Dimension Reduction

O dimension

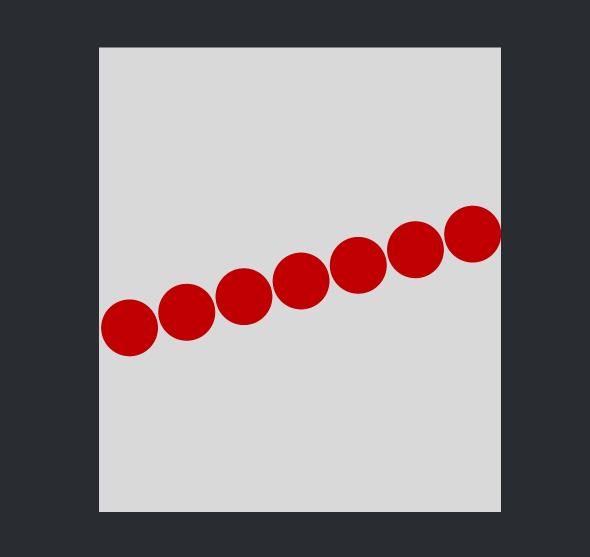
1 dimension

X
1
2
3
4
5
6
7



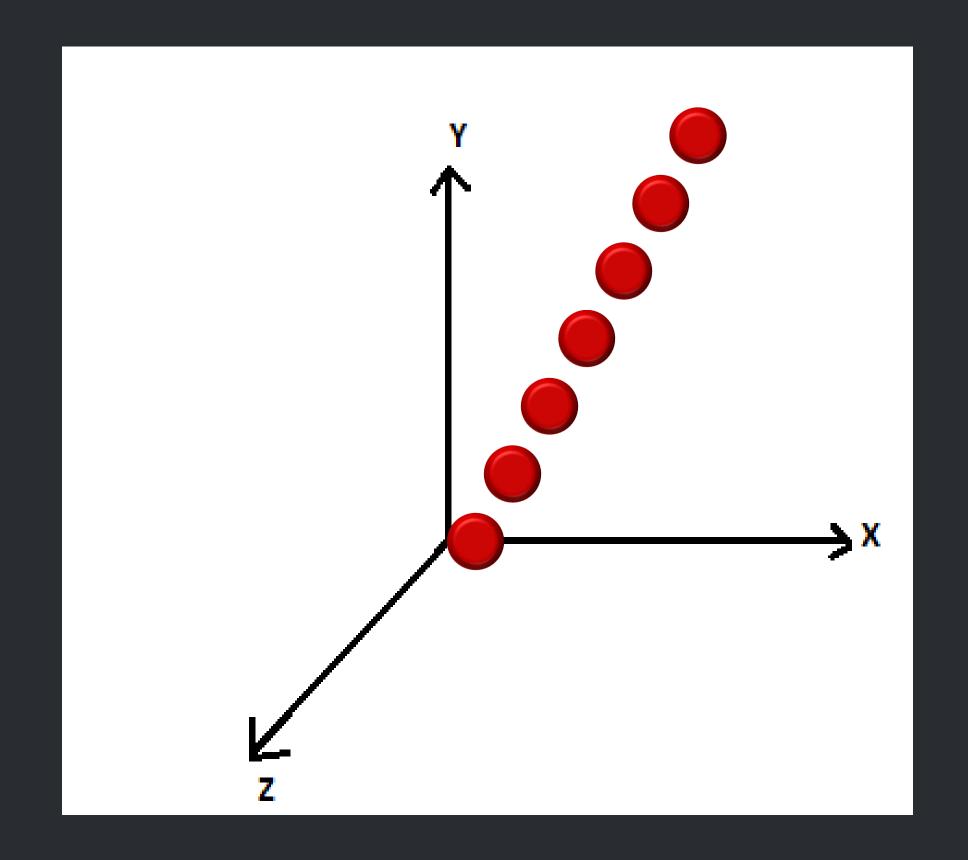
2 dimensions

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



3 dimensions

x2	x 1	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	x 1	y	хЗ	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!

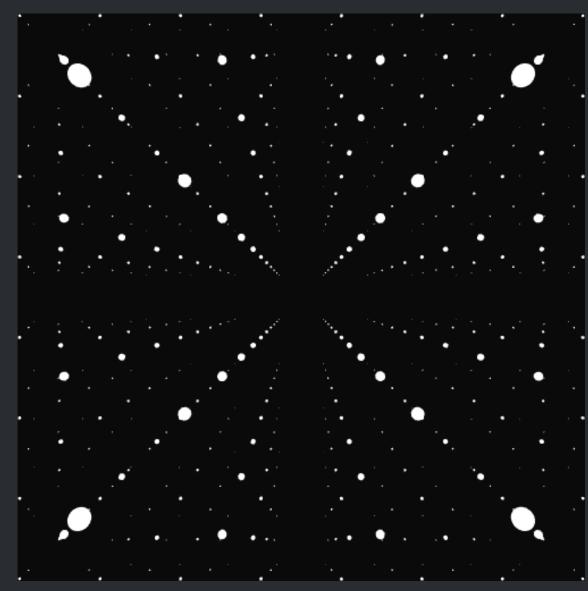
4 dimension?

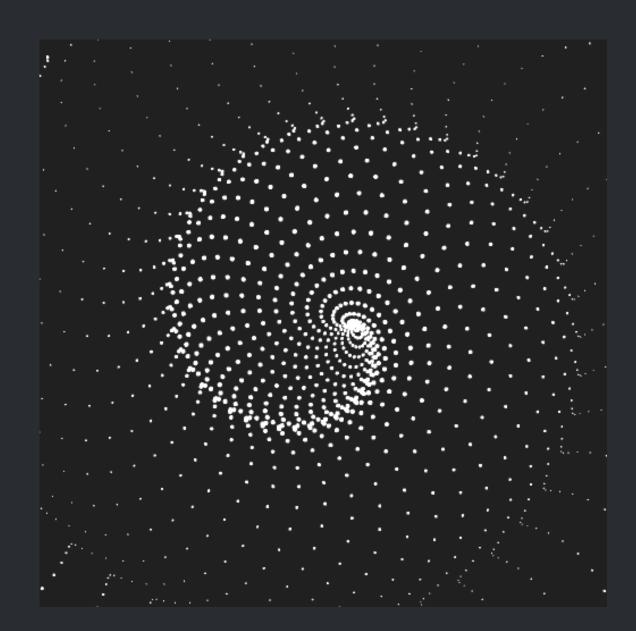
7 dimension?

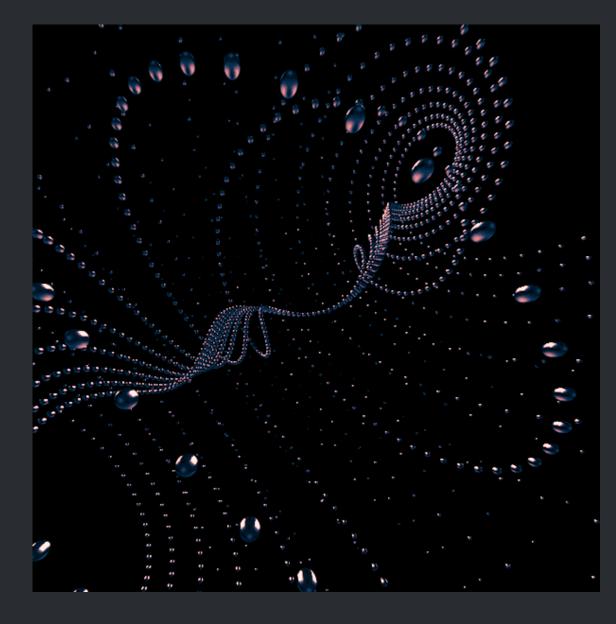
50 dimension?

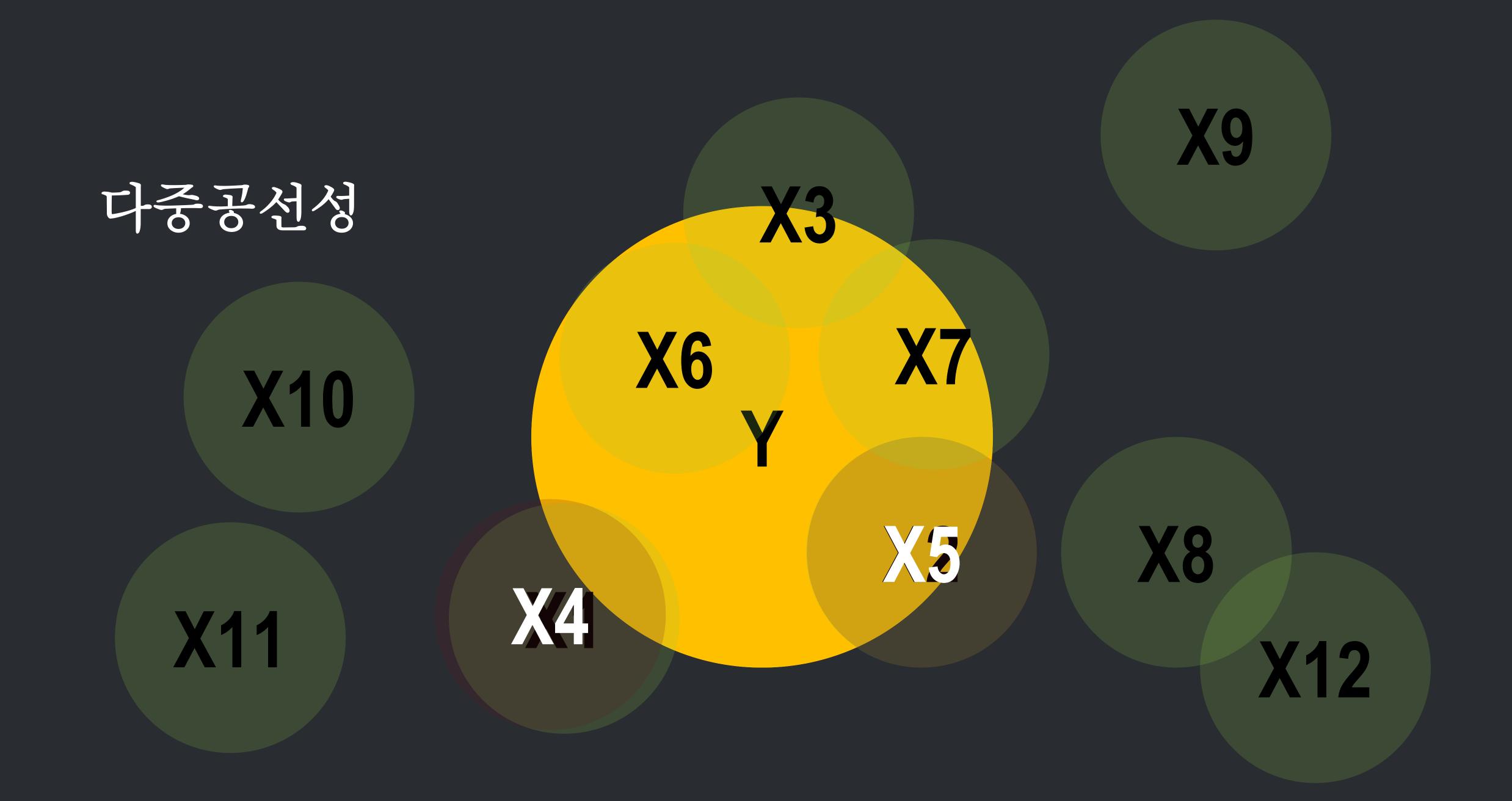
100 dimension?

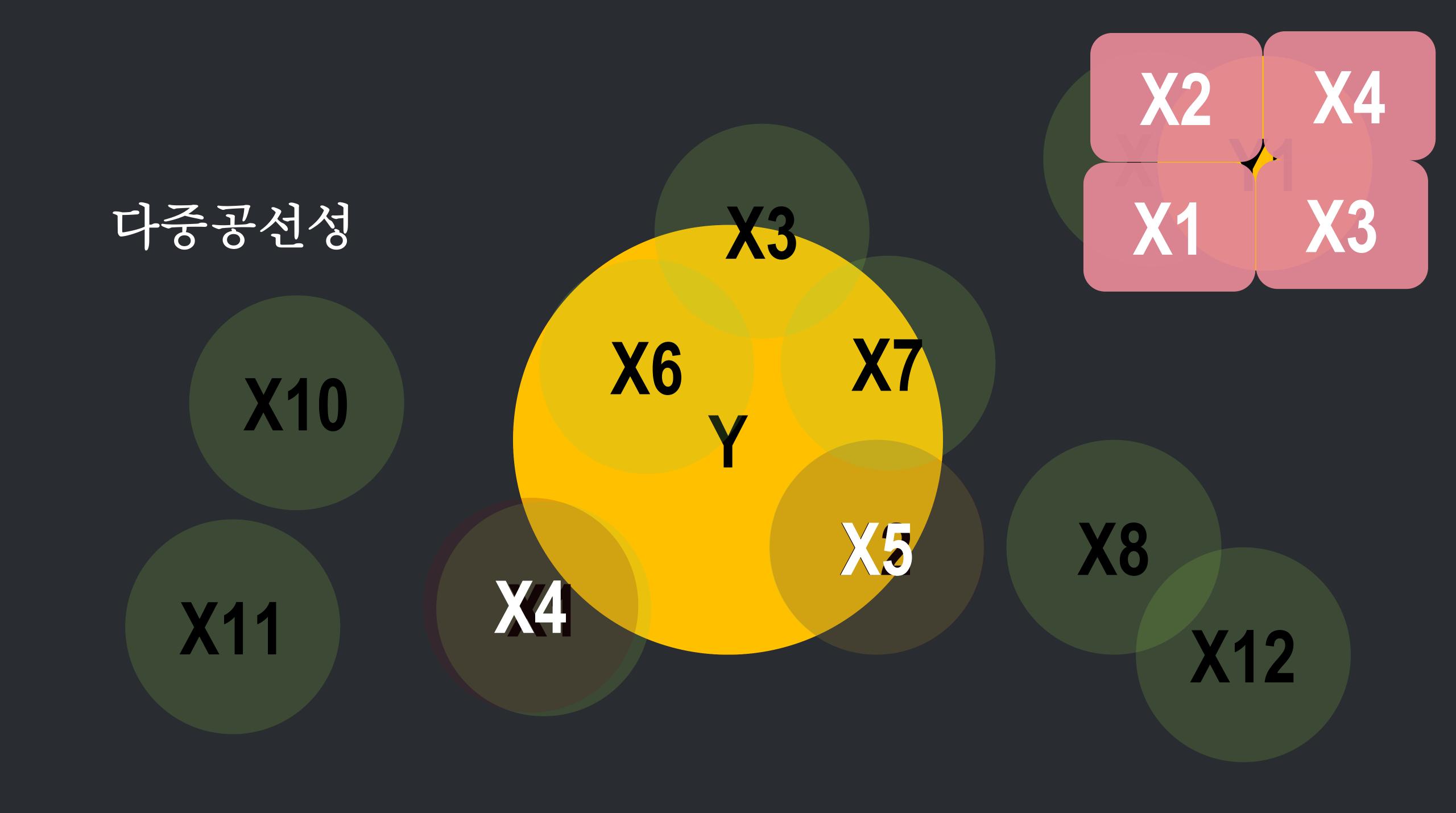


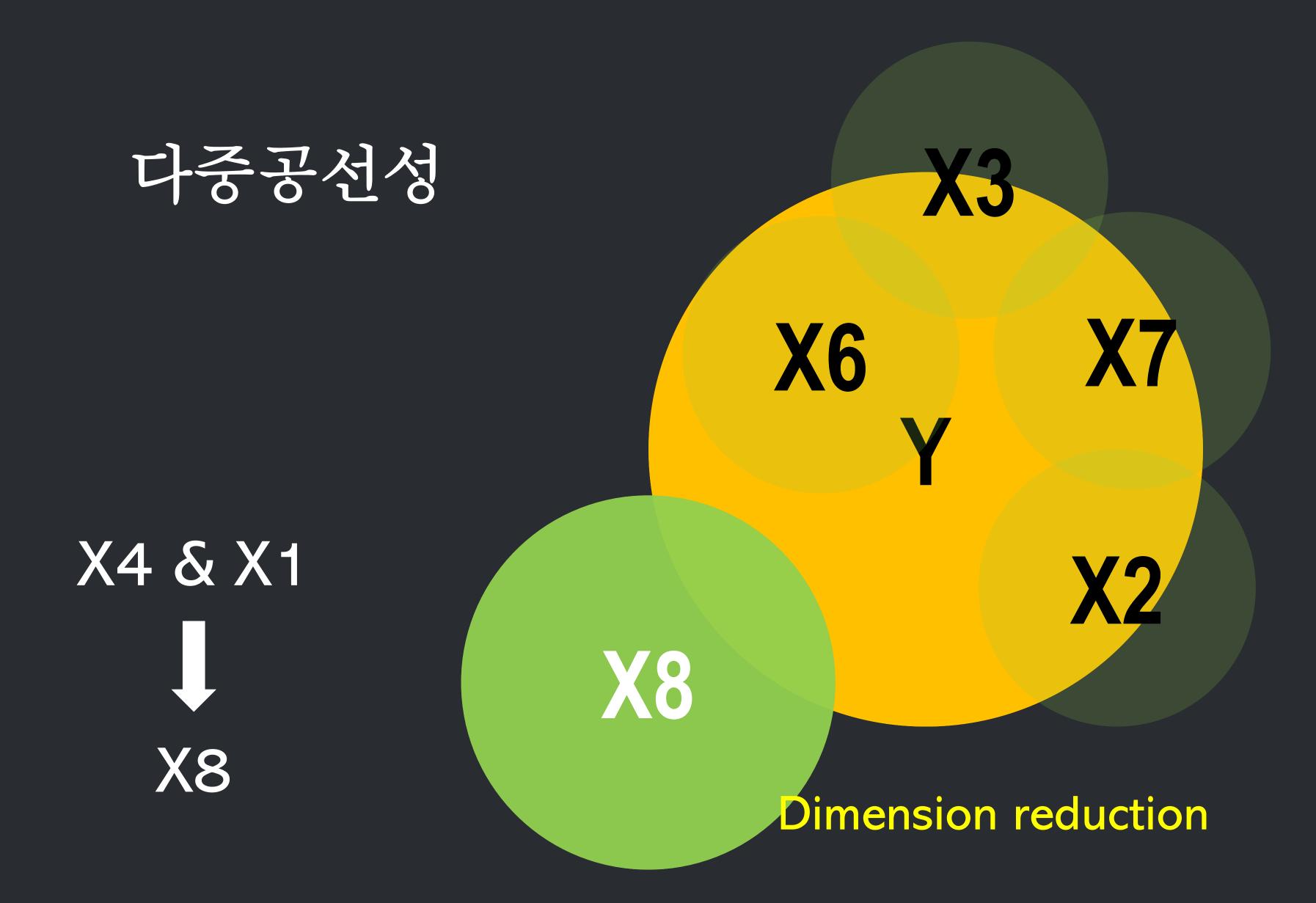




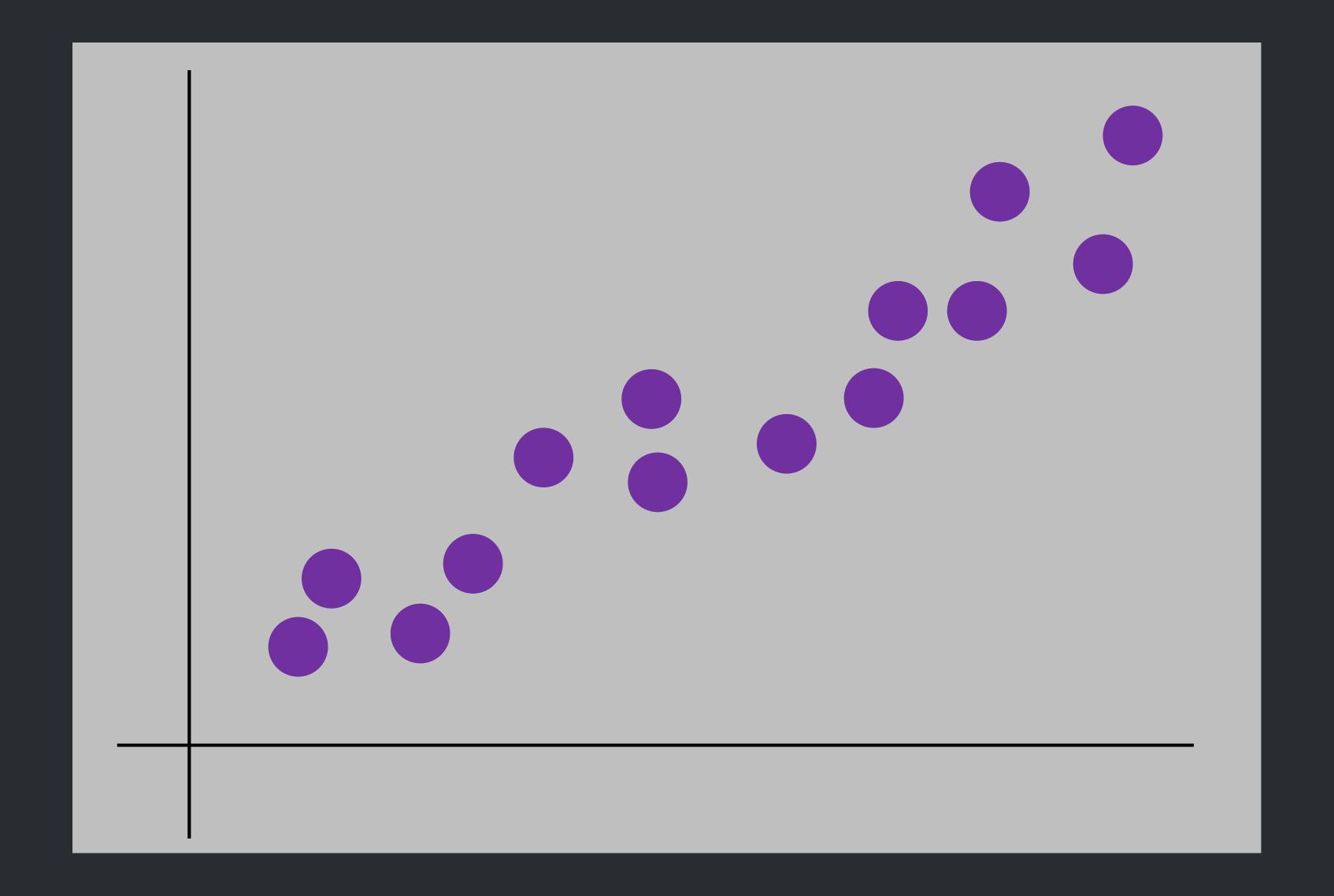


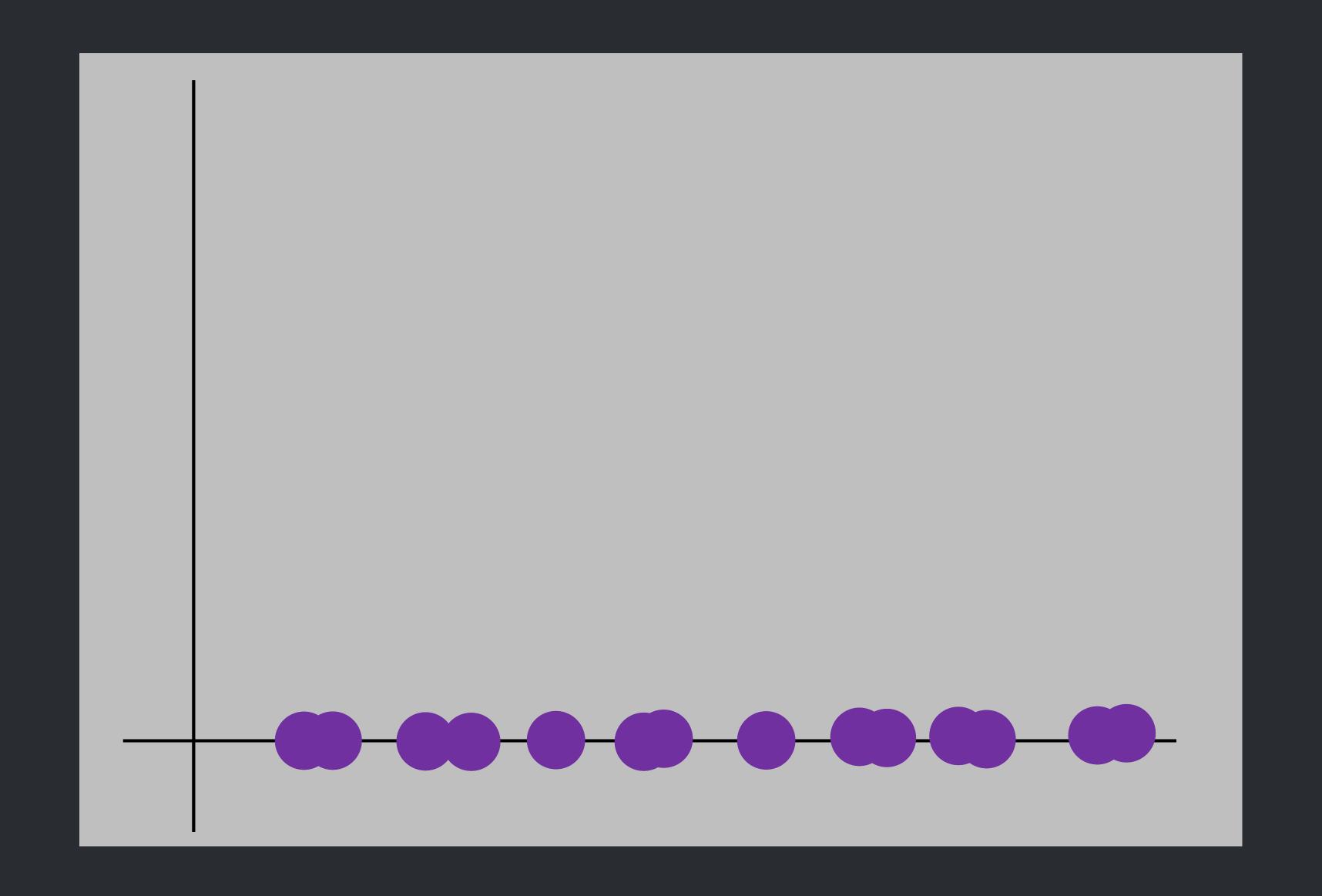


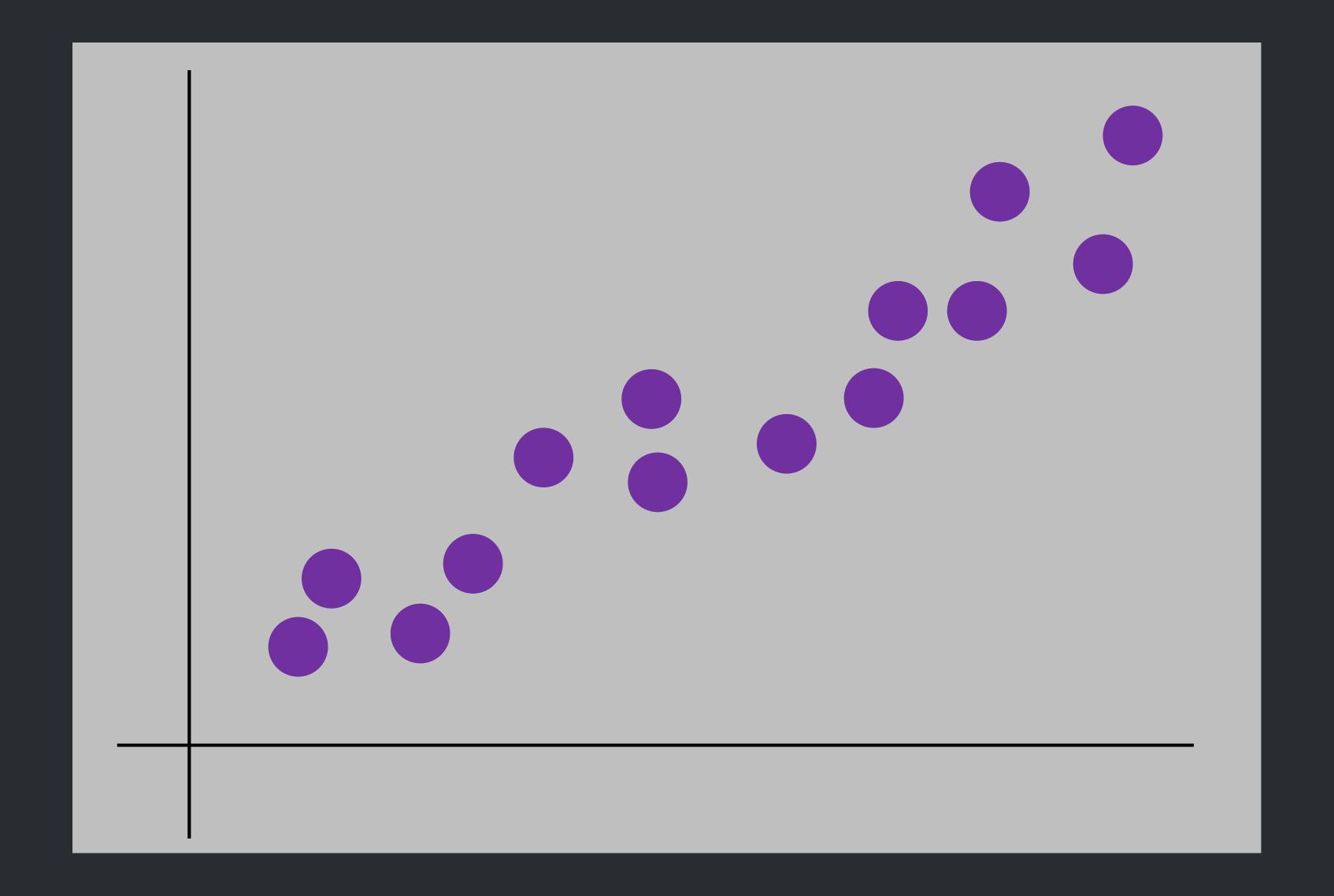


