Assignment.

Principal Component Analysis.

Aim: PCA Find principal components variance and standard deviation calculation of PC. (Using R)

Theory:

PCA is a method of extracting important variable (in form of component) form a large set of variables available in a dataset.

It extract low dimentional set of feature from a high diment data set with a as to capture as much information as possible. It is usefull for dealing with 3 or higher dimension.

Normalization.

The PC Principal Component are supplied with normalized version of original predictors.

Variance & covariance are a measure of the "spread" of a set of point around the center of mass.

Variance = (Standard Deviation)

~~Covariance~~ Covariance is a measure of how much each of the dimensions vary from the mean with respect to each other.

$$\text{Covariance}(x,y) = \sum_{i=1}^{n} \frac{(\bar{x}_i - X)\cdot(\bar{y}_i - Y)}{(n-1)}$$

Standard Deviation

$$\sigma = \sqrt{\frac{(x_i - mean)^2}{N}}$$

Conclusion

In this assignment we have studied and implement the Concept of PCA on the Big Mart sales dataset using R