Why we need the concept of covariance

Want to do is calculate the variance of a sum of random variables, where these random variables may or may not be independent.

$$\mathrm{Var}\left(\sum_{i=1}^{n} X_i\right) = E\left(\left(\sum_{i=1}^{n} X_i\right)^2\right) - \left(E\left(\sum_{i=1}^{n} X_i\right)\right)^2$$

$$= E\left(\left(\sum_{i=1}^{n} X_i\right)\left(\sum_{j=1}^{n} X_j\right)\right) - E\left(\sum_{i=1}^{n} X_i\right)E\left(\sum_{j=1}^{n} X_j\right)$$

$$= \sum_{i=1}^{n}\sum_{j=1}^{n} E(X_i X_j) - \sum_{i=1}^{n}\sum_{j=1}^{n} E(X_i)E(X_j)$$

$$= \sum_{i=1}^{n}\sum_{j=1}^{n} \underbrace{\left(E(X_i X_j) - E(X_i)E(X_j)\right)}$$

equivalent to the covariance of X_i and X_j

So the covariance of two random variables plays a really important role when taking a sum of random variables and finding the variance of such a sum.