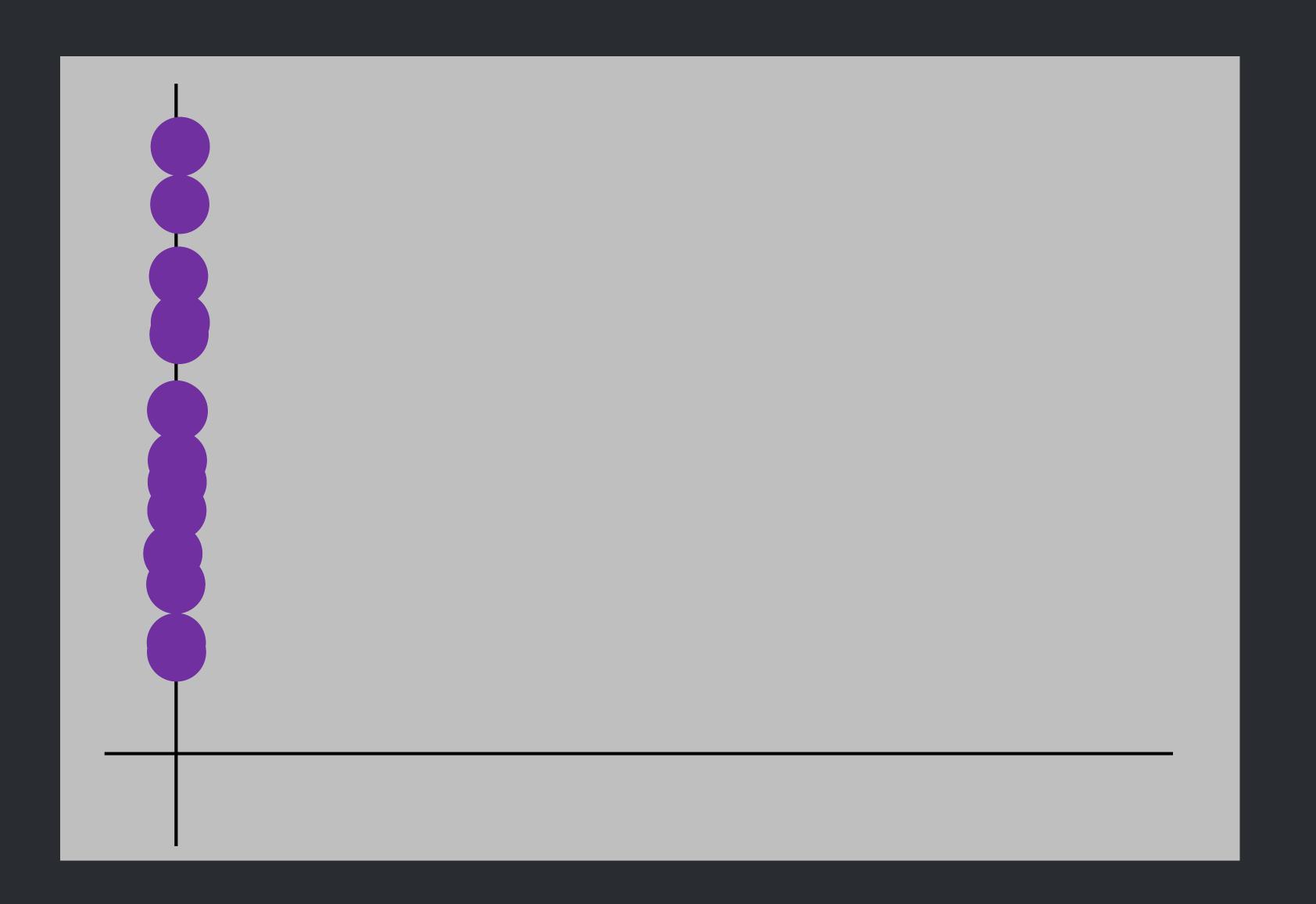


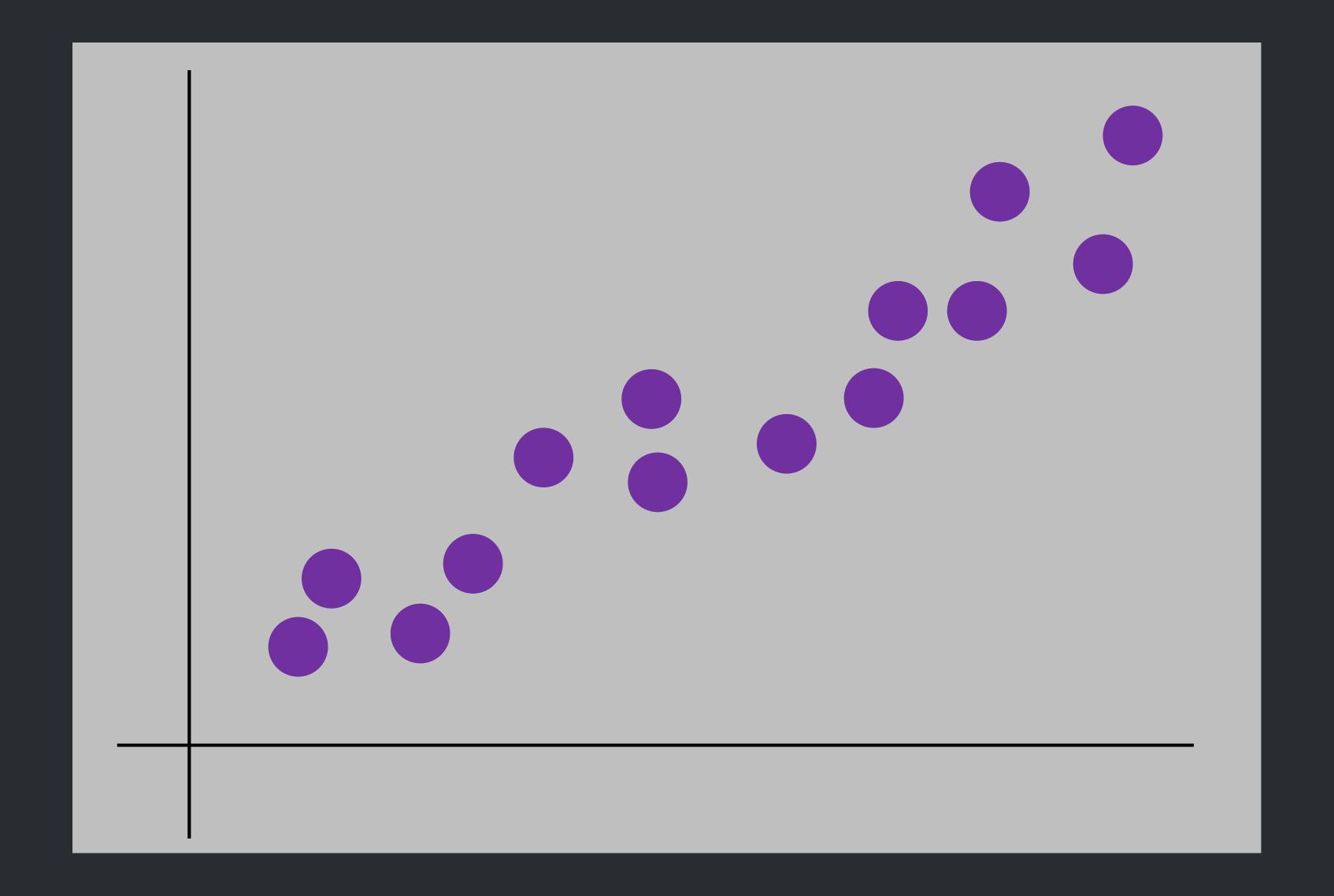
PCA



