Mini Project 3: Regression Dummy Dataset

Step 1: Import Library

```
import pandas as pd
```

```
import numpy as np
```

Step 2: Generating Dataset

```
from sklearn.datasets import make_regression
```

```
X, y = make_regression(n_samples=500, n_features=5, coef=False, bias=12, noise=10, random_state=2529)
```

```
X, y, z = make_regression(n_samples=500, n_features=5, coef=True, bias=12, noise=10, random_state=2529)
```

```
X.shape, y.shape
```

```
((500, 5), (500,))
```

```
z
```

```
array([29.45661718, 60.14529878, 61.7409438 , 13.32437893, 99.08122896])
```

```
X.shape, y.shape
```

```
((500, 5), (500,))
```

```
X[0:5]
```

```
array([[ 0.77913208, -1.09701784, -0.14239962,  1.02427891, -1.0708024 ],
       [-0.6925009 ,  0.45535977,  0.34707569, -0.32456746,  0.21970203],
       [-0.03901601, -0.3265115 ,  0.59793721,  0.61686653, -0.6237489 ],
       [-0.61566117, -0.11782129, -0.98234619, -0.78292727,  0.42713048],
       [ 1.30822207, -0.72541559,  0.60187975,  0.33285998,  1.48506184]])
```

```
y[0:5]
```

```
array([-136.21858395,   49.83118244,  -29.81097858,  -31.74001475,
        193.0687778 ])
```

```
X.shape, y.shape
```

```
((500, 5), (500,))
```

Step 3: Splitting Data

```
from sklearn.model_selection import train_test_split
```

```
X_train, X_test, y_train, y_test = train_test_split(X,y, test_size = 0.3, random_state=2529)
```

```
X_train.shape, X_test.shape, y_train.shape, y_test.shape
```

```
((350, 5), (150, 5), (350,), (150,))
```

Step 4: Creating model

```
from sklearn.linear_model import LinearRegression

model = LinearRegression()
```

Step 5: Training Model

```
model.fit(X_train, y_train)
```

```
LinearRegression()
```

Step 6: Getting Intercepts and Coefficients

```
model.intercept_
```

```
12.804677404011848
```

```
model.coef_
```

```
array([30.14690156, 59.8508539 , 61.00591796, 13.33058614, 98.52732683])
```

Step 7: Prediction Model

```
y_pred = model.predict(X_test)
```

```
y_pred.shape
```

```
(150,)
```

```
y_pred
```

```
array([ -32.2347268 ,  -91.78838198, -111.32428609,   32.28900762,
       -292.09184678,  232.80381314,   45.45589568, -206.64663509,
       -234.70619469,  -99.53722469,  109.51230639, -112.24192419,
        136.86734982,    9.50169605, -402.91778394,  -77.75468931,
         31.16247075,  -11.08354812,   31.85126688,   37.4989766 ,
        -33.17312782,  165.96248357,  136.44084324,   43.71369762,
        169.91060611, -255.66815548,   -6.97645701, -115.23137387,
        -27.09548272,   -2.27310513,  -97.06342353,   57.62031761,
        185.7350271 ,  -60.45930909,   11.20091455,  176.09294043,
         90.16023475,  121.22117973,  -12.7823069 ,  -46.97799373,
         84.74765683,   38.71436232,  -71.35539323,   -1.53368019,
        -30.73878351,   54.71398149,  -21.39863976, -272.82340302,
       -118.65913112,  -38.95749518,  -56.94266543,  -64.97208225,
         83.15562417,   84.93116766, -118.46836522,  181.58821651,
        -49.51641057,  -21.16415589,   24.61247158,   15.50449997,
        177.48102965, -280.00659035,  132.68190996, -146.3719296 ,
        181.40142785,  -99.13298792,   39.59393464,  116.66353758,
        179.36615051,  156.10519607,   57.69154717,   29.13217725,
       -102.086819  ,  -96.43086036,  -48.54067383,  102.07730091,
        109.27743273,  157.25553456,  135.70953968,  -22.19144017,
       -204.84997999,  132.85622533,   40.9548816 ,   93.87274125,
        -38.68637545,   20.16165644, -173.4777815 ,   31.78904817,
        174.6024829 ,  143.73624186,  144.25571404,   50.3249105 ,
        -66.49817321,  -61.60857001,   55.56455718,  -24.66893769,
       -129.90063813,   60.13995635,  134.52686195,  128.62980994,
        301.91544779,   11.76874944,  -11.48619142, -319.75425702,
        -89.22401612,  -89.77736186,  -22.84820519,  -56.45869267,
         90.15710992,  166.38819788,  145.02106406,  133.01449822,
        178.06301411,   27.83443113, -229.56624865,  -13.87055036,
        -65.16788489,   55.579956  ,  158.30138329,   23.80733476,
        -44.61754504, -104.29641738,   57.66550889, -177.80605374,
```

```
       -74.28729277,  127.77641272,   90.26103721,  302.49937215,
       -49.64748714,  -47.52519963,  -77.40654164,  -40.02137064,
        11.9445304 , -114.7616647 ,  220.84662992, -105.83580108,
      -124.76236231, -188.66177139,  117.19088414, -101.22112675,
       124.60506733, -282.53782645,   73.67016676,  -12.16238081,
      -161.68545386,   44.68598417,  172.95607293,    3.54588201,
        38.89340344, -169.02005187])
```

Step 8: Accuracy

```
from sklearn.metrics import mean_squared_error, mean_absolute_error, mean_absolute_percentage_error, r2_score
```

```
mean_squared_error(y_test, y_pred)
```

```
93.17298608096334
```

```
mean_absolute_error(y_test, y_pred)
```

```
7.861503882499261
```

```
mean_absolute_percentage_error(y_test, y_pred)
```

```
0.5394385232604642
```

```
r2_score(y_test, y_pred)
```

```
0.9944221597296871
```

Link of the same:

https://colab.research.google.com/drive/1fBrCDUFl0UNMnsxnY8ANZ21unKKR72a4?usp=sharing