py3-none-manylinux2014_x86_64.whl (88
3 kB)
   ━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━ 883.7/883.7 kB 53.3 MB/s eta 0:00:00
Downloading nvidia_cudnn_cu12-9.1.0.70-py3-none-manylinux2014_x86_64.whl (664.8 MB)
   ━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━ 664.8/664.8 MB 2.0 MB/s eta 0:00:00
Downloading nvidia_cufft_cu12-11.2.1.3-py3-none-manylinux2014_x86_64.whl (211.5 MB)
   ━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━ 211.5/211.5 MB 6.0 MB/s eta 0:00:00
Downloading nvidia_curand_cu12-10.3.5.147-py3-none-manylinux2014_x86_64.whl (56.3 M
B)
   ━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━ 56.3/56.3 MB 12.9 MB/s eta 0:00:00
Downloading nvidia_cusolver_cu12-11.6.1.9-py3-none-manylinux2014_x86_64.whl (127.9
MB)
   ━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━ 127.9/127.9 MB 8.7 MB/s eta 0:00:00
Downloading nvidia_cusparse_cu12-12.3.1.170-py3-none-manylinux2014_x86_64.whl (207.
5 MB)
   ━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━ 207.5/207.5 MB 6.3 MB/s eta 0:00:00
Downloading nvidia_nvjitlink_cu12-12.4.127-py3-none-manylinux2014_x86_64.whl (21.1
MB)
   ━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━ 21.1/21.1 MB 81.4 MB/s eta 0:00:00
Downloading torchmetrics-1.6.1-py3-none-any.whl (927 kB)
   ━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━ 927.3/927.3 kB 57.0 MB/s eta 0:00:00
Downloading formulaic-1.1.1-py3-none-any.whl (115 kB)
   ━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━ 115.7/115.7 kB 14.2 MB/s eta 0:00:00
Downloading lightning_utilities-0.12.0-py3-none-any.whl (28 kB)
Downloading interface_meta-1.3.0-py3-none-any.whl (14 kB)
Building wheels for collected packages: autograd-gamma
  Building wheel for autograd-gamma (setup.py) ... done
  Created wheel for autograd-gamma: filename=autograd_gamma-0.5.0-py3-none-any.whl
size=4031 sha256=e372c8a9d9eb1f113ab6bbac1d73b5ed24de76cd9c8e3f5bb4547a37c3f7e225
  Stored in directory: /root/.cache/pip/wheels/8b/67/f4/2caaae2146198dcb824f31a3038
33b07b14a5ec863fb3acd7b
Successfully built autograd-gamma
Installing collected packages: scipy, nvidia-nvjitlink-cu12, nvidia-curand-cu12, nv
idia-cufft-cu12, nvidia-cuda-runtime-cu12, nvidia-cuda-nvrtc-cu12, nvidia-cuda-cupt
i-cu12, nvidia-cublas-cu12, lightning-utilities, interface-meta, nvidia-cusparse-cu
12, nvidia-cudnn-cu12, autograd-gamma, pygam, nvidia-cusolver-cu12, formulaic, life
lines, torchmetrics, pytorch-lightning, ISLP
  Attempting uninstall: scipy
    Found existing installation: scipy 1.13.1
    Uninstalling scipy-1.13.1:
      Successfully uninstalled scipy-1.13.1
  Attempting uninstall: nvidia-nvjitlink-cu12
    Found existing installation: nvidia-nvjitlink-cu12 12.5.82
    Uninstalling nvidia-nvjitlink-cu12-12.5.82:
      Successfully uninstalled nvidia-nvjitlink-cu12-12.5.82
  Attempting uninstall: nvidia-curand-cu12
    Found existing installation: nvidia-curand-cu12 10.3.6.82
    Uninstalling nvidia-curand-cu12-10.3.6.82:
      Successfully uninstalled nvidia-curand-cu12-10.3.6.82
  Attempting uninstall: nvidia-cufft-cu12
    Found existing installation: nvidia-cufft-cu12 11.2.3.61
    Uninstalling nvidia-cufft-cu12-11.2.3.61:
      Successfully uninstalled nvidia-cufft-cu12-11.2.3.61
  Attempting uninstall: nvidia-cuda-runtime-cu12
    Found existing installation: nvidia-cuda-runtime-cu12 12.5.82
    Uninstalling nvidia-cuda-runtime-cu12-12.5.82:
      Successfully uninstalled nvidia-cuda-runtime-cu12-12.5.82
```

```
  Attempting uninstall: nvidia-cuda-nvrtc-cu12
    Found existing installation: nvidia-cuda-nvrtc-cu12 12.5.82
    Uninstalling nvidia-cuda-nvrtc-cu12-12.5.82:
      Successfully uninstalled nvidia-cuda-nvrtc-cu12-12.5.82
  Attempting uninstall: nvidia-cuda-cupti-cu12
    Found existing installation: nvidia-cuda-cupti-cu12 12.5.82
    Uninstalling nvidia-cuda-cupti-cu12-12.5.82:
      Successfully uninstalled nvidia-cuda-cupti-cu12-12.5.82
  Attempting uninstall: nvidia-cublas-cu12
    Found existing installation: nvidia-cublas-cu12 12.5.3.2
    Uninstalling nvidia-cublas-cu12-12.5.3.2:
      Successfully uninstalled nvidia-cublas-cu12-12.5.3.2
  Attempting uninstall: nvidia-cusparse-cu12
    Found existing installation: nvidia-cusparse-cu12 12.5.1.3
    Uninstalling nvidia-cusparse-cu12-12.5.1.3:
      Successfully uninstalled nvidia-cusparse-cu12-12.5.1.3
  Attempting uninstall: nvidia-cudnn-cu12
    Found existing installation: nvidia-cudnn-cu12 9.3.0.75
    Uninstalling nvidia-cudnn-cu12-9.3.0.75:
      Successfully uninstalled nvidia-cudnn-cu12-9.3.0.75
  Attempting uninstall: nvidia-cusolver-cu12
    Found existing installation: nvidia-cusolver-cu12 11.6.3.83
    Uninstalling nvidia-cusolver-cu12-11.6.3.83:
      Successfully uninstalled nvidia-cusolver-cu12-11.6.3.83
Successfully installed ISLP-0.4.0 autograd-gamma-0.5.0 formulaic-1.1.1 interface-me
ta-1.3.0 lifelines-0.30.0 lightning-utilities-0.12.0 nvidia-cublas-cu12-12.4.5.8 nv
idia-cuda-cupti-cu12-12.4.127 nvidia-cuda-nvrtc-cu12-12.4.127 nvidia-cuda-runtime-c
u12-12.4.127 nvidia-cudnn-cu12-9.1.0.70 nvidia-cufft-cu12-11.2.1.3 nvidia-curand-cu
12-10.3.5.147 nvidia-cusolver-cu12-11.6.1.9 nvidia-cusparse-cu12-12.3.1.170 nvidia-
nvjitlink-cu12-12.4.127 pygam-0.9.1 pytorch-lightning-2.5.0.post0 scipy-1.11.4 torc
hmetrics-1.6.1
```

In [ ]:
```
from ISLP import load_data
from ISLP.models import (ModelSpec as MS,
                         summarize,
                         poly)
```

In [ ]:
```
#Question 8
Auto = load_data("Auto")
Auto.columns
```

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

In [ ]:
```
#8a
X = pd.DataFrame({'intercept': np.ones(Auto.shape[0]),
                  'horsepower': Auto['horsepower']})
X[:4]
```

Out[ ]:

name	intercept	horsepower
chevrolet chevelle malibu	1.0	130
buick skylark 320	1.0	165
plymouth satellite	1.0	150
amc rebel sst	1.0	150

In [ ]:
```
y = Auto['mpg']
model = sm.OLS(y, X)
results = model.fit()
```

In [ ]:
```
summarize(results)
```

Out[ ]:

	coef	std err	t	P>|t|
intercept	39.9359	0.717	55.660	0.0
horsepower	-0.1578	0.006	-24.489	0.0

In [ ]:
```
results.summary()
```

Out[ ]:

OLS Regression Results

Dep. Variable:	mpg	**R-squared:**	0.606
Model:	OLS	**Adj. R-squared:**	0.605
Method:	Least Squares	**F-statistic:**	599.7
Date:	Tue, 18 Feb 2025	**Prob (F-statistic):**	7.03e-81
Time:	00:14:14	**Log-Likelihood:**	-1178.7
No. Observations:	392	**AIC:**	2361.
