jupyter linearreg Last Checkpoint: 6 days ago

File Edit View Run Kernel Settings Help Trusted

Code JupyterLab Python 3 (ipykernel)

[10]:
```
import numpy as np # linear algebra
import pandas as pd
```

[11]:
```
import matplotlib.pyplot as plt
from sklearn.linear_model import LinearRegression
```

[12]:
```
pip install matplotlib
```

```
Requirement already satisfied: matplotlib in c:\users\admin\appdata\local\programs\python\python311\lib\site-packages (3.8.0)
Requirement already satisfied: contourpy>=1.0.1 in c:\users\admin\appdata\local\programs\python\python311\lib\site-packages (from matplotlib) (1.1.1)
Requirement already satisfied: cycler>=0.10 in c:\users\admin\appdata\local\programs\python\python311\lib\site-packages (from matplotlib) (0.12.1)
Requirement already satisfied: fonttools>=4.22.0 in c:\users\admin\appdata\local\programs\python\python311\lib\site-packages (from matplotlib) (4.43.1)
Requirement already satisfied: kiwisolver>=1.0.1 in c:\users\admin\appdata\local\programs\python\python311\lib\site-packages (from matplotlib) (1.4.5)
Requirement already satisfied: numpy<2,>=1.21 in c:\users\admin\appdata\local\programs\python\python311\lib\site-packages (from matplotlib) (1.26.0)
Requirement already satisfied: packaging>=20.0 in c:\users\admin\appdata\local\programs\python\python311\lib\site-packages (from matplotlib) (23.2)
Requirement already satisfied: pillow>=6.2.0 in c:\users\admin\appdata\local\programs\python\python311\lib\site-packages (from matplotlib) (10.0.0)
Requirement already satisfied: pyparsing>=2.3.1 in c:\users\admin\appdata\local\programs\python\python311\lib\site-packages (from matplotlib) (3.1.1)
Requirement already satisfied: python-dateutil>=2.7 in c:\users\admin\appdata\local\programs\python\python311\lib\site-packages (from matplotlib) (2.8.2)
Requirement already satisfied: six>=1.5 in c:\users\admin\appdata\local\programs\python\python311\lib\site-packages (from python-dateutil>=2.7->matplotlib)
(1.16.0)
Note: you may need to restart the kernel to use updated packages.
```

[13]:
```
pip install scikit-learn
```

```
Requirement already satisfied: scikit-learn in c:\users\admin\appdata\local\programs\python\python311\lib\site-packages (1.4.0)Note: you may need to restar
t the kernel to use updated packages.

Requirement already satisfied: numpy>=1.19.5 in c:\users\admin\appdata\local\programs\python\python311\lib\site-packages (from scikit-learn) (1.26.0)
Requirement already satisfied: scipy>=1.6.0 in c:\users\admin\appdata\local\programs\python\python311\lib\site-packages (from scikit-learn) (1.11.3)
Requirement already satisfied: joblib>=1.2.0 in c:\users\admin\appdata\local\programs\python\python311\lib\site-packages (from scikit-learn) (1.3.2)
Requirement already satisfied: threadpoolctl>=2.0.0 in c:\users\admin\appdata\local\programs\python\python311\lib\site-packages (from scikit-learn) (3.2.0)
```

[14]:
```
import matplotlib.pyplot as plt
from sklearn.linear_model import LinearRegression
```

[15]:
```
df = pd.read_csv('Salary_Data.csv')
df.head()
```

[15]:

	YearsExperience	Salary
0	1.1	39343
1	1.3	46205
2	1.5	37731
3	2.0	43525

```
4              2.2   39891
```

[16]:
```
X = df.iloc[:, :-1].values    # Features => Years of experience => Independent Variable # X=df[["YearsExperience"]]
y = df.iloc[:, -1].values     # Target => Salary => Dependent Variable

X
```

[16]:
```
array([[ 1.1],
       [ 1.3],
       [ 1.5],
       [ 2. ],
       [ 2.2],
       [ 2.9],
       [ 3. ],
       [ 3.2],
       [ 3.2],
       [ 3.7],
       [ 3.9],
       [ 4. ],
       [ 4. ],
       [ 4.1],
       [ 4.5],
       [ 4.9],
       [ 5.1],
       [ 5.3],
       [ 5.9],
       [ 6. ],
       [ 6.8],
       [ 7.1],
       [ 7.9],
       [ 8.2],
       [ 8.7],
       [ 9. ],
       [ 9.5],
       [ 9.6],
       [10.3],
       [10.5]])
```

[17]:
```
y
```

[17]:
```
array([ 39343,  46205,  37731,  43525,  39891,  56642,  60150,  54445,
        64445,  57189,  63218,  55794,  56957,  57081,  61111,  67938,
        66029,  83088,  81363,  93940,  91738,  98273, 101302, 113812,
       109431, 105582, 116969, 112635, 122391, 121872], dtype=int64)
```

[18]:
```
# divide the dataset in some amount of training and testing data
from sklearn.model_selection import train_test_split

# random_state => seed value used by random number generator
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.3, random_state=0)
```

[19]:
```
from sklearn.linear_model import LinearRegression
model = LinearRegression()
model.fit(X_train, y_train)
```

[19]:
```
LinearRegression
```

```
LinearRegression()
```

[22]:
```
predictions = model.predict(X_test)
predictions
```

[22]:
```
array([ 40817.78327049, 123188.08258899,  65154.46261459,  63282.41035735,
       115699.87356004, 108211.66453108, 116635.89968866,  64218.43648597,
        76386.77615802])
```

[23]:
```
y_test
```

[23]:
```
array([ 37731, 122391,  57081,  63218, 116969, 109431, 112635,  55794,
        83088], dtype=int64)
```

[24]:
```
pip install seaborn
```

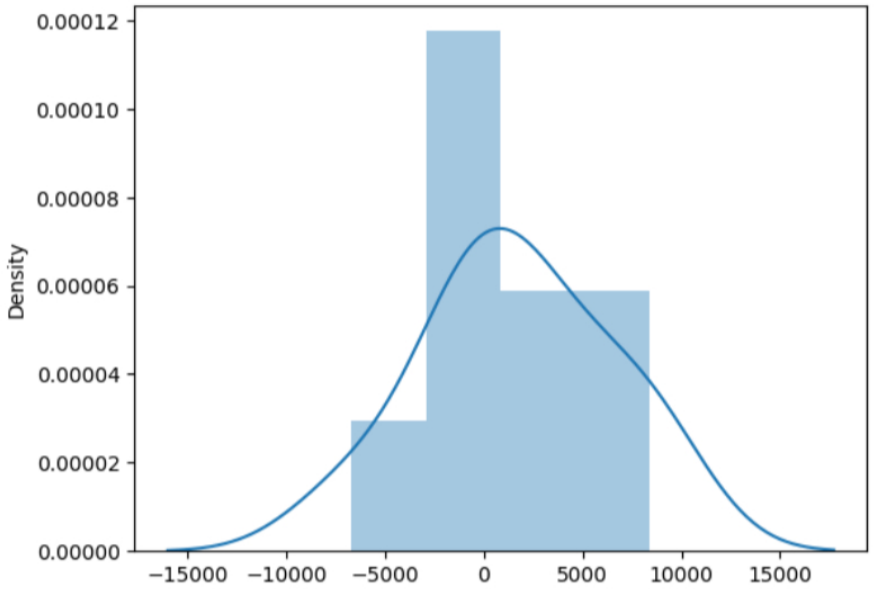
```
Requirement already satisfied: seaborn in c:\users\admin\appdata\local\programs\python\python311\lib\site-packages (0.13.1)
Requirement already satisfied: numpy!=1.24.0,>=1.20 in c:\users\admin\appdata\local\programs\python\python311\lib\site-packages (from seaborn) (1.26.0)
Requirement already satisfied: pandas>=1.2 in c:\users\admin\appdata\local\programs\python\python311\lib\site-packages (from seaborn) (2.1.4)
Requirement already satisfied: matplotlib!=3.6.1,>=3.4 in c:\users\admin\appdata\local\programs\python\python311\lib\site-packages (from seaborn) (3.8.0)
Requirement already satisfied: contourpy>=1.0.1 in c:\users\admin\appdata\local\programs\python\python311\lib\site-packages (from matplotlib!=3.6.1,>=3.4->
seaborn) (1.1.1)
Requirement already satisfied: cycler>=0.10 in c:\users\admin\appdata\local\programs\python\python311\lib\site-packages (from matplotlib!=3.6.1,>=3.4->seab
orn) (0.12.1)
Requirement already satisfied: fonttools>=4.22.0 in c:\users\admin\appdata\local\programs\python\python311\lib\site-packages (from matplotlib!=3.6.1,>=3.4-
>seaborn) (4.43.1)
Requirement already satisfied: kiwisolver>=1.0.1 in c:\users\admin\appdata\local\programs\python\python311\lib\site-packages (from matplotlib!=3.6.1,>=3.4-
>seaborn) (1.4.5)
Requirement already satisfied: packaging>=20.0 in c:\users\admin\appdata\local\programs\python\python311\lib\site-packages (from matplotlib!=3.6.1,>=3.4->s
eaborn) (23.2)
Requirement already satisfied: pillow>=6.2.0 in c:\users\admin\appdata\local\programs\python\python311\lib\site-packages (from matplotlib!=3.6.1,>=3.4->sea
born) (10.0.0)
Requirement already satisfied: pyparsing>=2.3.1 in c:\users\admin\appdata\local\programs\python\python311\lib\site-packages (from matplotlib!=3.6.1,>=3.4->
seaborn) (3.1.1)
Requirement already satisfied: python-dateutil>=2.7 in c:\users\admin\appdata\local\programs\python\python311\lib\site-packages (from matplotlib!=3.6.1,>=
3.4->seaborn) (2.8.2)
Requirement already satisfied: pytz>=2020.1 in c:\users\admin\appdata\local\programs\python\python311\lib\site-packages (from pandas>=1.2->seaborn) (2023.
3.post1)
Requirement already satisfied: tzdata>=2022.1 in c:\users\admin\appdata\local\programs\python\python311\lib\site-packages (from pandas>=1.2->seaborn) (202
3.4)
Requirement already satisfied: six>=1.5 in c:\users\admin\appdata\local\programs\python\python311\lib\site-packages (from python-dateutil>=2.7->matplotlib!
=3.6.1,>=3.4->seaborn) (1.16.0)
Note: you may need to restart the kernel to use updated packages.
```

