Section 4

Video: Space-time pattern mining

Print

Time	Caption
0:00	♪ [music] ♪
0:11	The Create Space Time Cube tools aggregate data
0:15	both spatially and temporally,
0:18	allowing you to perform analyses
0:20	that uncover patterns and trends over space and time.
0:24	When aggregating a set of points spatially,
0:27	we consider their location based on their x and y coordinates
0:31	and use a two-dimensional grid.
0:34	However, when working with spatiotemporal data,
0:37	we have a third dimension to consider -- time.
0:41	Conceptually, it can be helpful to use height as a way
0:44	to visualize the temporal dimension where points
0:47	at the top have the most recent timestamp.
0:51	Now, to aggregate this data in a way that
0:54	also incorporates the temporal dimension,
0:56	we can create a space time cube.
0:59	You can think of a space time cube as a three-dimensional,
1:03	grid-like structure where points are aggregated based on
1:06	their location on the ground and their place in time.
1:10	In this example, we can see that points falling within the same
1:13	10-kilometer by 10-kilometer spatial extent
1:16	and that took place within the same month
1:19	are aggregated together into a bin.
1:22	Bins are the individual aggregation units that make up
1:26	the space time cube, each with a unique spatiotemporal extent.
1:31	The space time cube counts up how many points fall into each bin
1:35	and can also summarize numeric attributes if the points have any.
1:40	Points can be aggregated into fishnet or hexagon grids
1:44	or into any polygon layer.
1:46	Alternatively, if your data does not require aggregation,
1:50	you can create a space time cube from defined locations,
1:54	where each point or polygon will become its own bin.

1:57	Once your data has been aggregated into a space time cube,
2:01	you're ready for analysis.
2:03	The Emerging Hot Spot Analysis tool finds spatiotemporal clusters
2:08	by extending the concept of what it means to be a neighbor
2:11	to include not only what is near in space,
2:14	but also what is recent in time.
2:17	Within the space time cube, each bin is evaluated in the context
2:21	of its neighboring bins, which includes the bin's spatial neighbors,
2:25	the ones that are closest geographically,
2:29	and its temporal neighbors, the ones that are closest in time as well.
2:33	So only bins that are both proximate in space
2:37	and recent in time are considered related.
2:40	With this three-dimensional conceptualization of proximity,
2:44	each bin's neighborhood is defined
2:46	and then compared to the study area.
2:48	If the neighborhood value is significantly higher than
2:52	the study area, then that bin is marked as a hot spot.
2:56	Just like in a two-dimensional hot spot analysis, each bin
3:00	is assigned a probability measuring how likely it is
3:03	to belong to a nonrandom cluster of high values, a hot spot,
3:07	or a nonrandom cluster of low values, a cold spot.
3:11	The type and intensity of clustering is then summarized
3:15	for each location and categorized based on
3:18	the location's pattern or trend in clustering over time.
3:22	There are 16 possible types of significance,
3:25	8 for hot and 8 for cold,
3:27	each describing a unique temporal pattern.
3:30	For example, this location has been marked as a Sporadic Hot Spot,
3:35	meaning that most recently it was hot, but over time,
3:38	it has switched back and forth between hot and not significant.
3:42	This location is marked as an Intensifying Hot Spot,
3:45	meaning that it has been hot at least 90% of the time,
3:49	and that there is a significant upward trend detected in
3:53	the clustering intensity, meaning that the clustering is becoming stronger.
3:57	The hot spot has been getting hotter over time.
4:00	And this location is marked as a New Hot Spot, meaning that
4:04	it had never been hot before until the most recent time period

4:08	when it became hot for the very first time.
4:11	Emerging hot spot analysis is just one way to identify
4:14	spatiotemporal clusters in our data.
4:17	Local outlier analysis uses the same spatiotemporal conceptualization
4:22	of what it means to be a neighbor to identify statistically significant
4:26	clusters and outliers in the context of both space and time.
4:31	Just like in an emerging hot spot analysis, a bin's neighborhood
4:35	is defined in terms of spatial and temporal proximity.
4:39	But in a local outlier analysis,
4:41	the bin is not included in its own neighborhood.
4:45	This allows for a different type of comparison, where both the bin value
4:49	and the neighborhood value are compared to the study area
4:52	to identify value clusters and detect local outliers.
4:57	The result includes four possible types of significance.
5:01	A bin with a high value surrounded by a neighborhood with a high value
5:04	is marked as a high-high cluster,
5:06	while a bin with a low value surrounded by a neighborhood
5:10	with a low value is marked as a low-low cluster.
5:13	Bins are considered outliers when their value
5:16	is very different from their neighbors.
5:18	So a bin with a high value surrounded by a neighborhood
5:21	with a low value is marked as a high-low outlier.
5:25	And a bin with a low value surrounded by a neighborhood
5:28	with a high value is marked as a low-high outlier.
5:31	The two-dimensional summary output of local outlier analysis
5:36	tells us if a location has ever been significant,
5:39	and if so, of which type.
5:42	For example, this location has been marked as only high-high cluster,
5:46	because over time it has been significant
5:49	and only as a high-high cluster.
5:51	While this location has been marked as multiple types,
5:54	because over time, it has been significant
5:57	as both a high-low outlier and a low-low cluster.
6:01	For both emerging hot spot analysis
6:03	and local outlier analysis combining
6:06	the 2D summary output with the 3D visualization
6:10	of the space time cube gives us valuable insights into where

6:14	and when significant patterns and trends are taking place.
6:18	Ultimately, by incorporating time into our cluster analysis,
6:22	we're able to understand our temporal data
6:25	in powerful new ways.
6:28	♪ [music] ♪