Df Residuals:	390	**BIC:**	2369.
Df Model:	1		
Covariance Type:	nonrobust		

	coef	std err	t	P>|t|	[0.025	0.975]
intercept	39.9359	0.717	55.660	0.000	38.525	41.347
horsepower	-0.1578	0.006	-24.489	0.000	-0.171	-0.145

Omnibus:	16.432	**Durbin-Watson:**	0.920
Prob(Omnibus):	0.000	**Jarque-Bera (JB):**	17.305
Skew:	0.492	**Prob(JB):**	0.000175
Kurtosis:	3.299	**Cond. No.**	322.

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

In [ ]:
```
np.sqrt(results.scale)/y.mean()
```

Out[ ]:
```
0.20923714066914834
```

In [ ]:
```
design = MS(['horsepower'])
X = design.fit_transform(Auto)
X[:4]
```

Out[ ]:

name	intercept	horsepower
chevrolet chevelle malibu	1.0	130
buick skylark 320	1.0	165
plymouth satellite	1.0	150
amc rebel sst	1.0	150

In [ ]:
```
new_df = pd.DataFrame({'horsepower':[98]})
newX = design.transform(new_df)
newX
```

Out[ ]:

	intercept	horsepower
0	1.0	98

In [ ]:
```
new_predictions = results.get_prediction(newX); new_predictions.predicted_mean
```

Out[ ]:
```
array([24.46707715])
```

In [ ]:
```
new_predictions.conf_int(alpha=0.05)
```

Out[ ]:
```
array([[23.97307896, 24.96107534]])
```

In [ ]:
```
new_predictions.conf_int(obs=True, alpha=0.05)
```

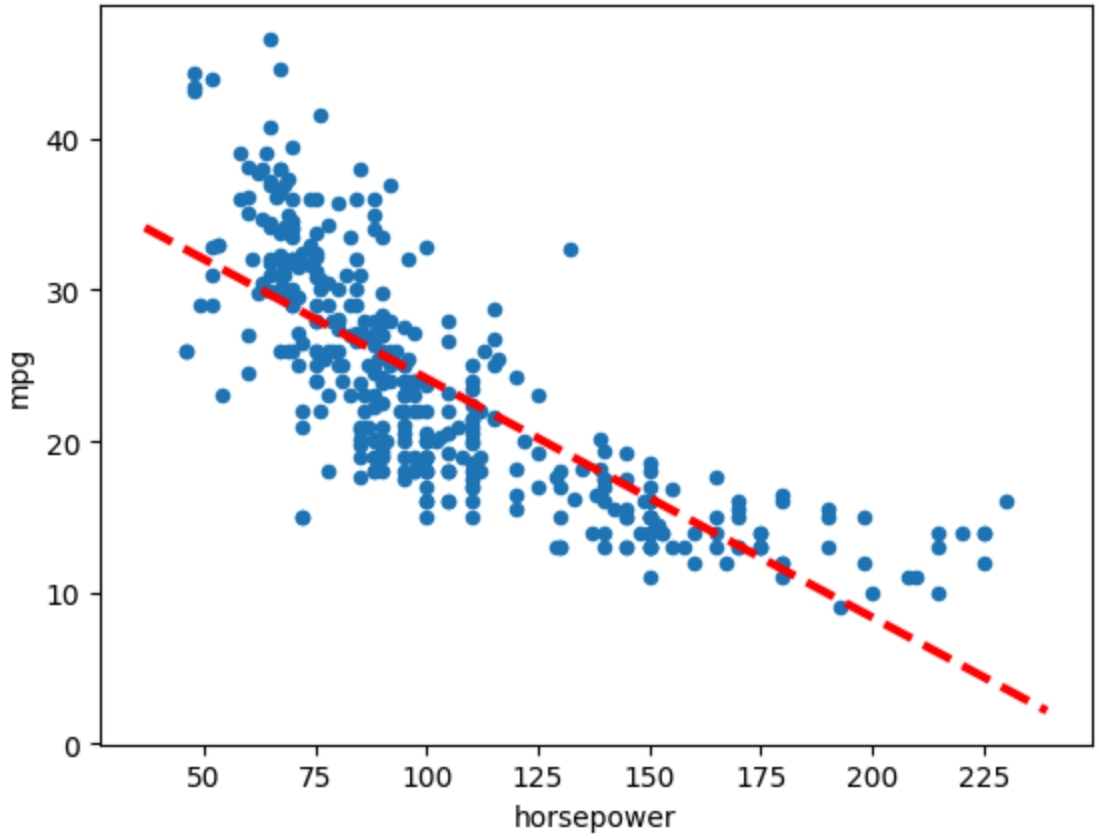
Out[ ]:
```
array([[14.80939607, 34.12475823]])
```

8a) i.There is a relation between the predictor and response since the t statistic is <0.05/2.
ii.There is a moderately strong relationship between the predictor and response since R^2=60.6% is high, and the percentage error of ~20% is low.
iii.The relation is **negative** between predictor and response since beta_1 estimate is less than 0.
iv.Preditcted mpg with 98 horsepower is 24.46707715.
Confidence interval: (23.97307896, 24.96107534)
Prediction interval: (14.80939607, 34.12475823)

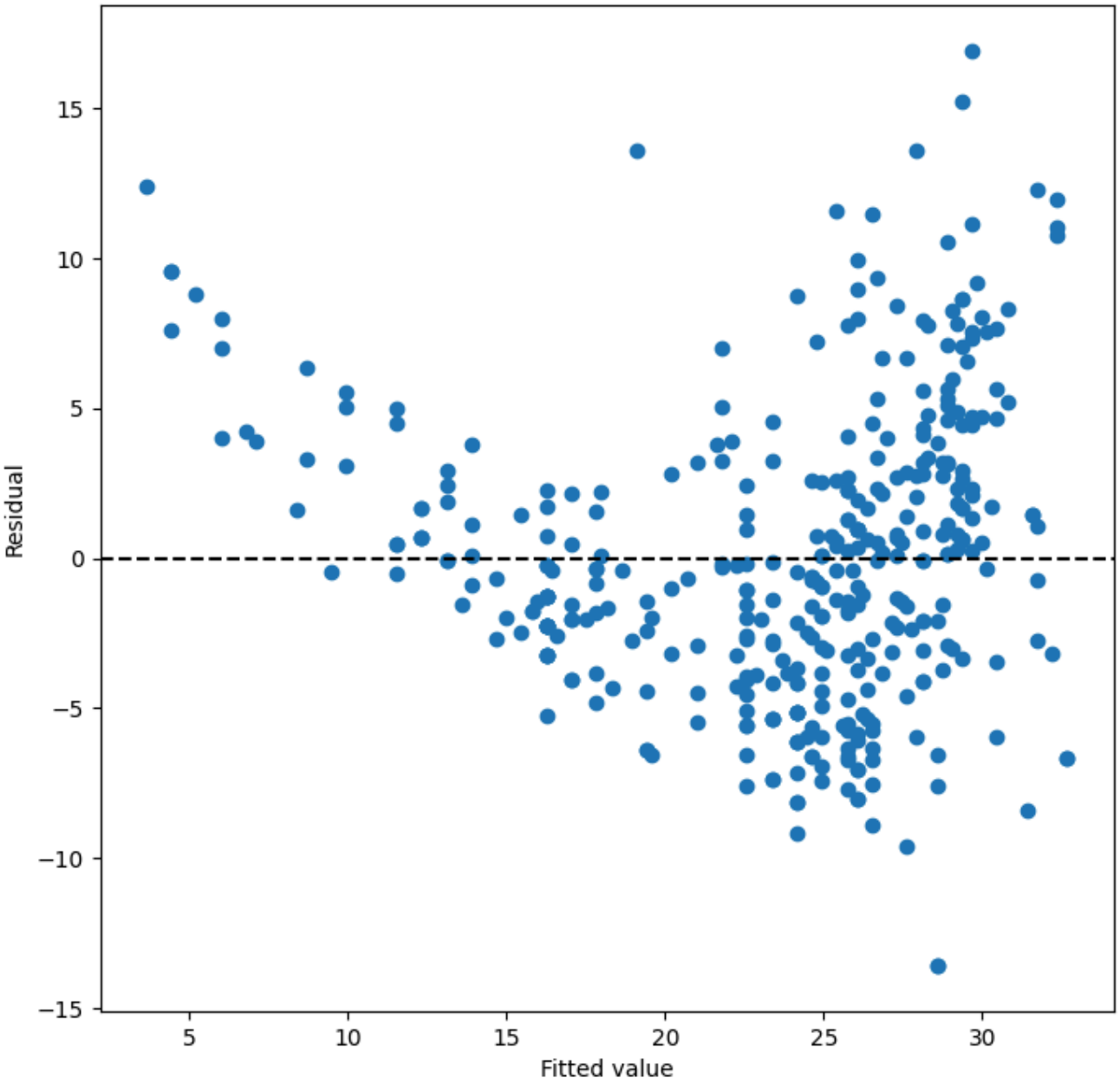
In [ ]:
```
#8b
def abline(ax, b, m, *args, **kwargs):
  "Add a line with slope m and intercept b to ax"
  xlim = ax.get_xlim()
  ylim = [m * xlim[0] + b, m * xlim[1] + b]
  ax.plot(xlim, ylim, *args, **kwargs)
```

In [ ]:
```
ax = Auto.plot.scatter('horsepower', 'mpg')
abline(ax,
        results.params[0],
        results.params[1],
       'r--',
       linewidth=3)
```

```
<ipython-input-20-79ad2517a0a4>:3: FutureWarning: Series.__getitem__ treating keys
as positions is deprecated. In a future version, integer keys will always be treate
d as labels (consistent with DataFrame behavior). To access a value by position, us
e `ser.iloc[pos]`
  results.params[0],
<ipython-input-20-79ad2517a0a4>:4: FutureWarning: Series.__getitem__ treating keys