[25]:
```
import seaborn as sns
sns.distplot(predictions-y_test)
```

```
C:\Users\ADMIN\AppData\Local\Temp\ipykernel_26332\468288966.py:2: UserWarning:

`distplot` is a deprecated function and will be removed in seaborn v0.14.0.

Please adapt your code to use either `displot` (a figure-level function with
similar flexibility) or `histplot` (an axes-level function for histograms).

For a guide to updating your code to use the new functions, please see
https://gist.github.com/mwaskom/de44147ed2974457ad6372750bbe5751
```

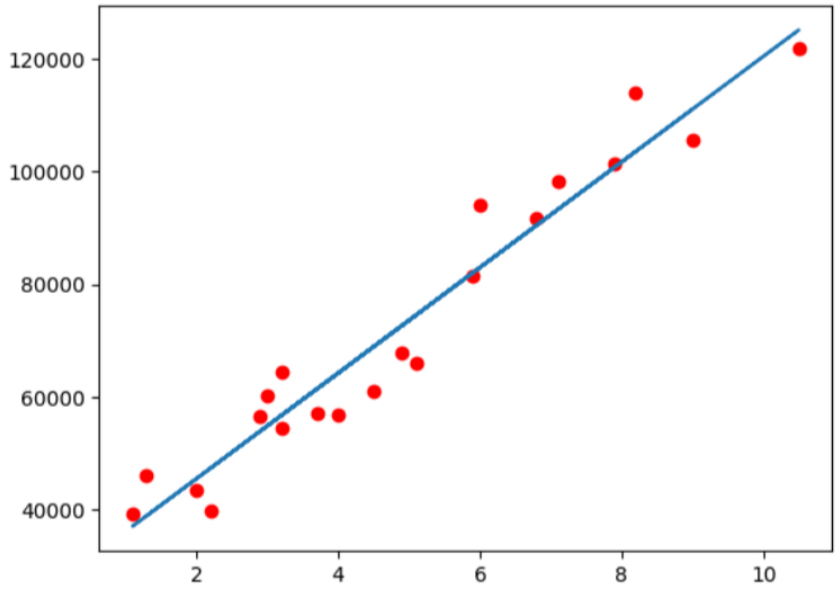
```
sns.distplot(predictions-y_test)
```

[25]:
```
<Axes: ylabel='Density'>
```



[18]:
```
plt.scatter(X_train, y_train, color='red')
plt.plot(X_train, model.predict(X_train))
```

[18]:
```
[<matplotlib.lines.Line2D at 0x1e148bf2190>]
```



[ ]: