

The SoVI Recipe

1. Collect variables: Depending on the level of geography, variables will come from the US Census Data Engine, City and County Databook, GNIS, Federal Elections project, or USA Counties.
2. Standardize where necessary all variables to either percentages, per capita, or density functions (per square mile).
3. Verify accuracy by using descriptive statistics such as the minimum and maximum values. Check for missing values for the unit of analysis. If some cells, but not all have a missing value, substitute the mean value for the variable in its place. The statistical procedure will not run properly with missing values. Census units with no population values should be omitted.
4. Normalize the input variables through the creation of z-scores (mean= 0 and standard deviation= 1).
5. Perform the principal components analysis (PCA) using a varimax rotation and Kaiser criterion for component selection (e.g. eigenvalues greater than 1.0). The varimax rotation tends to load each variable highly on only one component.
6. Examine the resulting components on 1) what they broadly represent and 2) how they may influence social vulnerability. This is done by examining the correlations between the variables and the components given in the loadings matrix output of the PCA to determine what characteristics are being represented, and if they have a tendency to increase or decrease social vulnerability.
7. The goal is to have all the components aligned so that higher positive values increase social vulnerability. Determine the direction (or cardinality) of the factor. If a factor tends to show strong positive values (greater than + .50) for variables that indicate low social vulnerability, the inverse of the corresponding factor score is used (in other words, the factor score is multiplied by -1). In some cases, notably the age and urban/rural components, there are strong positive and negative values on individual variables (e.g. loads positively on elderly, negatively on children or positive on urban and negative on rural). In those two instances the absolute value of the component is used. Finally, there are times when the phenomenon has a mix of loadings, but overall it shows strong positive values that indicate increasing social vulnerability then no change in directionality is required.
8. Save the component scores as a separate input variable file.
9. Place all the components with their directional (+, -, ||) adjustments into an additive model and sum to generate the overall SoVI score for the place.
10. SoVI scores are then mapped using a classification (tertile or quintile) based on standard deviations from the mean, with scores $\geq +1$ standard deviation indicating higher levels of social vulnerability, and scores ≤ -1 standard deviation showing lower levels of social vulnerability.
11. The following is an example of the 2000 County SoVI illustrating the variable loadings, naming of the factor, and the sign adjustment (directionality). The SoVI score is computed for each unit of analysis (e.g. county, census tract).

<i>Sign Adjustment</i>	<i>Factor</i>	<i>Name</i>	<i>Dominant Variables</i>	<i>Correlation</i>
-	1	Socioeconomic Status	QPOVTY	-0.86
			QED12LES	-0.85
			QCVLBR	0.81
			PERCAP	0.82
II	2	Age	MEDAGE	-0.89
			QKIDS	0.88
			QSSBEN	-0.83
			PPUNIT	0.83
			BRATE	0.80
			QPOP650	-0.80
+	3	Development	COMDEVDN	0.99
			MAESDEN	0.98
			EARNDEN	0.97
			HODENT	0.95
			HUPTDEN	0.89
			RPROPDEN	0.74
+	4	Rural Agriculture	QLANDFRM	0.73
			NNRESPC	0.65
			QPCHCG	-0.65
			QRFMR	0.63
			HOSPTEC	0.57
			MDHSEVAL	-0.50
+	5	Race and Gender	QFEMLBR	0.84
			QFHH	0.59
			QAGRI	-0.58
			QBLACK	0.56
			QVOTE	-0.55
+	6	Race	QASIAN	0.76
+	7	New Immigrants	QINDIAN	-0.76
			MIGRA	0.53
+	8	Ethnicity	QSPANISH	0.77
+	9	Gender	QFEMALE	0.81
			QSERV	-0.58
+	10	Infrastructure Employment	QTRAN	0.85
+	11	Economic Dependence	DEBTREV	0.98

$$\text{SoVI} = +(- \text{Factor 1}) + (\text{II Factor 2}) + \text{Factor 3} + \text{Factor 4} + \text{Factor 5} + \text{Factor 6} + \text{Factor 7} + \text{Factor 8} + \text{Factor 9} + \text{Factor 10} + \text{Factor 11}$$