NOTE: in the course of rewriting the dataset creation code in VBA for Excel, I discovered that the original worked example in my thesis had a couple of errors, and also needed some clarifications. I should have had more zones in the sample area, or used fewer variables, because the use of the zero eigenvalue's eigenvector was impossible (all the same values), and one has to use all different eigenvectors to construct the vectors $\mathbf{v}_i$. In addition, the last MC of 0.13 was a bit too small – the correct required value of the last MC was less than the smallest eigenvalue. I added two more zones and modified others, and boosted the MC of the last variable to 0.131. There are now enough eigenvectors to allow construction of the variables. Normally, this is not a problem, since the number of regions is usually far greater than the number of variables to create. I sincerely apologize if this has caused anybody any confusion!

Harold Reynolds, August 27, 2004.

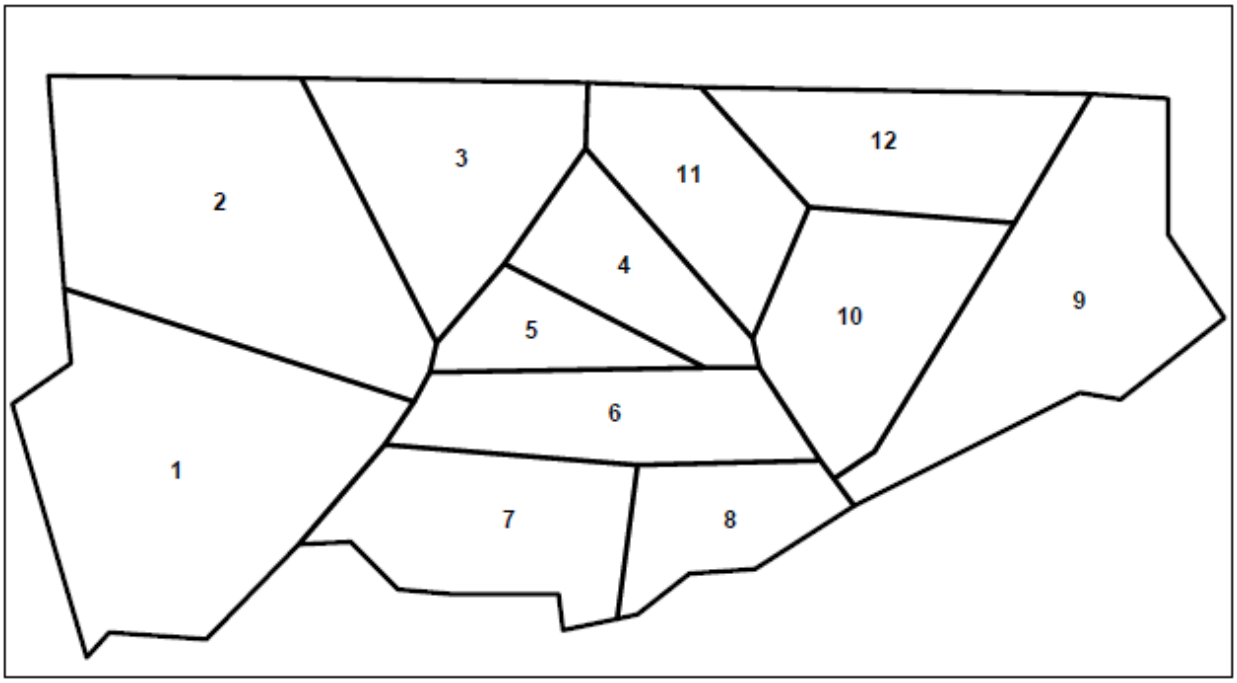
Postscript: I retyped this document (in Word 2007) because there were a couple of subscript glitches that needed correction, and the λs were showing as question marks (at least for me). January 18, 2018

3.1.4 Worked Example

The detailed description of the method below includes a worked example. The desired values of the statistics are:

Variable	Means	Variances	Moran Coeffs	Correlations				
1	20	6	0.4	1	-0.6	0.4	-0.4	-0.8
2	20	6	0.2	-0.6	1	0	0.8	0.6
3	20	6	-0.2	0.4	0	1	-0.2	0.2
4	20	6	0	-0.4	0.8	-0.2	1	0.3
5	20	6	0.131	-0.8	0.6	0.2	0.3	1

For the following set of regions:



1. Compute the eigensystem of MC_sM.

Eigenvalues

λ_1	λ_2	λ_3	λ_4	λ_5	λ_6	λ_7	λ_8	λ_9	λ_{10}	λ_{11}	λ_{12}
-0.5308	-0.4800	-0.4800	-0.4298	-0.2479	-0.1830	-0.0302	0.0000	0.0844	0.1540	0.5340	0.6093

Eigenvectors

$\mathbf{e}_1$	$\mathbf{e}_2$	$\mathbf{e}_3$	$\mathbf{e}_4$	$\mathbf{e}_5$	$\mathbf{e}_6$	$\mathbf{e}_7$	$\mathbf{e}_8$	$\mathbf{e}_9$	$\mathbf{e}_{10}$	$\mathbf{e}_{11}$	$\mathbf{e}_{12}$
-0.0230	0.2335	0.2938	0.3101	0.4481	0.1169	0.2292	0.2887	-0.3294	0.3644	0.3103	-0.2731
-0.3631	-0.2335	-0.2938	-0.1968	-0.2595	0.3840	0.1705	0.2887	0.2715	0.3609	-0.0247	-0.3960
0.3531	0.2335	0.2938	-0.3390	0.0704	0.2002	-0.5065	0.2887	0.0509	0.1101	-0.4197	-0.2092
-0.4240	0.1775	-0.2849	-0.1919	0.4434	-0.0950	0.1215	0.2887	-0.1686	-0.4699	-0.3393	-0.0518
-0.0322	-0.1775	0.2849	0.4381	-0.1165	-0.5087	0.0472	0.2887	0.3905	-0.1473	-0.2261	-0.3334
0.6161	0.1775	-0.2849	-0.1201	-0.2090	0.0048	0.3746	0.2887	-0.0141	-0.3404	0.2474	-0.2206
-0.0912	-0.4110	-0.0090	-0.3017	-0.0695	-0.3719	-0.4079	0.2887	-0.3709	-0.0314	0.4343	-0.1076
-0.2805	0.4110	0.0090	0.2870	-0.2424	0.2680	-0.3866	0.2887	0.2154	-0.2889	0.3753	0.1894
0.2440	-0.2335	-0.2938	0.0376	0.5167	-0.0403	-0.0971	0.2887	0.4742	0.1926	0.1328	0.3966
-0.0378	-0.3551	0.5698	-0.1930	-0.0628	0.3091	0.3450	0.2887	-0.0110	-0.2509	-0.0127	0.3815
0.1493	-0.2335	-0.2938	0.4944	-0.2152	0.1731	-0.1096	0.2887	-0.4732	0.0822	-0.3771	0.2192
-0.1106	0.4110	0.0090	-0.2246	-0.3037	-0.4402	0.2198	0.2887	-0.0352	0.4186	-0.1006	0.4050

2. Create the covariance matrix $\mathbf{\Sigma}_1$ by placing the variance of $\mathbf{e}_{12}$ $(0.8333)^1$ on the diagonal of a p×p matrix, where p is the number of variables. This can be done because the eigenvectors are uncorrelated, as well as orthonormal. In order to create a vector with a desired MC, a linear combination of 2 eigenvectors will be used, as discussed below. For 5 variables, 10 different eigenvectors must be used so that no correlation will exist between the variables.

[1] For variances here, I divide by n, not n-1.

