Principal Component Algorithm:

Setup: X is a dataset : $n \times p$ matrix.

Notation:

$E_{1:K}$ is the matrix created from E by discarding the last $n-K$ columns.

$$\begin{pmatrix} E \end{pmatrix} \quad \begin{pmatrix} E_{1:K} & \cdots \end{pmatrix}$$

$\uparrow$ discarded.

Step #1:

Center the matrix X by subtracting the column means.

Assumption: From now on, X is centered.

Step #2:

Compute the sample covariance matrix $\Omega = \frac{1}{n} X^t X$.

Step #3:

Compute the eigenvectors $\{e_1, e_2, \dots, e_p\}$
and eigenvalues $\{\lambda_1, \lambda_2, \dots, \lambda_p\}$ of Ω.

Return:

The matrix of eigenvectors $E = \overbrace{\begin{pmatrix} | & | & & | \\ e_1 & e_2 & \cdots & e_p \\ | & | & & | \end{pmatrix}}^{\text{each column is an eigenvector}}$

The vector of eigenvalues $\Gamma = (\lambda_1\ \lambda_2 \cdots \lambda_p)$

Properties:

(1) The projection of X onto the subspace spanned by $\{e_1, e_2, \dots, e_p\}$ preserves the maximum variance (out of all K-dim subspaces).

(2) The matrix product $X E_{1:K}$ gives the coordinates of X in the reduced basis $\{e_1, e_2, \dots, e_K\}$.

(3) The matrix product $(X E_{1:K}) E_{1:K}^t$ gives the best reconstruction of X using only K dimensions. (ie: the projection of X onto the "best" K-dim subspace)

(4) The sum of eigenvalues $\sum_{j=1}^{K} \lambda_j$ gives the total variance of the projection of X onto $E_{1:K}$

(5) The ratio of sums

$$\frac{\sum_{j=1}^{K} \lambda_j}{\sum_{j=1}^{p} \lambda_j}$$ gives the percentage of variance preserved when projecting X onto $E_{1:K}$.