Dimension Reduction

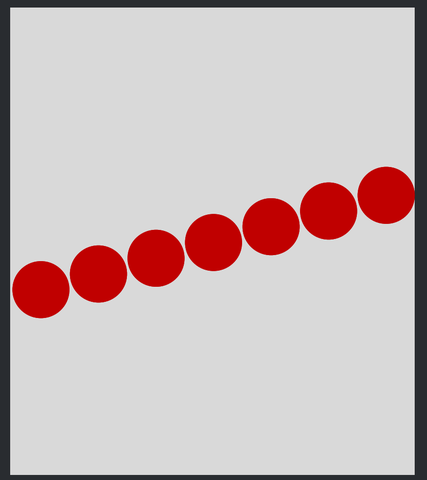
0 dimension

1 dimension

x
1
2
3
4
5
6
7

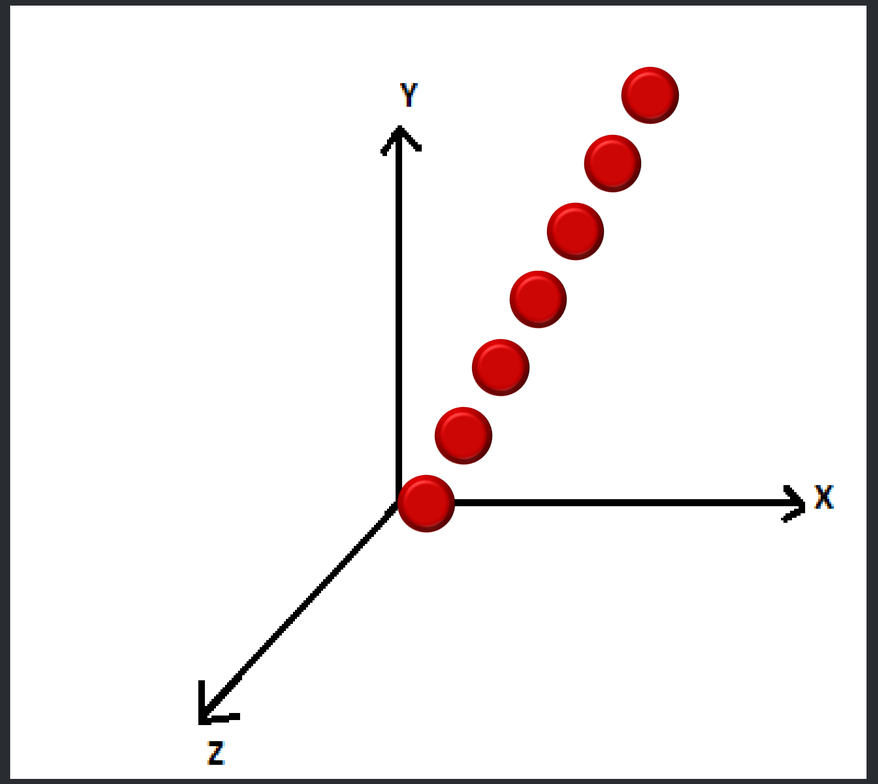
2 dimensions

x	y
1	1
2	2
3	3
4	4
5	5
6	6
7	7



3 dimensions

x2	x1	y
1	1	1
2	2	2
3	3	3
4	4	4
5	5	5
6	6	6
7	7	7

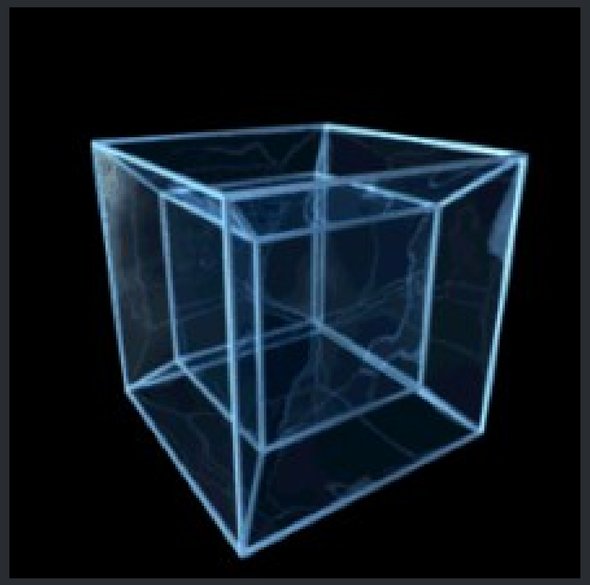


N dimensions

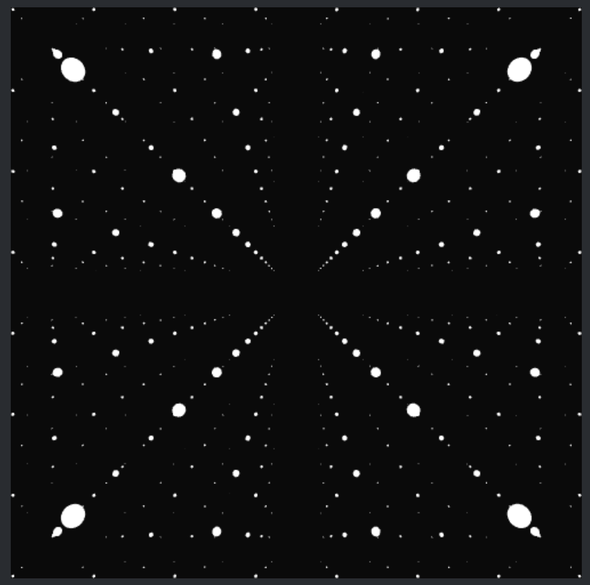
x2	x1	y	x3	x4	x5		xn
1	1	1	1	1	1		1
2	2	2	2	2	2		2
3	3	3	3	3	3		3
4	4	4	4	4	4	…	4
5	5	5	5	5	5		5
6	6	6	6	6	6		6
7	7	7	7	7	7		7

상상 불가!

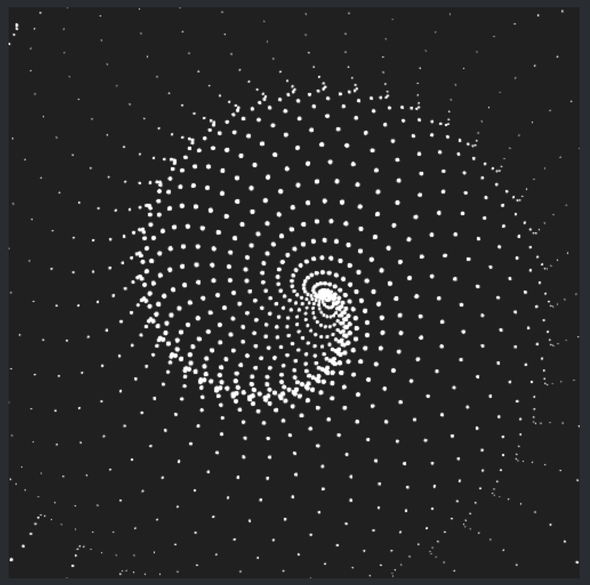
4 dimension?



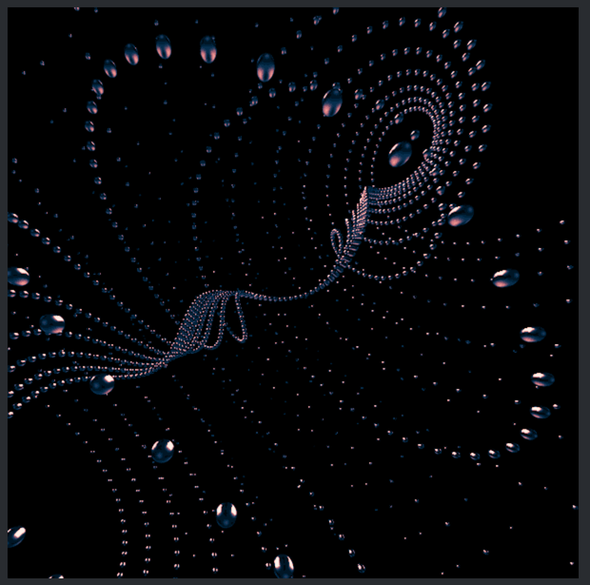
7 dimension?



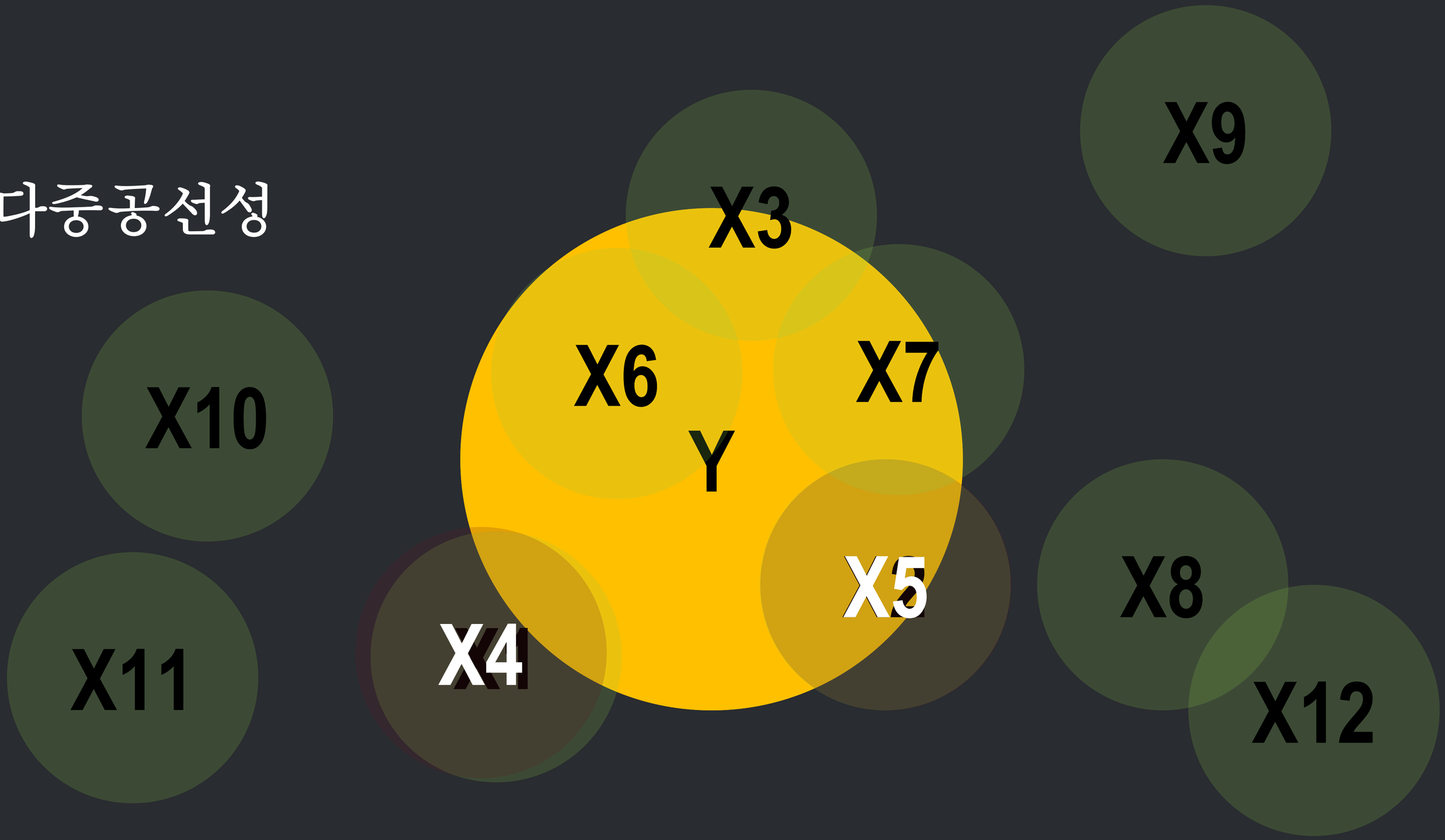
50 dimension?



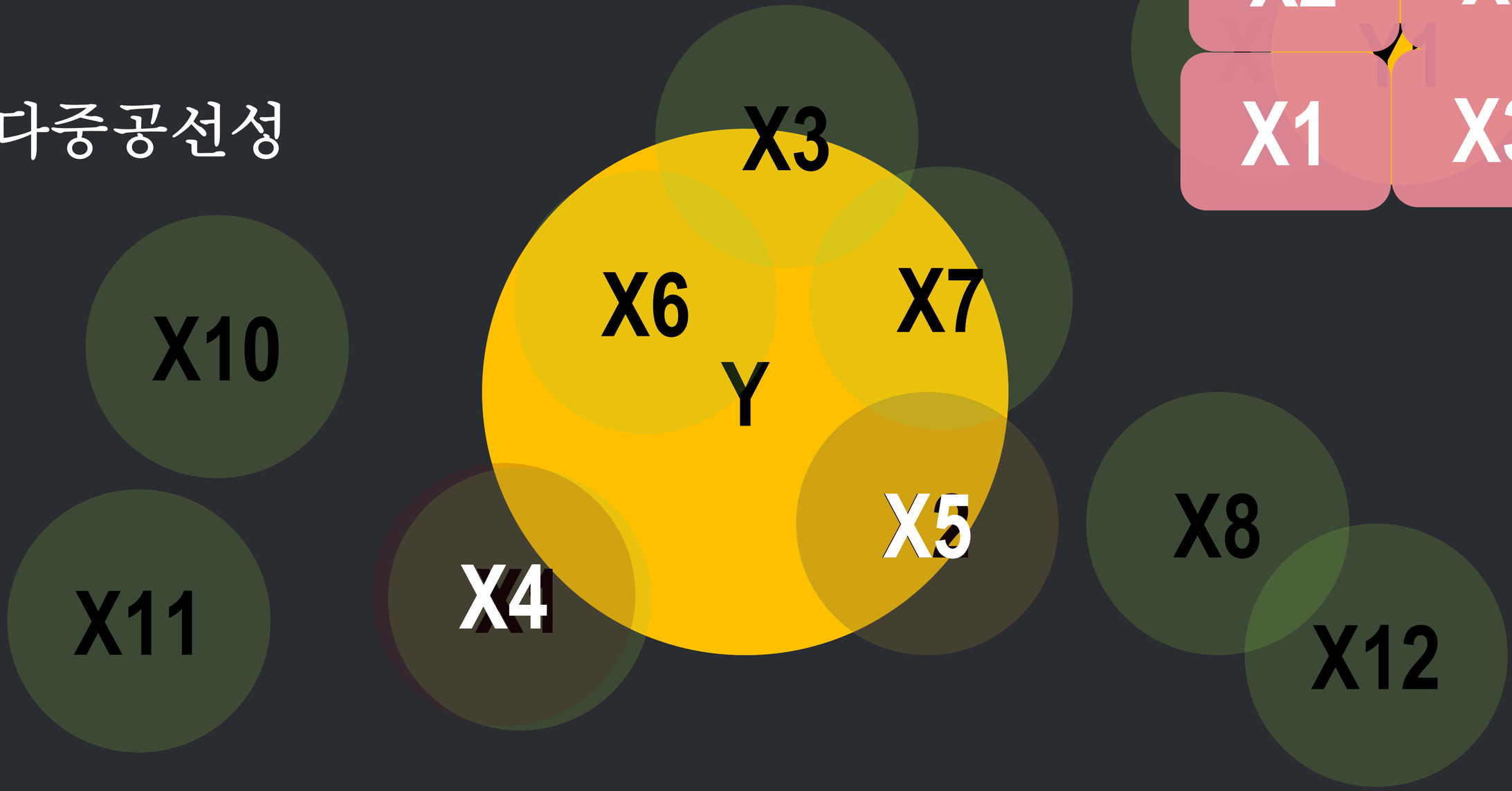
100 dimension?

