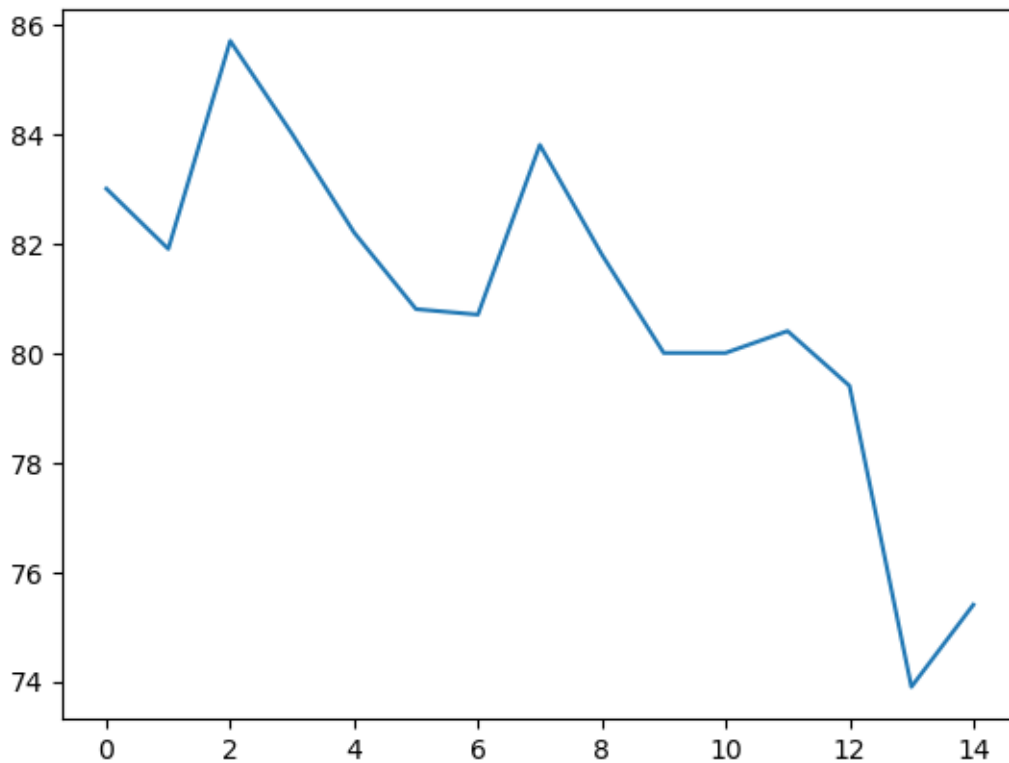# 300820241457

August 30, 2024

Occupazione per livello di istruzione in Valle d'Aosta dai dati oecd.

```python
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
from sklearn.linear_model import LinearRegression

df = pd.read_csv('oecd-va.csv')
df['OBS_VALUE'].plot()
```

[1]: <AxesSubplot: >

Regressione lineare $y = \beta_0 + \beta_1 x + \varepsilon$ per rilevare il tendenza.

```
[2]: model = LinearRegression()
     X = df['TIME_PERIOD'].values.reshape(-1, 1)
     y = df['OBS_VALUE'].values
     model.fit(X, y)
     plt.plot(df['TIME_PERIOD'], df['OBS_VALUE'], 'o')
     plt.plot(df['TIME_PERIOD'], model.predict(X), 'r')
     plt.xlabel('Data')
     plt.ylabel('Valore')
     plt.title('Regressione lineare')
     plt.show()
```


[ ]: