

Nama	: Adhimas Aryo Bimo
NIM	: 13523052
Model	: PCA

1. PCA merupakan teknik unsupervised untuk mereduksi dimensi data. Hal ini digunakan dengan mempertahankan informasi dalam hal ini varian sebanyak mungkin. Pada implementasi yang saya buat, saya menggunakan matriks kovarians yang merupakan matriks berisi varians. Pada perkuliahan, kita sempat diajari juga hal ini namun menggunakan metode SVD. Metode SVD dan matriks kovarians memiliki cara yang berbeda dalam menentukan eigenvalues dan eigenvector namun hasilnya mirip. Berikut langkah-langkah algoritma PCA:
  - a. Hitung mean dan pusatkan data
    - Tujuannya agar PCA fokus ke variasi, bukan ke posisi absolut data
  - b. Hitung matriks covarians
  - c. Hitung eigenvalues dan eigenvectors
  - d. Urutkan berdasarkan eigenvalues terbesar ke n komponen
  - e. Lalu transformasi

## 2. Berikut perbandingan antara kedua PCA

```
Custom PCA - Shape: (960, 2)
Sklearn PCA - Shape: (960, 2)
Explained Variance (Sklearn): [0.06492089 0.0618286 ]
Explained Variance (Custom): [0.06493249 0.06187601
0.05412557 0.05291852 0.05157233 0.05031413
0.04966853 0.04841004 0.04801879 0.04659844 0.04569713
0.04489845
0.04403721 0.04215603 0.04197984 0.04050279 0.03986843
0.03858368
0.03784353 0.03702494 0.03115339 0.02781974]
```

Secara output, dimensi yang dihasilkan sama. Perbedaannya hanya pada tampilan saja. Saya menampilkan semua (22) nilai variance rasionya sedangkan pada sklearn hanya 2. Perbedaannya juga tidak terlalu jauh antara 2 nilai variance ratio terbesar pada custom dengan sklearn