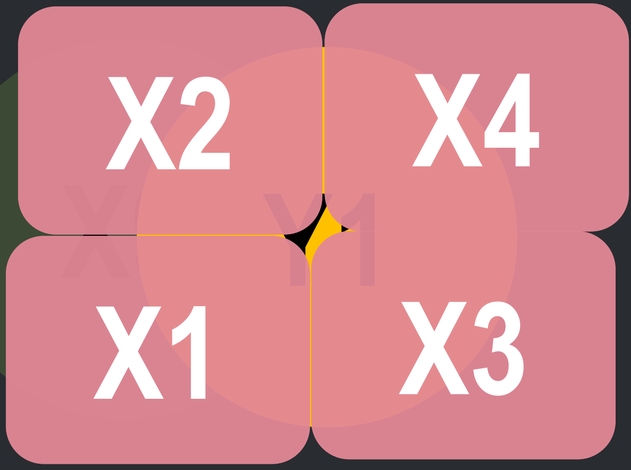


다중공선성



다중공선성



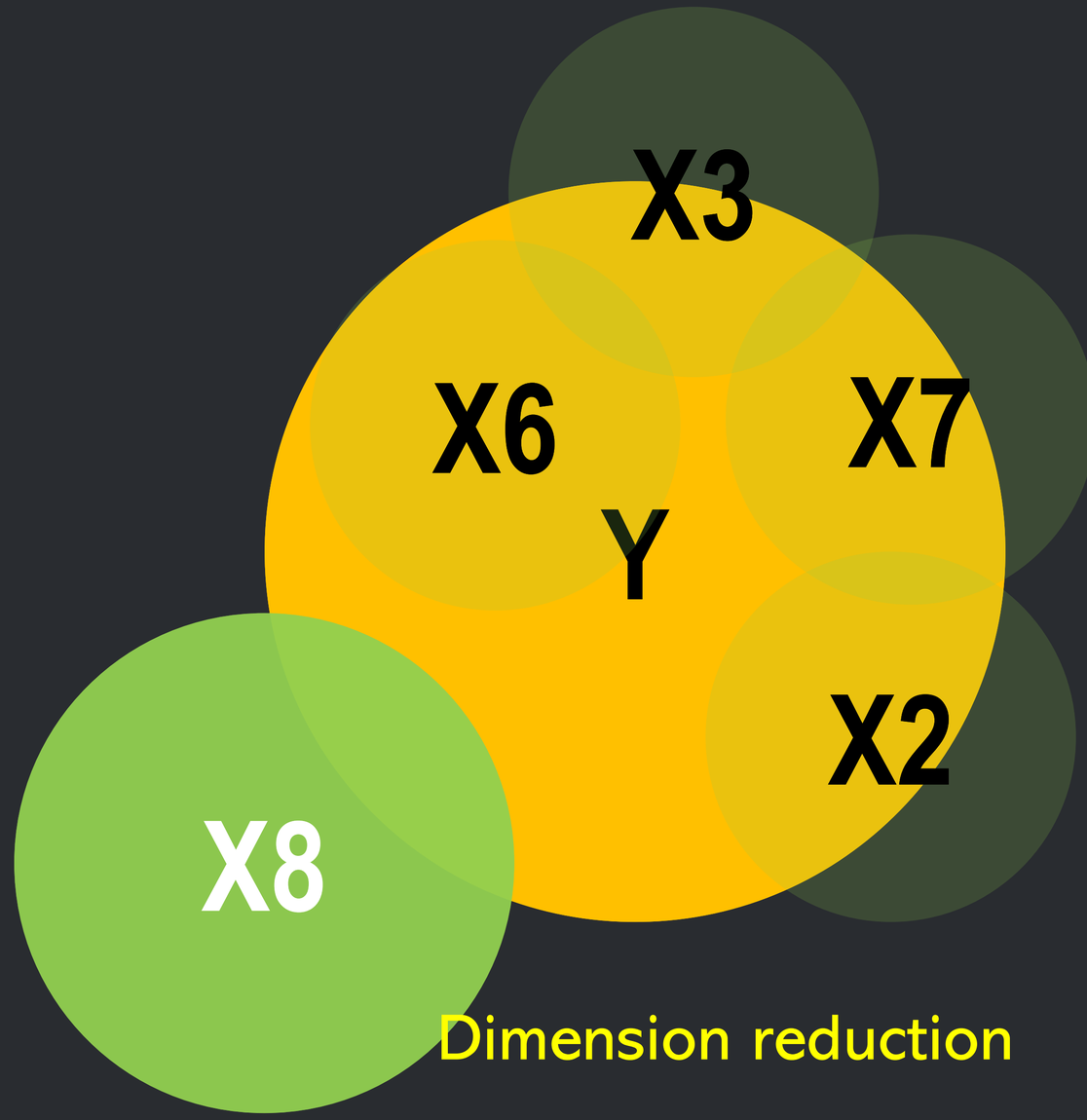


다중공선성

X4 & X1

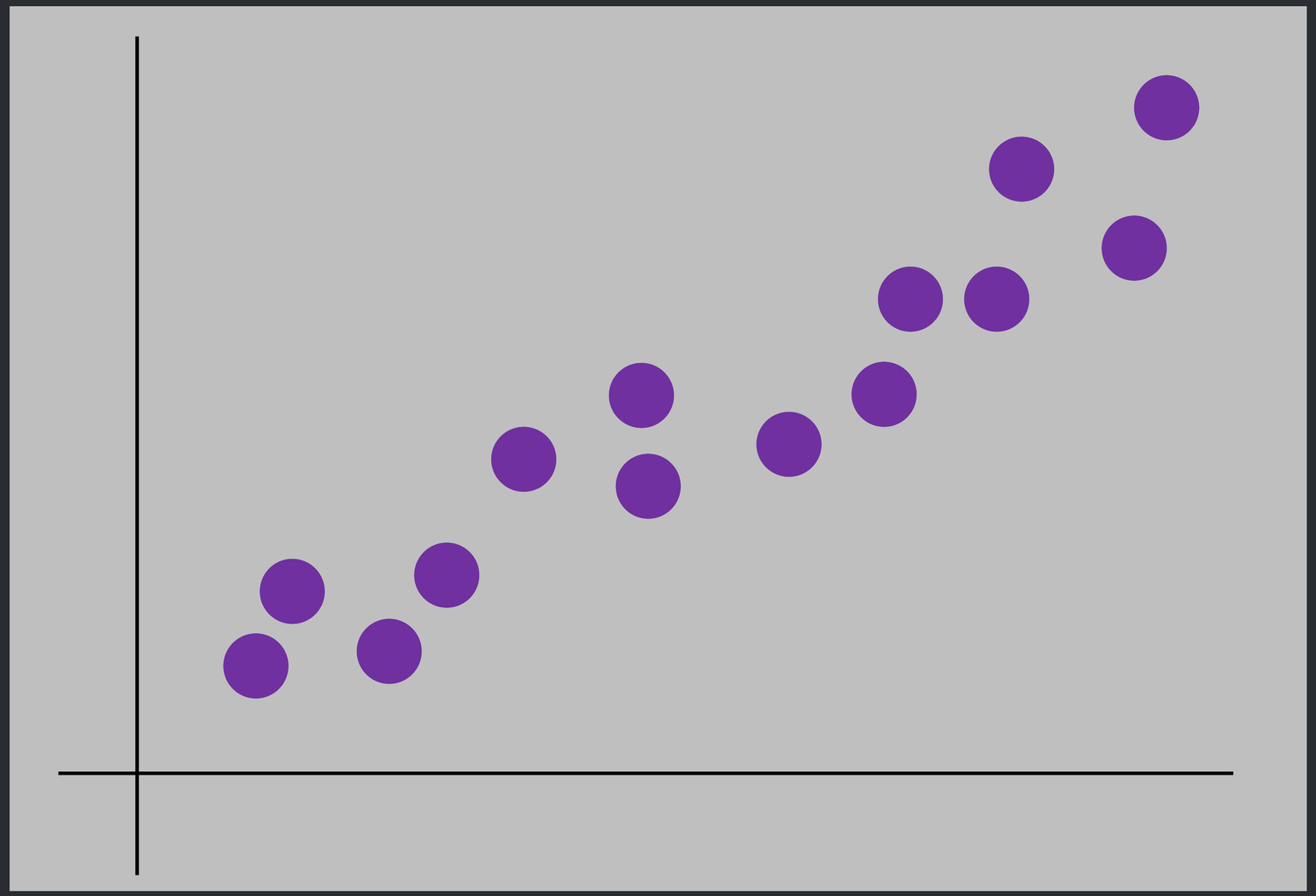
⬇

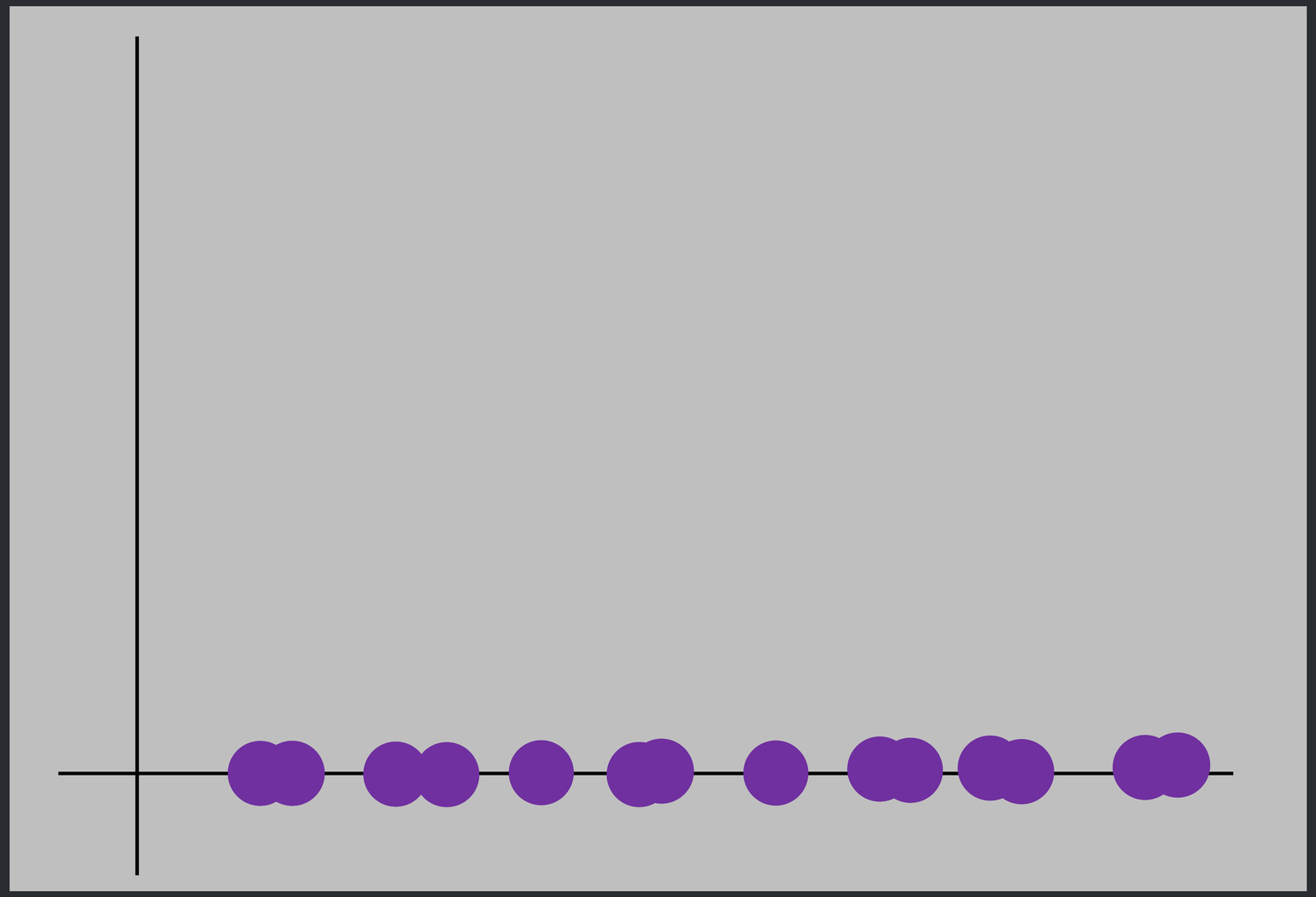
X8

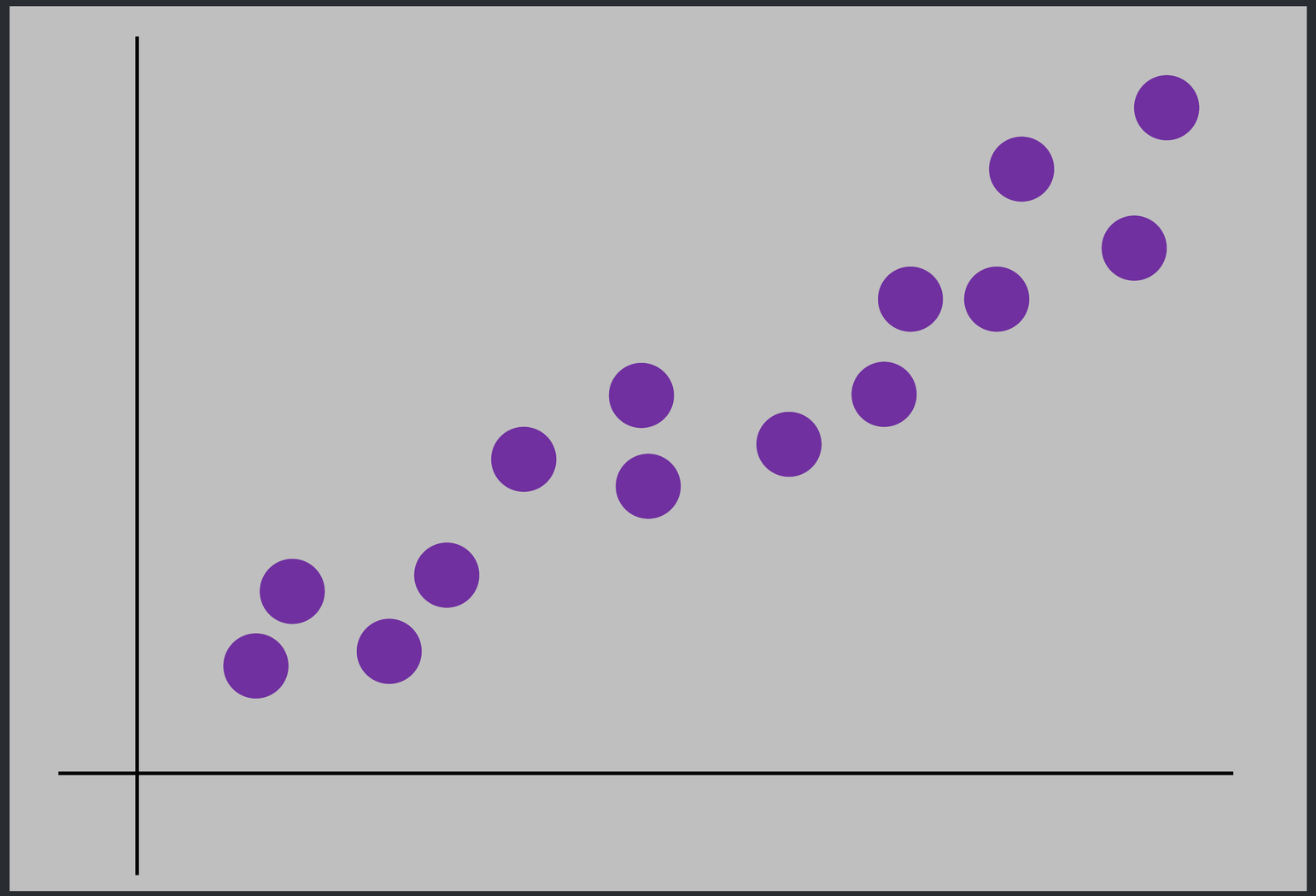


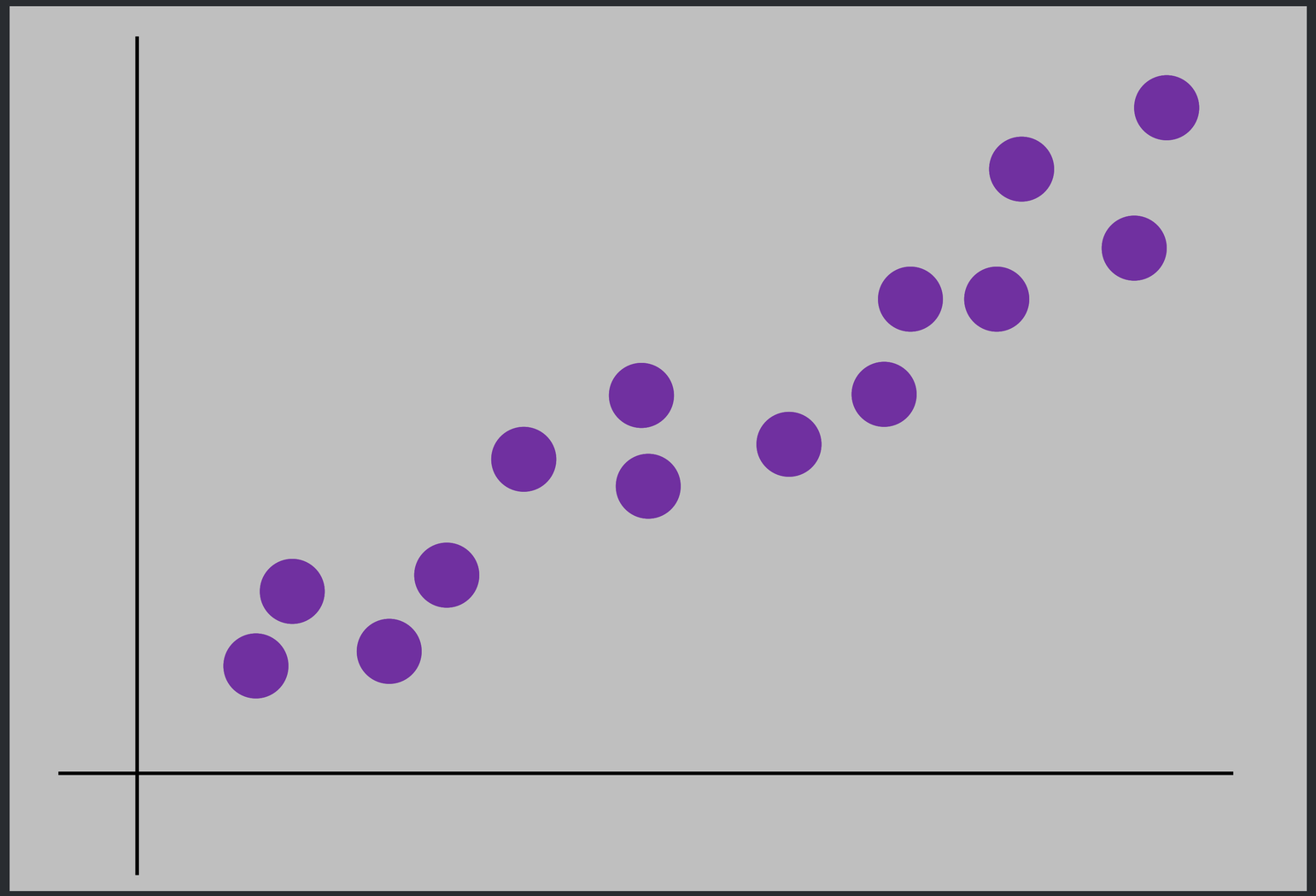
Dimension reduction

PCA

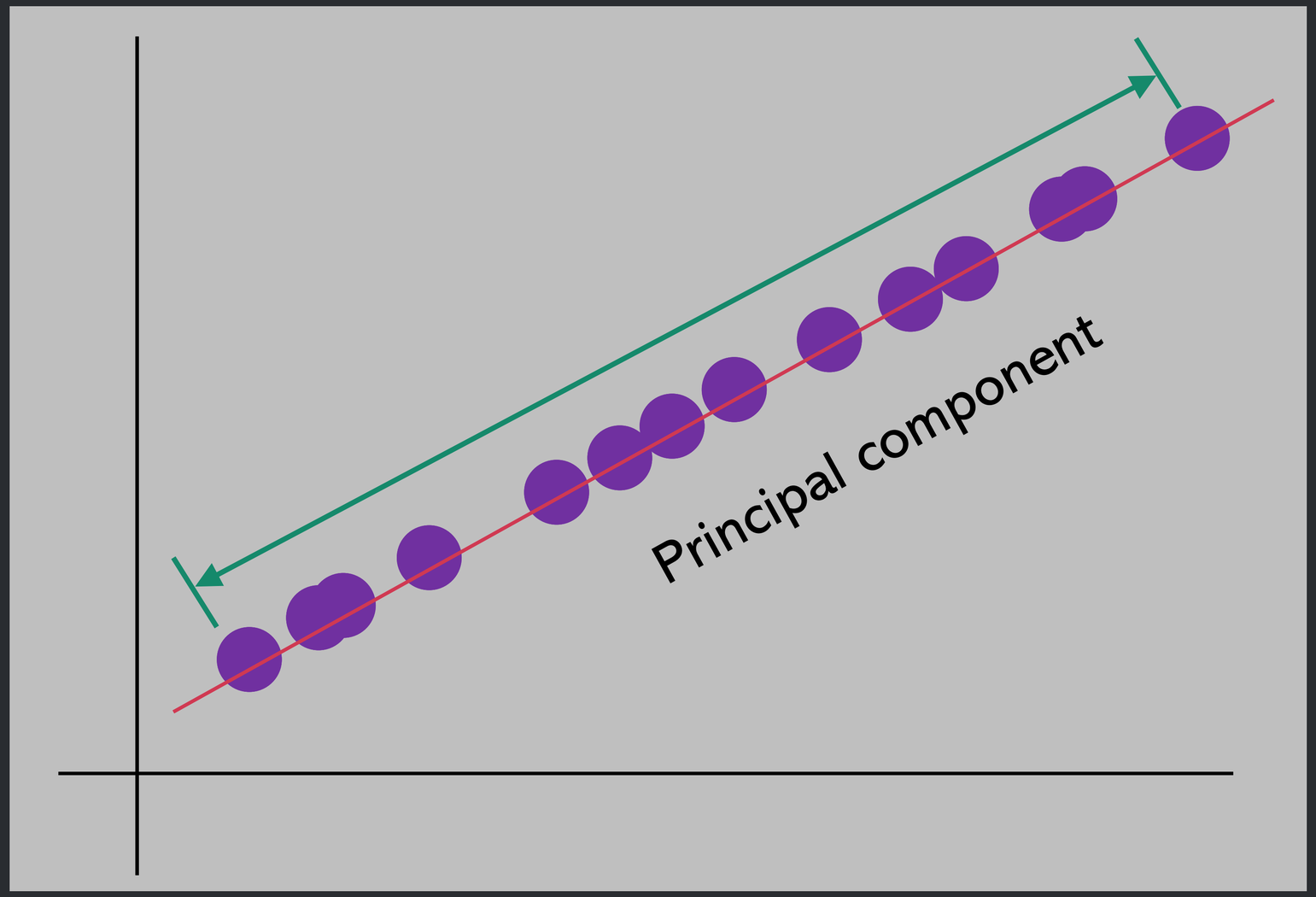








Find maximum eigen value



어떻게 컴퓨터가 자동으로 찾을 수 있을까?

Find maximum eigen value

$$\boldsymbol{Av} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} = \begin{pmatrix} a_{11}v_1 + a_{12}v_2 + a_{13}v_3 \\ a_{21}v_1 + a_{22}v_2 + a_{23}v_3 \end{pmatrix}$$

$$\boldsymbol{A} = \begin{pmatrix} 2 & 0 & -2 \\ 1 & 1 & -2 \\ 0 & 0 & 1 \end{pmatrix}$$ 이고 $\boldsymbol{v} = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$ 라면,

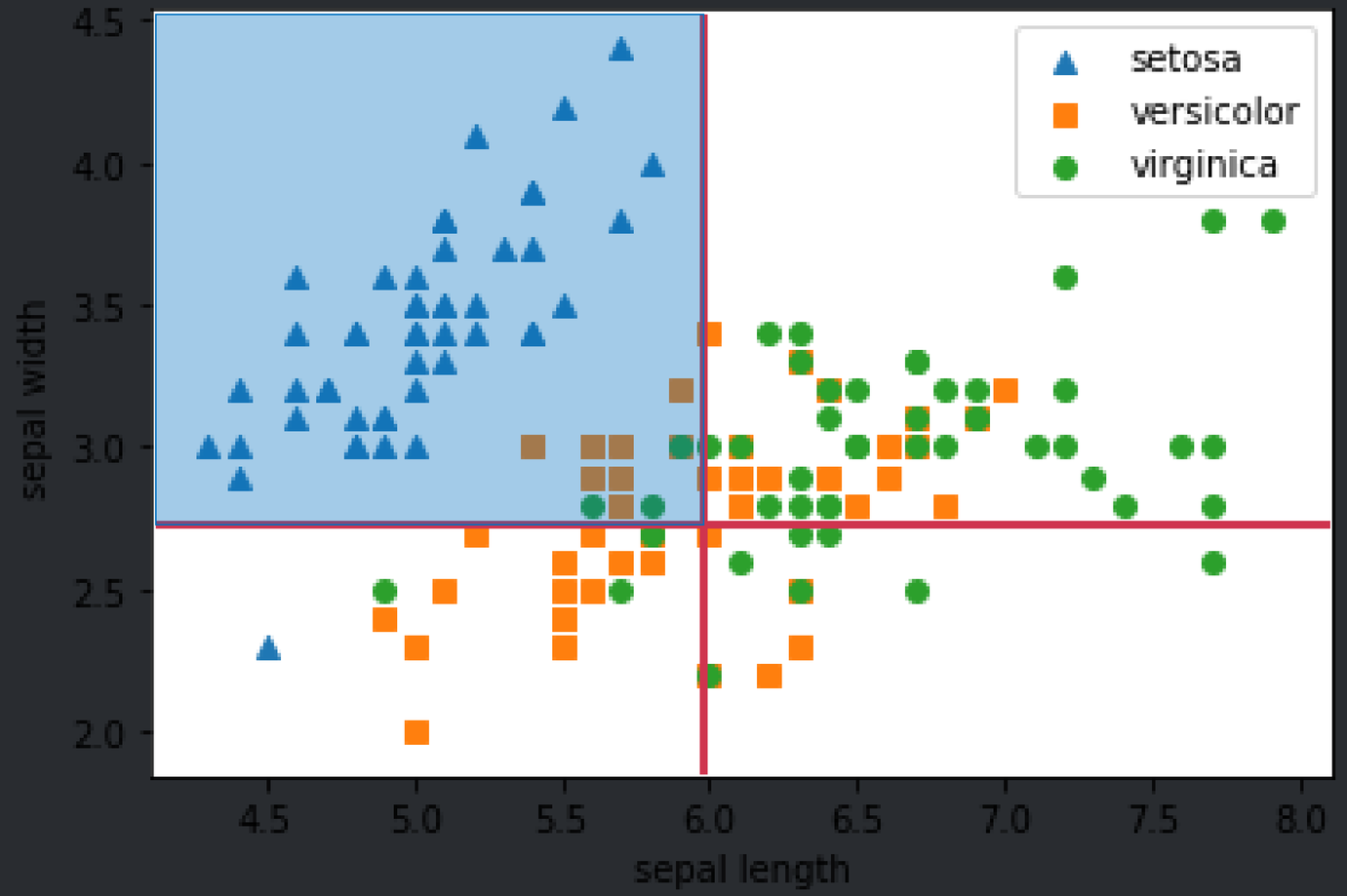
eigenvector

$$\boldsymbol{Av} = \begin{pmatrix} 2 & 0 & -2 \\ 1 & 1 & -2 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 2 \\ 2 \\ 0 \end{pmatrix} = 2\boldsymbol{v}$$

