# 2.1 Questions about finding and extracting objects from data

Thursday, September 09, 2021 3:02 PM

Geospatial objects include concrete or abstract (conceptual or latent) objects with spatial properties.

1. Concrete geospatial objects: examples and characters?
2. Abstract geospatial objects: examples and characters?

Concrete objects are generally visible and less ontological uncertainty.
Abstract objects are conceptual constructs made up for functional or reasoning purposes.
Epistemologically, abstract objects represent higher level concepts than concrete objects.
Nevertheless, concrete objects can be difficult to identify as well.
From this perspective, many of our problems are **classification or labeling problems**.

a bowl, a cup, or a mug?

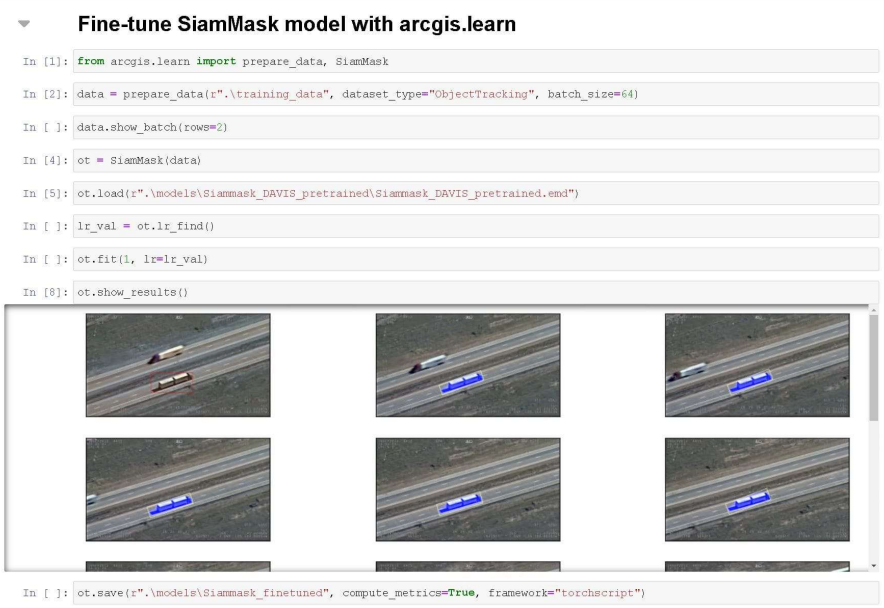
Q#1: Which of the following is (are) common abstract geospatial objects? Check all correct answers.
A) crime incidents
B) crime hotspots
C) GPS locations
D) movement trajectories
E) wealthy neighborhoods
F) activity space

Q#2: Which of the following problems can be considered classification problems or labeling problems? Check all correct answers.
A) remember names of state capitols
B) identify the size of a building
C) delineate the boundary of a neighborhood
D) differentiate bare grounds from vegetation covered areas
E) rank areas of different levels of flood risk
F) estimate the total population in the city of Dallas
G) determine the shore of Lake Texoma
H) extract trees from a satellite image

Procedures are available at https://www.esri.com/arcgis-blog/products/api-python/analytics/how-we-did-it-integrating-arcgis-and-machine-learning-at-uc-2018/

## Object Tracking in Full Motion Video

https://www.esri.com/arcgis-blog/products/arcgis/analytics/object-tracking-in-full-motion-video/



**Fine-tune SiamMask model with arcgis.learn**

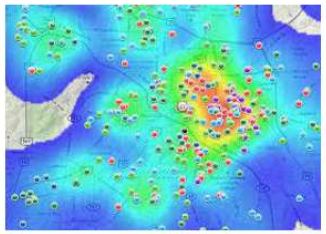
```
In [1]: from arcgis.learn import prepare_data, SiamMask
In [2]: data = prepare_data(r".\training_data", dataset_type="ObjectTracking", batch_size=64)
In [ ]: data.show_batch(rows=2)
In [4]: ot = SiamMask(data)
In [5]: ot.load(r".\models\Siammask_DAVIS_pretrained\Siammask_DAVIS_pretrained.emd")
In [ ]: lr_val = ot.lr_find()
In [ ]: ot.fit(1, lr=lr_val)
In [8]: ot.show_results()
```

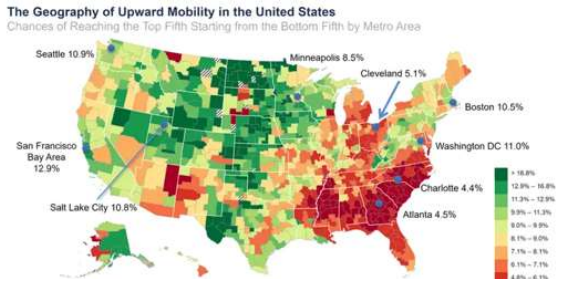
```
In [ ]: ot.save(r".\models\Siammask_finetuned", compute_metrics=True, framework="torchscript")
```

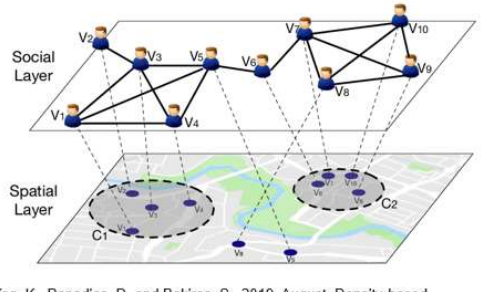
In data science and spatial data science projects, we often start with data from concrete objects to conclude properties or relationships among abstract objects.

For example: using items (concrete objects) of what I bought in a grocery store to guess that I am a vegetarian (an abstract object).
Extend the line of thinking, a data science project may be to estimate how many customers are vegetarians in a grocery store.
Below are additional examples of conceptual objects from concrete objects:

Tropical Storm Nicholas
in the Gulf of Mexico. NOAA



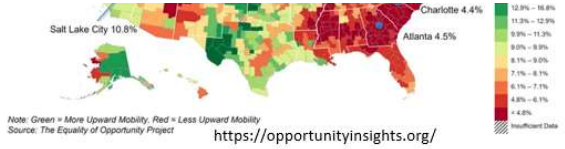


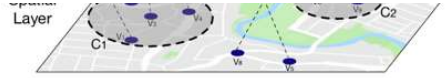

Q#3 Name a concrete object and an abstract object in each example included in the left figure:

Bison: individual bison / herd (herd movement trajectory)
Tropical storm Nicholas: climatic measurements / hurricane path
Crime incidents: individual crime incident / crime hotspot
Upward mobility: individual county mobility/ regional mobility
Geo-social network: individual node / social network

Q#4: In our project on 311 service calls in Dallas, explain how we use




https://opportunityinsights.org/




Yao, K., Papadias, D. and Bakiras, S., 2019, August. Density-based community detection in geo-social networks. In *Proceedings of the 16th International Symposium on Spatial and Temporal Databases* (pp. 110-119).

**Q#4: In our project on 311 service calls in Dallas, explain how we use concrete geospatial objects to draw insights about abstract geospatial objects?**

For a data science project, concrete objects have features (which are observations or measurements that we use to characterize these objects).
We hypothesize how these concrete objects can be aggregated to form abstract objects, and
the abstract objects help us draw insights and build knowledge.

Spatial data science involves three stages to find and extract objects from data:

1. What concrete objects or conceptual objects (or abstract objects) serve our goal , and what concrete objects do we have?
2. How to characterize the conceptual objects?
3. How to relate the concrete objects we have to form the conceptual objects?

Both concrete and abstract objects are subject to issues with

1. Boundary: pre-determined, indetermined, post-determined
2. Properties: continuous or discrete; homogeneous or organized components; static or vary over time

Either concrete or abstract geospatial objects can be individuals or collectives

1. Geospatial objects that are individuals
2. Geospatial objects that are collectives

Objects serve the foundation for any databases, including geospatial databases.
In essence, we "objectify" reality in databases (relational, object-oriented, graph, or file databases)

1. An object is what is countable
   (in GIS databases, an object is a digital representation of a geospatial entity; objects exist in databases to represent entities in reality)
2. An instance is a realization of a class
3. A class is an object type

Q#5: Which of the following issues that we will encounter when we deal with abstract geospatial objects with pre-determined boundaries.
Choose all that apply. B,C,D,E

A) Their corresponding concrete geospatial objects must have pre-determined boundaries
B) The knowledge out of these abstract geospatial objects is only valid with the pre-determined boundaries.
C) These abstract geospatial objects are fixed at locations (not movable)
D) The pre-determined boundaries limit the precision of these abstract geospatial objects.
E) If we choose different pre-determined boundaries, we may end up with abstract geospatial objects in shapes, sizes, and locations)

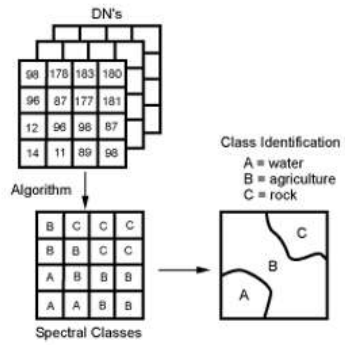
different Modifiable Area Unit problem

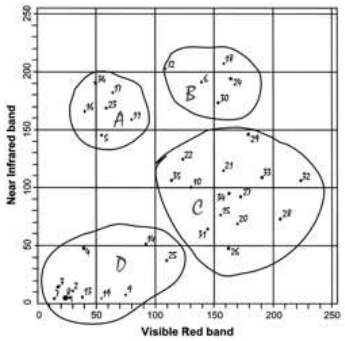
Q#6: Which of the following database statement is false? D
A) It is possible that multiple objects in a database may represent the same entity in reality.
B) It is possible to have more than one instance in a class.
C) A class is by definition a collective.
D) It is possible that a class consists of objects of different types.
E) None of the above

