

File Home Insert Draw Design Transitions Animations Slide Show Review View Help Tell me what you want to do

New Slide Slides Table Tables Pictures Photo Album Images Shapes Icons Illustrations 3D Models SmartArt Chart Add-ins Get Add-ins My Add-ins Zoom Links Action Comments Comments Text Box Header & Footer Text Symbols Video Audio Screen Recording Media

Handwritten notes on the left margin:

- $d=2,3 \rightarrow$ visualization
- $d=4,5,6 \rightarrow$ better
- $d=10,12,1289$

Why to reduce dim?

- ① Visualization
- ② efficient

How? PCA

50×90

$X = \begin{bmatrix} 1 & 12 \\ 2 & 60 \end{bmatrix}$ $S = X^T X$

$X' = \begin{bmatrix} 1 & 12 \\ 2 & 60 \end{bmatrix}$

Weight $40-60$ $2D \rightarrow 1D$ Age

QUESTION: CLARIFY THE PROBLEMS AS CLASSIFICATION OR REGRESSION

- 1. How many attributes used to read 5 minutes?
- 2. The device this used is a game console?
- 3. The device I play is a game console?
- 4. The game console is a game console?

Handwritten notes on the main slide:

$X = \begin{bmatrix} 1 & 12 \\ 2 & 60 \end{bmatrix}$ $S = X^T X$

60×100

60×90

$$S = X^T X$$

$$X = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}_{2 \times 3}$$

$$(3 \times 2)(2 \times 3)$$

$$\boxed{3 \times 3}$$

$$X^T = \begin{bmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{bmatrix}_{3 \times 2}$$

Handwritten notes on the left margin:

- Why to reduce dim?
① Visualization
② Efficient
Flow ① PCA
- Handwritten matrix and calculations.
- Handwritten notes about eigen values and vectors.
- Handwritten notes about PCA.

Handwritten notes on the main slide:

$X = \begin{bmatrix} x_1 & x_2 & \dots & x_{100} \\ 1 & 2 & \dots & 100 \\ \vdots & \vdots & \ddots & \vdots \\ 60 & \vdots & \vdots & \vdots \end{bmatrix}$

$S = X^T X$ 100×100

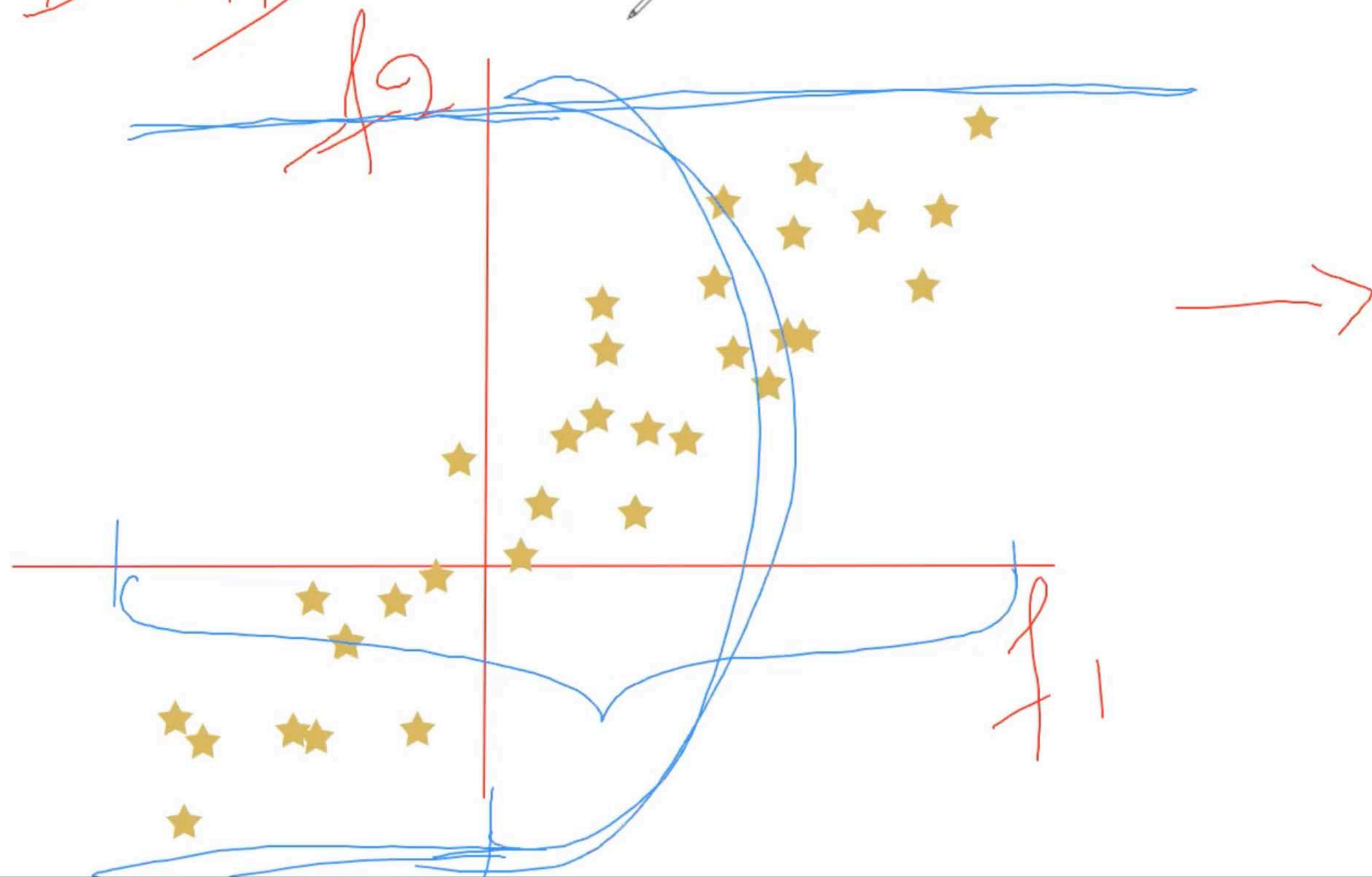
③ Eigen value
direction with
max var 100

60×100 Eigen vector

60×20

$(60 \times 100) \times (100 \times 20) \rightarrow 20 \times 100$

2D \rightarrow 1D



PCA technology (maximum variance feature){example of $60 \times 100 \rightarrow 60 \times 20$ feature reduction say dimensions reduction}

1. take data set in array form
2. find covariance of that matrix
3. find eigen value and corresponding eigen vector
4. take eigen vector corresponding to top 20 eigen value
5. compute X -transpose of input matrix (dot product of input matrix $\{60 \times 100\}$ and eigen vectors matrix $\{100 \times 20\}$)

\rightarrow covariance = $X(\text{transpose}) * X$ {matrix multiplication get 100×100 }

\rightarrow column standardization = the means of a column should be zero.

\rightarrow eigen value give the direction with maximum variance. {variance is also understands as spread.}

\rightarrow eigen vector give vector corresponding to that eigen value(that have maximum variance).

\rightarrow no of eigen value ,eigen vector = no of column(each column represent one vector or say one feature).

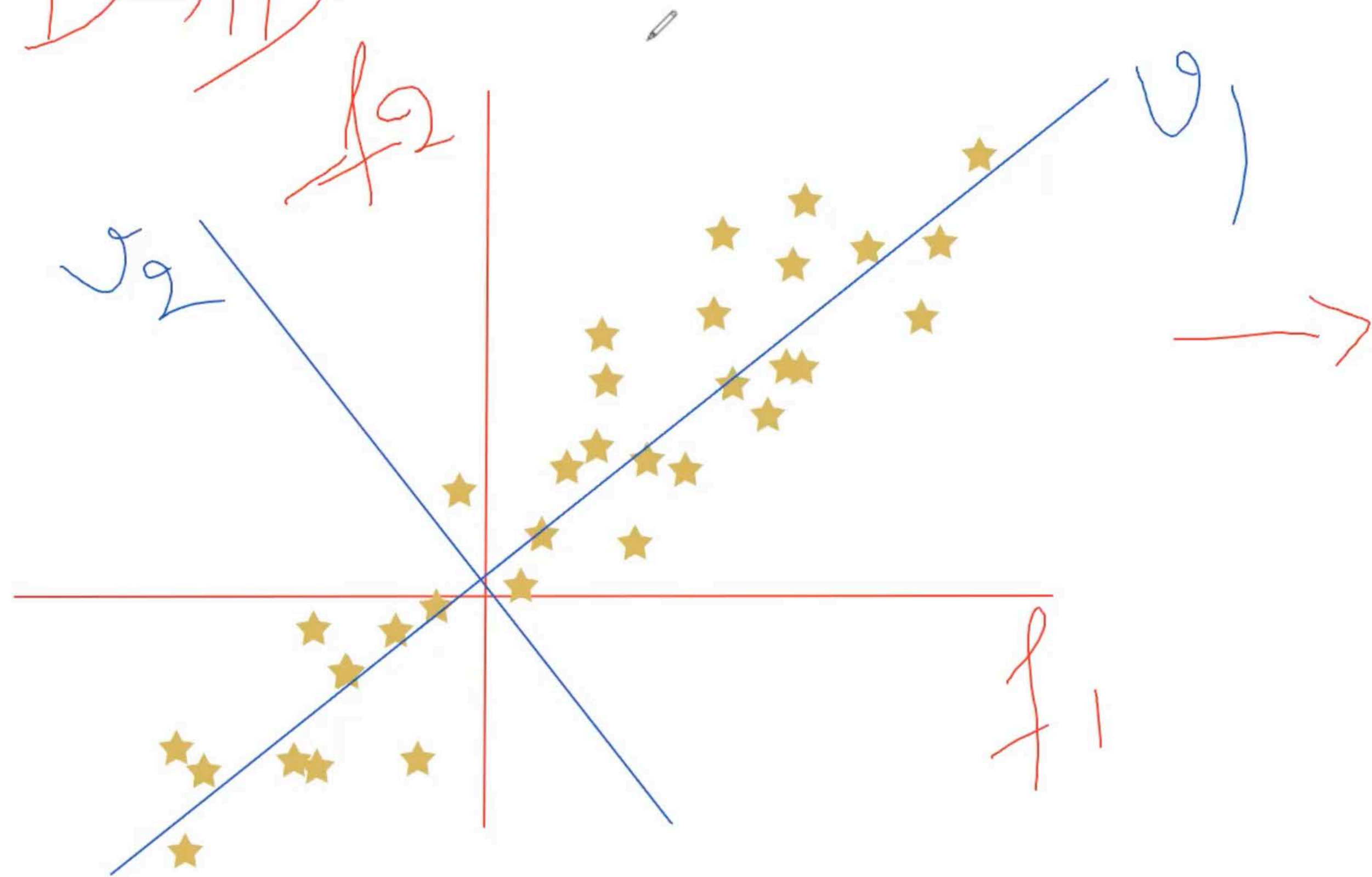
\rightarrow make a matrix of those eigen vectors $\{100 \times 100 \rightarrow 100 \times 20\}$

\rightarrow no of eigen vector = no of dimension

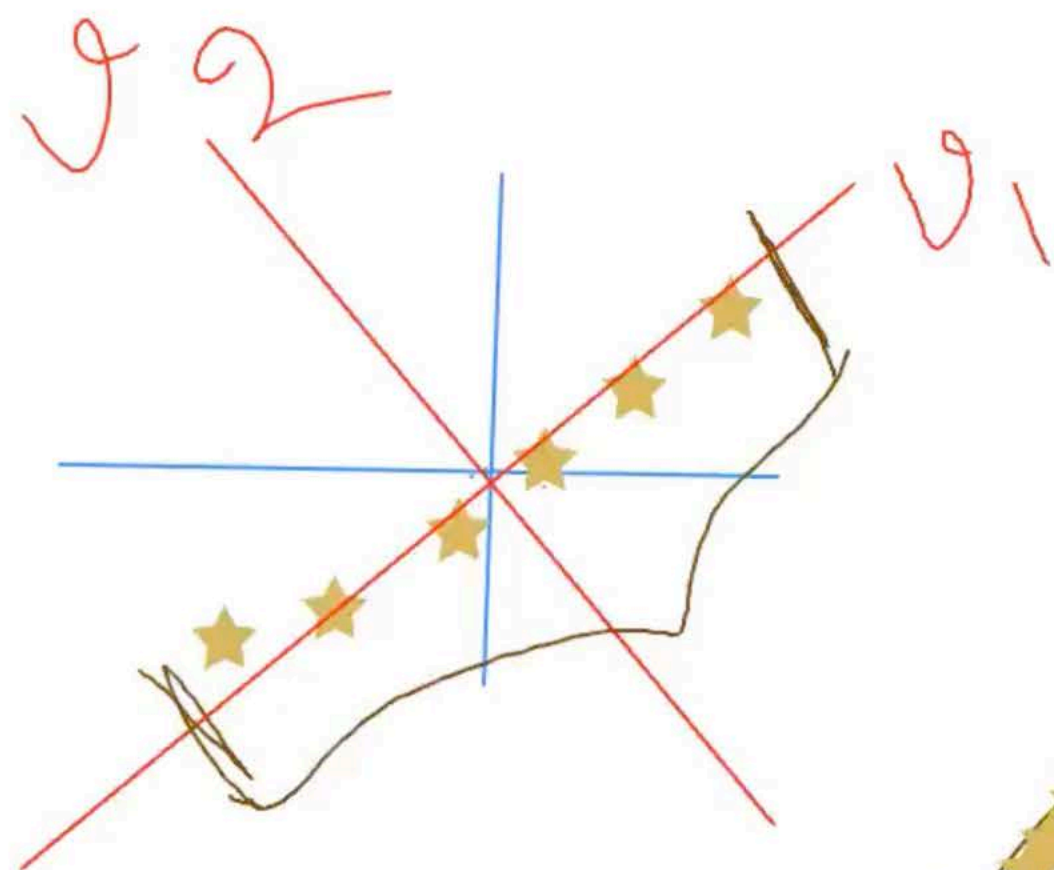
\rightarrow every eigen vector(corresponding to different dimension) is perpendicular to each other(this is the feature of eigen vector.)

2D \rightarrow 1D

2D
eigen



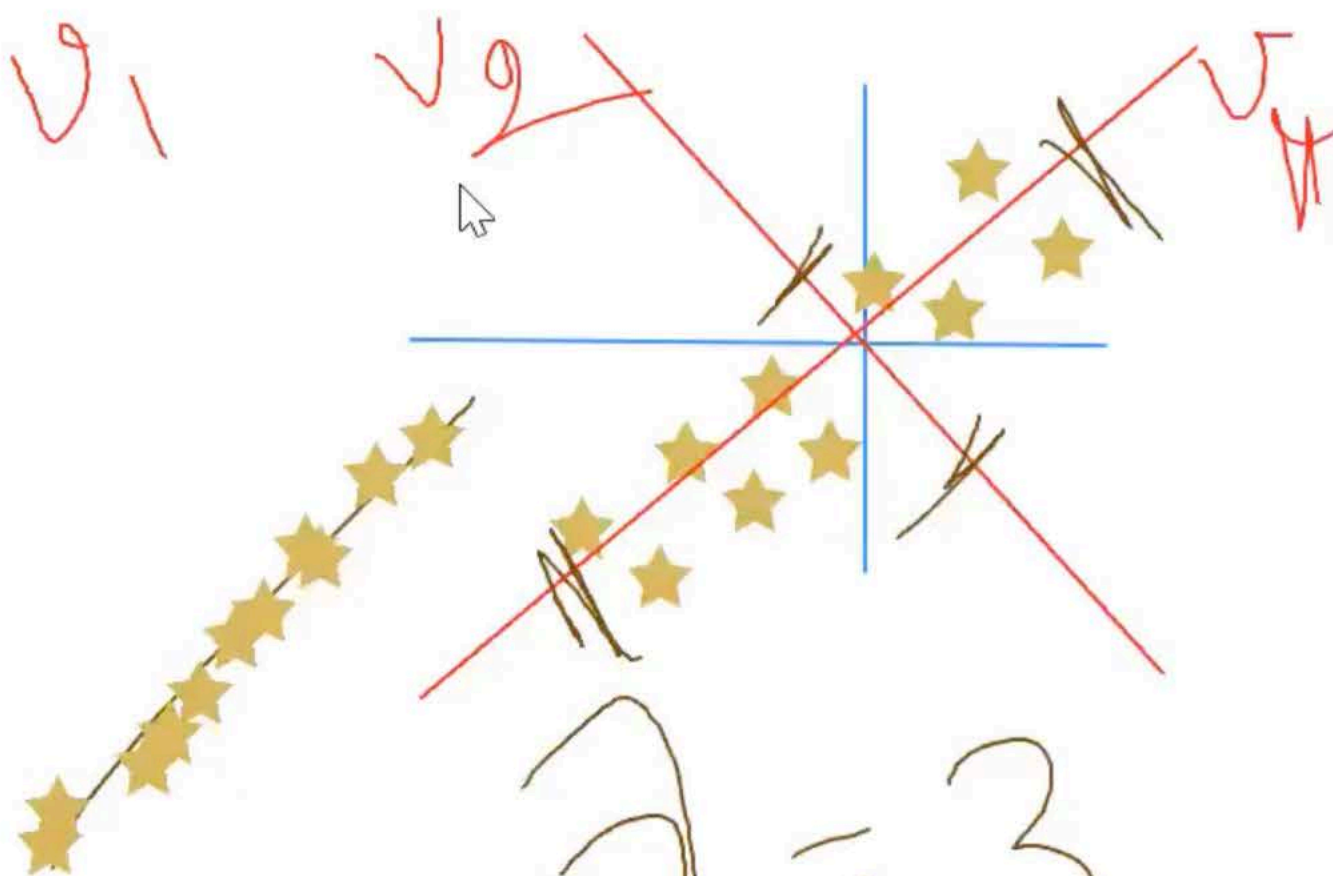
2



$$\lambda_1 = 3$$

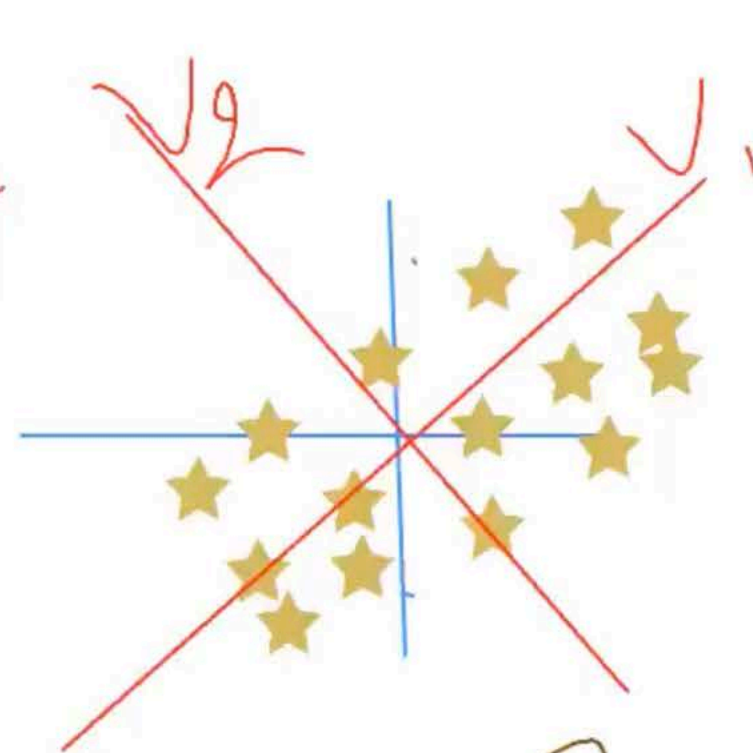
$$\lambda_2 = 0$$

$$\frac{\lambda_1}{\lambda_1 + \lambda_2} = 1$$



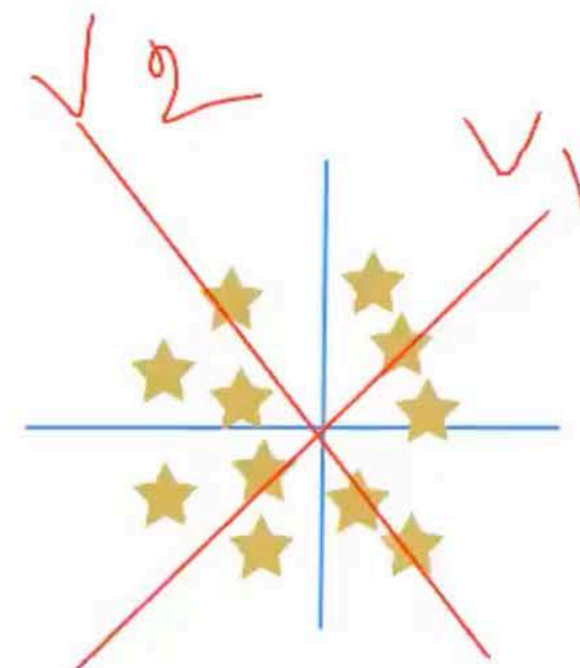
$$\lambda_1 = 3$$

$$\lambda_2 = 1$$



$$\lambda_1 = 3$$

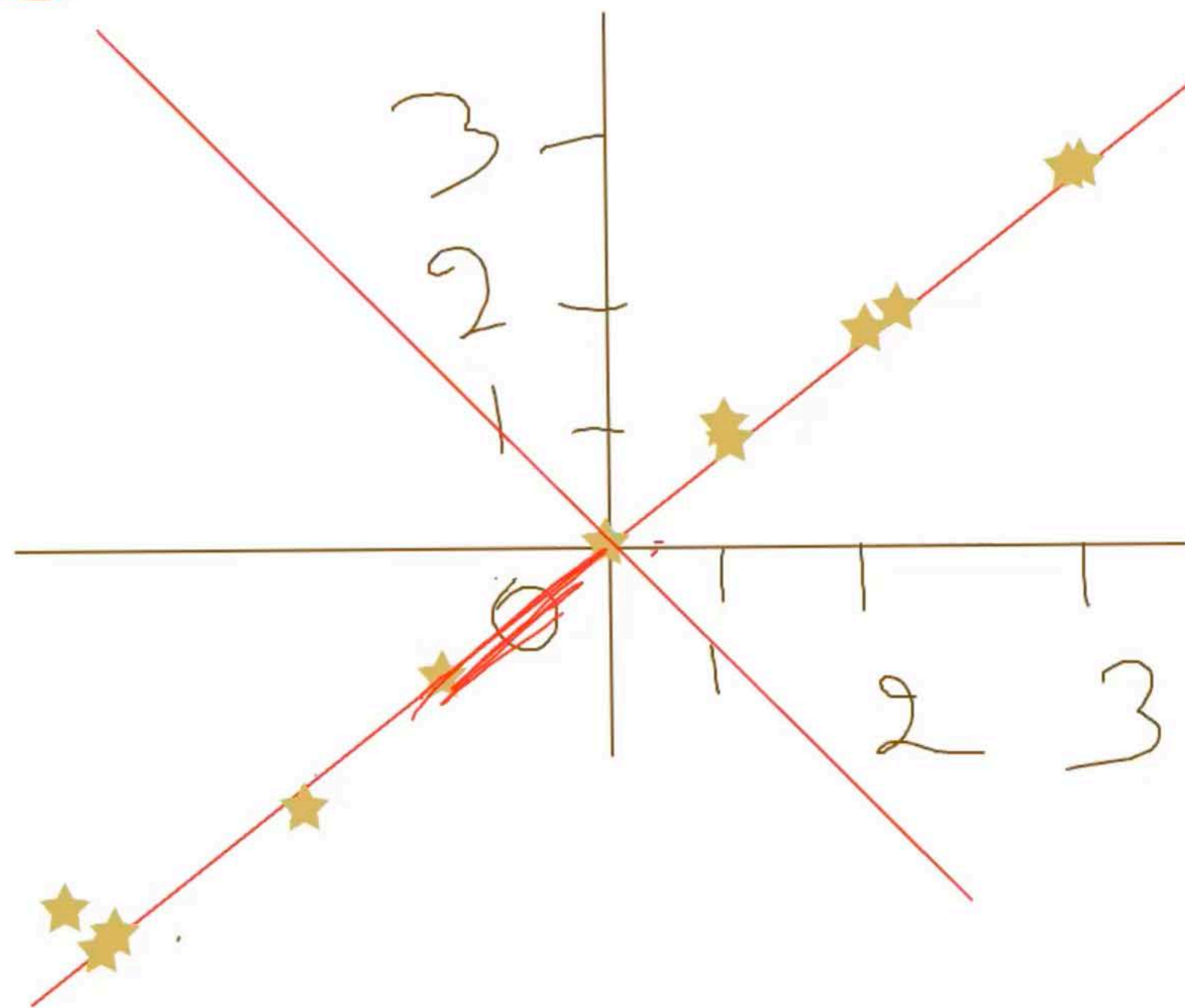
$$\lambda_2 = 2$$



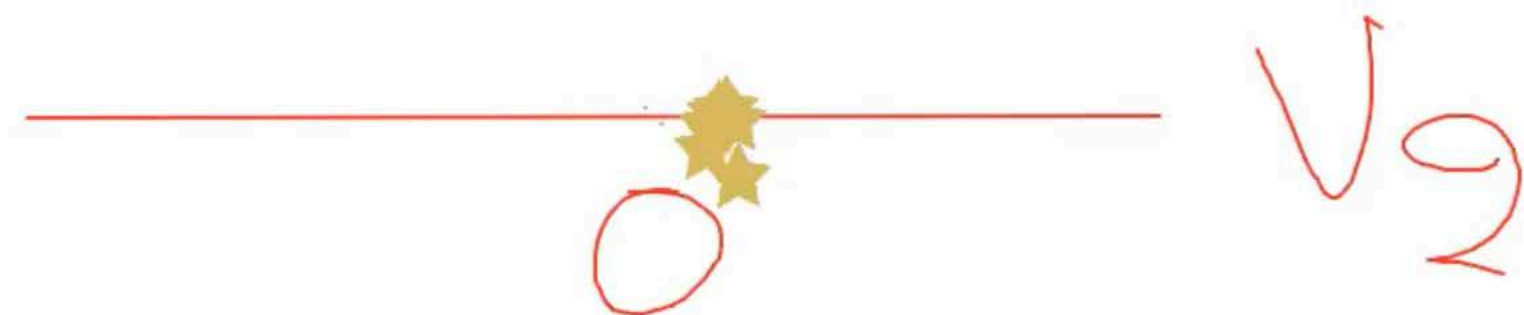
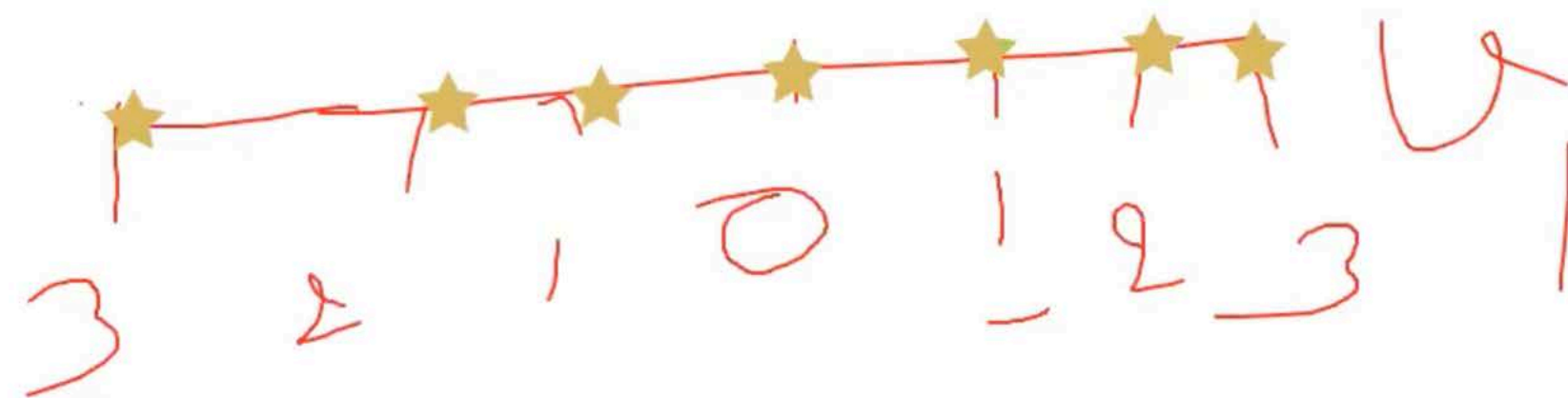
$$\lambda_1 = 3$$

$$\lambda_2 = 3$$

59



$V_1, 2D \rightarrow 1D$



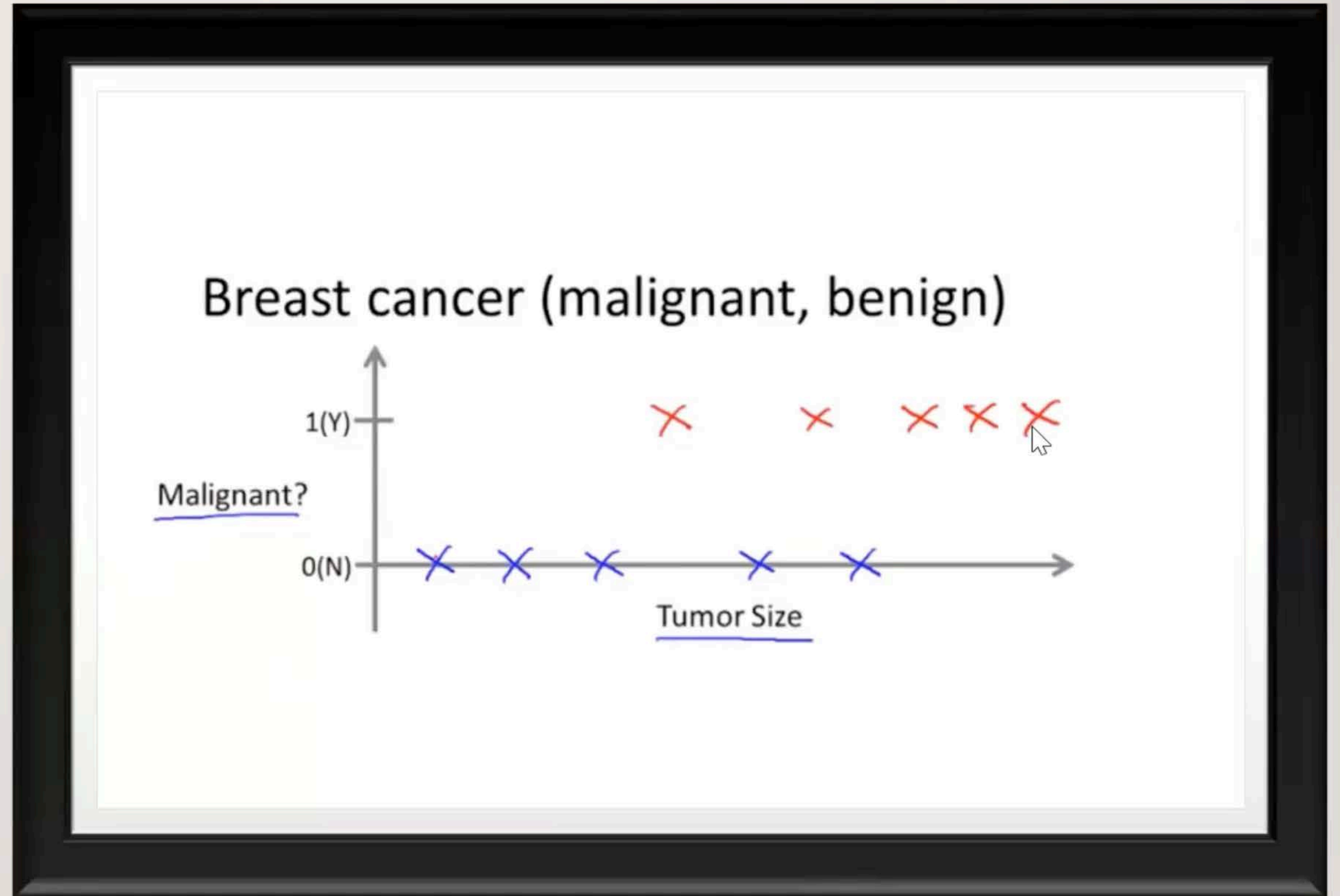
Supervised Learning

```
graph TD; A[Supervised Learning] --> B[Regression : Predict continuous valued output eg. Price]; A --> C[Classification : Discrete valued output (0/1)]
```

Regression :
Predict continuous
valued output eg. Price

Classification :
Discrete valued output
(0/1)

CLASSIFICATION BREAST CANCER



Tools: Select, Draw with Touch, Eraser, Lasso Select

Pens: Green, Orange, Blue, Grey, Pink, Red, Add Pen

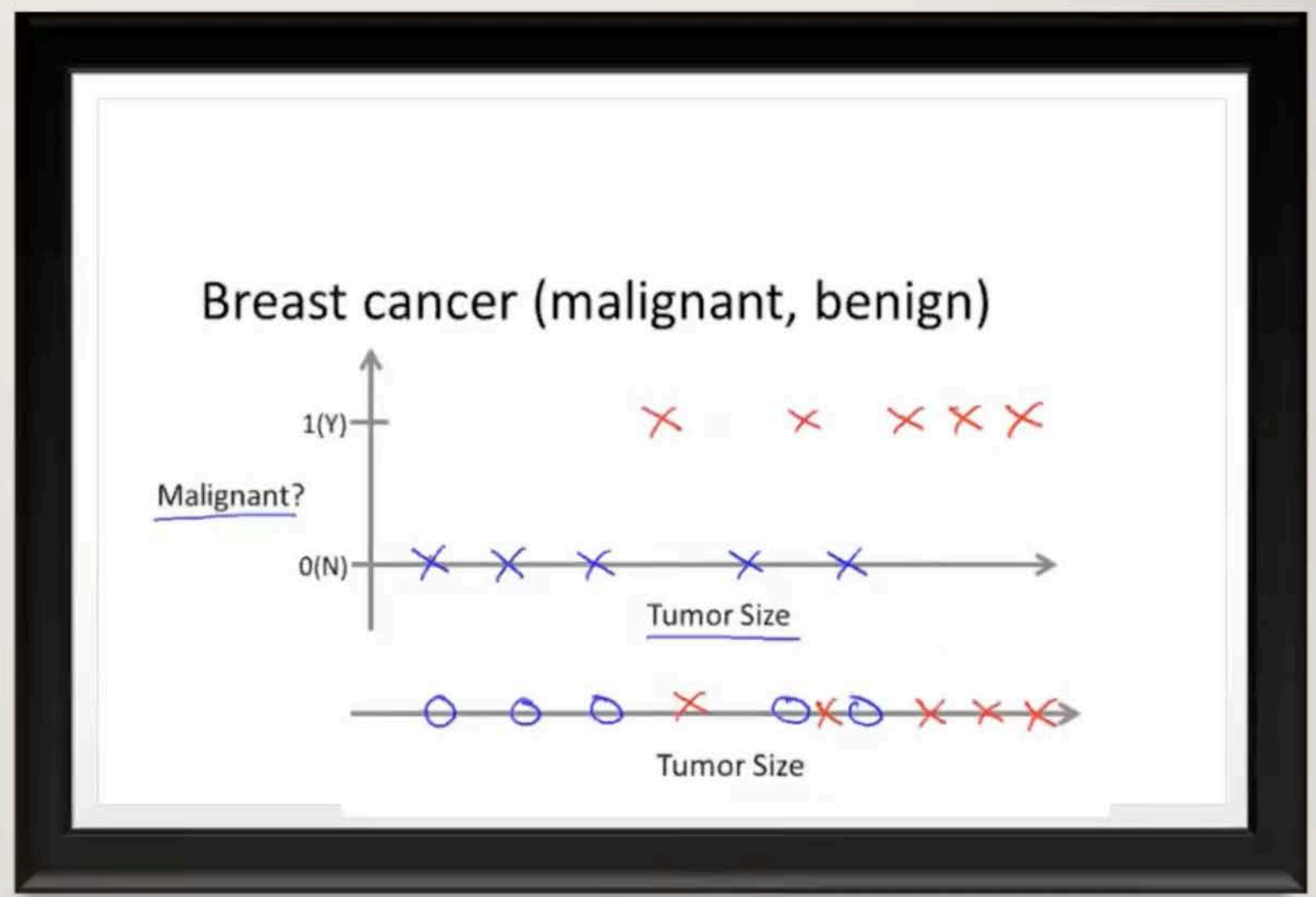
Stencils: Ruler

Convert: Ink to Shape, Ink to Math

Slide thumbnails:

- Slide 1: SUPERVISED LEARNING
- Slide 2: Supervised Learning
 - Regression: Predict continuous valued output (e.g. Price)
 - Classification: Discrete valued output (0/1)
- Slide 3: REGRESSION: HOUSING PREDICTION
- Slide 4: CLASSIFICATION: BREAST CANCER (malignant, benign)
- Slide 5: CLASSIFICATION: SPAM FILTERING
- Slide 6: CLASSIFICATION: SPAM FILTERING

CLASSIFICATION BREAST CANCER



Tools: Select, Draw with Touch, Eraser, Lasso Select

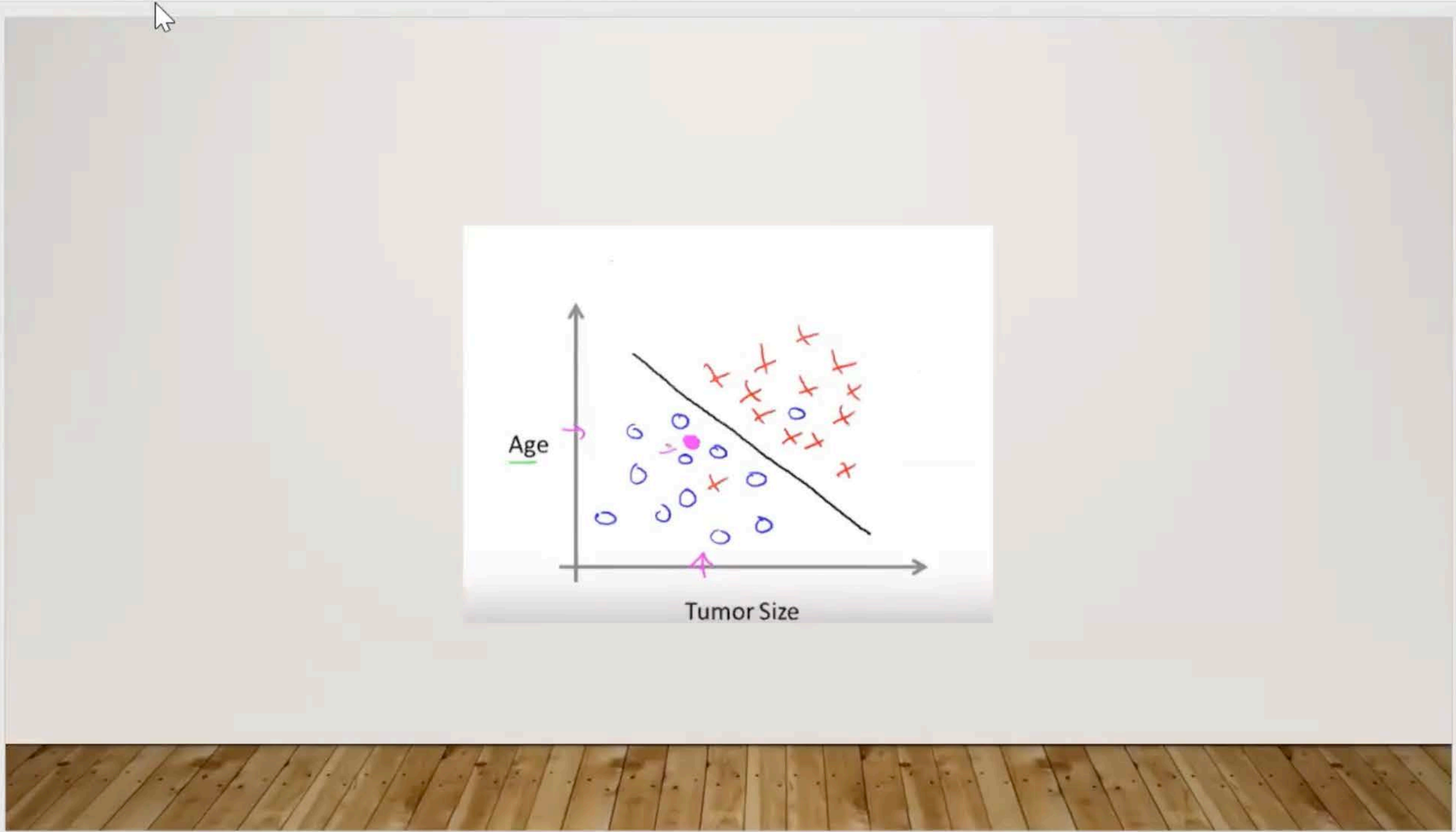
Pens: Green, Orange, Blue, Grey, Pink, Red, Add Pen

Stencils: Ruler

Convert: Ink to Shape, Ink to Math

Slide thumbnails:

- REGRESSION: HOUSING PREDICTION
- QUESTION: SPASTIC DANCER
- QUESTION: SPASTIC DANCER
- QUESTION: CLASSIFY THE PROBLEMS AS CLASSIFICATION OR REGRESSION



Tools: Select, Draw with Touch, Eraser, Lasso Select

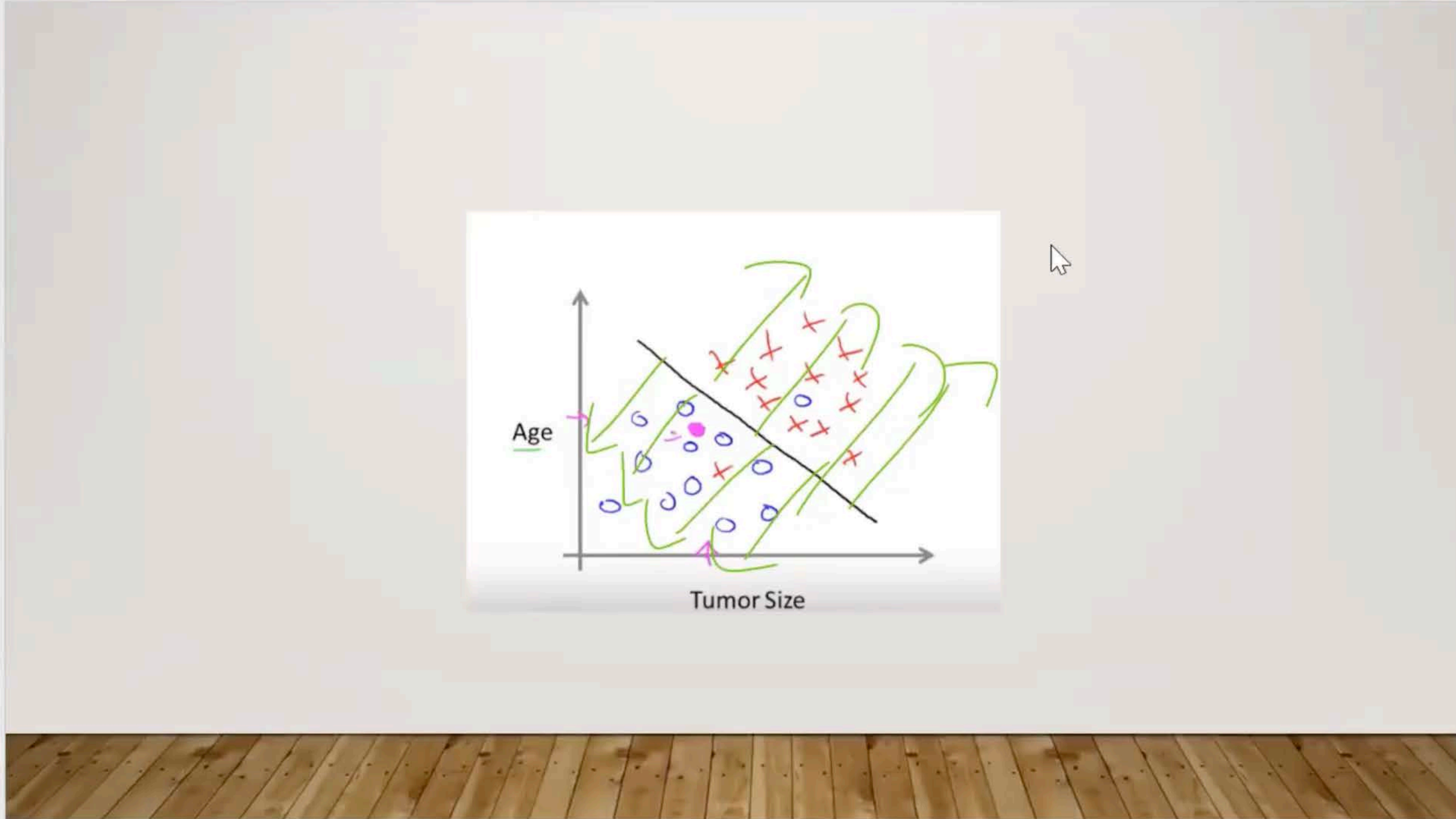
Pens: Green, Orange, Blue, Grey, Pink, Red, Add Pen

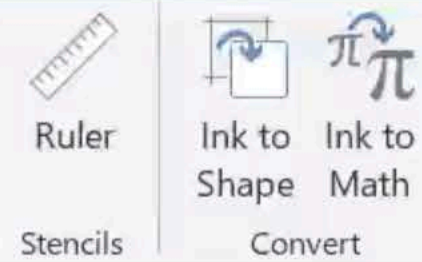
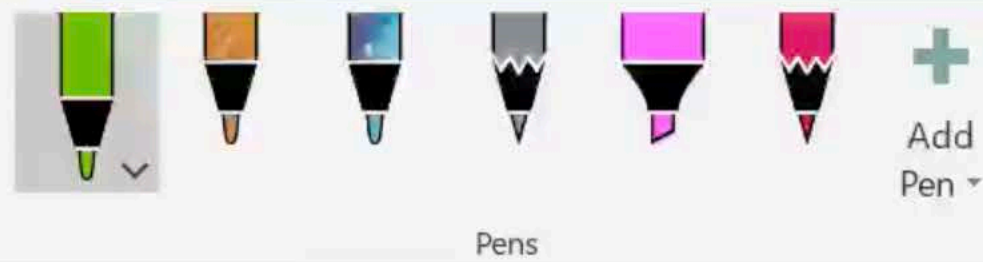
Stencils: Ruler

Convert: Ink to Shape, Ink to Math

Slide thumbnails:

- REGRESSION: HOUSING PREDICTION
- QUESTION: BREAST CANCER
- QUESTION: BREAST CANCER
- QUESTION: CLASSIFY THE PROBLEMS AS CLASSIFICATION OR REGRESSION





A/B ✗
 $A/B \quad B/C$
 $d=2, 3 \Rightarrow$ Scatter Plot
 $d=4, 5, 6 \Rightarrow$ Pair plots
 $d=10, 12,$
 $\begin{matrix} m \\ 2 \end{matrix}$
 $\begin{matrix} 784 \end{matrix}$

Tools: Select, Draw with Touch, Eraser, Lasso Select

Pens: Green, Orange, Blue, Grey, Pink, Red, Add Pen

Stencils: Ruler

Convert: Ink to Shape, Ink to Math

Slide thumbnails:

- Slide 1: Title slide
- Slide 2: Question: What is PCA?
- Slide 3: Scatter plot with 60 data points
- Slide 4: Handwritten notes on PCA
- Slide 5: Handwritten notes on PCA
- Slide 6: Handwritten notes on PCA
- Slide 7: Handwritten notes on PCA
- Slide 8: Question: Classify the problems as classification or regression

Why to reduce dim?

① Visualization

② efficient

How

① PCD

② max var features

100D → 20D

QC

Tools: Select, Draw with Touch, Eraser, Lasso Select

Pens: Green, Orange, Blue, Grey, Pink, Red, Add Pen

Stencils: Ruler

Convert: Ink to Shape, Ink to Math

Slide thumbnails:

- Slide 1: Title slide
- Slide 2: Scatter plot with a regression line
- Slide 3: Handwritten notes on dimensionality reduction
- Slide 4: Handwritten notes on PCA
- Slide 5: Handwritten notes on PCA
- Slide 6: Handwritten notes on PCA
- Slide 7: Handwritten notes on PCA
- Slide 8: Handwritten notes on PCA
- Slide 9: Handwritten notes on PCA
- Slide 10: Handwritten notes on PCA