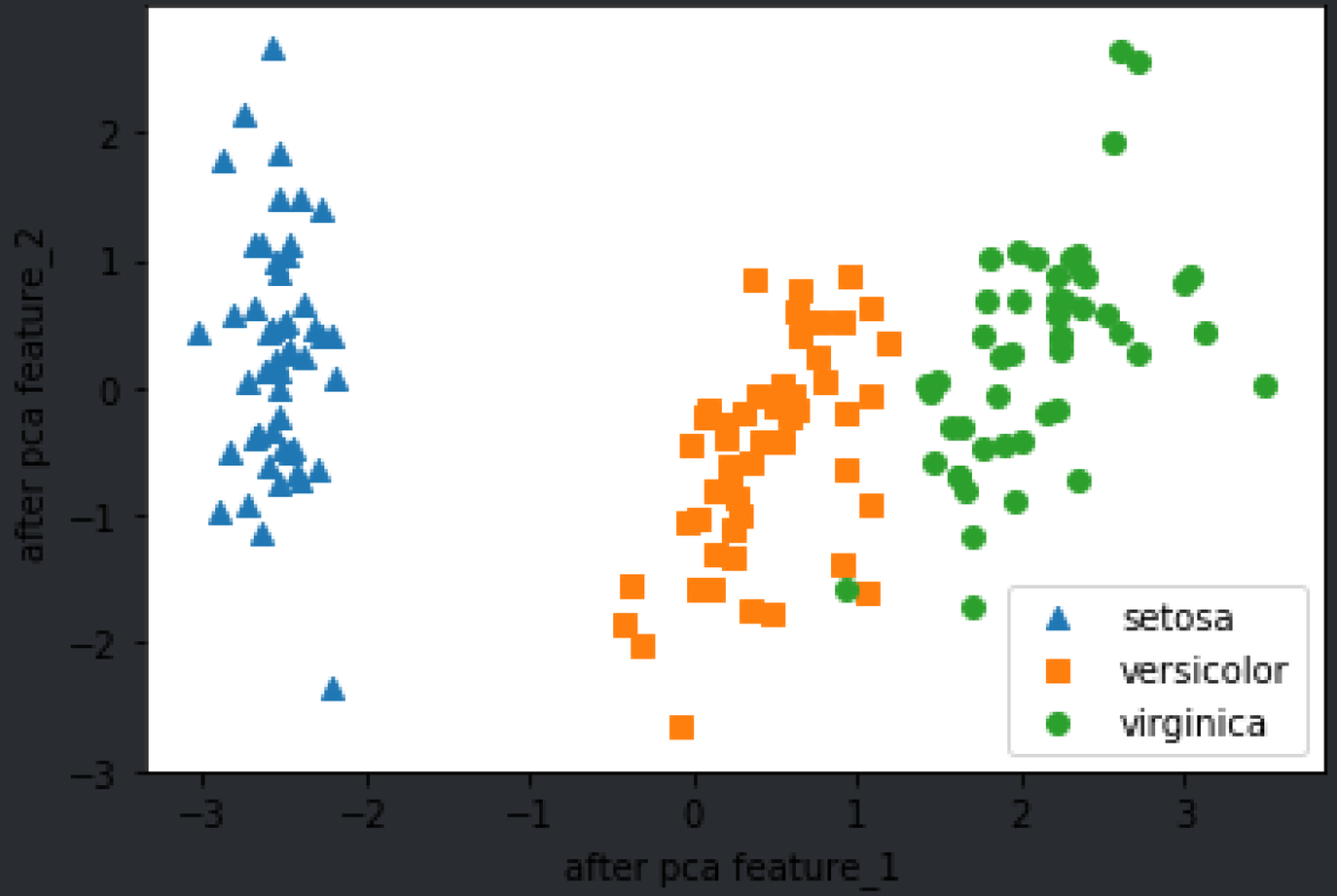
eigenvalue

Jupyter notebook

n_components = 2

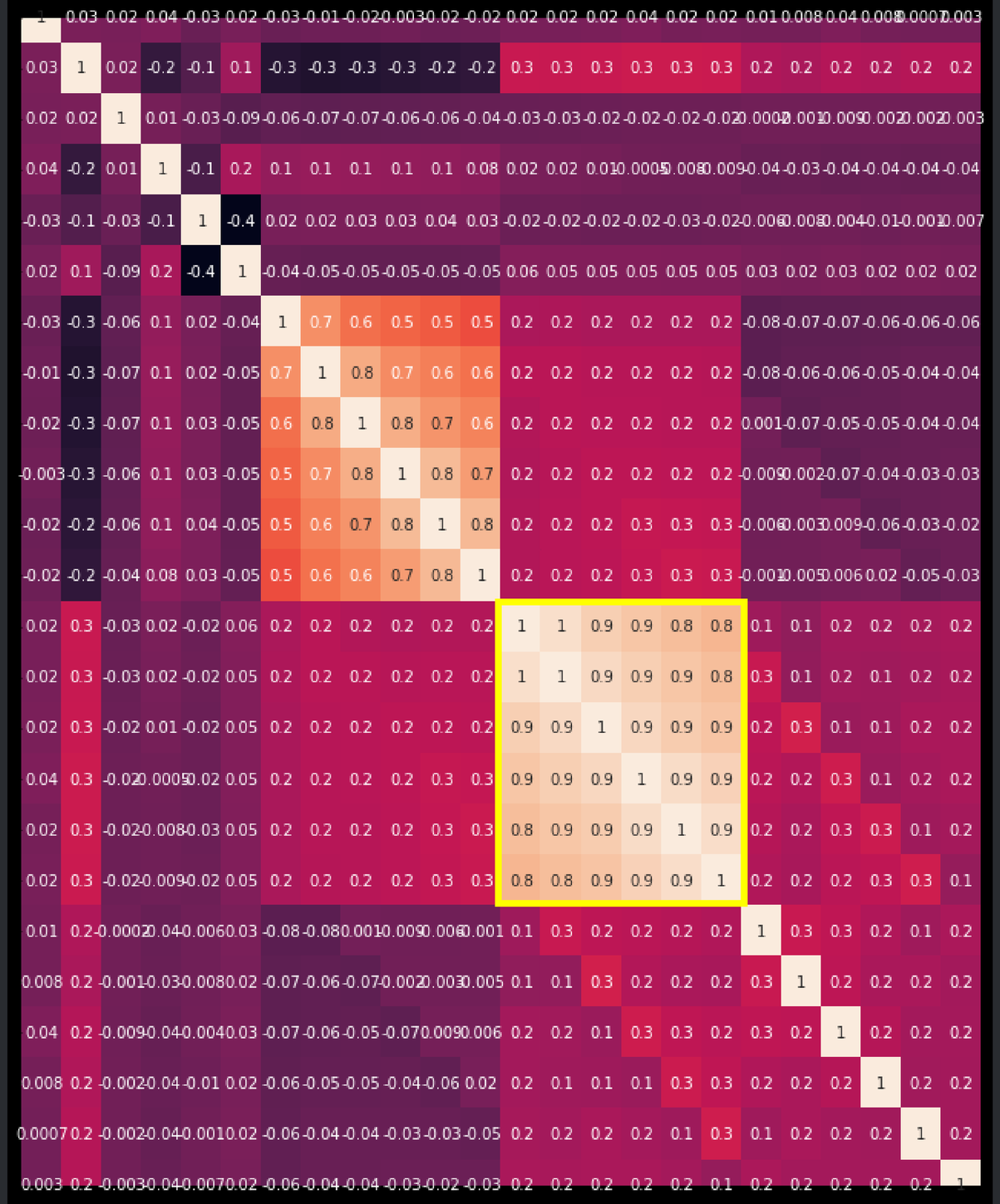


	sepal length (cm)	sepal width (cm)	petal length (cm)	petal width (cm)
0	5.1	3.5	1.4	0.2
1	4.9	3.0	1.4	0.2
2	4.7	3.2	1.3	0.2



	pca후 feature_1	pca후 feature_2
0	-2.576120	0.474499
1	-2.415322	-0.678092
2	-2.659333	-0.348282


syrup
HANA CARD
VISA
KOREAN AIR
KE017
KAL
KDR



상관관계가 매우 높은,

다중공선성이 보이는

피처들 확인