education-2

February 21, 2024

[4]:
```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
from sklearn.model_selection import train_test_split
from sklearn.linear_model import LinearRegression
from sklearn.metrics import mean_squared_error
```

[5]:
```
data=pd.read_csv('/content/student.csv')
print(data.head(10))
```

```
   s.no   Attendance  gender study habits socioeconomic background  \
0     1           90  female         high                     high
1     2           80    male       medium                   medium
2     3           85  female       medium                     high
3     4           70    male       medium                     high
4     5           60  female          low                      low
5     6           65    male          low                   medium
6     7           80  female       medium                      low
7     8           55  female          low                     high
8     9           75    male       medium                      low
9    10           80  female       medium                     high

   teacher experience(years)  class size school resources  math score  \
0                          4          25             high          59
1                          5          30           medium          96
2                          7          35           medium          57
3                          6          40           medium          70
4                          2          20              low          83
5                          7          45              low          68
6                          4          40           medium          82
7                          6          55              low          46
8                          2          30           medium          80
9                          5          35           medium          57

   reading score  writing score  Average(%)
0             70             78          69
1             93             87          92
2             76             77          70
```

3	70	63	68
4	85	86	85
5	57	54	59
6	83	80	62
7	61	58	55
8	75	73	76
9	69	77	68

[6]:
```
X = data[['reading score']]
y = data['writing score']
```

[7]:
```
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2,␣
↪random_state=42)
```

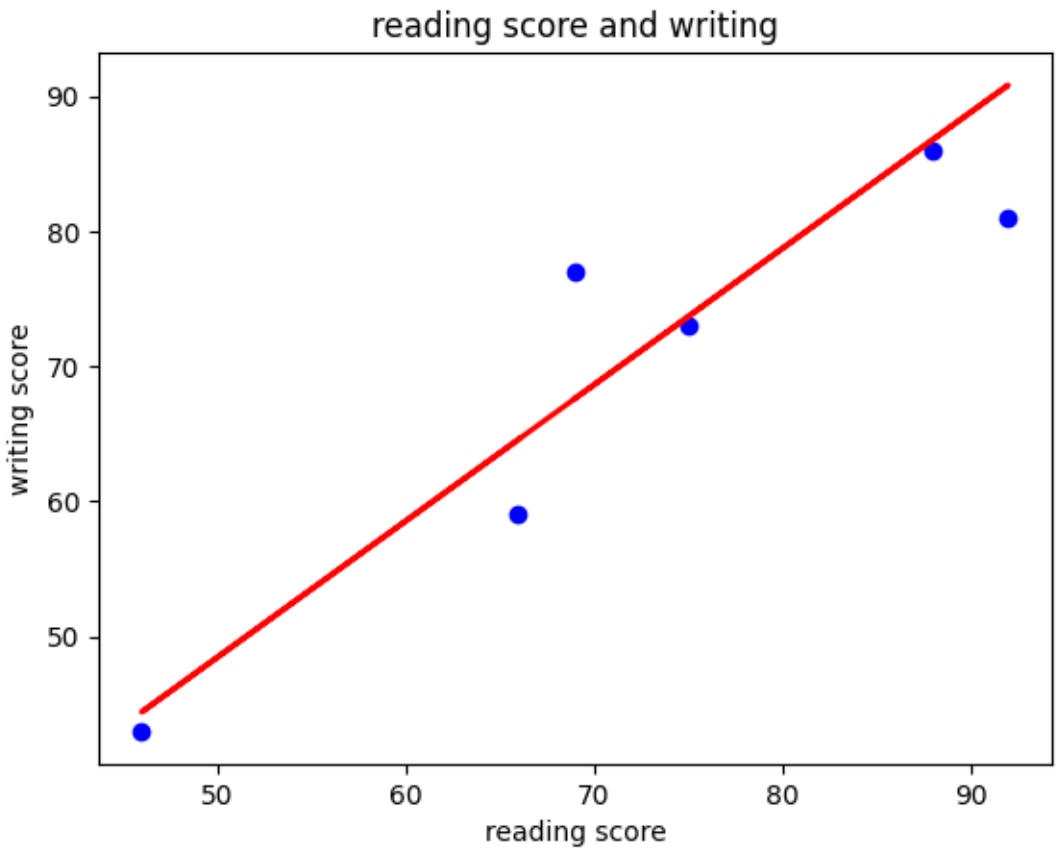
[8]:
```
model = LinearRegression()
model.fit(X_train, y_train)
```

[8]:
```
LinearRegression()
```

[9]:
```
y_pred = model.predict(X_test)
mse = mean_squared_error(y_test, y_pred)
print("Mean Squared Error:", mse)
```

