

▼ Dane z Google Trends

Dane zostały zebrane od 01 czerwca 2019 do 31 grudnia 2019, z godzinny interwałem = 5136

Pojedynczy wektor danych zawiera

- datę (dzień i godzina)
- atrybuty: popularność wyszukiwania 20 haseł z Google Trends (znormalizowane względem
- wartość bitcoina wyrażona w euro

Poniżej dla przykładu pierwsze 6 instancji.

```
date,facebook,google,youtube,bmw,amazon,iphone,apple,samsung,btc,twitter,netflix
2019-06-01 00:00:00,90.0,62.0,71.789474,4.486842,24.677631,17.62688,10.13039,13.
2019-06-01 01:00:00,89.0,60.0,72.328766,3.7671232,22.312962,15.937829,8.976479,1
2019-06-01 02:00:00,86.0,60.0,71.35135,3.7162163,23.84572,17.032657999999998,9.5
2019-06-01 03:00:00,83.0,60.0,70.540540000000002,3.7521565,21.069801,15.58588,8.4
2019-06-01 04:00:00,81.0,60.0,69.72973,3.7489102,19.898062,14.418885,7.389678999
2019-06-01 05:00:00,80.0,60.0,71.50685,3.8035557,17.93105,13.312446,6.742667,13.
```

▼ Przeprowadzone wstępne badania

Do wstępnego sprawdzenia zależności wydzielono zbiór treningowy i testowy z powyżej opisan instancji, czyli danych od czerwca do października włącznie. Zbiór testowy to instancje z dwóch

Regresja liniowa:

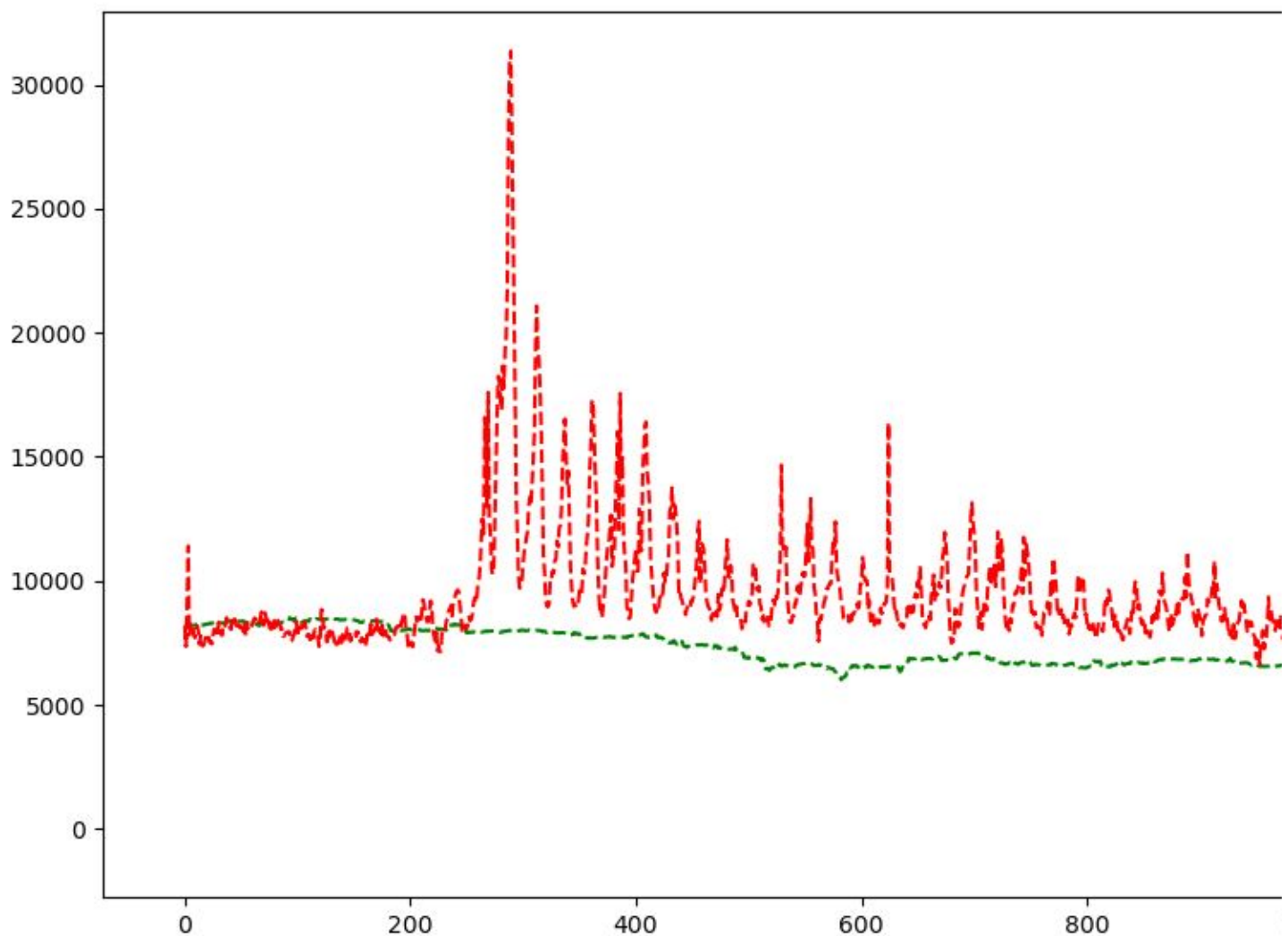
```
from sklearn import linear_model

def multiple_linear_regression(train, test):
    X_train, y_train = split_into_x_and_y(train)
    X_test, y_test = split_into_x_and_y(test)

    model = linear_model.LinearRegression()
    model.fit(X_train, y_train)
    print('Intercept: \n', model.intercept_)
    print('Coefficients: \n', model.coef_)

    y_predict = model.predict(X_test)
    return y_predict
```