as positions is deprecated. In a future version, integer keys will always be treate
d as labels (consistent with DataFrame behavior). To access a value by position, us
e `ser.iloc[pos]`
  results.params[1],
```

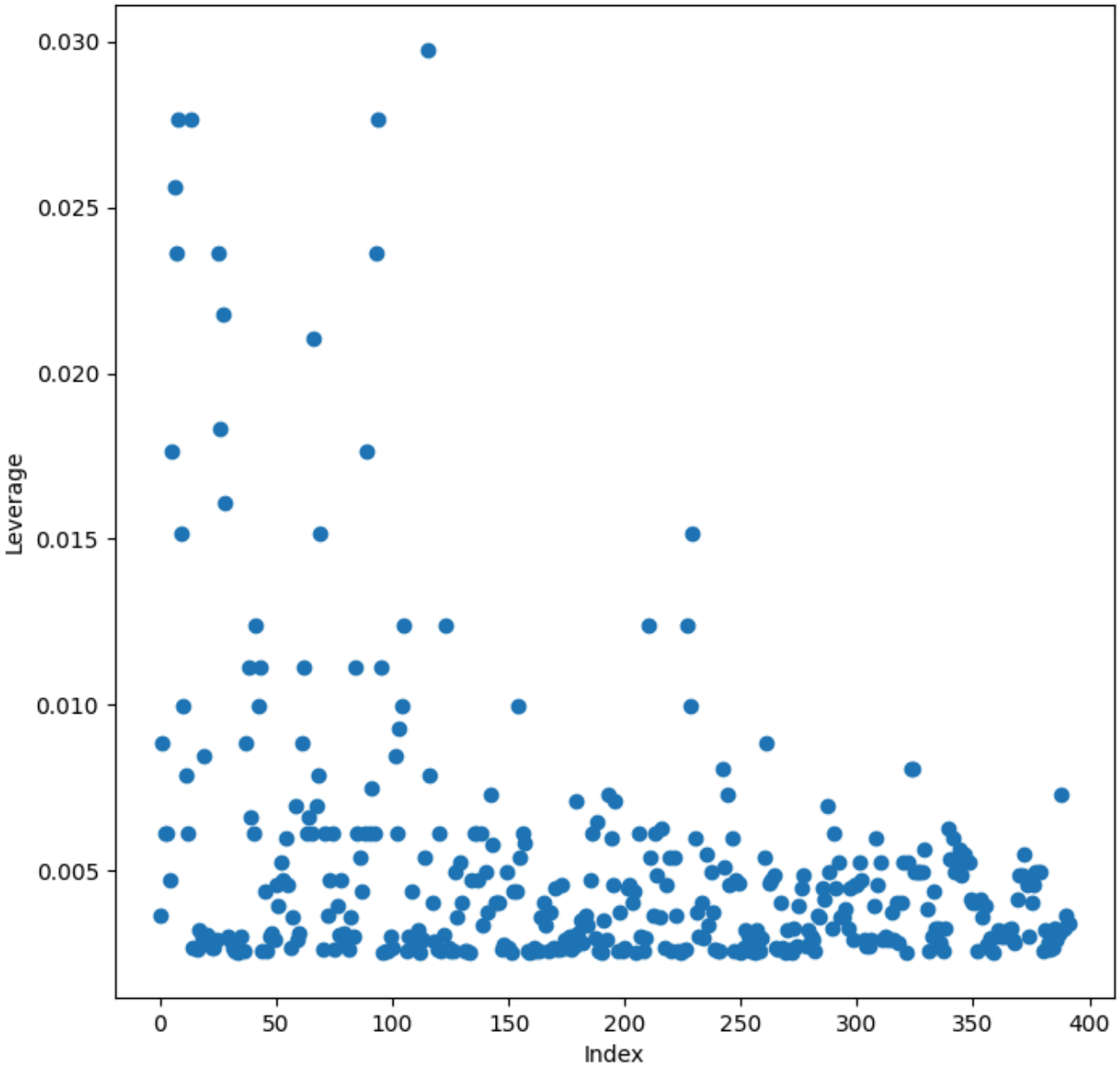


In [ ]:
```
#8c
ax = subplots(figsize=(8,8))[1]
ax.scatter(results.fittedvalues , results.resid)
ax.set_xlabel('Fitted value')
ax.set_ylabel('Residual')
ax.axhline(0, c='k', ls='--');
```



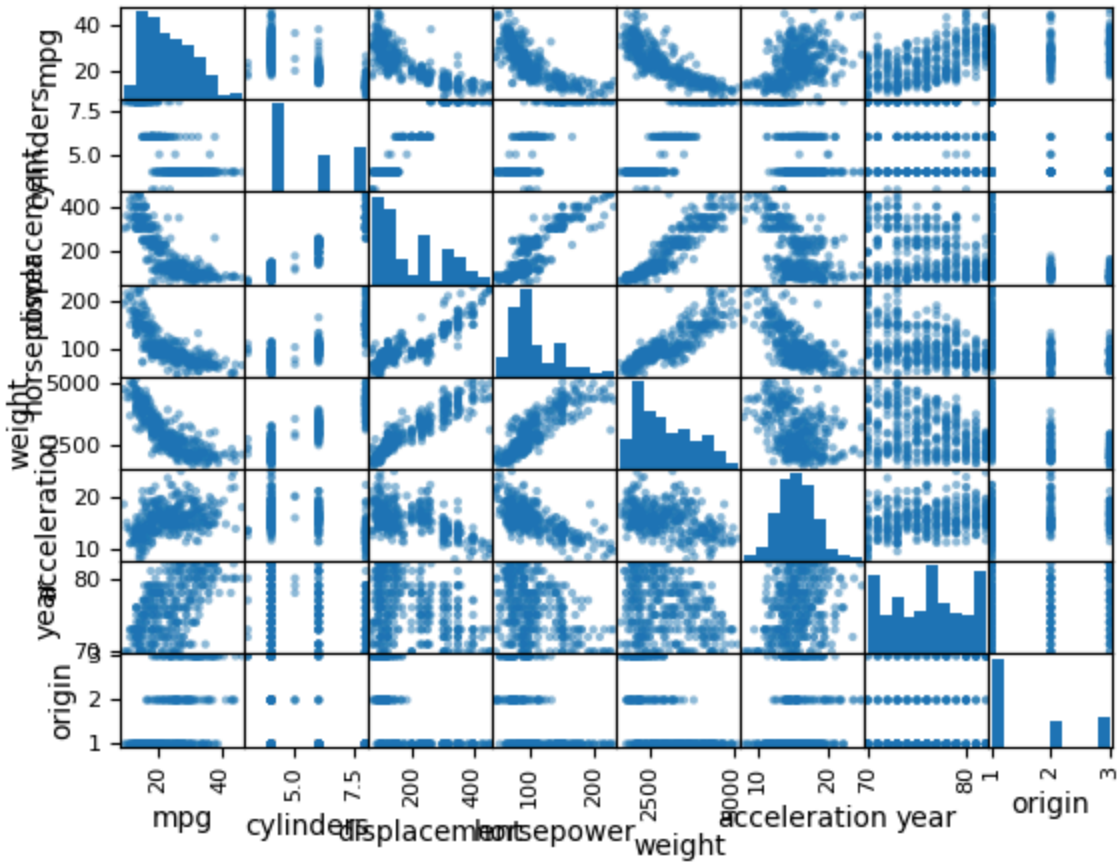
In [ ]:
```
infl=results.get_influence()
ax = subplots(figsize=(8,8))[1]
ax.scatter(np.arange(X.shape[0]), infl.hat_matrix_diag)
ax.set_xlabel('Index')
ax.set_ylabel('Leverage')
np.argmax(infl.hat_matrix_diag)
```

Out[ ]: 115



8c) I noticed that the residuals have heteroscedascity, and there are lots of leverage values higher than (p+1)/n=2/392=0.005.

