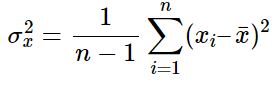
The Covariance Matrix

**Variance:**Variance helps us understand how far our random variable is spread out from the mean.

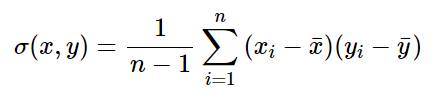
Let's use a hypothetical income distribution of people as an example for reference. The formula for variance is given by:



where **n** is the number of samples (ex: the number of people in the income distribution) and **x̄** is the mean of the random variable **x** (ex: the mean of the income).

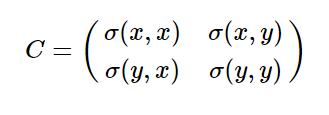
**Covariance:**Covariance, on the other hand, measures the extent to which two random variables vary together.

For a similar example, let's consider the income of a person and the expenditure of that person in a population. The formula for covariance is given by:



where **n** is the number of samples (ex: the number of people) and **x̄** is the mean of the random variable **x** (represented as a vector). The variance, **σ2(x)**, of a random variable, **x**, can also be expressed as the covariance with itself, i.e, **σ(x,x)**.

**Covariance Matrix:**Following from the previous equations, the covariance matrix for two dimensions is given by:



In this matrix, the variances appear along the diagonal and the covariances appear in the off-diagonal elements.

**Note:**The function **numpy.cov()**in Python's Numpy variable can be used to get the covariance matrix in Python.