Lab 1 Linear Regression

No.	Program

1	`# Import tools`
	`import numpy as np`
	`import matplotlib.pyplot as plt`
	`import pandas as pd`

2	`# Import the dataset`
	`dataset = pd.read_csv('gaji.csv')`
	`dataset.head()`

3	`# split menjadi data dan target`
	`X = dataset.iloc[:, :-1].values # data`
	`y = dataset.iloc[:, 1].values # target`

4	`# Split dataset kea Training set and Test set`
	`from sklearn.model_selection import train_test_split`
	`X_train, X_test, y_train, y_test = train_test_split(X, y, test_size = 1/3,`
	`random_state = 0)`

5	`# Proses training/fit`
	`from sklearn.linear_model import LinearRegression`
	`regressor = LinearRegression()`
	`regressor.fit(X_train, y_train)`

6	`# Prediksi` `pred = regressor.predict(X_test)`

7	`# Visualisasi hasil training`
	`plt.scatter(X_train, y_train, color = 'red')`
	`plt.plot(X_train, regressor.predict(X_train), color = 'blue')`

8	`# Visualising hasil tes prediksi`
	`plt.scatter(X_test, y_test, color = 'red')`
	`plt.plot(X_train, regressor.predict(X_train), color = 'blue')`
	`plt.title('Salary vs Experience (Test set)')`
	`plt.xlabel('Years of Experience')`
	`plt.ylabel('Salary')`
	`plt.show()`

9	`# evaluasi`
	`# R-Squared, mendekati 1 lebih baik`
	`from sklearn.metrics import r2_score`
	`r2_score(y_test, pred)`

10	`# MSE, semakin kecil semakin baik`
	`from sklearn.metrics import mean_squared_error`
	`mean_squared_error(y_test, pred)`

11	`# RMSE, nilai error rata2 asli`
	`from math import sqrt`
	`rmse = sqrt(mean_squared_error(y_test, pred))`
	`rmse`