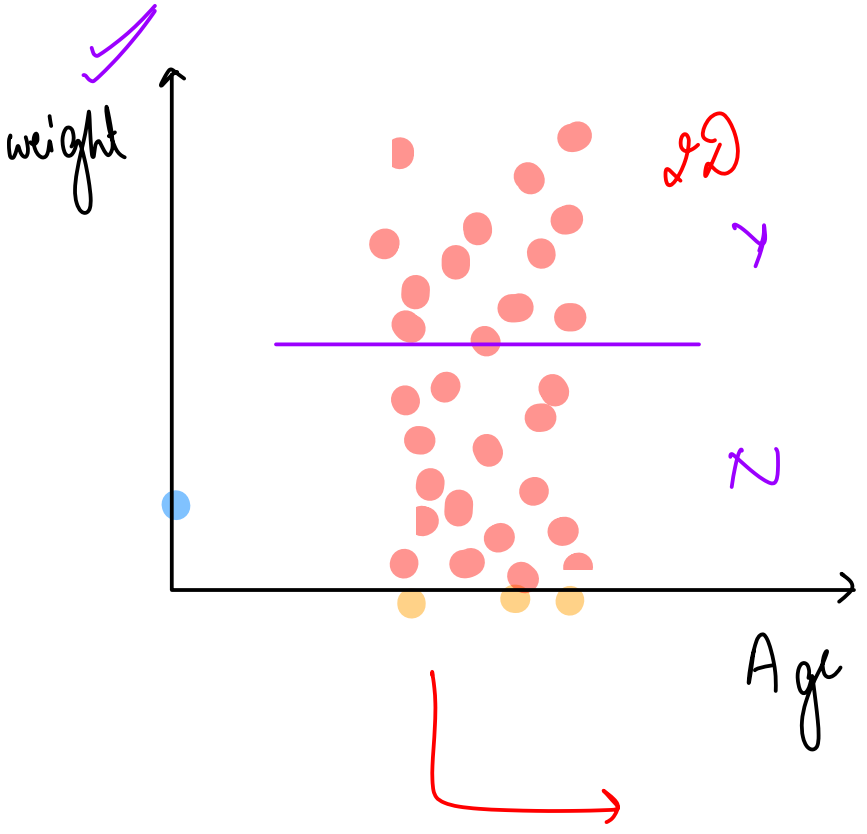
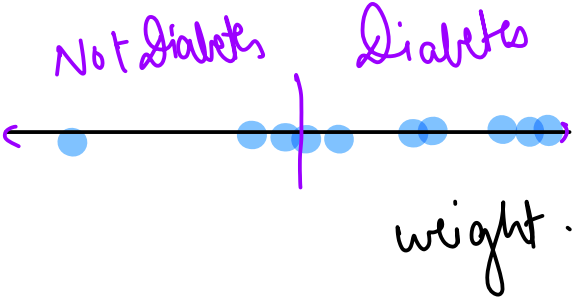
Higher Dimensional Visualisations

1. PCA - Principle Component Analysis
2. TSNE - T-distributed Stochastic Neighbour Embedding.



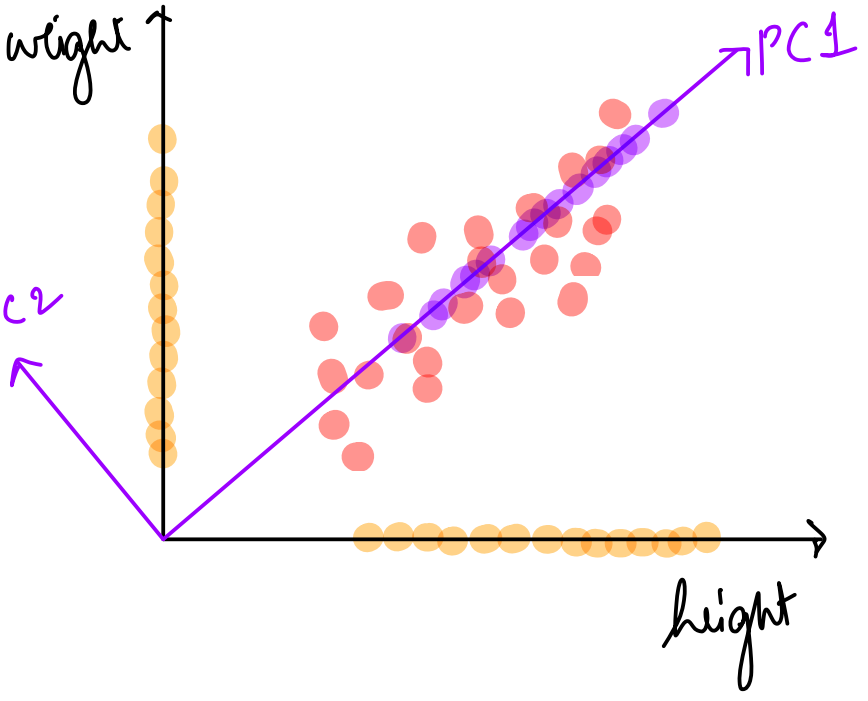
Weight	Age	Diabetes
-	-	Y
-	-	N
-	-	Y
-	-	N
-	-	Y

1D



Q Which features has most information?