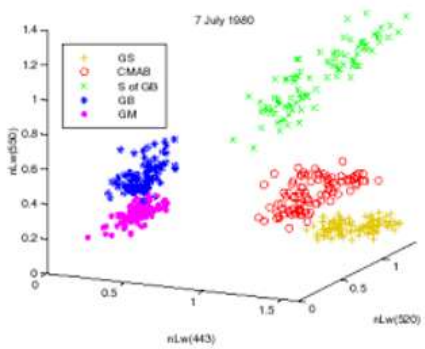
Common strategies to identify and extract spatial objects --> classification

a. Unsupervised classification – based on data values to identify groups
b. Supervised classification – based on labelled data (or classed data)

1. Object detection based on Attribute, feature, or spectral combination. Commonly used in image classification, land use classification, zoning, community detection.
   The goal is to **determine boundaries or thresholds that separate classes**.




Q#7: Select all correct statements: A,B,C,D,E
A) A clustering or hotspot analysis is unsupervised classification.
B) The results of unsupervised classification can be significantly different if we use distinctive features for the input objects.
C) For supervised classification, it is important to have similar numbers of objects in each class among input objects.
D) For supervised classification, the labeled data are considered ground-truth data.
E) All examples in the left figure are unsupervised classification cases.

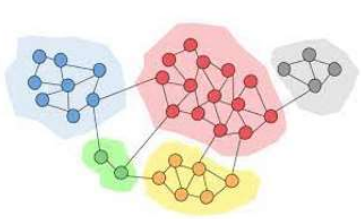
Common data science methods for classification:

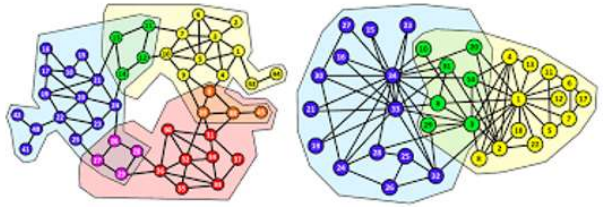
- Unsupervised classification >> clustering

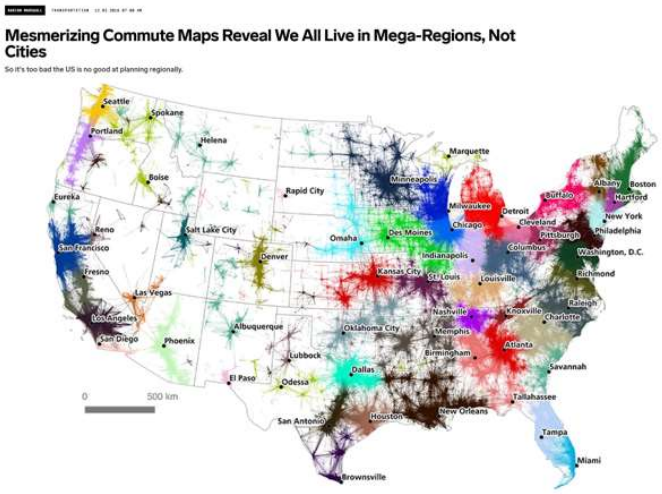
Q#8: What is the essential measure used in unsupervised classification?

Distance definition

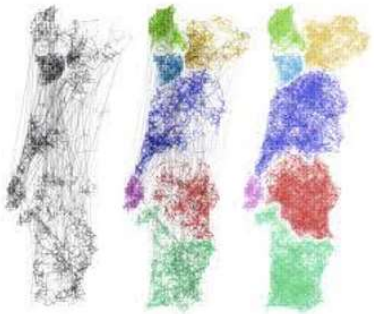
- Unsupervised classification >> clustering
  - K-means
  - Agglomerative Clustering
  - Affinity Propagation
  - Self Organizing Map (SOM)
  - Network Connectivity to determine groups (or regions): more connected within groups and less connected between groups. Connectivity can be physical links (such as transportation networks), social relationships (such as kinships, organizational structures, memberships), information flows (such as messages, calls, or social activities – likes, followers), activities (movements, participations, shopping, financial dependence).
- Supervised classification
  - Logistic Regression
  - Support Vector Machine (SVM)
  - Decision trees and Random Forest
  - Neural Network








https://www.wired.com/2016/12/mesmerizing-commute-maps-reveal-live-mega-regions-not-cities/



https://senseable.mit.edu/community_detection/

Q#9: Are examples in the left figure supervised or unsupervised classification? C

A) unsupervised classification
B) supervised classification
C) can be both unsupervised and supervised classification
D) neither unsupervised or supervised classification