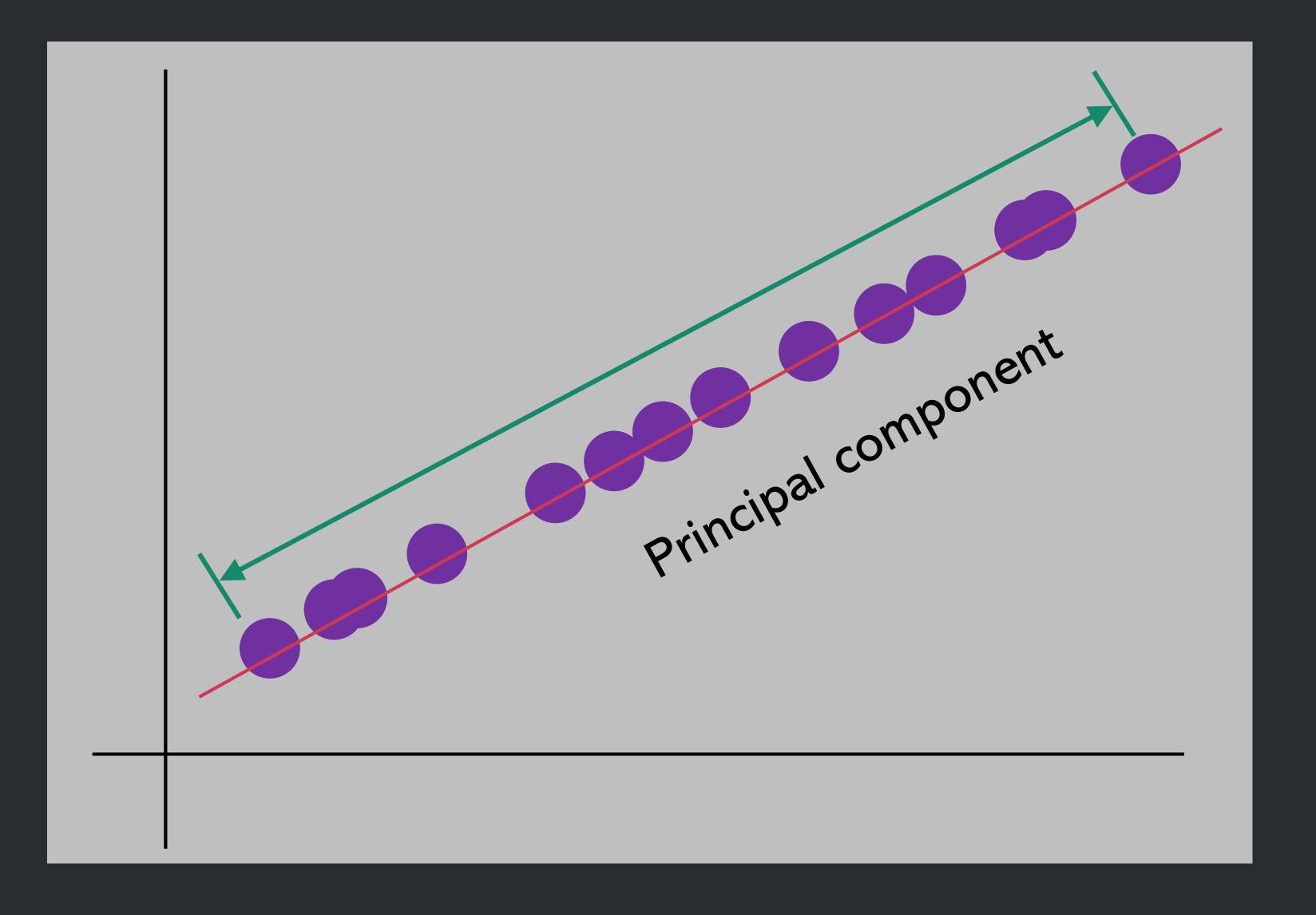








Find maximum eigen value



어떻게 컴퓨터가 자동으로 찾을 수 있을까?