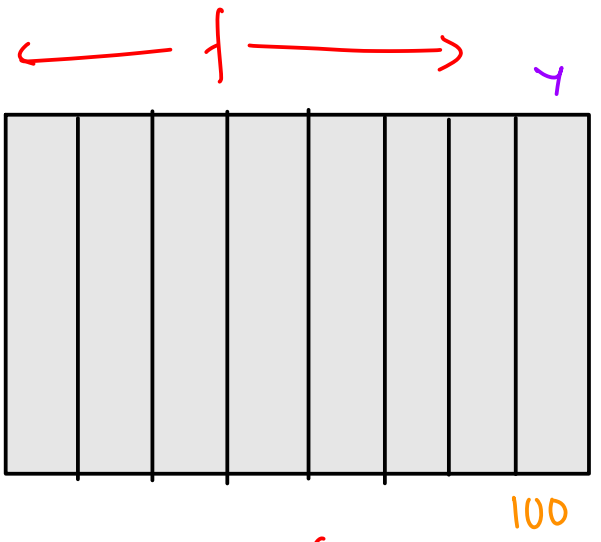
1. Weight
2. Height
3. α Weight + β Height.



Q Which PC is better / having most info preserved?

① PC1 ② PC2

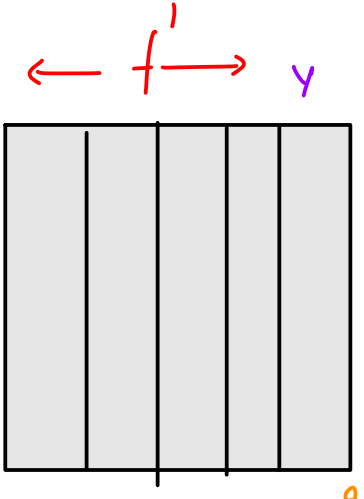
Max variance of projection.



Original Features space

100% variance/info

PCA



new feature space

90% variance/info

100% variance

100% variance

(90% - 95% Variance)

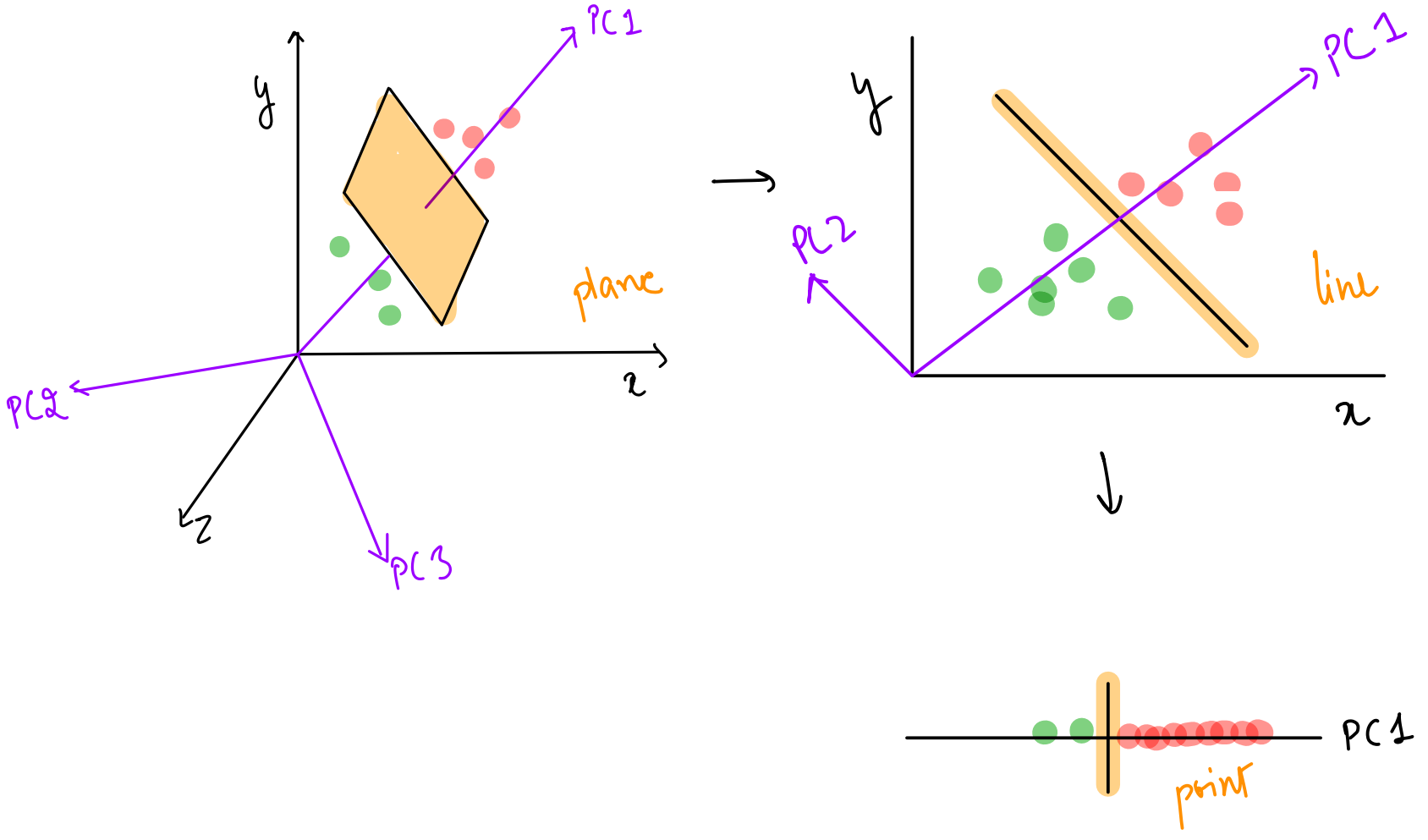
m features $\xrightarrow{\text{PCA}}$ m PCs $\xrightarrow{\text{Retain}}$ m' PC

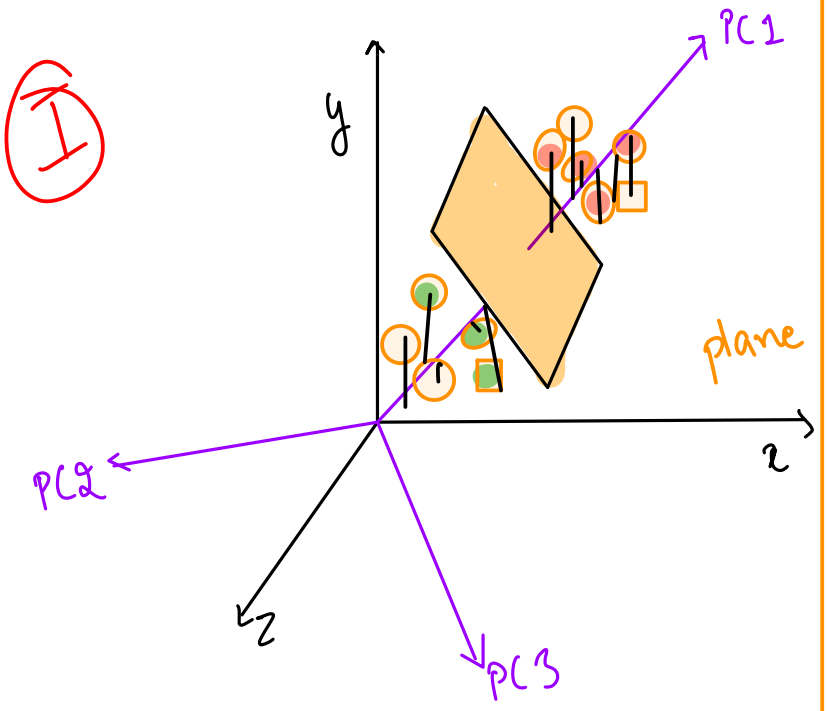
$(m' << m)$

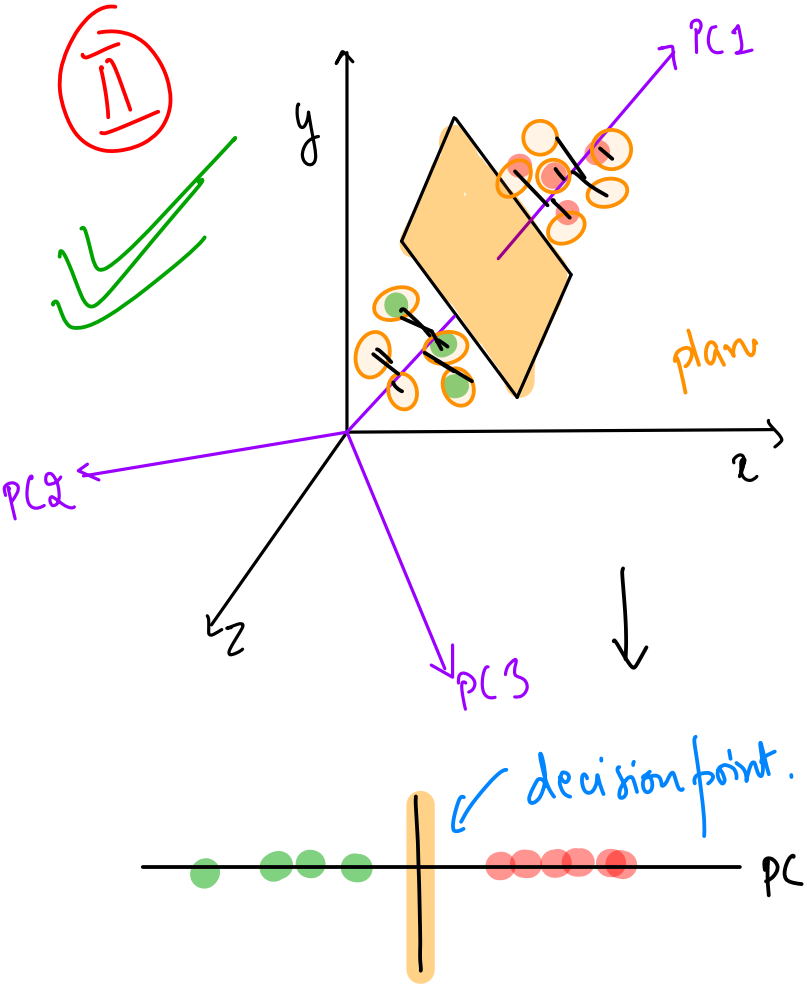
Why are we reducing dimension?

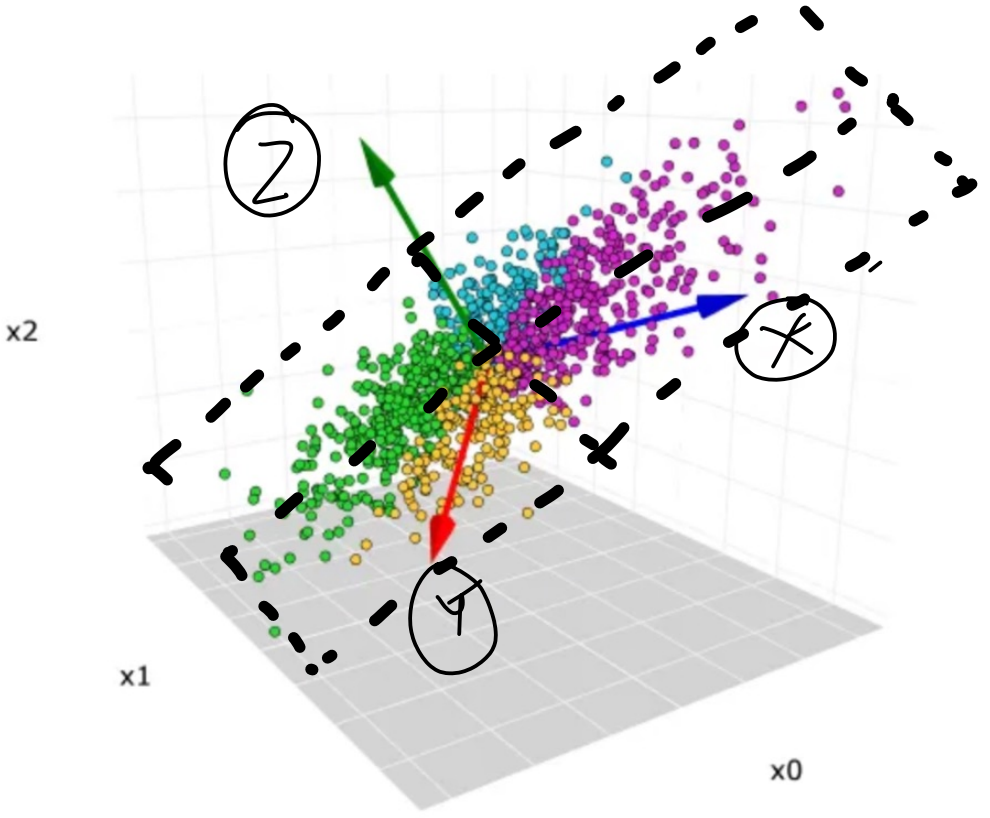
(1) Visualisation ⇒ help me take better decision
↳ Business
↳ EDA

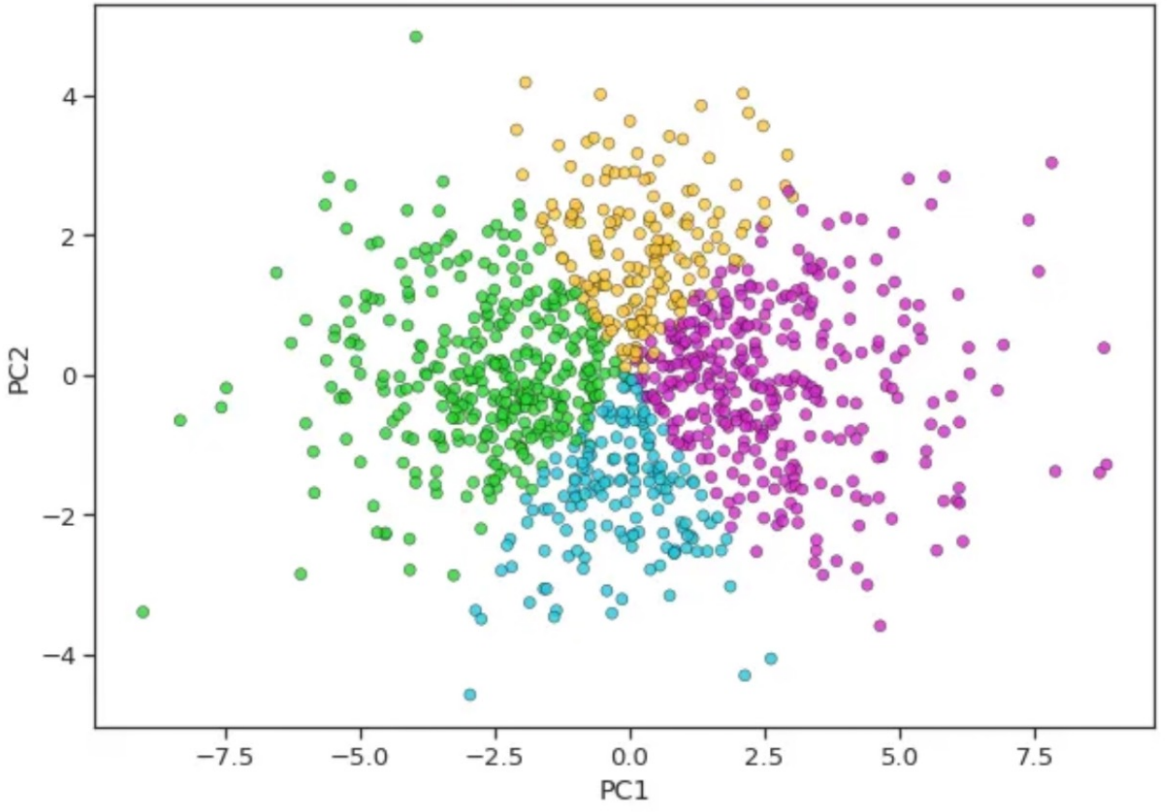
(2) Compression ⇒ Faster Training.


PC1
y
plane
x
PC2
z
PC3
y
PC1
PC2
line
x
PC1
point


I
PC1
y
plane
x
PC2
z
PC3


II
PC1
y
plan
x
PC2
z
PC3
decision point.
PC1


x2
x1
x0
Z
X
Y

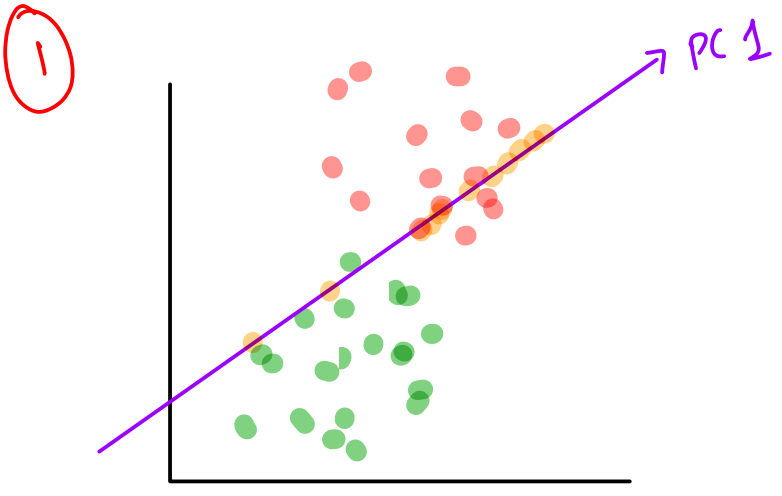

PC2
4
2
0
−2
−4
−7.5
−5.0
−2.5
0.0
2.5
5.0
7.5
PC1

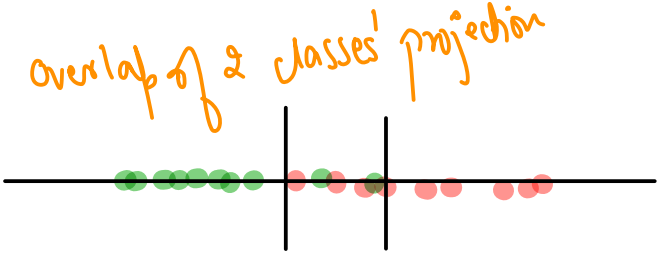
Conclusion :-

① Find new axis / dimension where max info is present.