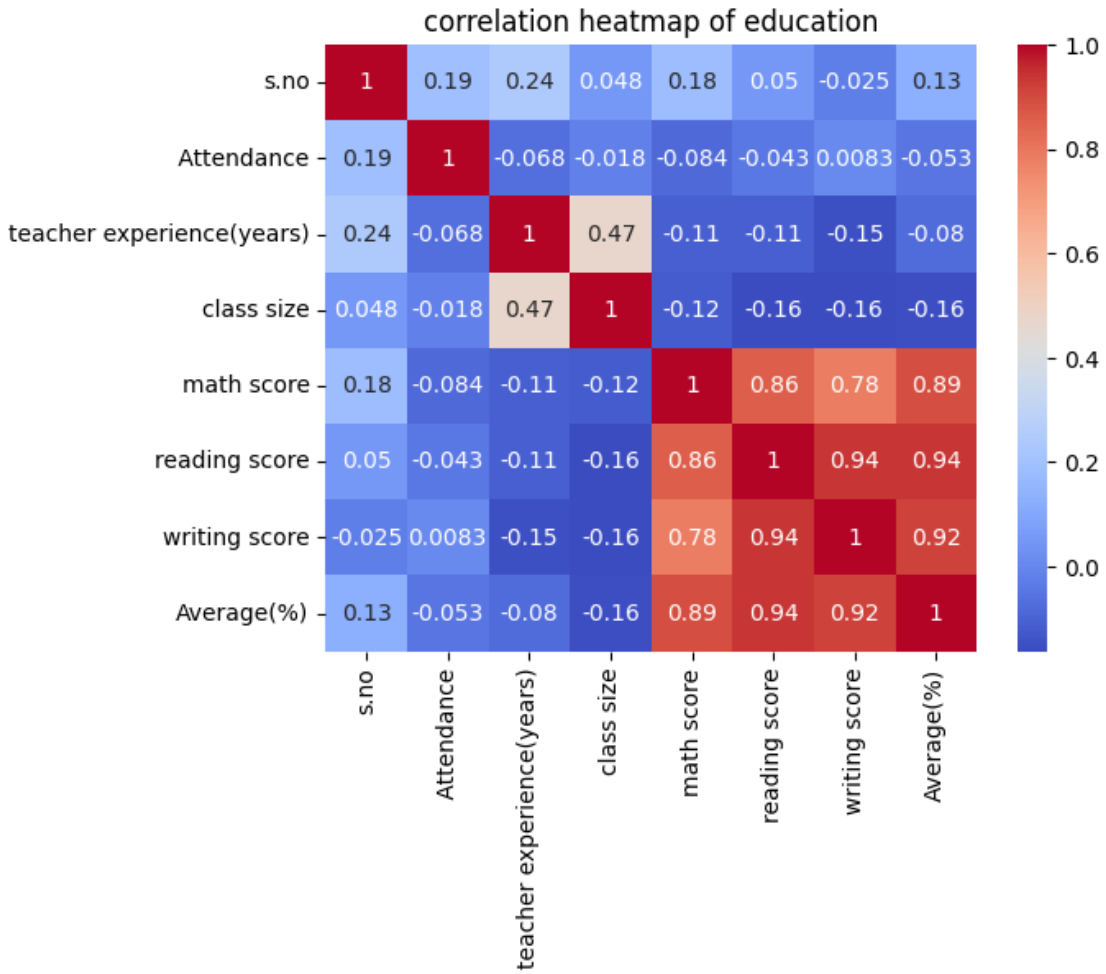
```
Mean Squared Error: 36.39560217326471
```

[10]:
```
plt.scatter(X_test, y_test, color='blue')
plt.plot(X_test, y_pred, color='red', linewidth=2)
plt.xlabel('reading score')
plt.ylabel('writing score')
plt.title('reading score and writing')
plt.show()
```



[11]:
```
import seaborn as sns
correlation_matrix=data.corr()
sns.heatmap(correlation_matrix,annot=True,cmap='coolwarm')
plt.title('correlation heatmap of education')
plt.show()
```

```
<ipython-input-11-54953a91aac3>:2: FutureWarning: The default value of
numeric_only in DataFrame.corr is deprecated. In a future version, it will
default to False. Select only valid columns or specify the value of numeric_only
to silence this warning.
  correlation_matrix=data.corr()
```



[12]:
```
from sklearn.linear_model import LinearRegression
Average= float(input("Enter the Average(%)"))
X_test = [[Average]]
predicted_student_grade= model.predict(X_test)
print("predicted student grade:", predicted_student_grade )
if(predicted_student_grade>=91):
    print("A+ grade Performance")
elif(predicted_student_grade>80 and predicted_student_grade<=90):
    print("A grade Performance")
elif(predicted_student_grade>70 and predicted_student_grade<=80):
    print("B grade Performance")
elif(predicted_student_grade>60 and predicted_student_grade<=70):
    print("C grade Performance")
elif(predicted_student_grade>50 and predicted_student_grade<=60):
    print("D grade Performance")
else:
```

```
    print("Bad Performance")
```

```
Enter the Average(%)80
predicted student grade: [78.69723709]
B grade Performance
```

```
/usr/local/lib/python3.10/dist-packages/sklearn/base.py:439: UserWarning: X does
not have valid feature names, but LinearRegression was fitted with feature names
  warnings.warn(
```