Find maximum eigen value

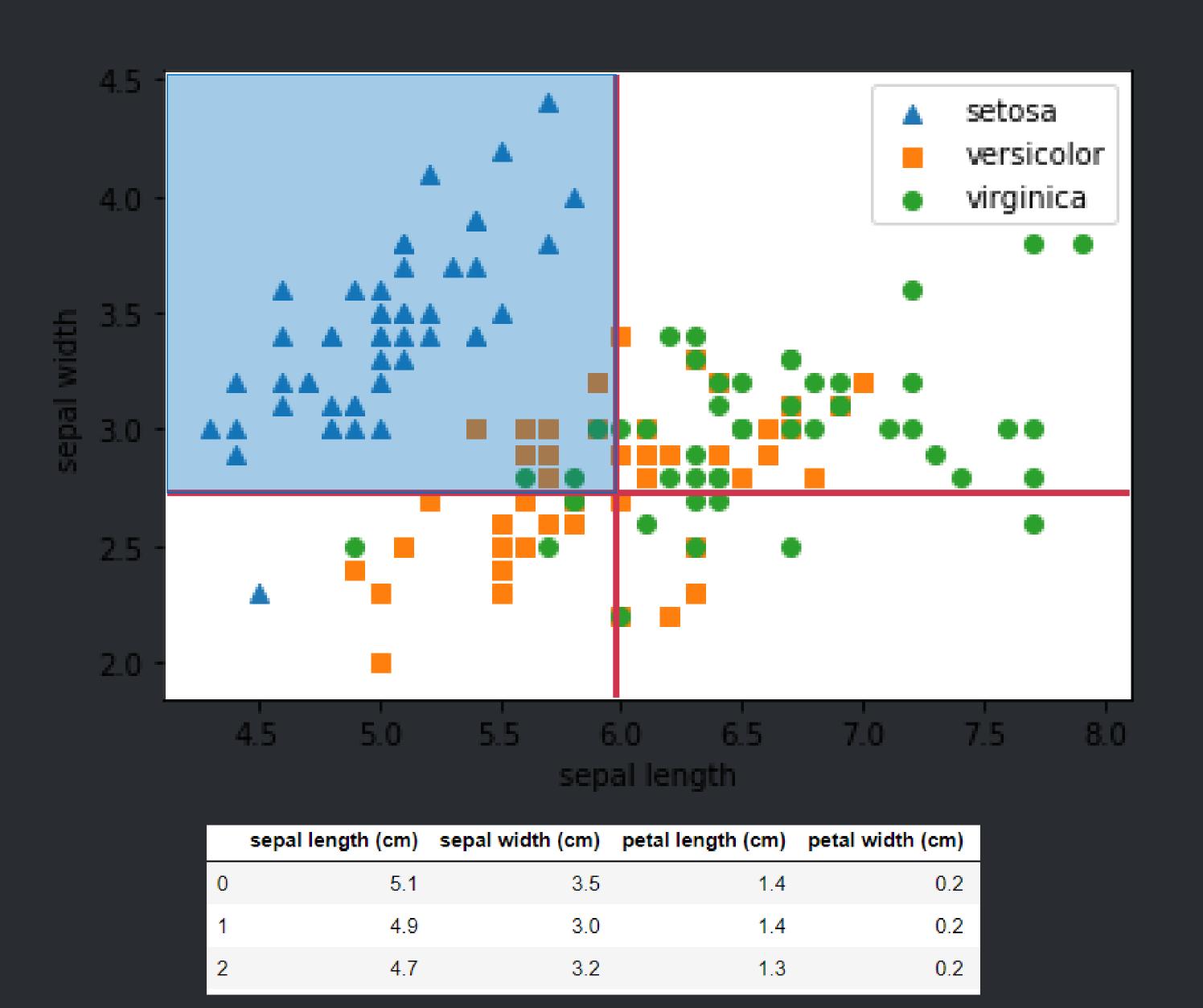
$$A oldsymbol{v} = egin{pmatrix} a_{11} & a_{12} & a_{13} \ a_{21} & a_{22} & a_{23} \end{pmatrix} egin{pmatrix} v_1 \ v_2 \ v_3 \end{pmatrix} = egin{pmatrix} a_{11} v_1 + a_{12} v_2 + a_{13} v_3 \ a_{21} v_1 + a_{22} v_2 + a_{23} v_3 \end{pmatrix}$$

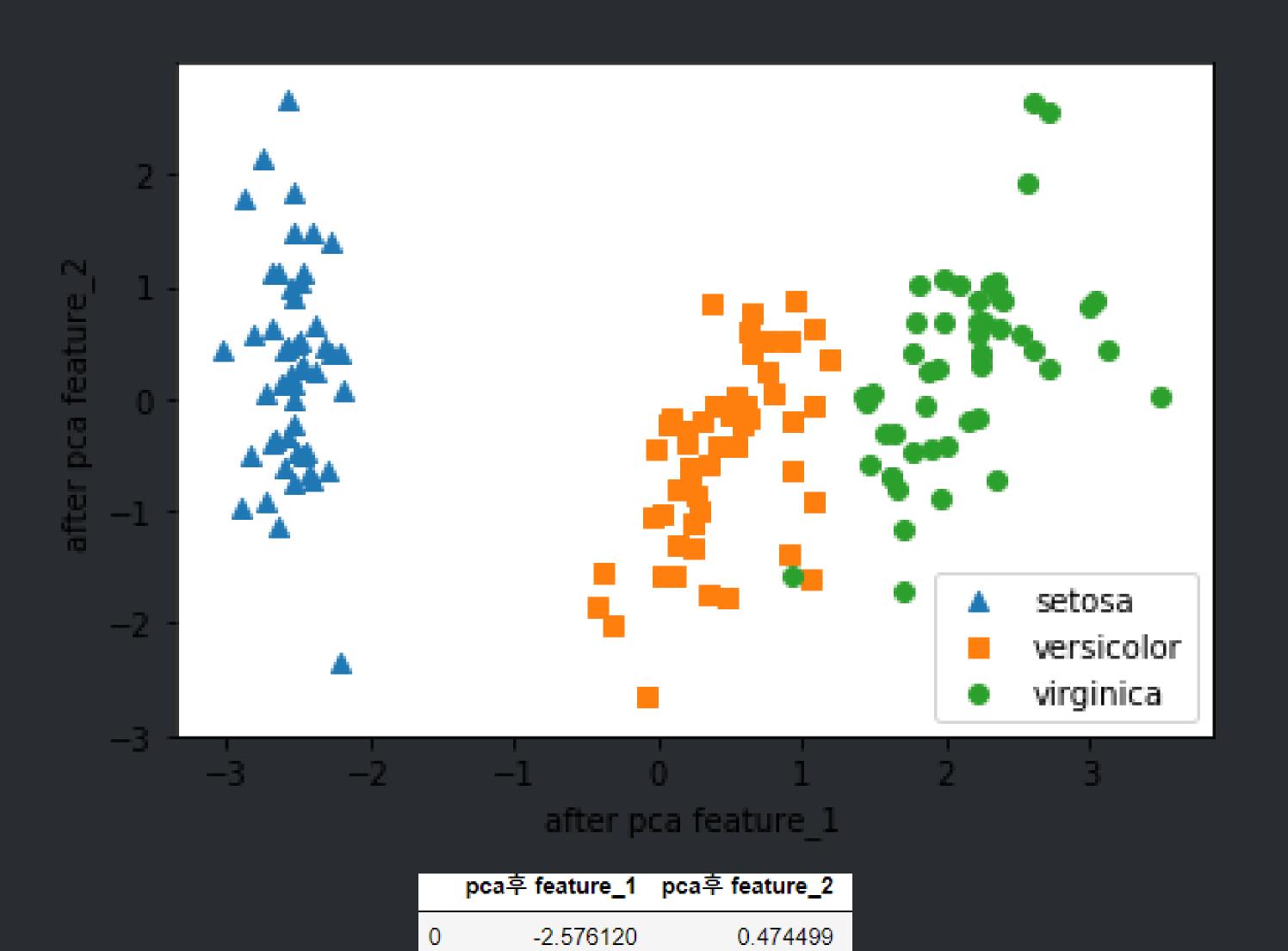
$$A=egin{pmatrix} 2&0&-2\1&1&-2\0&0&1 \end{pmatrix}$$
이고 $oldsymbol{v}=egin{pmatrix} 1\1\1\end{pmatrix}$ 라면,

