Below is the formula for Moran’s I:

The formula has several components:

* is the mean of the variable X
* *Xi* is the variable value at a particular location *i*
* *Xj* is the variable value at another location *j*
* *wij* is a weight indexing location of *i* relative to *j*
* *n* is the number of observations (points or areal units)

Here are the steps to calculate Moran’s I:

1. First, we create the weight matrix *W* – for example, it might be a Queen, Rook or a distance-based matrix. Looking at the value of *W* in row *i* and column *j*, denoted as *wij*, will tell us whether observations *i* and *j* are neighbors.
   * Imagine we’re using Queen weights. In that case, if observations *i* and *j* are Queen neighbors, *wij* will take on the value of 1; otherwise, it will take on the value of 0.
2. Using all observations in our data set, we calculate .
3. For each pair of locations *i* and *j*, we do the following:
   1. Calculate how far away X at location *i* is from the mean . This is
   2. Calculate how far away X at location *j* is from the mean . This is
   3. Multiply by .
   4. Multiply by *wij*.
      1. So, if *i* and *j* are neighbors, *wij* will take on the value of 1, and will be .
      2. On the other hand, if *i* and *j* are *not* neighbors, wij will take on the value of 0, and will be 0.
4. We sum for all pairs of locations *i* and *j*, and divide that sum by , which is how many pairs of observations there are that are neighbors.
   1. So, gives us an idea of how much values of X at nearby locations co-vary.
5. Note that the term in the denominator is simply the variance of X, which is the sum of squared deviations of X from the global mean , without any regard to values of X at locations that are near it.
6. So when we divide by , we see how much nearby values of X co-vary relative to the overall variation in X. And this is exactly what we get at with Moran’s I!