Often it is preferrable to compress data. This can be done with minimal information loss using projection. This is related to how similar two datapoints are.

In projection we would be interested in maintaining two properties:

1. distance between vectors
2. angles between vectors

(*) This can be done via the inner product.
It is a tool to measure angles, lengths and distances.

The Dot Product

$$a \cdot b = a^T b = \sum^n a_i b_i$$

The length of a vector u is,

$$\|u\| = \sqrt{u \cdot u} = \sqrt{\sum^n x_i^2}$$

To calculate the distance between two vectors can apply the norm to the difference of the vectors.

$$d(x, y) = \|x - y\| = (x-y) \cdot (x-y) = (x-y)^T (x-y)$$

To calculate the angle between two vectors can use the dot product.

$$\cos \alpha = \frac{x^T y}{\|x\| \|y\|}$$

(*) based on a manipulation of the cosine rule.

$$s \cdot r = \|r\| \|s\| \cos \theta.$$