

# ADVANCED STATISTICS ( Week 3)

# DSBA CURRICULUM DESIGN

## FOUNDATIONS

**Python For Data  
Science**

**Statistical Methods  
for Decision  
Making**

## CORE COURSES

**Advanced  
Statistics(Week-3/5)**

**Data Mining**

**Predictive  
Modelling**

**Machine Learning**

**Time Series  
Forecasting**

**Data Visualization**

**SQL**

## DOMAIN APPLICATIONS

**Financial Risk  
Analytics**

**Marketing Retail  
Analytics**



# LEARNING OBJECTIVE OF THIS MODULE

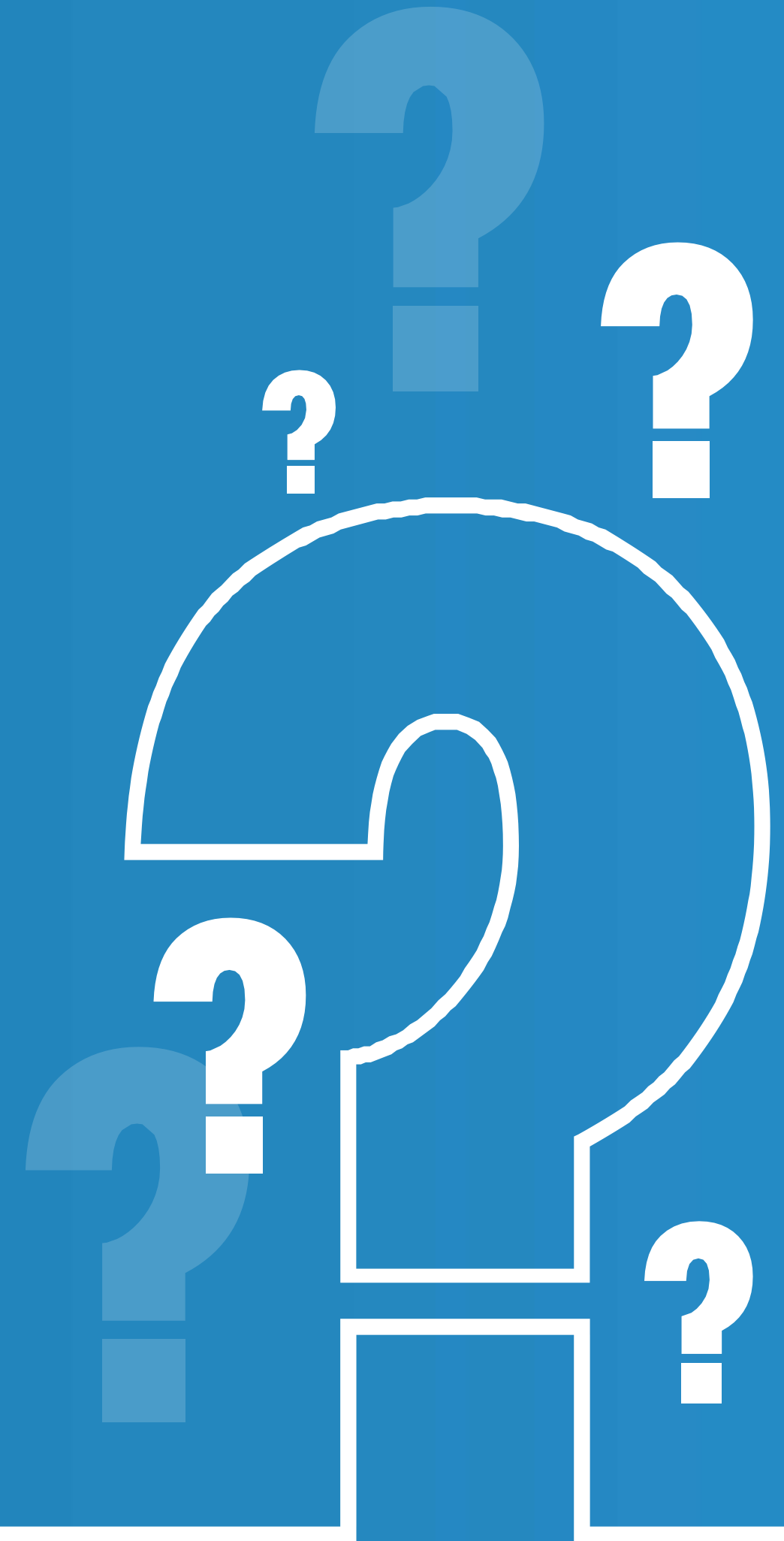
- ANOVA
- EDA
- **PCA**

# LEARNING OBJECTIVES OF THIS SESSION -

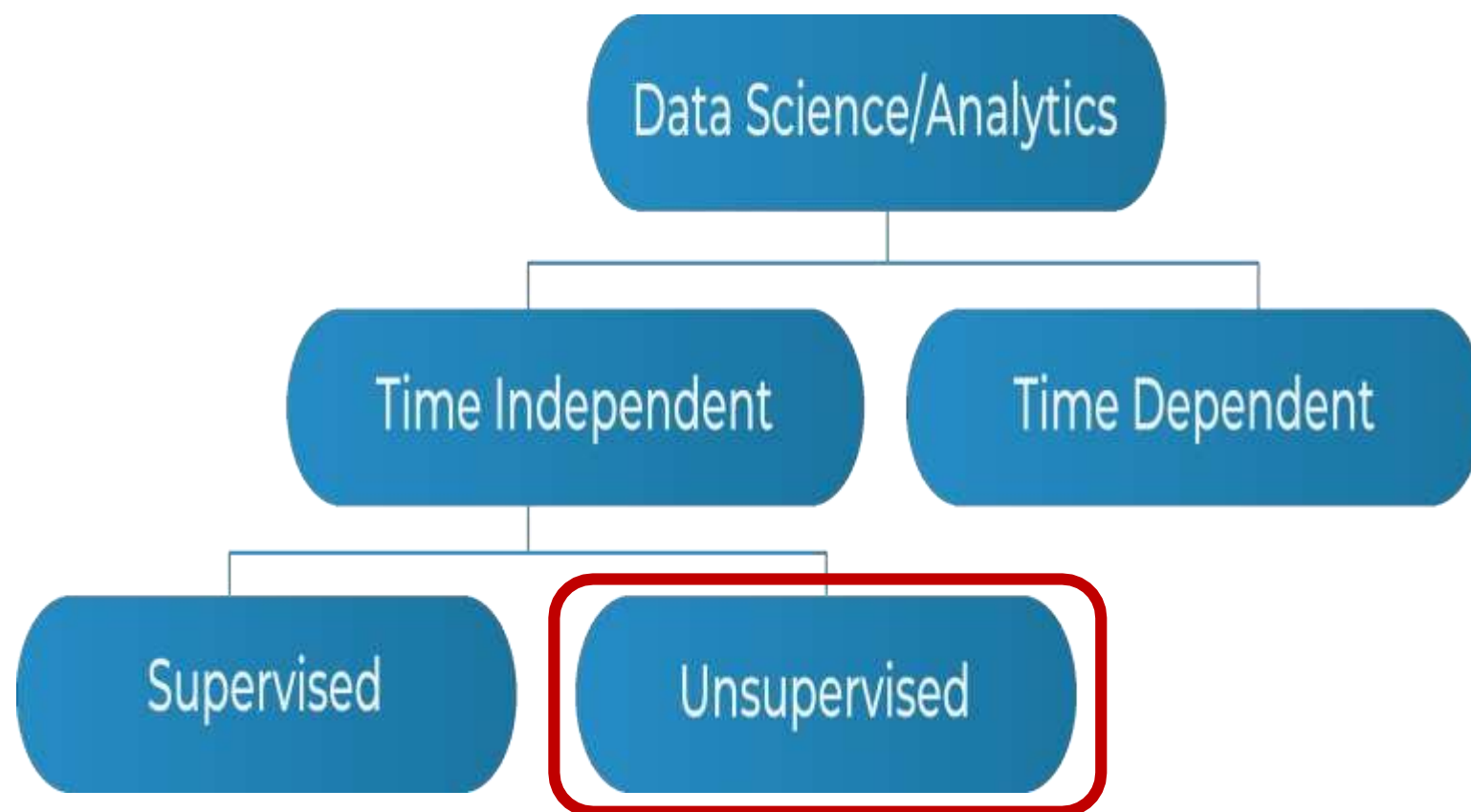
- Covariance Matrix
- Eigen Value
- Variance and Cumulative Variance
- Scree Plot

# TRY ANSWERING THE FOLLOWING

- Can we use PCA for dimensionality reduction?
- What is the purpose of using Scree Plot in PCA?
- Does different Principal Components orthogonal to each other?



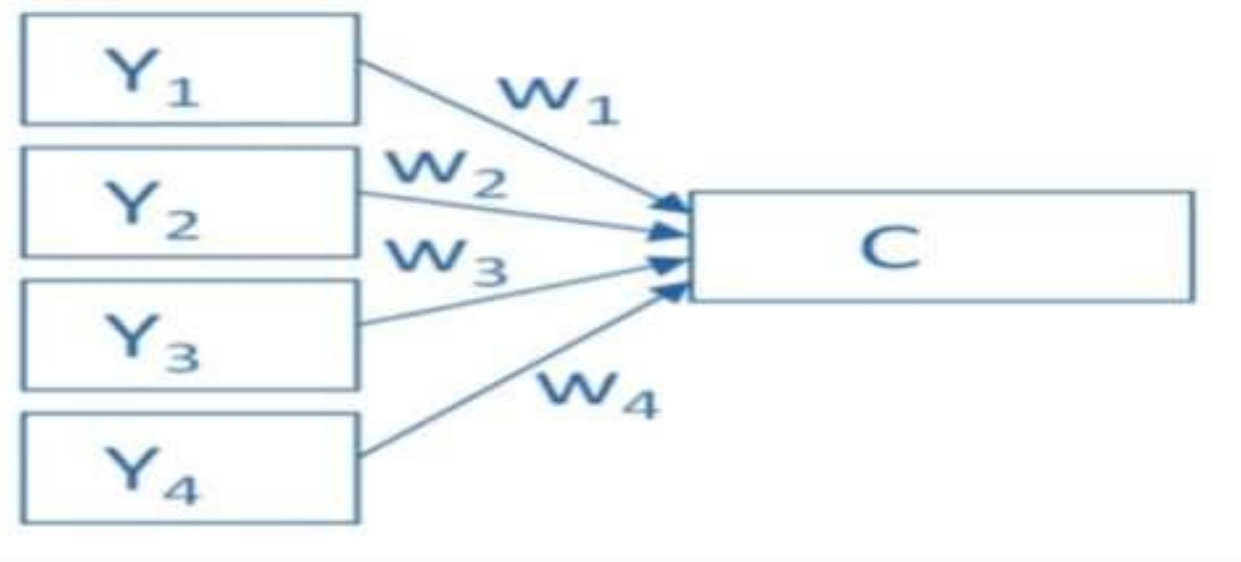
# BROAD OVERVIEW: Principal Component Analysis (PCA)



- Dimension reduction technique
- Pattern Recognition (Based on Features)
- Resolve Multicollinearity



# PCA Technique



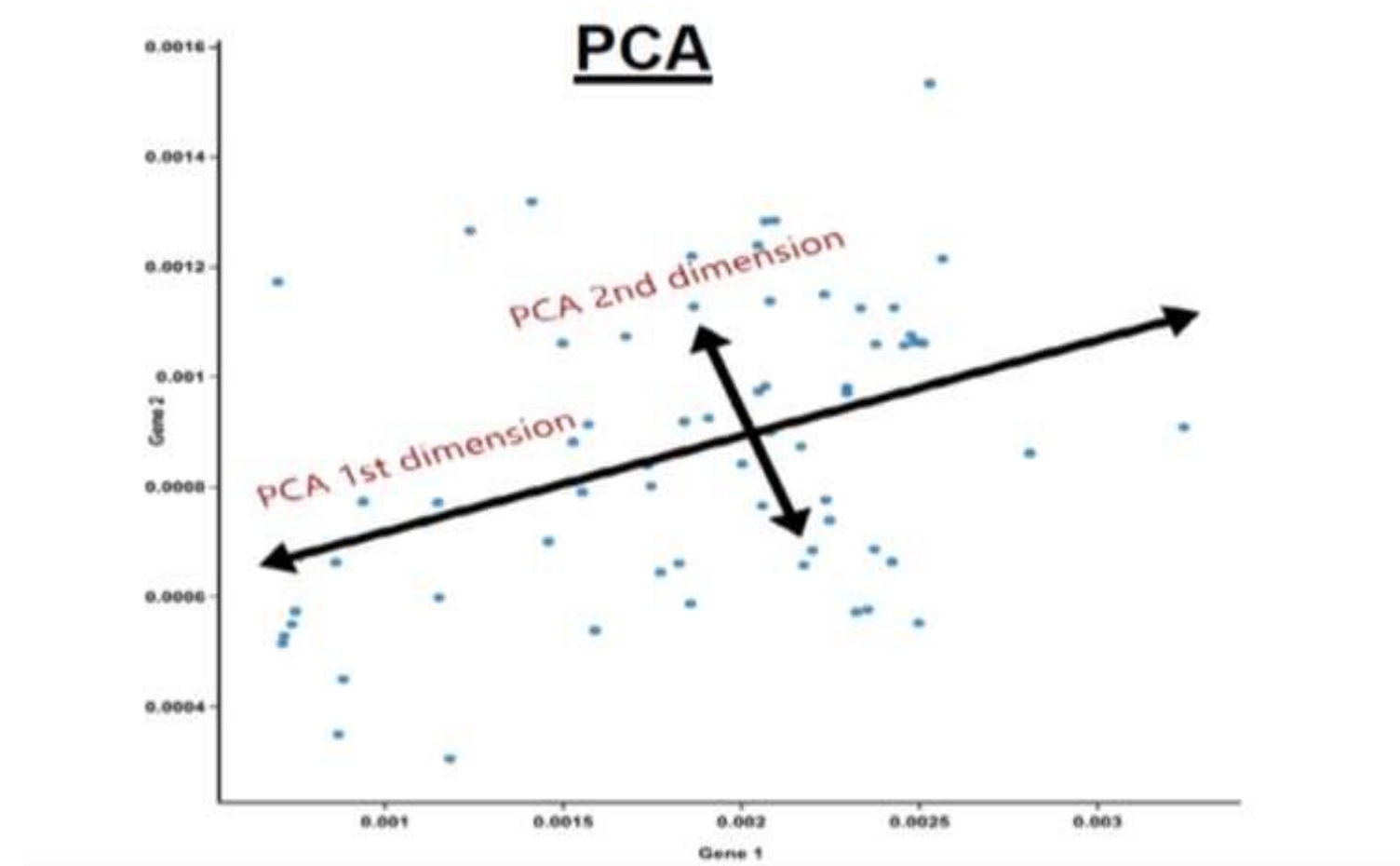
- *Linear combination of variables*

$$C = w_1(Y_1) + w_2(Y_2) + w_3(Y_3) + w_4(Y_4)$$

C: Component

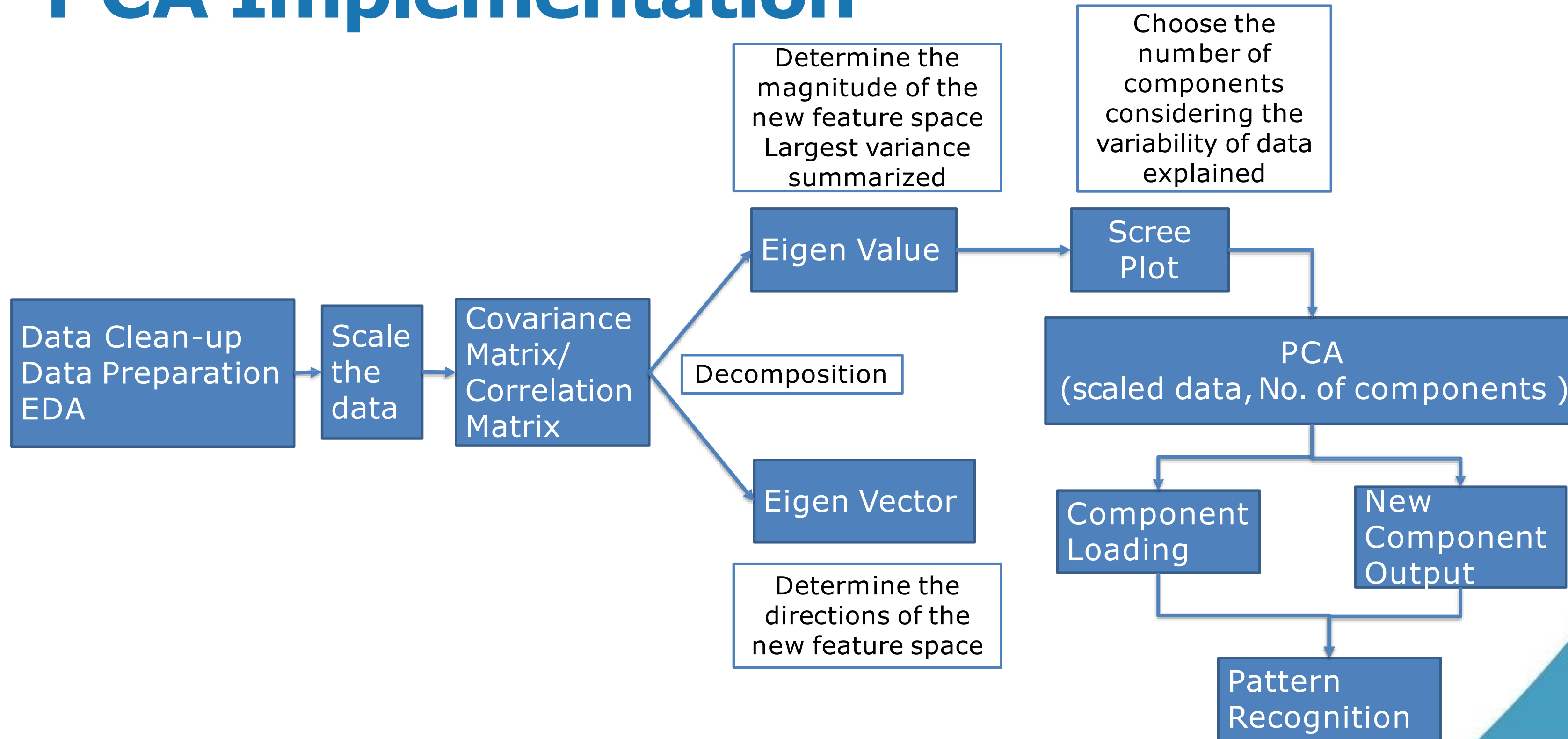
$w_1, w_2, w_3, w_4$ : pca components loadings

$Y_1, Y_2, Y_3, Y_4$ : Features



- PCA1 *accounts for as much variation* in the data as possible
- PCA2 accounts for as much as remaining variation in the data as possible, with the constraint that the *correlation between PCA1 and PCA2 is zero*
- All subsequent components have the same property

# PCA Implementation





# Let's Learn Together – A Unique Platform for Peer to Peer Learning

## Next Week's Theme:

### Mind-Map of Advance Statistics Course

Reference Link-<https://www.mindmapping.com/>



## Benefits of Creating Mind-Map:

- ❖ Quick Revision of the course
- ❖ This document will prove very handy later in the course
- ❖ Opportunity to show your creativity

## What all can be discussed in a Discussion forum?

- ❖ Analytical Concepts
- ❖ Issues in Code
- ❖ Real Time/Industry Examples



# Data Science @ Work

Apply **Data Science at your workplace** to gain some instant benefits:

- Get noticed by your management with your outstanding analysis backed by data science.
- Create an impact in your organization by taking up small projects/initiatives to solve critical issues using data science.
- Network with members from the data science vertical of your organization and seek opportunities to contribute in small projects.
- Share your success stories with us and the world to position yourself as a subject matter expert in data science.



# Case Study: PCA on Loan Transactions



Lending is one of the critical functions of any financial institution. Customers are provided loans across different products at competitive interest rates for an acceptable tenure. There is always a risk that a customer may

- default on the loan or
- may try and repay the loan in advance

which leads to financial losses to the business.

There may not be enough data or evidence available from the past which will help the firm to predict and mitigate the above possible risk.

However, it does have information about the customer demographics, loan details, EMI transactions etc. which may contribute to 30+ features related to a customer's loan account.

Let's apply PCA on this high dimension data, identify patterns and help the bank with some useful information to monitor the risk





**ANY QUESTIONS**

# Excelerate CareerPrep:

## Career services by Great Learning.

The CareerPrep is an organized way to learn about all of the topics asked in interviews for Data Science jobs.

- Complete Excelerate career prep and learn how to crack an interview with ease.
- Understand the different kind of roles available in the data science ecosystem along with their skill set requirements
- Build a resume that has a higher probability of getting shortlisted by recruiters
- Prepare for different data science roles by reviewing 400+ questions asked in actual interviews
- Understand and implement frameworks to solve guesstimates, logical and aptitude problems often used in interviews
- Learn from the experience of alumni who have successfully transitioned
- Speed up your path towards a transition to a Data science role.

Complete Career Prep today to speed up your path towards DS transition. You'll have a 40% better chance of getting hired than other candidates.

Note: CareerPrep is mandatory step in unlocking the job board.



# HAPPY LEARNING