# 1 data availability

# 2 separating independent and dependent

# 3 identifying algorithms

# 4 training

# 5 evaluation

import matplotlib.pyplot as plt

import numpy as np

import seaborn as sns

import pandas as pd

import statsmodels.api as sm #https://www.statsmodels.org/stable/generated/statsmodels.regression.linear\_model.OLS.html#statsmodels.regression.linear\_model.OLS

aa=pd.read\_csv('/content/slr.csv')   # step 1

aa.head()

| **Exam** | **GPA** |
| --- | --- |
| **0** | 1714 | 2.40 |
| **1** | 1664 | 2.52 |
| **2** | 1760 | 2.54 |
| **3** | 1685 | 2.74 |
| **4** | 1693 | 2.83 |

# define independent and dependent variable       # step 2

x1=aa['Exam']  #independent

y=aa['GPA']    #dependent

sns.set()      # easy to understand background grid lines                                  # step 4

plt.scatter(x1,y)  # show scatter

x=sm.add\_constant(x1) # create to constant value(1)         # step 3

model=sm.OLS(y,x) # creating a model by using indepentant and dependent through OLS Methode      # step 5

result=model.fit() #traing the model           # step 6

result.summary()                               # step 6 result

|  |  |  |  |
| --- | --- | --- | --- |
| OLS Regression Results | | | |
| **Dep. Variable:** | GPA | **R-squared:** | 0.406 |
| **Model:** | OLS | **Adj. R-squared:** | 0.399 |
| **Method:** | Least Squares | **F-statistic:** | 56.05 |
| **Date:** | Fri, 15 Mar 2024 | **Prob (F-statistic):** | 7.20e-11 |
| **Time:** | 04:40:15 | **Log-Likelihood:** | 12.672 |
| **No. Observations:** | 84 | **AIC:** | -21.34 |
| **Df Residuals:** | 82 | **BIC:** | -16.48 |
| **Df Model:** | 1 |  |  |
| **Covariance Type:** | nonrobust |  |  |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **coef** | **std err** | **t** | **P>|t|** | **[0.025** | **0.975]** |
| **const** | 0.2750 | 0.409 | 0.673 | 0.503 | -0.538 | 1.088 |
| **Exam** | 0.0017 | 0.000 | 7.487 | 0.000 | 0.001 | 0.002 |

|  |  |  |  |
| --- | --- | --- | --- |
| **Omnibus:** | 12.839 | **Durbin-Watson:** | 0.950 |
| **Prob(Omnibus):** | 0.002 | **Jarque-Bera (JB):** | 16.155 |
| **Skew:** | -0.722 | **Prob(JB):** | 0.000310 |
| **Kurtosis:** | 4.590 | **Cond. No.** | 3.29e+04 |

Notes:  
[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.  
[2] The condition number is large, 3.29e+04. This might indicate that there are  
strong multicollinearity or other numerical problems.

yhat=0.275+0.0017\*x1  # y=c+mx