In [ ]:
```
#9a
pd.plotting.scatter_matrix(Auto);
```



In [ ]:
```
#9b
print(Auto.corr())
```

```
                    mpg  cylinders  displacement  horsepower    weight  \
mpg            1.000000  -0.777618     -0.805127   -0.778427 -0.832244
cylinders     -0.777618   1.000000      0.950823    0.842983  0.897527
displacement  -0.805127   0.950823      1.000000    0.897257  0.932994
horsepower    -0.778427   0.842983      0.897257    1.000000  0.864538
weight        -0.832244   0.897527      0.932994    0.864538  1.000000
acceleration   0.423329  -0.504683     -0.543800   -0.689196 -0.416839
year           0.580541  -0.345647     -0.369855   -0.416361 -0.309120
origin         0.565209  -0.568932     -0.614535   -0.455171 -0.585005

              acceleration      year    origin
mpg               0.423329  0.580541  0.565209
cylinders        -0.504683 -0.345647 -0.568932
displacement     -0.543800 -0.369855 -0.614535
horsepower       -0.689196 -0.416361 -0.455171
weight           -0.416839 -0.309120 -0.585005
acceleration      1.000000  0.290316  0.212746
year              0.290316  1.000000  0.181528
origin            0.212746  0.181528  1.000000
```

In [ ]:
```
Auto.columns
```

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

In [ ]:
```
#9c
allvars = list(Auto.columns.drop('mpg'))
y = Auto['mpg']
```

```
final = allvars
X = MS(final).fit_transform(Auto)
model = sm.OLS(y, X)
summarize(model.fit())
```

Out[ ]:

	coef	std err	t	P>|t|
intercept	-17.2184	4.644	-3.707	0.000
cylinders	-0.4934	0.323	-1.526	0.128
displacement	0.0199	0.008	2.647	0.008
horsepower	-0.0170	0.014	-1.230	0.220
weight	-0.0065	0.001	-9.929	0.000
acceleration	0.0806	0.099	0.815	0.415
year	0.7508	0.051	14.729	0.000
origin	1.4261	0.278	5.127	0.000

In [ ]:
```
anova_lm(results,model.fit())
```

Out[ ]:

	df_resid	ssr	df_diff	ss_diff	F	Pr(>F)
0	390.0	9385.915872	0.0	NaN	NaN	NaN
1	384.0	4252.212530	6.0	5133.703341	77.267308	5.376746e-63

In [ ]:
```
model.fit().summary()
```

Out[ ]:

OLS Regression Results

Dep. Variable:	mpg	**R-squared:**	0.821
Model:	OLS	**Adj. R-squared:**	0.818
Method:	Least Squares	**F-statistic:**	252.4
Date:	Tue, 18 Feb 2025	**Prob (F-statistic):**	2.04e-139
Time:	00:14:22	**Log-Likelihood:**	-1023.5
No. Observations:	392	**AIC:**	2063.
Df Residuals:	384	**BIC:**	2095.
Df Model:	7		
Covariance Type:	nonrobust		

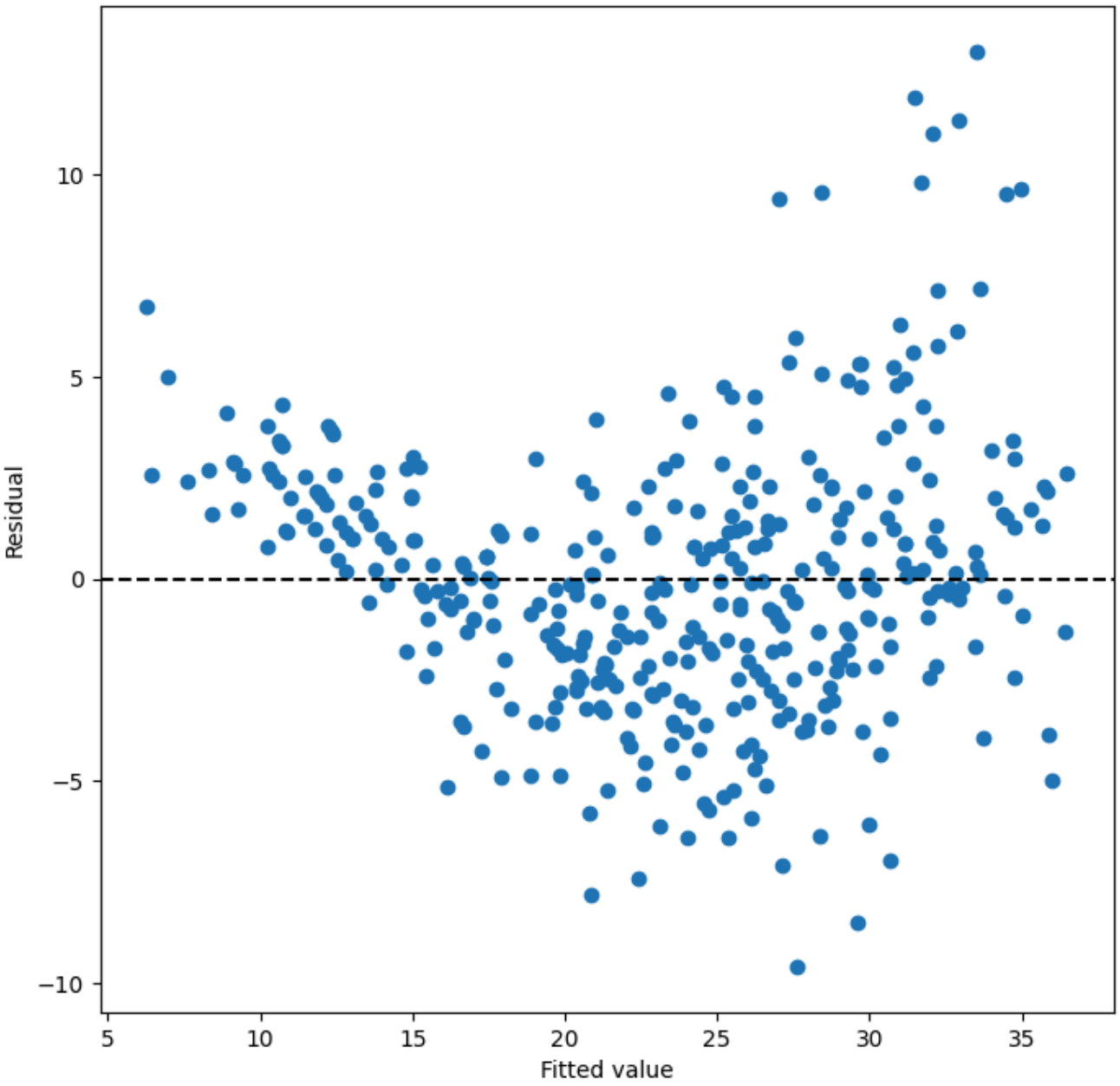
	coef	std err	t	P>|t|	[0.025	0.975]
intercept	-17.2184	4.644	-3.707	0.000	-26.350	-8.087
cylinders	-0.4934	0.323	-1.526	0.128	-1.129	0.142
displacement	0.0199	0.008	2.647	0.008	0.005	0.035
horsepower	-0.0170	0.014	-1.230	0.220	-0.044	0.010
weight	-0.0065	0.001	-9.929	0.000	-0.008	-0.005
acceleration	0.0806	0.099	0.815	0.415	-0.114	0.275
year	0.7508	0.051	14.729	0.000	0.651	0.851
origin	1.4261	0.278	5.127	0.000	0.879	1.973

Omnibus:	31.906	**Durbin-Watson:**	1.309
Prob(Omnibus):	0.000	**Jarque-Bera (JB):**	53.100
Skew:	0.529	**Prob(JB):**	2.95e-12
Kurtosis:	4.460	**Cond. No.**	8.59e+04

Notes:
[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.
[2] The condition number is large, 8.59e+04. This might indicate that there are
strong multicollinearity or other numerical problems.