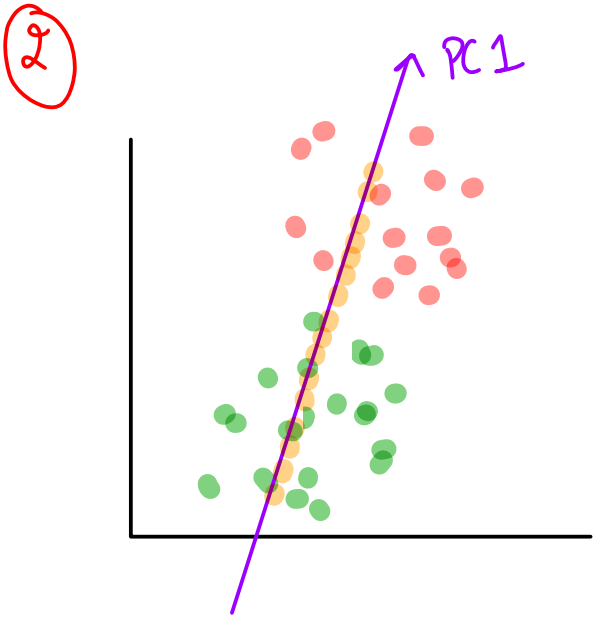
⇒ information can be stored with less # of features.

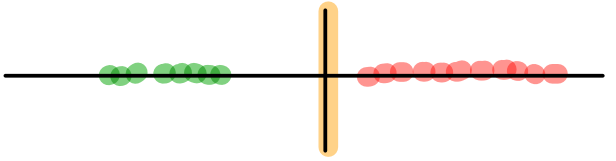
There is always some loss of information.