$$A oldsymbol{v} = egin{pmatrix} 2 & 0 & -2 \ 1 & 1 & -2 \ 0 & 0 & 1 \end{pmatrix} egin{pmatrix} 1 \ 1 \ 0 \end{pmatrix} = egin{pmatrix} 2 \ 2 \ 0 \end{pmatrix} = egin{pmatrix} 2 \ v \ eigenvalue \end{pmatrix}$$

Jupyter notebook

n_components = 2





-2.415322

-2.659333

2

-0.678092

-0.348282



- 1	0.03	0.02	0.04	-0.03	0.02	-0.03	-0.01	-0.02	0.003	30.02	-0.02	0.02	0.02	0.02	0.04	0.02	0.02	0.01	0.008	0.04	0.008	0.000	0.003
0.03	1	0.02	-0.2	-0.1	0.1	-0.3	-0.3	-0.3	-0.3	-0.2	-0.2	0.3	0.3	0.3	0.3	0.3	0.3	0.2	0.2	0.2	0.2	0.2	0.2
0.02	0.02	1	0.01	-0.03	-0.09	-0.06	-0.07	-0.07	-0.06	-0.06	-0.04	-0.03	-0.03	-0.02	-0.02	-0.02	-0.02	0.000	2 00:	10.009	90.002	20.002	0.003
0.04	-0.2	0.01	1	-0.1	0.2	0.1	0.1	0.1	0.1	0.1	0.08	0.02	0.02	0.010	0.000	5 0.008	30 .009	30.04	-0.03	-0.04	-0.04	-0.04	-0.04
-0.03	-0 1	-0 03	-0.1	1	-0 4	0.02	0 02	0.03	0.03	0 04	0.03	-0 02	-0 02	-0 02	-0 02	-0 03	-0 02	-0 006	an oos	30 004	10 01	0 003	0.007
																							0.02
-0.03																							
-0.01	-0.3	-0.07	0.1	0.02	-0.05	0.7	1	0.8	0.7	0.6	0.6	0.2	0.2	0.2	0.2	0.2	0.2	-0.08	-0.06	-0.06	-0.05	-0.04	-0.04
-0.02	-0.3	-0.07	0.1	0.03	-0.05	0.6	0.8	1	0.8	0.7	0.6	0.2	0.2	0.2	0.2	0.2	0.2	0.001	0.07	-0.05	-0.05	-0.04	-0.04
-0.003	-0.3	-0.06	0.1	0.03	-0.05	0.5	0.7	0.8	1	0.8	0.7	0.2	0.2	0.2	0.2	0.2	0.2	0.009	90.002	20.07	-0.04	-0.03	-0.03
-0.02	-0.2	-0.06	0.1	0.04	-0.05	0.5	0.6	0.7	0.8	1	0.8	0.2	0.2	0.2	0.3	0.3	0.3	0.00€	3 0.003	30.009	-0.06	-0.03	-0.02
-0.02	-0.2	-0.04	0.08	0.03	-0.05	0.5	0.6	0.6	0.7	0.8	1	0.2	0.2	0.2	0.3	0.3	0.3	0.003	10.00	30.006	0.02	-0.05	-0.03
0.02	0.3	-0.03	0.02	-0.02	0.06	0.2	0.2	0.2	0.2	0.2	0.2	1	1	0.9	0.9	0.8	0.8	0.1	0.1	0.2	0.2	0.2	0.2
0.02	0.3	-0.03	0.02	-0.02	0.05	0.2	0.2	0.2	0.2	0.2	0.2	1	1	0.9	0.9	0.9	0.8	0.3	0.1	0.2	0.1	0.2	0.2
0.02	0.3	-0.02	0.01	-0.02	0.05	0.2	0.2	0.2	0.2	0.2	0.2	0.9	0.9	1	0.9	0.9	0.9	0.2	0.3	0.1	0.1	0.2	0.2
0.04	0.3	-0 020	000	5 0 02	0.05	0.2	0.2	0.2	0.2	0.3	0.3	0.9	0.9	0.9	1	0.9	0.9	0.2	0.2	0.3	0.1	0.2	0.2
						0.2																	
						0.2																	
0.01	0.24	0.000	20.04	0.006	50.03	-0.08	-0.08	0.001	200.00	0.00€	0.001	0.1	0.3	0.2	0.2	0.2	0.2	1	0.3	0.3	0.2	0.1	0.2
0.008	0.2	0.001	-0.03	0.008	30.02	-0.07	-0.06	-0.07	0.002	0.003	0.005	0.1	0.1	0.3	0.2	0.2	0.2	0.3	1	0.2	0.2	0.2	0.2
0.04	0.2	0.009	0.04	0.004	10.03	-0.07	-0.06	-0.05	-0.07	0.009	0.006	0.2	0.2	0.1	0.3	0.3	0.2	0.3	0.2	1	0.2	0.2	0.2
0.008	0.2	0.002	0.04	-0.01	0.02	-0.06	-0.05	-0.05	-0.04	-0.06	0.02	0.2	0.1	0.1	0.1	0.3	0.3	0.2	0.2	0.2	1	0.2	0.2
0.0007	70.2 -	0.002	0.04	0.001	10.02	-0.06	-0.04	-0.04	-0.03	-0.03	-0.05	0.2	0.2	0.2	0.2	0.1	0.3	0.1	0.2	0.2	0.2	1	0.2
0.003	0.2	0.003	0.04	0.007	0.02	-0.06	-0.04	-0.04	-0.03	-0.02	-0.03	0.2	0.2	0.2	0.2	0.2	0.1	0.2	0.2	0.2	0.2	0.2	1

상관관계가 매우 높은, 다중공선성이 보이는

피처들 확인