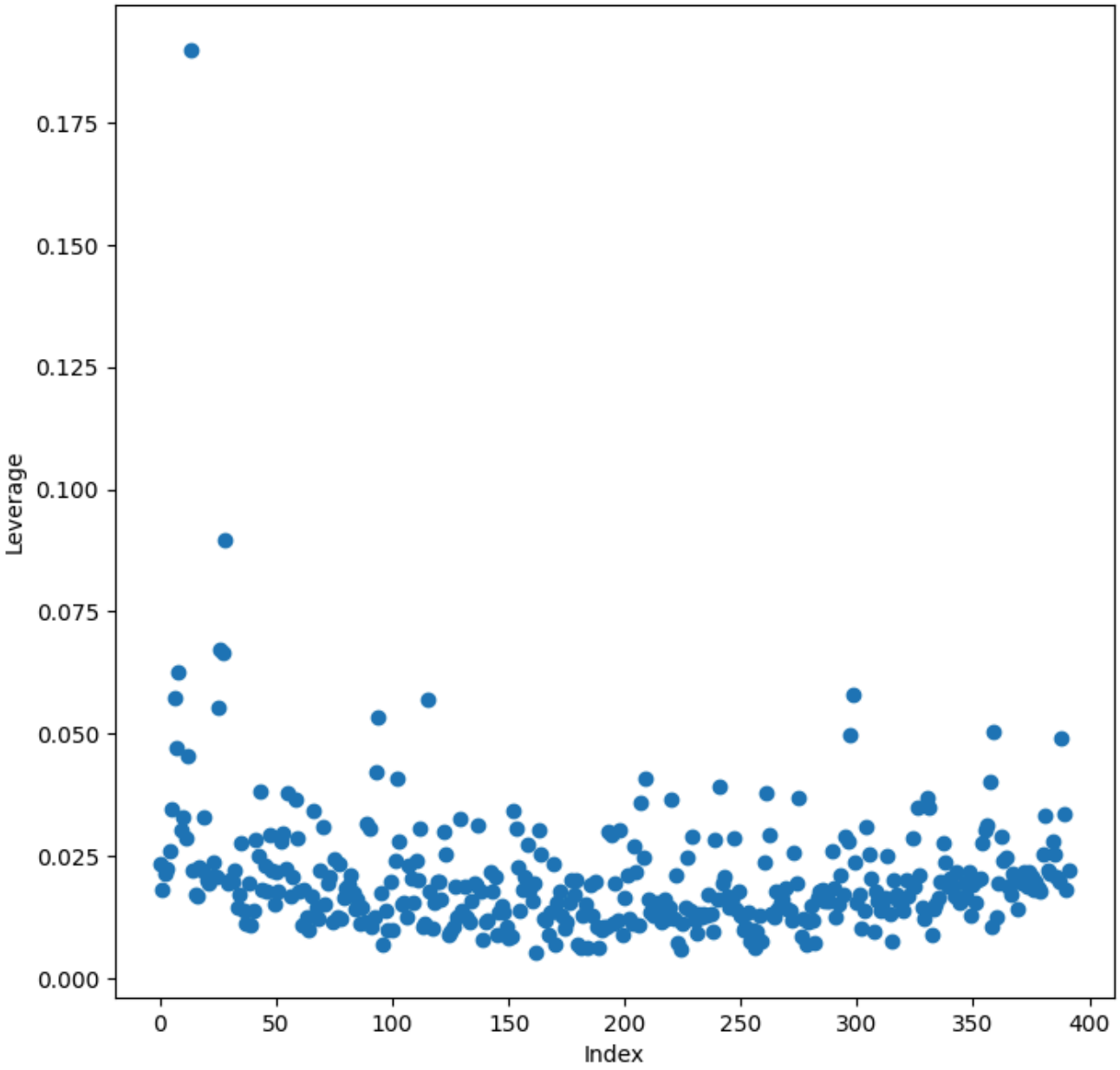
9c) Some of these predictors have a t-value greater than 0.025, showing insignificance.
i)Since the p-value in anova_lm is near 0, we can conclude the bigger model is superior and that there seems to be a relationship between the response and the predictors.
ii)Only displacement, weight,year, and origin are statistically significant.
iii)The coefficient for "year" suggests that each increase in (single) car model year, holding all other predictors constant, the mpg for a car increases by 0.7508.\

In [ ]:
```
#9d
ax = subplots(figsize=(8,8))[1]
ax.scatter(model.fit().fittedvalues , model.fit().resid)
ax.set_xlabel('Fitted value')
ax.set_ylabel('Residual')
ax.axhline(0, c='k', ls='--');
```



In [ ]:
```
infl=model.fit().get_influence()
ax = subplots(figsize=(8,8))[1]
ax.scatter(np.arange(X.shape[0]), infl.hat_matrix_diag)
ax.set_xlabel('Index')
ax.set_ylabel('Leverage')
np.argmax(infl.hat_matrix_diag)
```

Out[ ]: 13



9d)I notice that there are some residuals of -10 or 10 (showing high outliers), and that there is one extremely high leverage point.