drop in accuracy





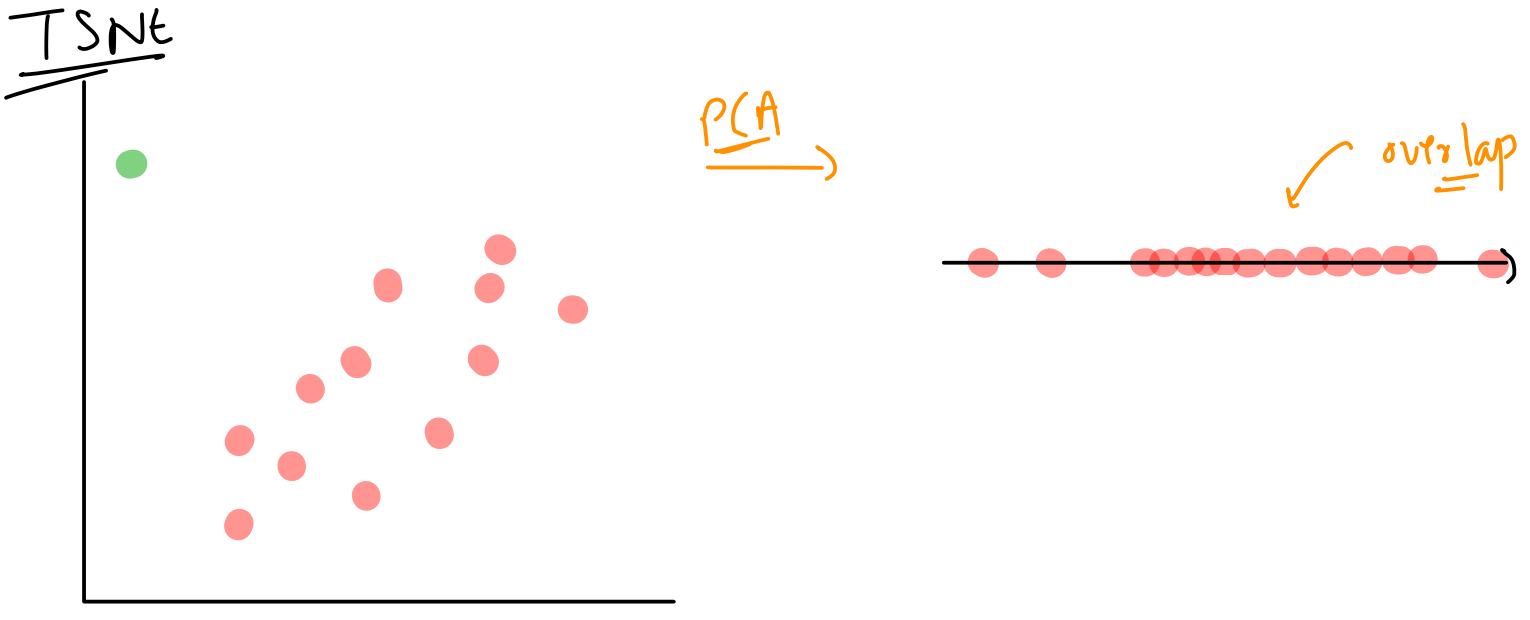
100% accuracy

5 dims $\xrightarrow{\text{PCA}}$ 5 PCs

→	10	PC 1
→	7	PC 2
→	6	PC 3
→	3	PC 4
→	1	PC 5

Eigenvalues