Do porównania narysowano wykres z wartościami testowymi (kolor zielony) bitcoina oraz tymi czasie.



y - wartość bitcoina w euro x - numer próbek (po kolei w czasie)

Na pierwszy rzut oka wiadać, że po jakimś czasie (ponad 200 próbek) przewidywana wartość b historycznych kryptowaluty.

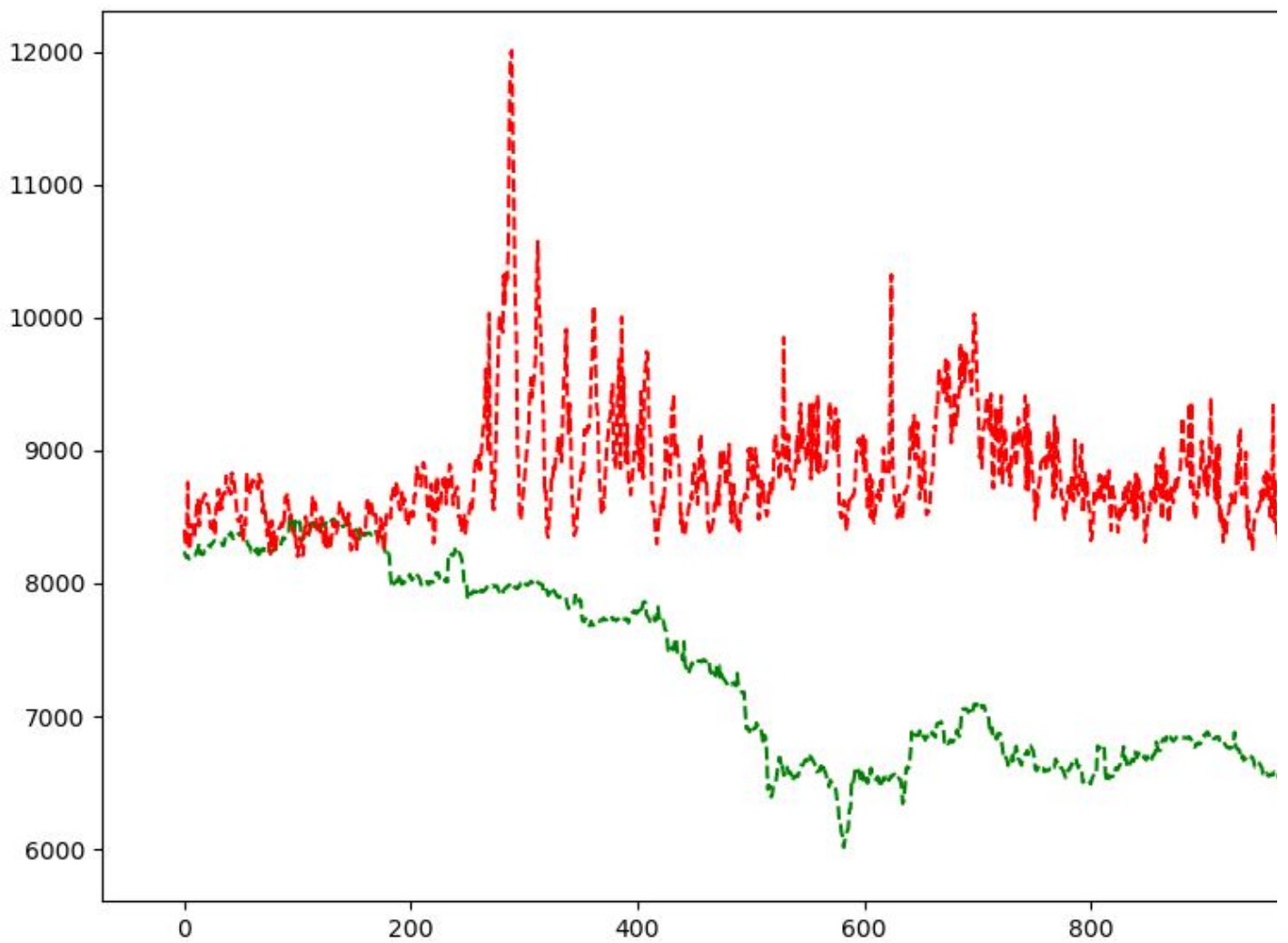
```
Intercept:
8403.023200987647
Coefficients:
[ 50.02985492 -7.26022611 -76.11379925 26.97418722
 14.17426255 -85.72554693 243.26735804 41.03217806 -
-29.69676632 548.7557647 -11.15847117 -53.28479878 -1
445.08359795 35.8871241 -441.72582551 -284.01626205 -5
```

Wypisano także parametry modelu:

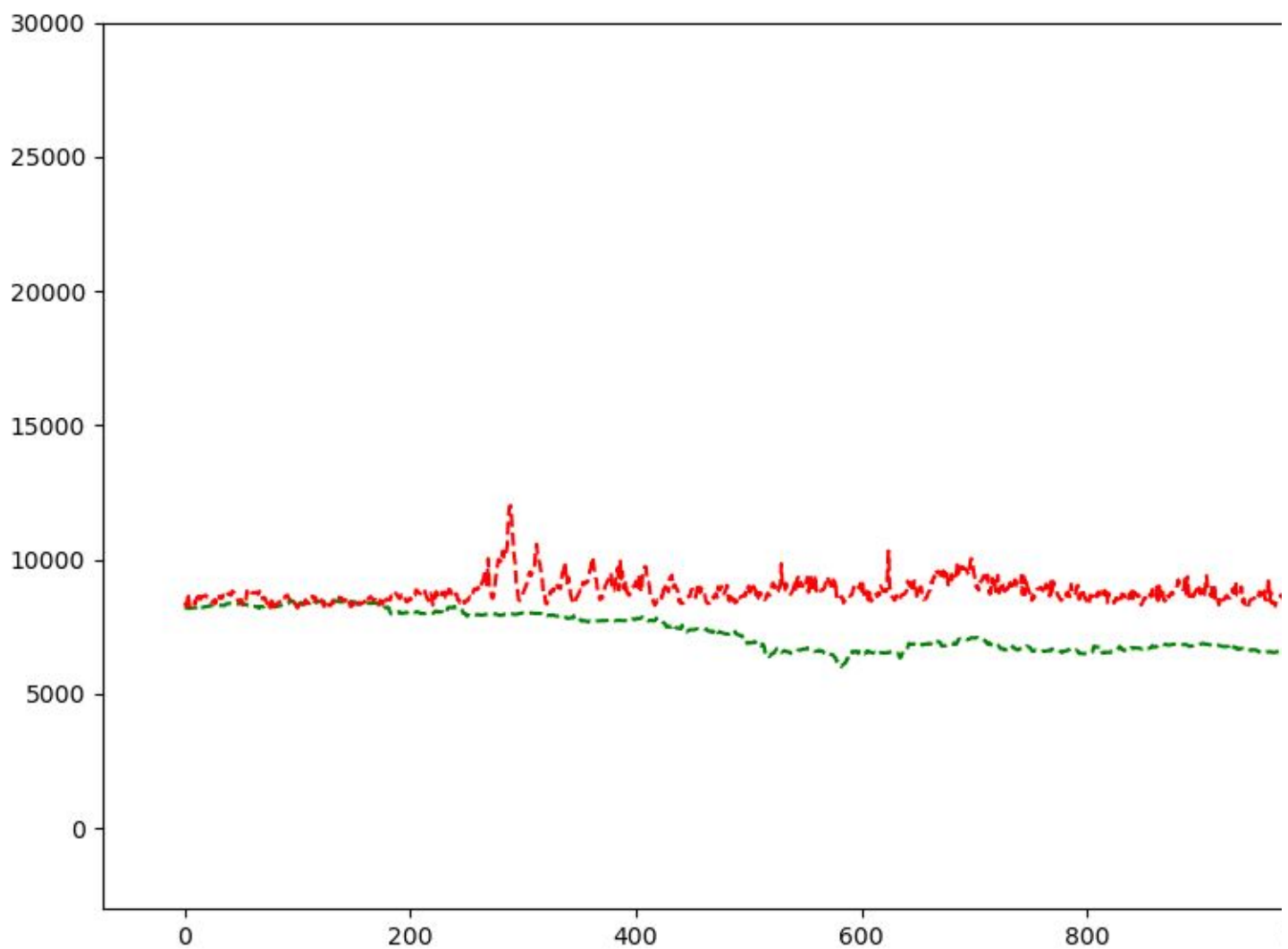
Czy otrzymane wyniki parametrów są dobre, ciężko mi ocenić.

Bazując na podstawowej wiedzy o współczynnikach przeprowadziłam kolejne pomiary. Ze zbioru współczynników były liczbą ujemną.

Otrzymałam wykres:



Ten sam wykres w takiej samej skali, jak z poprzedniego eksperymentu:

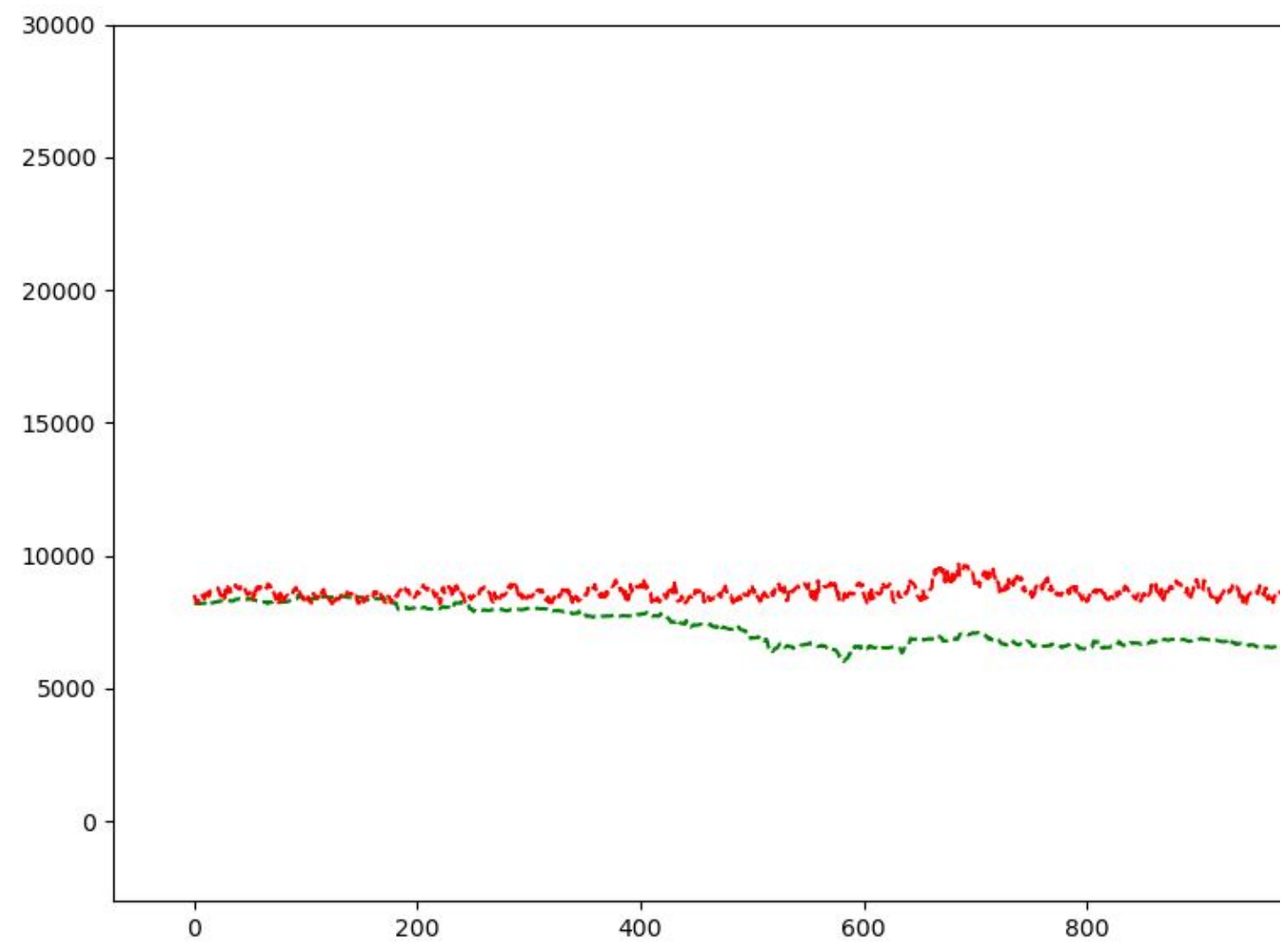


Chyba trochę lepiej.

Otrzymane parametry modelu w tym przypadku.

```
Intercept:  
6573.129066296269  
Coefficients:  
[ 11.72717786  24.27720521 -65.48777945 147.17337247 -44.1931787  
 83.36271538 -2.21617527 -27.79279183]  
|
```

Trzeci eksperyment: pozostawione jedynie atrybuty o indeksach: 0, 3, 7, 11 tj. facebook, bmw, s; pierwszego eksperymentu). Poniżej otrzymany wykres oraz parametry modelu



```
5772.8440402813885
Coefficients:
[ 18.5912208    8.87819815 100.36037656    4.82346864]
```