In [ ]:
```
#9e
#After some changing of predictors (based on collinearity), I found this model whe
final1 = allvars + [('horsepower','acceleration'),('displacement','weight','cylind
X1 = MS(final).fit_transform(Auto)
model1 = sm.OLS(y, X1)
model1.fit().summary()
```

Out[ ]:

OLS Regression Results

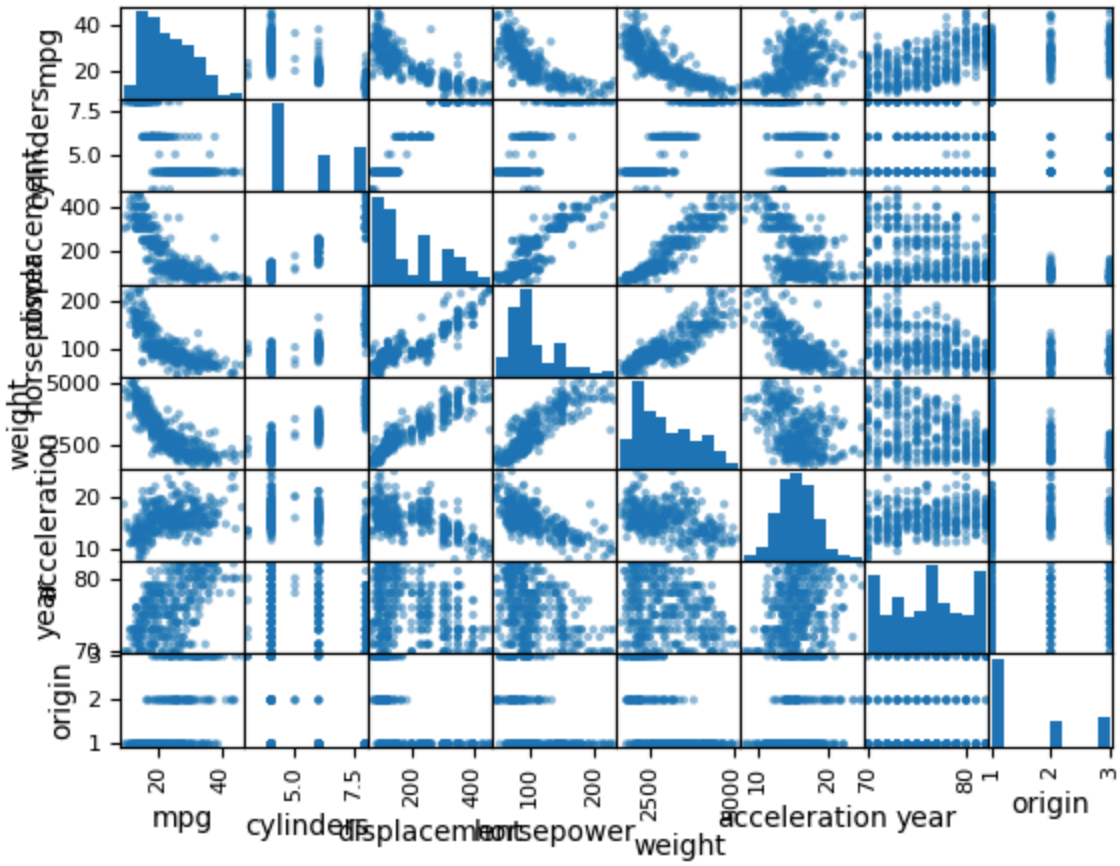
Dep. Variable:	mpg	**R-squared:**	0.821
Model:	OLS	**Adj. R-squared:**	0.818
Method:	Least Squares	**F-statistic:**	252.4
Date:	Tue, 18 Feb 2025	**Prob (F-statistic):**	2.04e-139
Time:	00:14:23	**Log-Likelihood:**	-1023.5
No. Observations:	392	**AIC:**	2063.
Df Residuals:	384	**BIC:**	2095.
Df Model:	7		
Covariance Type:	nonrobust		

	coef	std err	t	P>|t|	[0.025	0.975]
intercept	-17.2184	4.644	-3.707	0.000	-26.350	-8.087
cylinders	-0.4934	0.323	-1.526	0.128	-1.129	0.142
displacement	0.0199	0.008	2.647	0.008	0.005	0.035
horsepower	-0.0170	0.014	-1.230	0.220	-0.044	0.010
weight	-0.0065	0.001	-9.929	0.000	-0.008	-0.005
acceleration	0.0806	0.099	0.815	0.415	-0.114	0.275
year	0.7508	0.051	14.729	0.000	0.651	0.851
origin	1.4261	0.278	5.127	0.000	0.879	1.973

Omnibus:	31.906	**Durbin-Watson:**	1.309
Prob(Omnibus):	0.000	**Jarque-Bera (JB):**	53.100
Skew:	0.529	**Prob(JB):**	2.95e-12
Kurtosis:	4.460	**Cond. No.**	8.59e+04

Notes:
[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.
[2] The condition number is large, 8.59e+04. This might indicate that there are strong multicollinearity or other numerical problems.

In [ ]:
```
pd.plotting.scatter_matrix(Auto);
```

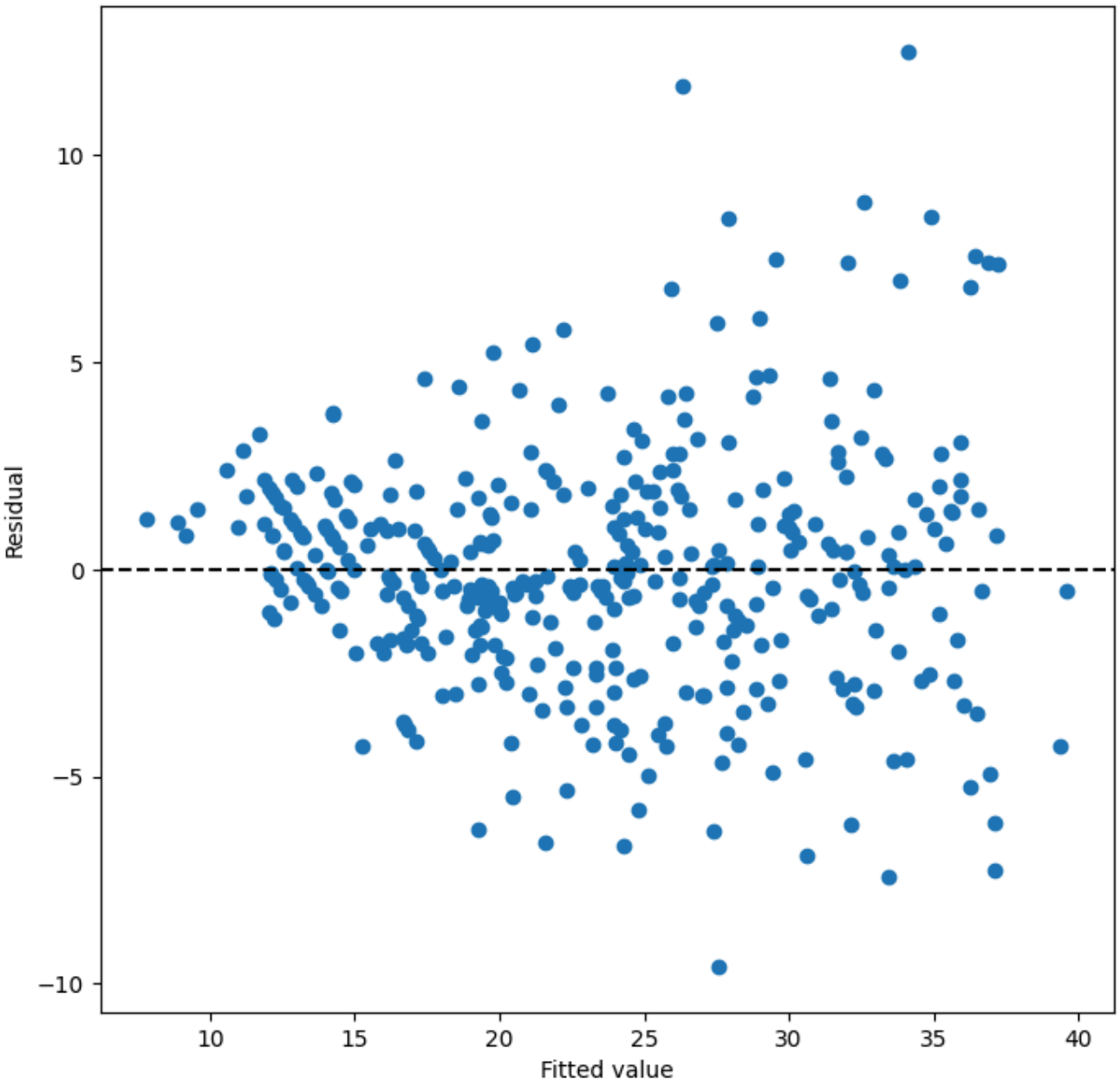


In [ ]:
```
print(Auto.corr())
```

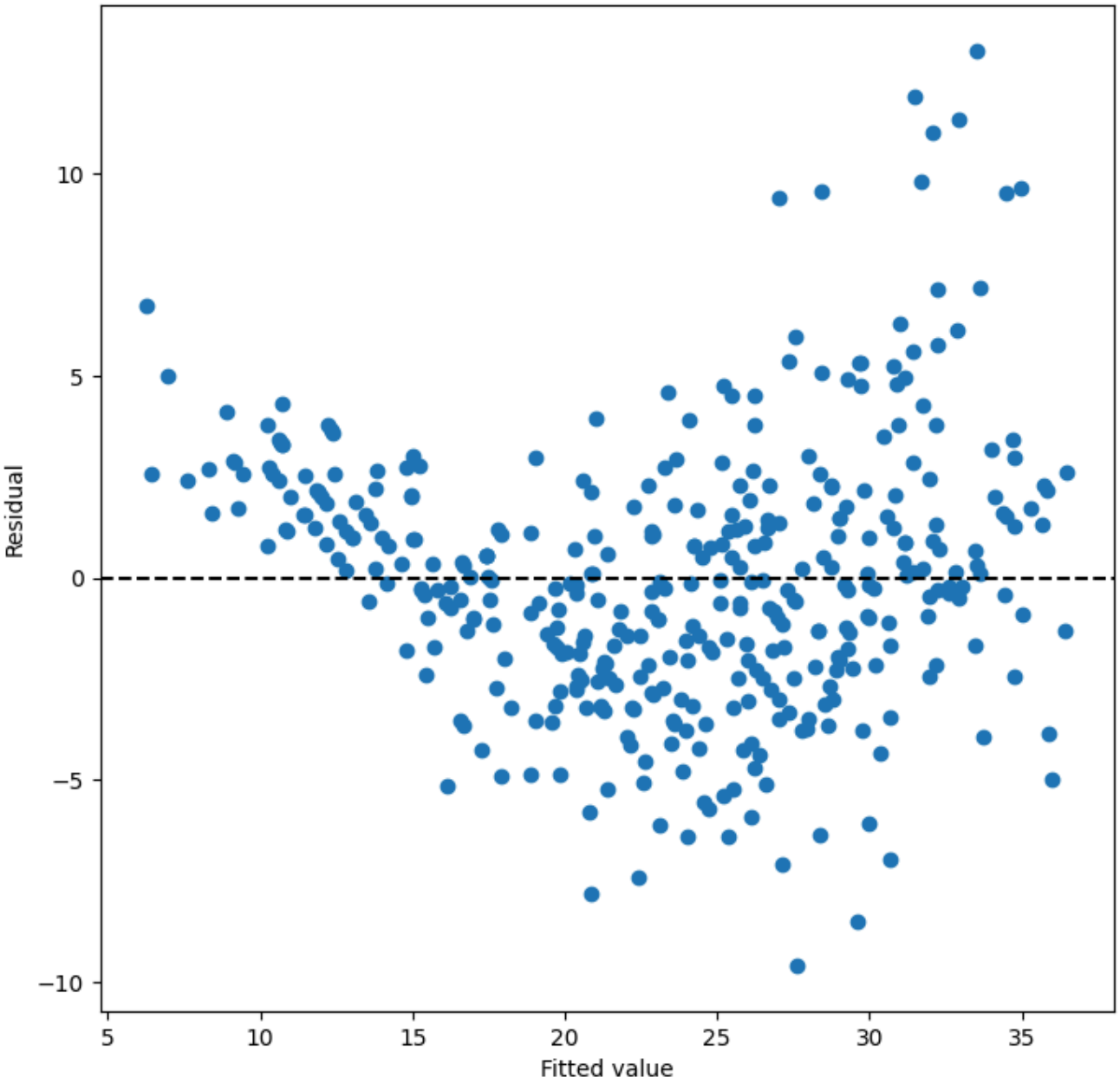
```
                     mpg  cylinders  displacement  horsepower    weight  \
mpg             1.000000  -0.777618     -0.805127   -0.778427 -0.832244
cylinders      -0.777618   1.000000      0.950823    0.842983  0.897527
displacement   -0.805127   0.950823      1.000000    0.897257  0.932994
