Linear Regression on Randomly Created Dataset

```
import numpy as np
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, r2_score

# Step 1: Create a random dataset
np.random.seed(42) # for reproducibility
x=2* np.random.rand(100, 1)  # 100 random points for the feature
y=4 + 3 * x + np.random.randn(100, 1) # Linear relationship with some
noise

x
```

```
array([[0.74908024],
       [1.90142861],
       [1.46398788],
       [1.19731697],
       [0.31203728],
       [0.31198904],
       [0.11616722],
       [1.73235229],
       [1.20223002],
       [1.41614516],
       [0.04116899],
       [1.9398197 ],
       [1.66488528],
       [0.42467822],
       [0.36364993],
       [0.36680902],
       [0.60848449],
       [1.04951286],
       [0.86389004],
       [0.58245828],
       [1.22370579],
       [0.27898772],
       [0.5842893 ],
       [0.73272369],
       [0.91213997],
       [1.57035192],
       [0.39934756],
       [1.02846888],
       [1.18482914],
       [0.09290083],
```

```
       [1.2150897 ],
       [0.34104825],
       [0.13010319],
       [1.89777107],
       [1.93126407],
       [1.6167947 ],
       [0.60922754],
       [0.19534423],
       [1.36846605],
       [0.88030499],
       [0.24407647],
       [0.99035382],
       [0.06877704],
       [1.8186408 ],
       [0.51755996],
       [1.32504457],
       [0.62342215],
       [1.04013604],
       [1.09342056],
       [0.36970891],
       [1.93916926],
       [1.55026565],
       [1.87899788],
       [1.7896547 ],
       [1.19579996],
       [1.84374847],
       [0.176985  ],
       [0.39196572],
       [0.09045458],
       [0.65066066],
       [0.77735458],
       [0.54269806],
       [1.65747502],
       [0.71350665],
       [0.56186902],
       [1.08539217],
       [0.28184845],
       [1.60439396],
       [0.14910129],
       [1.97377387],
       [1.54448954],
       [0.39743136],
       [0.01104423],
       [1.63092286],
       [1.41371469],
       [1.45801434],
       [1.54254069],
       [0.1480893 ],
       [0.71693146],
```

```
       [0.23173812],
       [1.72620685],
       [1.24659625],
       [0.66179605],
       [0.1271167 ],
       [0.62196464],
       [0.65036664],
       [1.45921236],
       [1.27511494],
       [1.77442549],
       [0.94442985],
       [0.23918849],
       [1.42648957],
       [1.5215701 ],
       [1.1225544 ],
       [1.54193436],
       [0.98759119],
       [1.04546566],
       [0.85508204],
       [0.05083825],
       [0.21578285]])
```

y

```
array([[ 6.33428778],
       [ 9.40527849],
       [ 8.48372443],
       [ 5.60438199],
       [ 4.71643995],
       [ 5.29307969],
       [ 5.82639572],
       [ 8.67878666],
       [ 6.79819647],
       [ 7.74667842],
       [ 5.03890908],
       [10.14821022],
       [ 8.46489564],
       [ 5.7873021 ],
       [ 5.18802735],
       [ 6.06907205],
       [ 5.12340036],
       [ 6.82087644],
       [ 6.19956196],
       [ 4.28385989],
       [ 7.96723765],
       [ 5.09801844],
       [ 5.75798135],
       [ 5.96358393],
       [ 5.32104916],
       [ 8.29041045],
```

```
       [ 4.85532818],
       [ 6.28312936],
       [ 7.3932017 ],
       [ 4.68275333],
       [ 9.53145501],
       [ 5.19772255],
       [ 4.64785995],
       [ 9.61886731],
       [ 7.87502098],
       [ 8.82387021],
       [ 5.88791282],
       [ 7.0492748 ],
       [ 7.91303719],
       [ 6.9424623 ],
       [ 4.69751764],
       [ 5.80238342],
       [ 5.34915394],
       [10.20785545],
       [ 6.34371184],
       [ 7.06574625],
       [ 7.27306077],
       [ 5.71855706],
       [ 7.86711877],
       [ 7.29958236],
       [ 8.82697144],
       [ 8.08449921],
       [ 9.73664501],
       [ 8.86548845],
       [ 6.03673644],
       [ 9.59980838],
       [ 3.4686513 ],
       [ 5.64948961],
       [ 3.3519395 ],
       [ 7.50191639],
       [ 5.54881045],
       [ 5.30603267],
       [ 9.78594227],
       [ 4.90965564],
       [ 5.91306699],
       [ 8.56331925],
       [ 3.23806212],
       [ 8.99781574],
       [ 4.70718666],
       [10.70314449],
       [ 7.3965179 ],
       [ 3.87183748],
       [ 4.55507427],
       [ 9.18975324],
       [ 8.49163691],
```

```
       [ 8.72049122],
       [ 7.94759736],
       [ 4.67652161],
       [ 6.44386684],
       [ 3.98086294],
       [11.04439507],
       [ 8.21362168],
       [ 4.79408465],
       [ 5.03790371],
       [ 4.89121226],
       [ 6.73818454],
       [ 9.53623265],
       [ 7.00466251],
       [10.28665258],
       [ 7.24607048],
       [ 5.53962564],
       [10.17626171],
       [ 8.31932218],
       [ 6.61392702],
       [ 7.73628865],
       [ 6.14696329],
       [ 7.05929527],
       [ 6.90639808],
       [ 4.42920556],
       [ 5.47453181]])
```

Spliting

```
x_train, x_test,y_train,y_test=train_test_split(x,y,test_size=0.2,
random_state=42)

model-LinearRegression()
model.fit(X_train,y_train)

LinearRegression()

y_train_pred=model.predict(x_train)
y_test_pred=model.predict(x_test)

train_mse = mean_squared_error(y_train,y_train_pred)
test_mse = mean_squared_error(y_test,y_test_pred)
train_r2 = r2_score(y_train,y_train_pred)
test_r2 = r2_score(y_test,y_test_pred)

print("Training MSE:", train_mse)
print("Testing MSE:", test_mse)
print("Training R2 score:", train_r2)
print("Testing R2 score:", test_r2)

Training MSE: 0.8476788564209705
Testing MSE: 0.6536995137170021
```

```
Training R2 score: 0.7582381034538057
Testing R2 score: 0.8072059636181392
```

```
plt.figure(figsize=(11,6))
plt.scatter(x, y, color="blue", label="original data")
plt.plot(x, model.predict(x),color="red", linewidth=2,
label="regression line")
plt.xlabel("X(, feature)")
plt.ylabel("y, (target)")
plt.title("Linear Regression on random data")
plt.legend()
plt.grid(True)
plt.show()
```