explained variance by Top 3 PC $= \frac{10 + 7 + 6}{10 + 7 + 6 + 3 + 1} = 85\%$

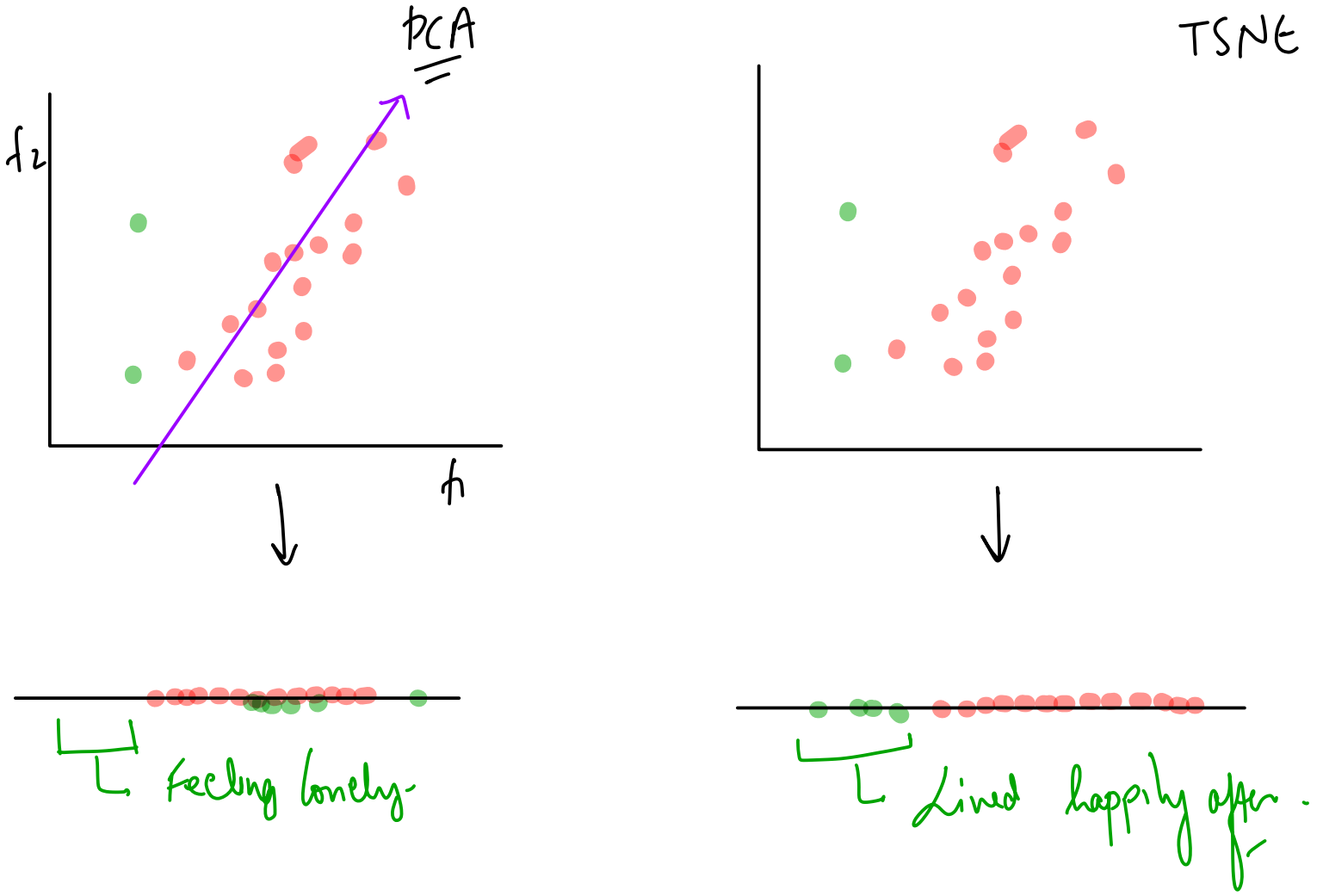


TSNE : Preserve spatial arrangement of datapoint in lower dimension.

