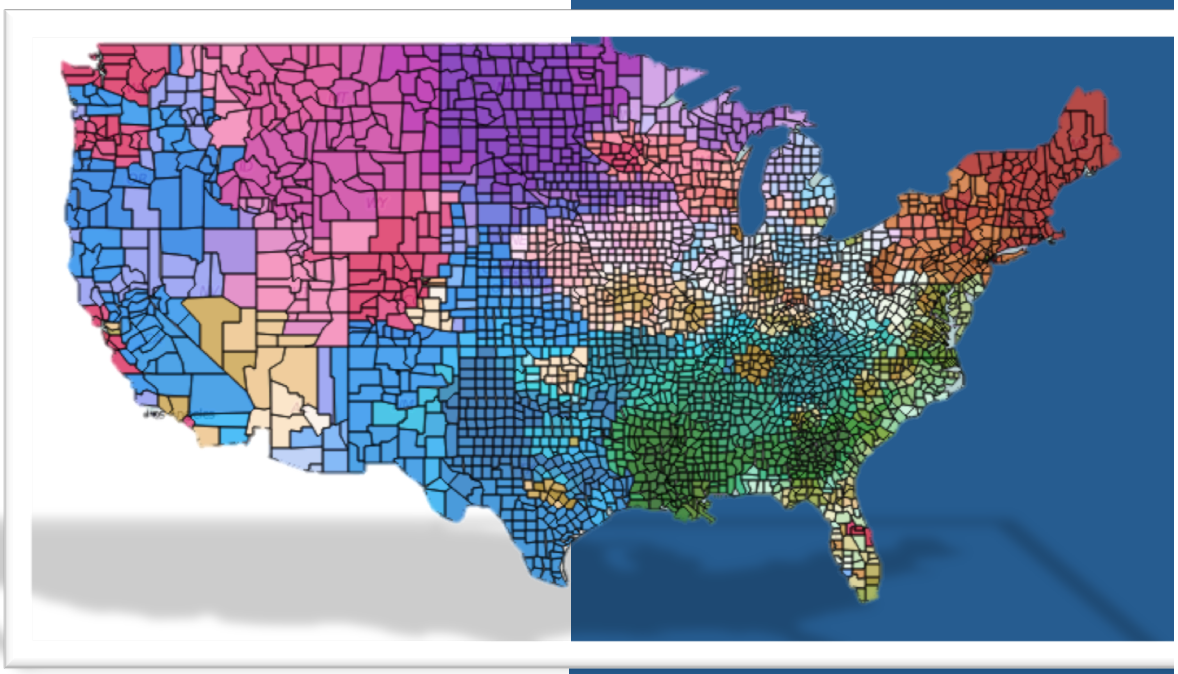
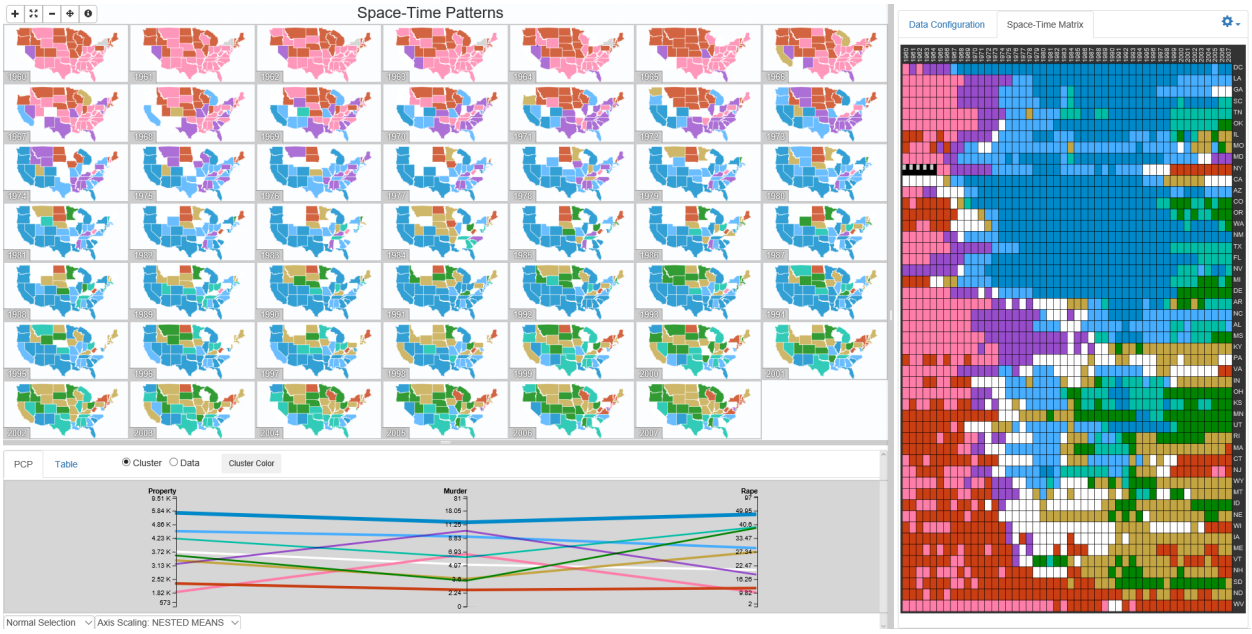
VIS-STAMP



Introduction

VIS-STAMP is a software package to explore multivariate spatio-temporal data, discover interesting and unknown complex patterns, and present them in an easy-to-understand form to support human interpretation, analytical reasoning, and/or decision-making. The software tool integrates computational, visual, and cartographic methods together to detect and visualize **multivariate spatio-temporal** patterns. It is able to:

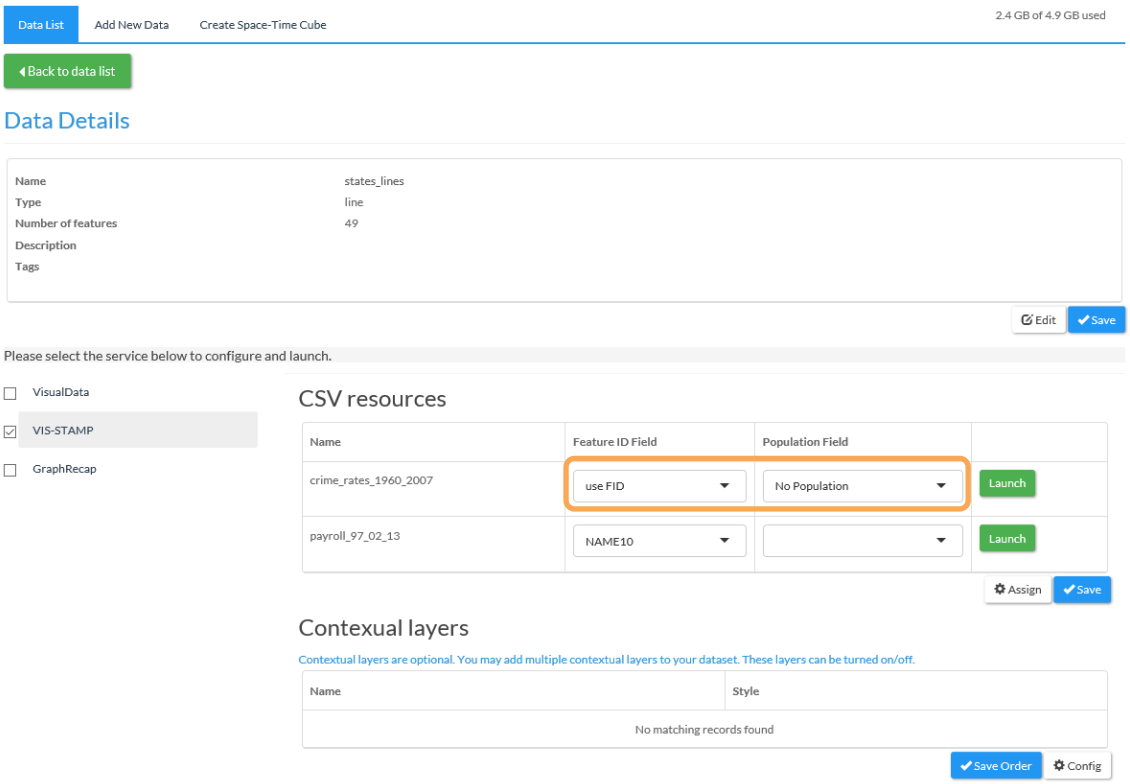
- Perform multivariate clustering and abstraction (including time-series clustering) with a Self Organizing Map (SOM);
- Encode SOM result with colors derived from a two-dimensional diverging-diverging color scheme;
- Visualize the multivariate patterns with an enhanced Parallel Coordinate Plot (PCP) display, which serves as a multivariate "legend" in the integrated system;
- Visualize the spatio-temporal variations of multivariate patterns, or the space-variable variations of temporal patterns in a hierarchical, computationally sortable matrix and a temporally or geographically ordered map matrix; and
- Support human interactions to explore patterns from different perspectives and at different detail levels.



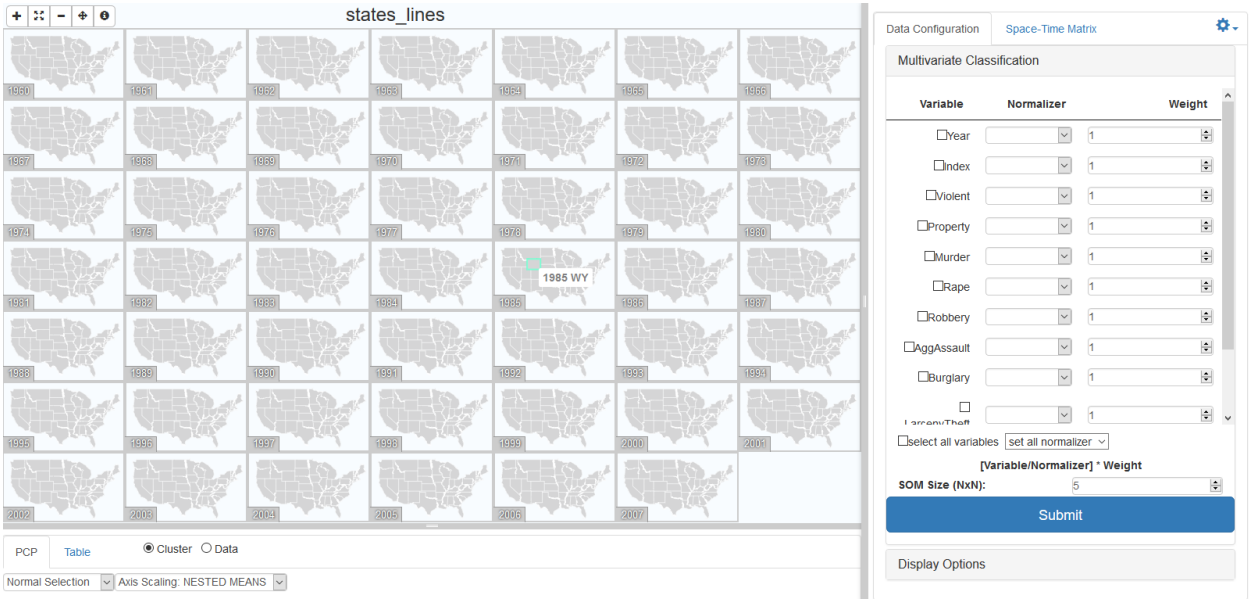
Configure Data

Before conducting analysis using VIS-STAMP, users need to configure data properly. The following figure shows data details and CSV resources. FID should be selected for Feature ID Field from the drop-down

menu, which is used to join CSV and shapefile. No Population should be chosen for Population Field. The next figure shows the result after data is launched.



The following figure includes three parts. The first part contains maps to view the dataset in different space-time patterns. The right part is to set variables for the multivariate classification. And the bottom part is used to show the results in PCP charts and tables. Anyway conducting multivariate classification analysis is the first step.



Multivariate Classification

1. *Choose Variables*

- Click the checkbox to select variables
- Multiple variables can be selected.
- All of the variables can be normalized.*
- User can assign weights for selected variables. **

***Normalizers are used to reduce variables originally measured on different scales to a standard scale using a common factor to make them easily comparable. Select a variable that contains a scale you wish to normalize with such as total population, total votes, etc. Any variable can be used as long as there are no negative or zero values.**

****The weight is the importance factor in comparison to other factors. The default weight is 1. Any number between 1 and 100 can be used.**

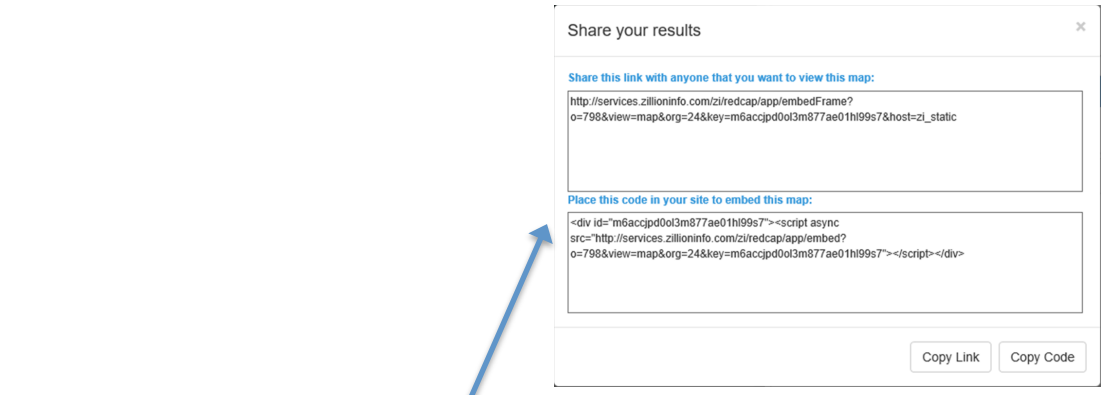
2. *Set Number of Classes*

Enter the number of SOM size. The SOM size determines how many clusters will be calculated. For example, because the SOM dimensions are N x N, if the number 3 is chosen 9 total clusters will be calculated.

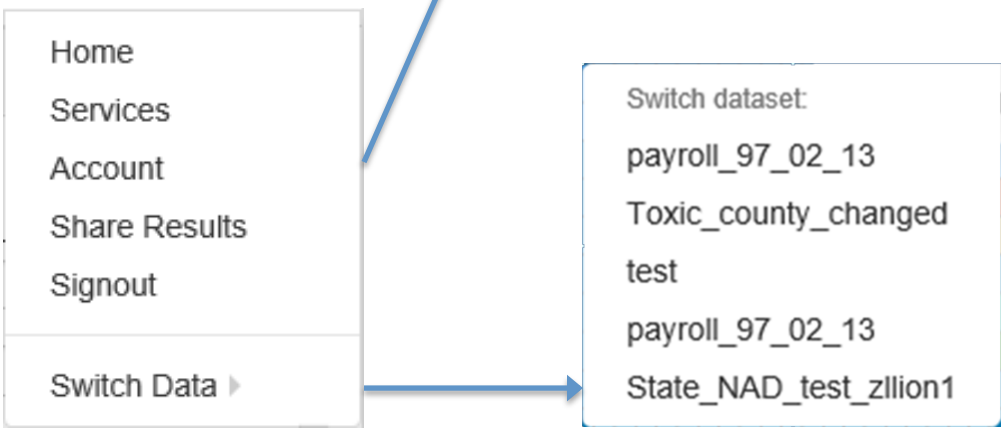
3. *Submit*

Click "**Submit**" and wait for the classification to be calculated and be shown in the map and in the PCP.

4. *Setting and share*



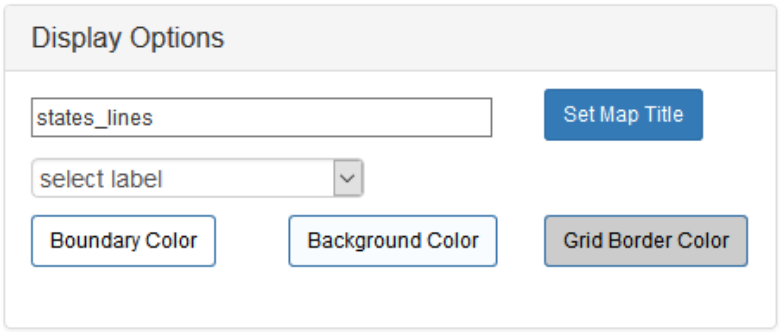
Share the result by link or code.



Switch to a different dataset for the current service

Share Results allows user selecting two different methods of sharing this map. The first one is to share this map link, which ensures that users can view this map in the same way in a new web page. The second one uses codes to represent this map that allows users embedding this map to any website.

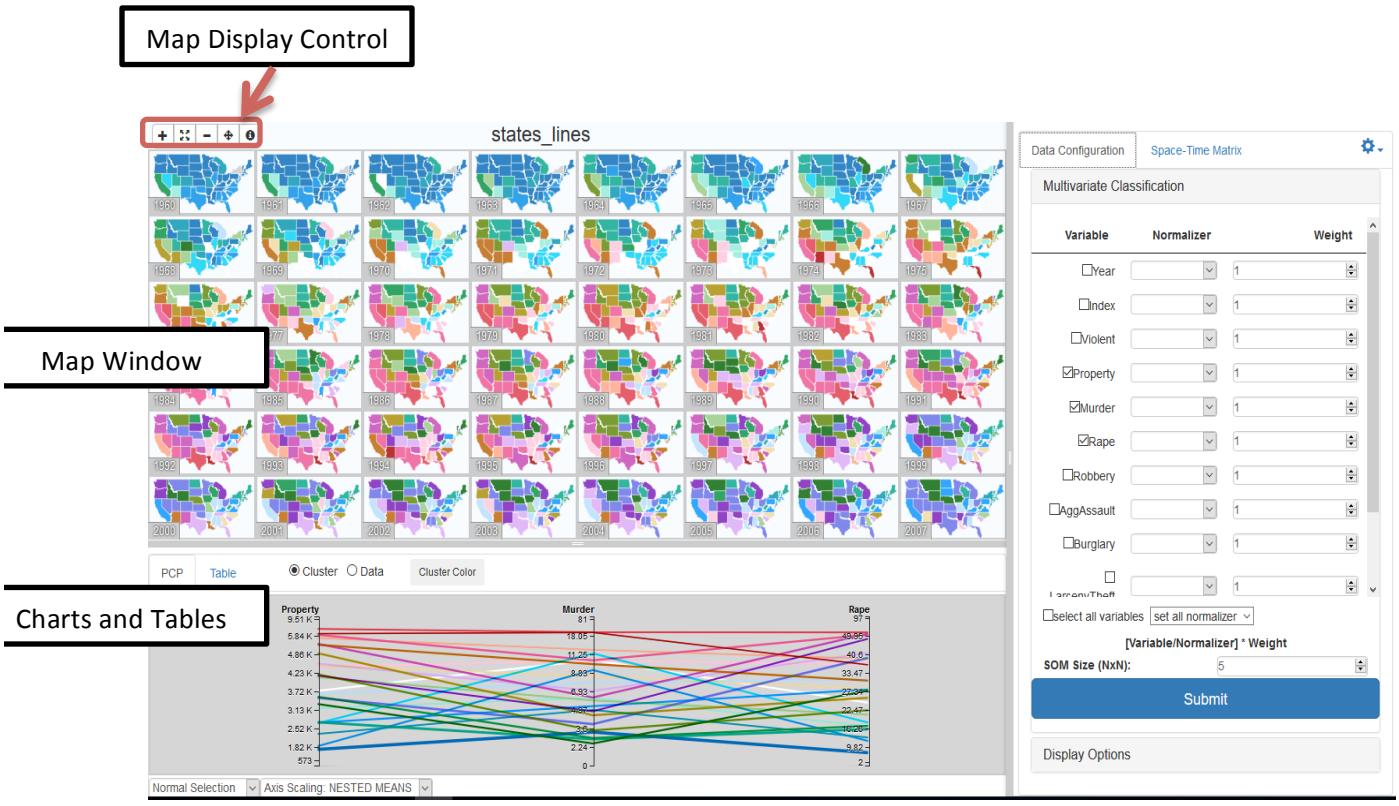
5. *Display Option*



Users can set map title, change boundary color, background color, grid border color, and select different labels in the panel of Display Option.

Map

Results of the classification can be viewed in the map window. The map display can be used as a control tool to change the map display (see details below).



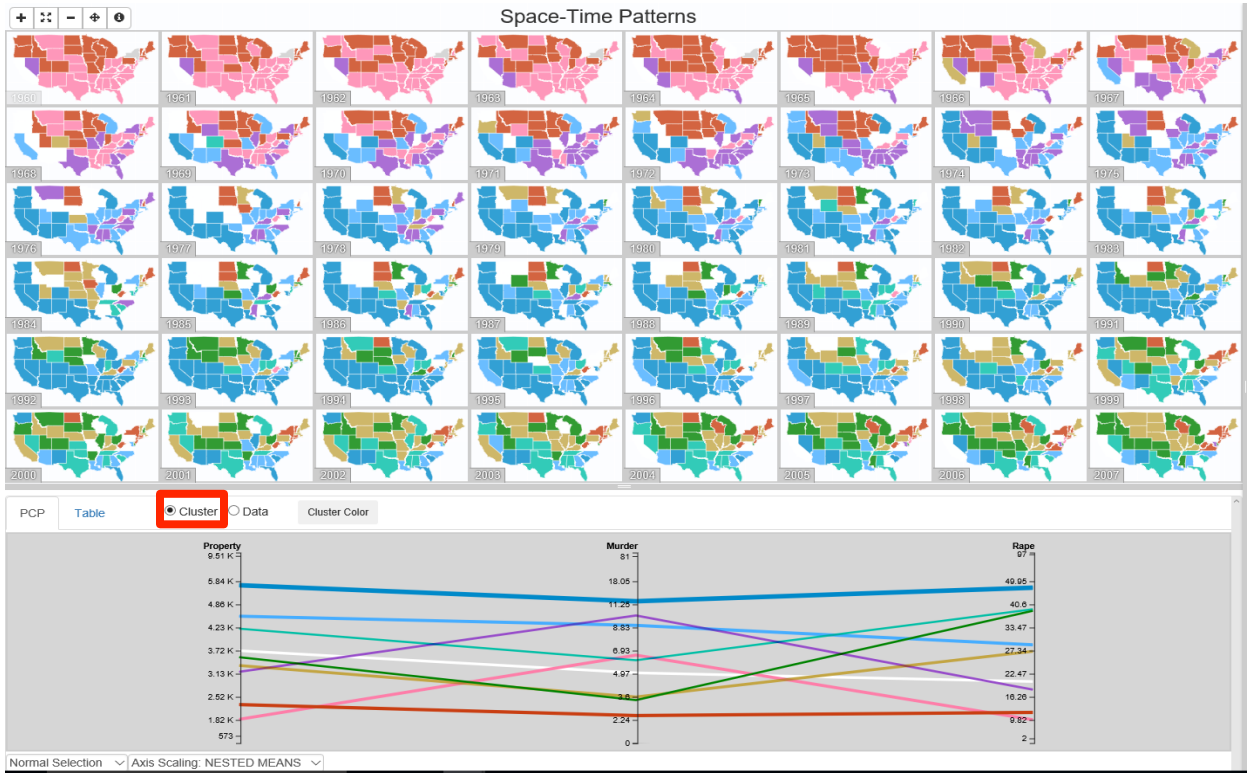
Function	Description
Zoom In	Zoom in on areas of the map – you can also use the mouse wheel to zoom in by scrolling up.
Full Extent	Get a full view of the dataset.
Zoom out	Zoom out to areas of the map – you can also use the mouse wheel to zoom out by scrolling down.
Drag Maps	Drag the map to other places.
Identify Features	Identify all attributes for each polygons

Charts and tables

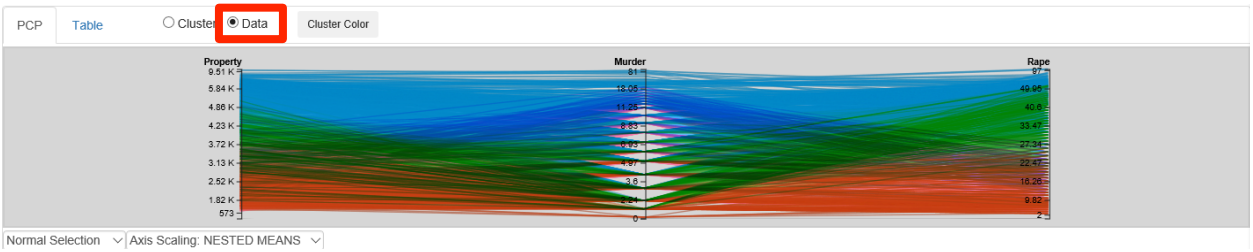
There are three charts and tables to visualize data and results: **PCP chart**, and **Data Table**.

PCP Chart:

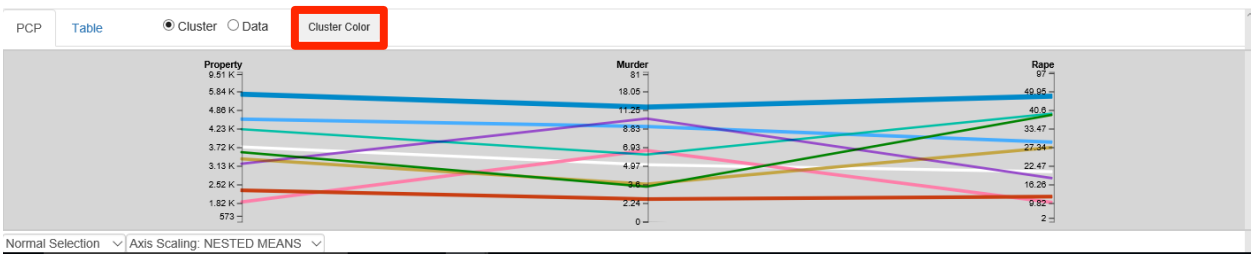
Click “PCP Chart” to check the Parallel Coordinate Plot (PCP). Please check Appendix I for more details.



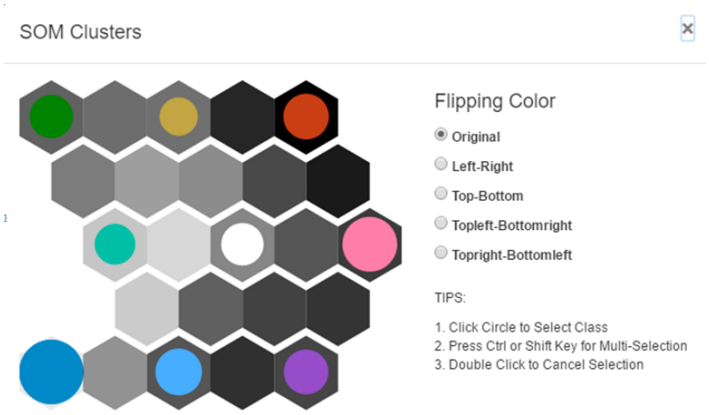
In the PCP Chart, each line represents one cluster in the map.



If Data is selected, each line will represent a data unit that contains the numbers respectively for property crime, murder, and rape in this figure. In contrast in the previous figure, if you select Cluster, each line will represent one cluster, which is the average numbers respectively for property crime, murder, and rape.



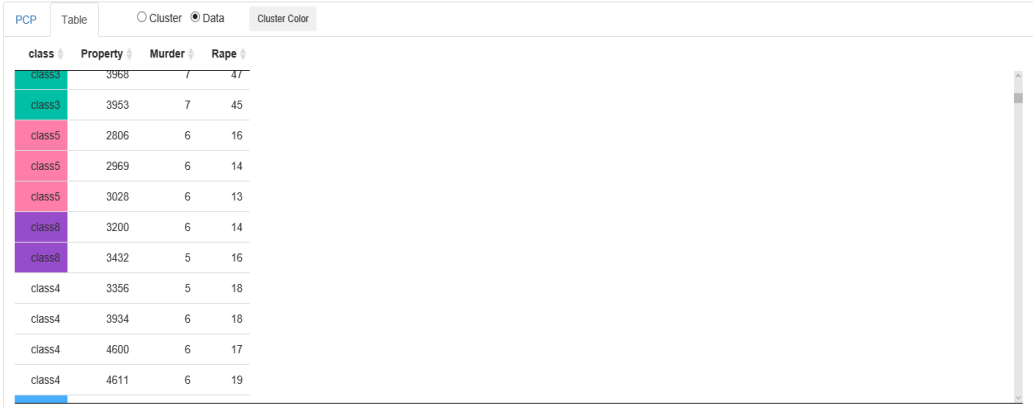
If the button Cluster Color is clicked, the following panel will show up, which totally includes nine colors.



Nine colors in this figure respectively represent nine clusters in the map. If user clicks any one color, the corresponding line will be highlighted. If more colors need to be selected, users can press Ctrl or Shift Key for Multi-Selection.

Data Table:

Click “Data table” to check details of classification by table. The mean value of each class will be displayed in the data table.

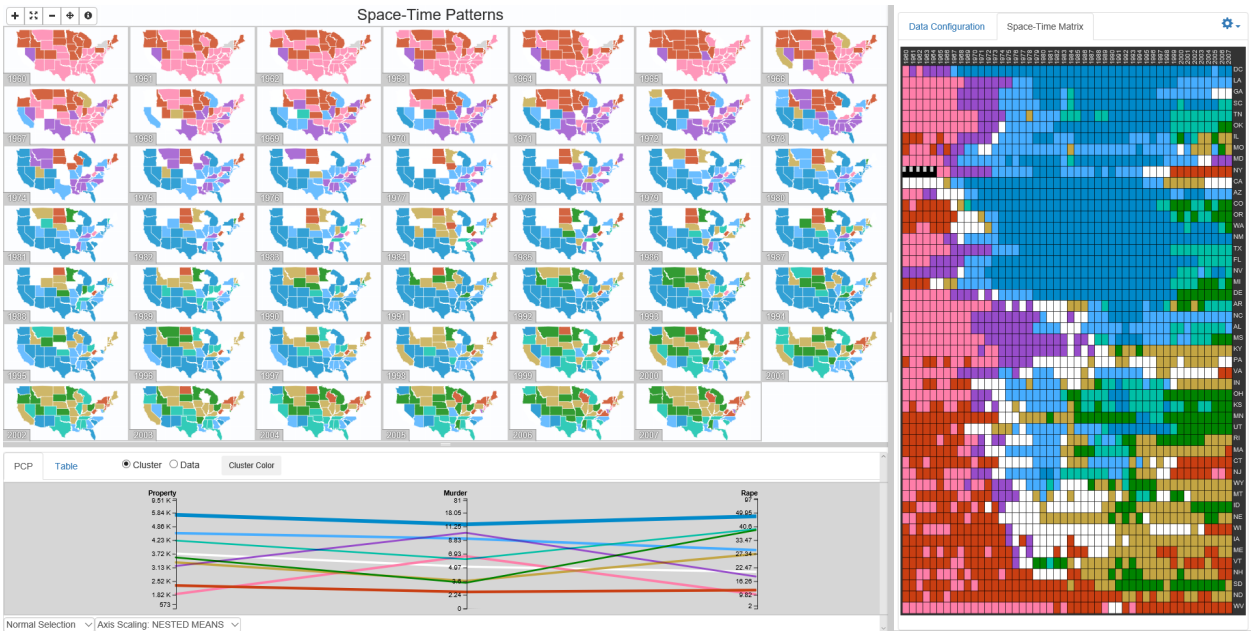


class	Property	Murder	Rape
class3	3953	7	45
class5	2806	6	16
class5	2969	6	14
class5	3028	6	13
class8	3200	6	14
class8	3432	5	16
class4	3356	5	18
class4	3934	6	18
class4	4600	6	17
class4	4611	6	19

One color type corresponds to a unique cluster. A cluster contains similar number in property crime, murder, and rape, but there is difference between random two clusters.

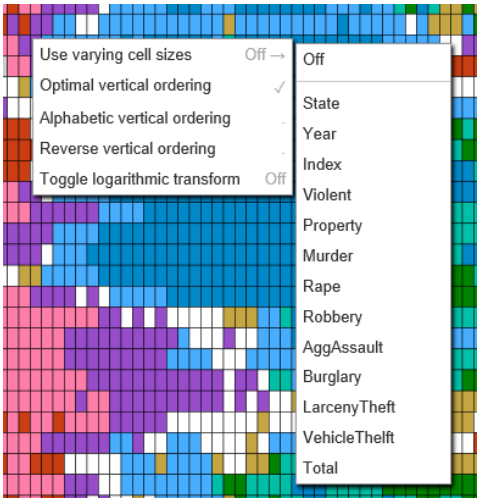
Spatio-Temporal Visualization of Multivariate Patterns

A form of reorderable matrix (top-right in the figure below) is developed to organize multivariate patterns (represented with colors) across space and time. This reorderable matrix, when one of the two dimensions represents geography, will be accompanied with orderable map matrix (top-left in the figure below) and PCP chart (bottom-left in the figure below). In the application shown in the figure below, columns represent time (years) and rows represent places (states of the US). Ordering of time is fixed (for these applications) in normal temporal order. Ordering of places is computationally derived with several cluster-based ordering methods. After the reordering, states that have similar crime patterns over time are next to each other in the matrix and form homogeneous spatio-temporal "regions". The reorderable map matrix essentially converts each column in the reorderable matrix to a map and these maps are arranged in the same order as that of the columns.



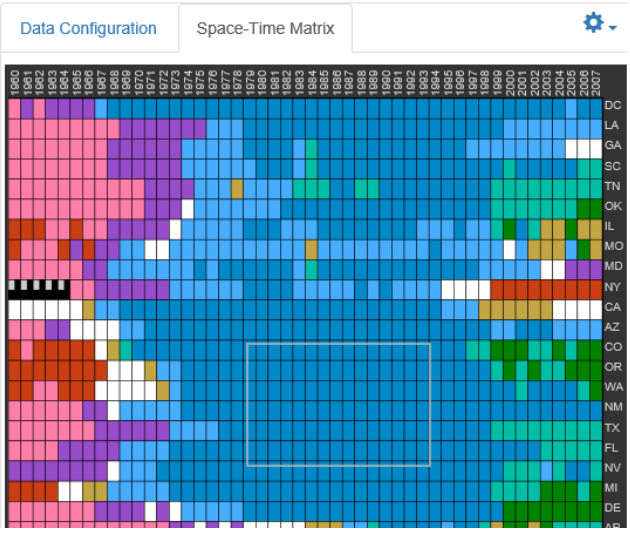
Interactive User Exploration

In addition to constructing a holistic view of patterns in the spatio-temporal and multivariate data cube, VIS-STAMP also supports a variety of user interactions that allow the analyst to examine patterns in detail. Each view is able to support user selections and the selection made in one view will be highlighted in other views. The user can make a selection in one component and then refine that selection in the same or another visual component by adding or subtracting new selection(s). Each component should be able to respond to selections made at different levels (i.e., data elements or clusters). The user can right click the reorderable matrix, and use varying-size cells based on the values of a specific variable.



Making Selections

- Click and drag rectangle with mouse over desired cells.
- You can select across rows, columns, or both.
- Selections are viewable in all windows including.



Refining Selections

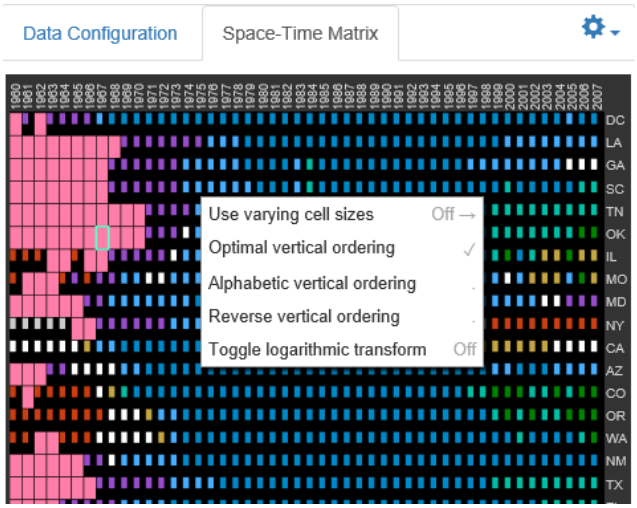
Users can refine selections made in the S-T Matrix by using the PCP:

- Click and drag rectangle with mouse over desired cells in S-T Matrix
- Using the PCP,
    - To add to selection made in S-T Matrix:
        - Use the first drop down menu to select “Union Selection”
        - Click and drag vertically to select **additional** clusters
    - Use the first drop down menu to select “Intersect Selection”

- Click and drag vertically within the previously, selected clusters to select within the selection that has already been made

S-T Matrix Options:

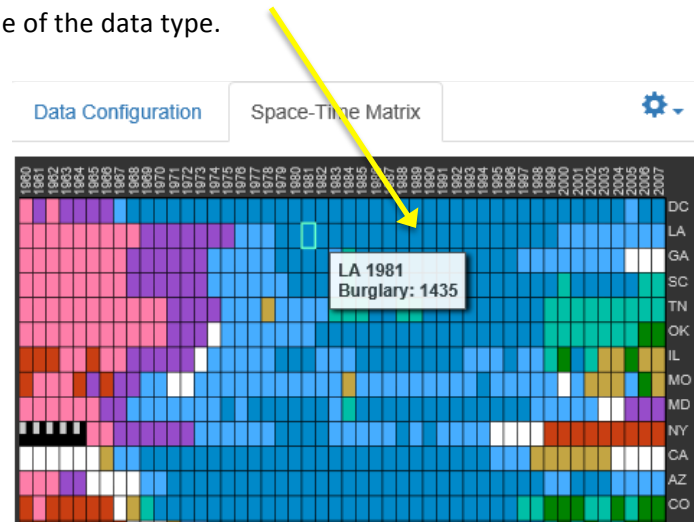
To access the options menu, right click in the S-T Matrix Window.



Use varying cell sizes

Selecting this option **creates cell sizes within the S-T matrix that represent the numerical value** of the type of data. For example, selecting Burglary created different sized cells based off the number of Burglary crimes in that place during that time.

To view the information for each cell, **hover your mouse** over the desired cell to determine time, place, and the numerical value of the data type.



Using this option also allows for the S-T Matrix to represent a certain type of data.

For example, selecting Burglary causes all cells to represent the values of Burglary. If one desires to view the matrix with Burglary values but with homogenous cells, one can unselect using the right-click menu and the values will remain Burglary.

Optimal vertical ordering:

States that have similar crime patterns over time are next to each other in the matrix and form homogeneous spatio-temporal "regions".

Alphabetic vertical ordering:

Selecting this feature will order rows in the alphabetical order.

Reverse vertical ordering:

Selecting this feature reverses the order of rows from top to bottom.