3. Create the scaling matrix $\mathbf{A} = \mathbf{B}^{-1}\mathbf{D}$, where **B** and **D** are the Cholesky decompositions of $\mathbf{\Sigma}_1$ and $\mathbf{\Sigma}_2$ respectively. $\mathbf{\Sigma}_2$ is computed from the desired correlations using $S_{ij} = corr_{ij} * (var_i * var_j)^{1/2}$, where var_i and var_j are the desired variances for variables i and j.

8.4853	-5.0912	3.3941	-3.3941	-6.7882
0.0000	6.7882	2.5456	5.9397	1.2728
0.0000	0.0000	7.3485	-2.4495	4.6540
0.0000	0.0000	0.0000	4.3818	0.5477
0.0000	0.0000	0.0000	0.0000	1.5297

4. Compute the MCs that each variable $\mathbf{v}_i$ must have in order for the equivalent $\mathbf{x}_i$ to have the desired MC. This must be done because multiplying **VA** will change the MCs for all but the first variable. The procedure is as follows. Recalling that **X** and **A** are composed of p vectors of length n, write $\mathbf{X} = \mathbf{VA} \Rightarrow (\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \mathbf{x}_4) = (\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4)\mathbf{A}$. Using the upper-triangular form of **A** to simplify, we get $\begin{cases} x_1 = a_{11}\mathbf{v}_1 \\ x_2 = a_{21}\mathbf{v}_1 + a_{22}\mathbf{v}_2 \\ x_3 = a_{31}\mathbf{v}_1 + a_{32}\mathbf{v}_2 + a_{33}\mathbf{v}_3 \\ x_4 = a_{41}\mathbf{v}_1 + a_{42}\mathbf{v}_2 + a_{43}\mathbf{v}_3 + a_{44}\mathbf{v}_4 \end{cases}$, etc.

Since the $\mathbf{v}_j$ are eigenvectors[2], the MCs of the $\mathbf{x}_j$ are, using the relations previously defined,

$$M_1 = \lambda_1$$
$$M_2 = (a_{12}^2\lambda_1 + a_{22}^2\lambda_2)/(a_{12}^2 + a_{22}^2)$$
$$M_3 = (a_{13}^2\lambda_1 + a_{23}^2\lambda_2 + a_{33}^2\lambda_3)/(a_{13}^2 + a_{23}^2 + a_{33}^2)$$
$$M_4 = (a_{14}^2\lambda_1 + a_{24}^2\lambda_2 + a_{34}^2\lambda_3 + a_{44}^2\lambda_4)/(a_{14}^2 + a_{24}^2 + a_{34}^2 + a_{44}^2)$$

etc, where M_j is the Moran Coefficient for variable j, and λ_j is the MC that $\mathbf{v}_j$ must have so that $\mathbf{x}_j$ will have the MC that is desired. Solving for λ_j gives:

$$\lambda_1 = M_1$$
$$\lambda_2 = [M_2(a_{12}^2 + a_{22}^2) - a_{12}^2\lambda_1]/a_{22}^2$$
$$\lambda_3 = [M_3(a_{13}^2 + a_{23}^2 + a_{33}^2) - (a_{13}^2\lambda_1 + a_{23}^2\lambda_2)]/a_{33}^2$$
$$\lambda_4 = [M_4(a_{14}^2 + a_{24}^2 + a_{34}^2 + a_{44}^2) - (a_{14}^2\lambda_1 + a_{24}^2\lambda_2 + a_{34}^2\lambda_3)]/a_{44}^2$$
$$\lambda_j = \left[M_j \sum_{i=1}^{j} a_{ij}^2 - \sum_{i=1}^{j-1} a_{ij}^2\lambda_i \right] / a_{jj}^2$$

As can be seen, the required MC for variable j depends on the values of the MCs of the previous variables. If a value exceeds the bounds of $\lambda_1 \le MC \le \lambda_n$, it means that the desired MC is not attainable with the current configuration of correlations and MCs.

Variable	**1**	**2**	**3**	**4**	**5**
Required MC	0.4000	0.0875	-0.3625	-0.2875	-0.5144

5. Randomly select the eigenvalues λ_{li} and λ_{ui} that bracket each of the required MCs listed above. (The subscripts l and u refer to lower and upper bound respectively.) In order for the variables $\mathbf{v}_i = a\mathbf{e}_{li} + b\mathbf{e}_{ui}$ to be uncorrelated (because the covariance matrix $\mathbf{\Sigma}_1$ was specified to have no correlations), no eigenvector can be used to construct more than one variable.

[2] A linear combination of eigenvectors is also an eigenvector.

Furthermore, some eigenvectors with zero eigenvalues will have elements that are all identical. These must not be used, because the MC of **x**+c**1** is the same as the MC of **x**. The value of b can be chosen randomly; the required value of a is computed using the formula

$$a^2 = \left(\frac{\lambda_2\text{-MC}}{\text{MC-}\lambda_1}\right) b^2$$

(hence the need to have the desired MC bracketed by the eigenvalues).

	Lower Eigenvalue		Upper Eigenvalue			
RequiredMC	**Index**	**Value**	**Index**	**Value**	**a**	**b**
0.4000	10	0.1540	12	0.6093	0.4842	0.5249
0.0875	9	0.0844	11	0.5340	9.2117	0.7671
-0.3625	4	-0.4298	5	-0.2479	0.0698	0.0535
-0.2875	3	-0.4800	6	-0.1830	0.4364	0.5925
-0.5144	1	-0.5308	2	-0.4800	0.6796	0.4687

6. Create the variables $\mathbf{v}_i$ and scale the values so that their variances match that of $\mathbf{e}_2$.

Zone	$\mathbf{v}_1$	$\mathbf{v}_2$	$\mathbf{v}_3$	$\mathbf{v}_4$	$\mathbf{v}_5$
1	0.0464	-0.3025	0.5187	0.2684	0.1136
2	-0.0463	0.2686	-0.3141	0.1349	-0.4315
3	-0.0791	0.0158	-0.2263	0.3355	0.4232
4	-0.3567	-0.1962	0.1174	-0.2454	-0.2482
5	-0.3450	0.3704	0.2769	-0.2406	-0.1273
6	-0.3929	0.0065	-0.2225	-0.1651	0.6080
7	-0.1004	-0.3336	-0.2818	-0.3047	-0.3085
8	-0.0567	0.2458	0.0804	0.2211	0.0024
9	0.4221	0.4836	0.3441	-0.2068	0.0683
10	0.1103	-0.0121	-0.1914	0.5868	-0.2327
11	0.2168	-0.5028	0.2615	-0.0349	-0.0097
12	0.5815	-0.0434	-0.3630	-0.3491	0.1423

7. Compute **X = VA** and shift the values of the $\mathbf{x}_j$ so that their means equal the desired means by adding the difference between the desired mean and the current mean to each $\mathbf{x}_j$.

Zone	$\mathbf{x}_1$	$\mathbf{x}_2$	$\mathbf{x}_3$	$\mathbf{x}_4$	$\mathbf{x}_5$
1	20.3934	17.7105	23.1992	17.9511	22.0353
2	19.6068	22.0590	18.2186	23.1128	18.6087
3	19.3289	20.5102	18.1090	22.3868	20.3350
4	16.9735	20.4841	19.1525	18.6822	22.2037
5	17.0730	24.2704	21.8068	21.6383	23.7752
6	16.6659	22.0446	17.0481	21.1939	22.4799
7	19.1477	18.2467	16.7390	17.7144	18.3070
8	19.5189	21.9571	21.0240	22.4241	21.1966
9	23.5815	21.1336	25.1924	19.6907	19.3432
10	20.9361	19.3565	18.9371	22.5940	18.3104
11	21.8399	15.4829	21.3777	15.4839	19.0713
12	24.9346	16.7444	19.1955	17.1277	14.3339