If you are my 3rd nearest neighbour in d dimensions

I want you to be my 3rd nearest neighbour in d' dimensions

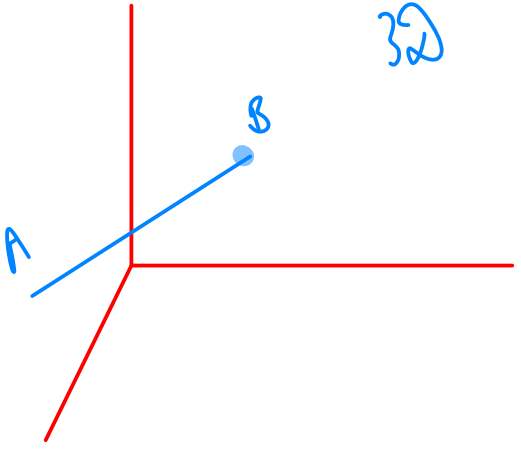
$[d' << d]$


PCA
TSNE
f2
f1
Feeling lonely.
Lived happily after.

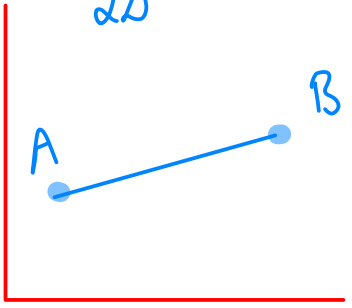
Q Can we compare distance b/w 2 points in diff. dimensions?

(a) Yes (b) No

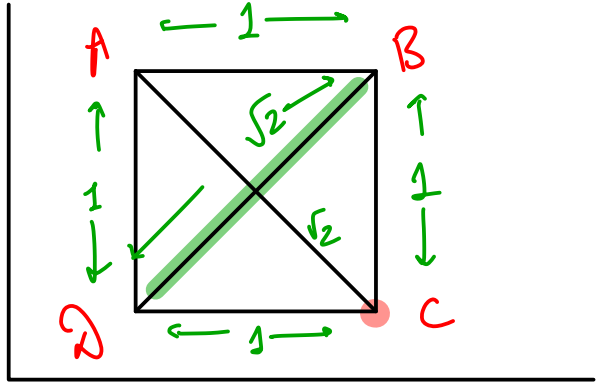
3D



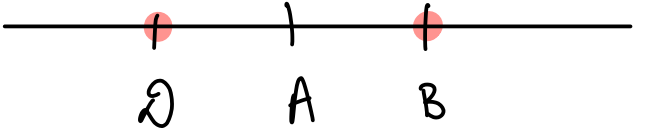
2D



1D



$\Rightarrow$

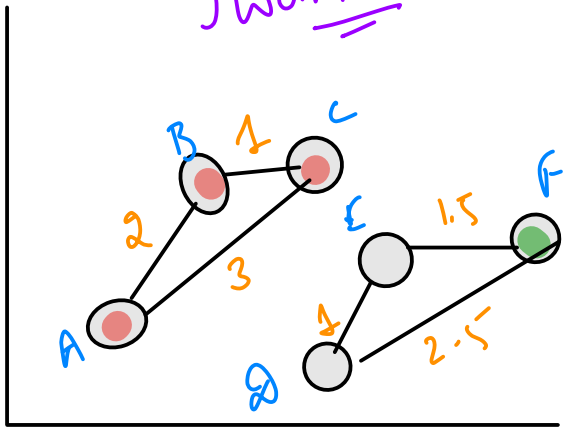


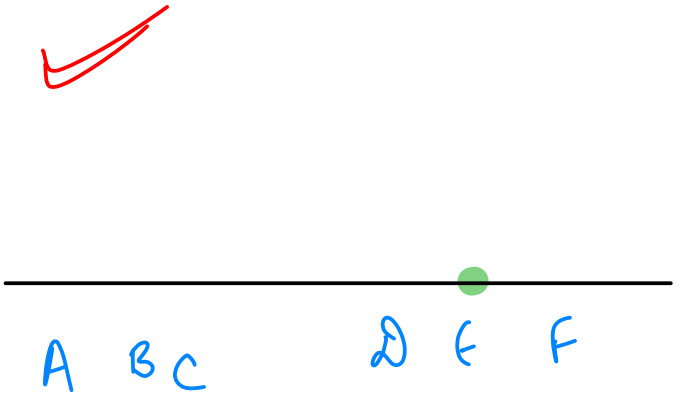
d dimension = 2

d' dimension = 1

Wherever you place point 'C', it will violate some of the condition.

I want this





What do we do?

How do we define probability that point x_i is in neighbourhood of point x_j?

$$P_{ij} \propto \frac{1}{\text{distance } i\,j}$$

$$P_{ij} \propto \frac{1}{\text{dist}(x_i, x_j)} = \frac{1}{\| x_i - x_j \|_2}$$

$$\|x_i - x_j\|_2^{-1} \Rightarrow \boxed{\text{dist}^{-1}}$$