Toggle log transformation:

The log transformation can be used to make highly skewed distributions less skewed by lowering the number scale. This can be valuable both for making patterns in the data more interpretable and for helping to meet the assumptions of inferential statistics.

Selecting this option can further reveal pattern if your data has a wide distribution or very large numbers.

Appendix I: Parallel Coordinate Plot (PCP)

The PCP supports two detail levels: Clusters vs. Data Items

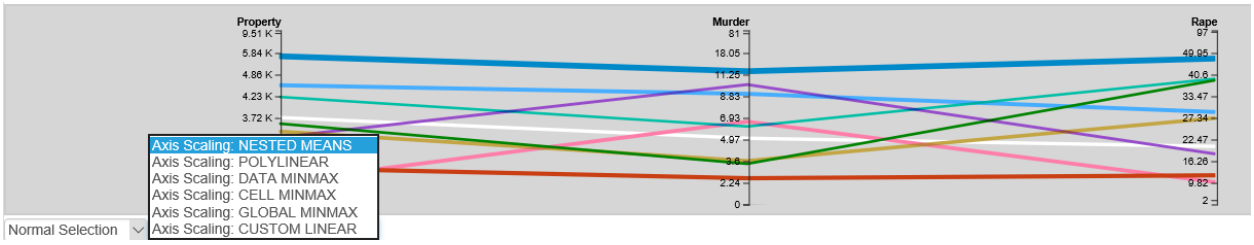
View Clusters: Each string represents a cluster with its mean vector. The thickness of each string represents the cluster size (i.e., the number of data items in the cluster).

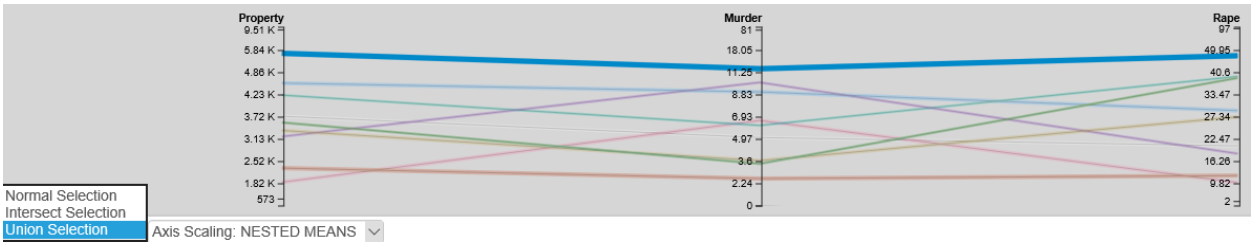
View Data Items: Each string represents a data item with its multivariate vector.

For either choice, each cluster (or data item) has the same color as it does in the SOM view.

The PCP supports five axis-scaling methods:

- **Nested-Means:** scaling on each axis using nested means and thus adjust the spacing of intervals according to data distribution. This method can alleviate the overlapping problem in PCP for skewed data distribution. Specifically, nested-means is a non-linear scaling method that recursively calculates a number of mean values (and sub-means) and uses these values as break points to divide each axis into equal-length segments. Therefore, nested-means scaling always puts the mean value at the center of each axis and thus makes axes defined by different units and data ranges comparable.
- **Polylinear:** each axis is marked with the min and max value of different variables, but not scaled.
- **Data Min-Max**: each axis is linearly scaled using its min and max values.
- **Cluster Min-Max** (or Cell Min-Max): each axis is linearly scaled using cluster centroid min and max values.
- **Global Min-Max:** this option is only useful when all the variables are comparable to each other, for example percentage values. Axes will be scaled linearly using the global min and max values (for all variables).
- **Customized Linear:** this option is only useful when all the variables are comparable to each other, for example percentage values. The user will define the min and max (same for all variables) to linearly scale each axis. In future versions, the user may be able to define the min/max differently for each axis.





The PCP can *optimally* order dimensions:

Optimal Ordering of Axes: dimensions are ordered using an optimal hierarchical ordering method based on the mutual correlations among dimensions

Original Ordering of Axes: dimensions are in their original order as in the data file.

The PCP supports different types of selection at different levels:

The PCP at the cluster level presents a global view of the overall patterns. A user can select one or more clusters in the PCP (or in the SOM), then switch to the data item level (instead of the cluster level), and examine all the data items in the cluster(s).

Selection can also be made at the data item level. For example, one can show data at the item level in the PCP and then select a single data item to read its exact variable values. One can also switch back to the cluster level and see which cluster the selected item belongs to. If that cluster also contains other items, its circle will become a wedge to show the partial selection.

By selection "Intersect Sel." or "Union Sel.", the user may also combine two different selections or select within a selection.

Interacting with the PCP

Function	How to:
Select Clusters	1. Use your mouse to **click and drag vertically** to select clusters of interest 2. When selecting with mouse, **do not drag horizontally across vertical** axes 3. When features are selected they remain colored and all other clusters turn to gray(default)
Select within a selection	1. Use the dropdown menu in the top left-hand corner and click "Normal Selection" 2. **Click and drag vertically** to select clusters of interest 3. After selected, use the dropdown menu **and click "Intersect Sel."** 4. You can now click and drag vertically to select clusters within the selection already made.
Combine two different selections	1. Use the dropdown menu to click **"Normal Selection"** 2. Click and drag vertically to select clusters of interest 3. After selected, use the dropdown menu and click "Union Sel."

	4. Click and drag vertically over other clusters you did not previously select to add to the selection.
Combine two different selections (Different Way)	5. Use the dropdown menu to click **"Normal Selection"** 6. Click and drag vertically to select clusters of interest 7. After selected, **hold down shift key** AND 8. **Click and drag vertically** over other clusters you did not previously select to add to the selection.
Clear Selection	1. Click and drag vertically over gray space within PCP 2. All clusters will become unselected