horsepower     -0.778427   0.842983      0.897257    1.000000  0.864538
weight         -0.832244   0.897527      0.932994    0.864538  1.000000
acceleration    0.423329  -0.504683     -0.543800   -0.689196 -0.416839
year            0.580541  -0.345647     -0.369855   -0.416361 -0.309120
origin          0.565209  -0.568932     -0.614535   -0.455171 -0.585005

               acceleration      year    origin
mpg                0.423329  0.580541  0.565209
cylinders         -0.504683 -0.345647 -0.568932
displacement      -0.543800 -0.369855 -0.614535
horsepower        -0.689196 -0.416361 -0.455171
weight            -0.416839 -0.309120 -0.585005
acceleration       1.000000  0.290316  0.212746
year               0.290316  1.000000  0.181528
origin             0.212746  0.181528  1.000000
```

In [37]:
```
#9f
Autonew=Auto
Autonew['transdisplacement']=1/(Auto['displacement'])
Autonew['transcylinders']=1/(Auto['cylinders'])
Autonew['transweight']=-1/(Auto['weight'])
Autonew['transacceleration']=1/(Auto['acceleration'])
Autonew['transhorsepower']=-1/(Auto['horsepower'])
Autonew['transorigin']=np.log(Auto['origin'])
Autonew['transyear']=np.log(Auto['year'])
Autonew=Autonew.drop(['displacement','cylinders','weight','acceleration','horsepow
```

```
y2 = Autonew['mpg']
allvars2 = list(Autonew.columns.drop(['mpg']))
final2 = allvars2
X2 = MS(final2).fit_transform(Autonew)
modelnew = sm.OLS(y2, X2)
ax = subplots(figsize=(8,8))[1]
ax.scatter(modelnew.fit().fittedvalues , modelnew.fit().resid)
ax.set_xlabel('Fitted value')
ax.set_ylabel('Residual')
ax.axhline(0, c='k', ls='--');
#This was the best I could do, but I was able to clump the data together.
```



In [ ]:
```
ax = subplots(figsize=(8,8))[1]
ax.scatter(model.fit().fittedvalues , model.fit().resid)
ax.set_xlabel('Fitted value')
ax.set_ylabel('Residual')
ax.axhline(0, c='k', ls='--');
```


10
5
0
−5
−10
Residual
5
10
15
20
25
30
35
Fitted value