Lecture 7 Attribute-based Operations

Learning Objectives
7.1 What is meant by "reduce information content" using an attribute-based operation? Why would one want to reduce information?
7.2 What is meant by "increase information content" by an attribute-based operation?
7.3 How do "operations on attributes" differ for vector and raster structures?
7.4 What are some operations that work on pairs of data values?
7.5 What are some coastal management issues that can be addressed by attribute operations?

Data potentially carry information content; e.g., relationships previously unknown. Operations (labeled procedures below) "reveal" information content. Generally, attribute operations work *within one measurement framework*. Attribute operations work on attribute data values, not coordinate data values. Operations on attributes can be organized by the **level of measurement** of the input and the output. Sometimes **external information** is required to 'reveal' information.

7.1 What is meant by "reduce information content" using an attribute-based operation? Why would one want to reduce information? Assume data values at a particular level of measurement for input, then apply operation to derive (reveal) output result.

Operations that REDUCE information, simplify information and increase insight.
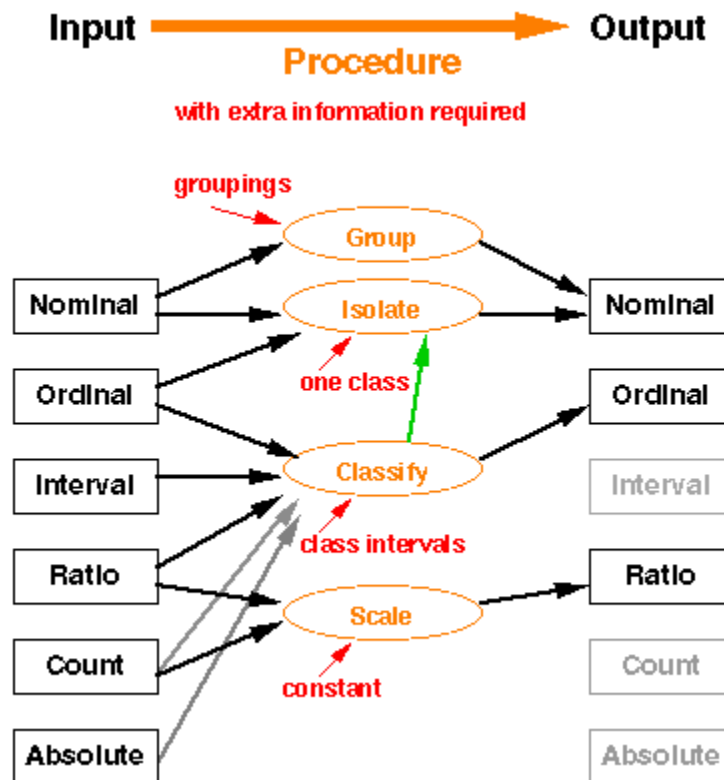
Group, Isolate, Classify, and Scale Operations used in procedure *reduce extraneous information content* by collapsing **many** categories to **fewer or one**. The red text above identifies the external information (which can be thought of as "rules") required to make this happen.

Group and Isolate working on nominal data values result in nominal data values.

Classify reduces higher level measurements into ordinal measurements.

Some **classification** rules are:
- Equal Intervals (depends solely on range of attribute; applies only to Interval or higher)
- Quantiles (depends on counting the objects; depends on measurement framework)
- Jenk's Iterative (minimizes within class variance; depends on Interval or higher)
- "Natural" Breaks (an aesthetic, experiential judgment about gaps, what Esri calls Jenk's)
- Thresholds with external meaning (freezing point of water, gain/loss, legal requirements, etc.)

Isolation of a single category can change from an exhaustive categorical coverage to an isolated object view, but the decisions about the original coverage provide the geometric representation.

Scale reduces information only if the *resolution* of the scale is altered (e.g., adopt coarser grained measurement unit), otherwise a scaling changes no information.

Even these simple operations produce results that influence the apparent measurement framework. When the aspatial groupings of attributes require a geometric process to remove boundaries which are internal to one category we call this **Aggregation**. In ArcGIS software it is referred to as **'dissolve'** or **'drop-line'** aggregation.

7.2 What is meant by "increase information content" by an attribute-based operation?

Operations that INCREASE information, to enrich insight
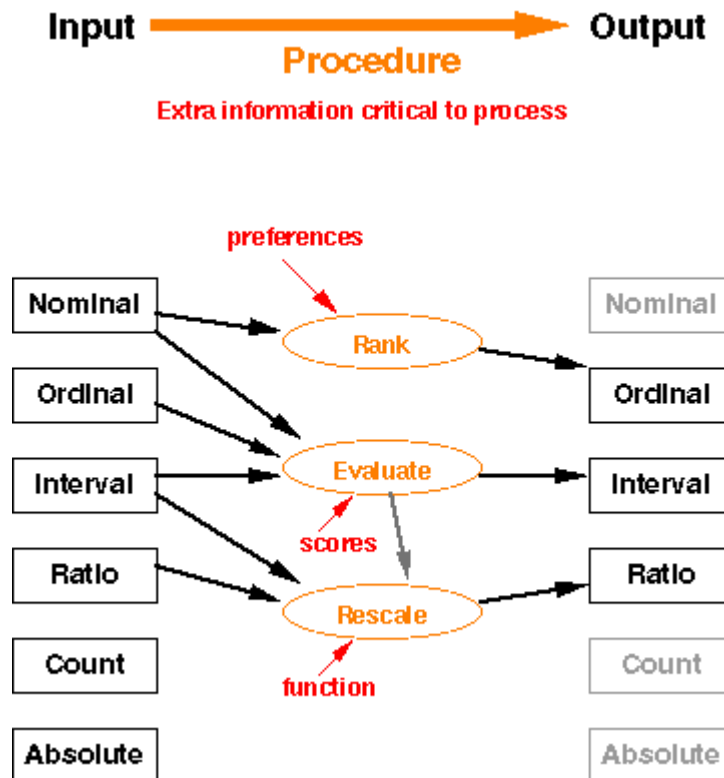
**Rank, Evaluate, Rescale operations** increase information content.
Increase in information content comes from **external assumptions** or **judgments**. With a source of *preferences,* you can order categories. With a set of *scores*, you can assign numbers to categories, e.g. winners and losers. With a non-linear function, you can rescale numbers, for every number increase (domain) of "1", make the output a number increase (range) of "two". The external information comes from additional insight ("external assumptions") one might have about the problem under study.

For example, for "wetlands rating system" categorized into four groups as identified by local planners working with the Washington Dept of Ecology. The rationale for wetlands categories starts on document page 6 (pdf here).   The emphasis is on identifying those wetlands:
- where our ability to replace them is low,
- that are sensitive to adjacent disturbance,
- that are rare in the landscape,
- that perform many functions well, and
- that are important in maintaining biodiversity.

7.3 How do "operations on attributes" differ for vector and raster structures?

Operations on attribute data values take one set of attributes and give you new "columns". How this is implemented depends on the data format (assuming a similar structure).

Operation

Input   Output
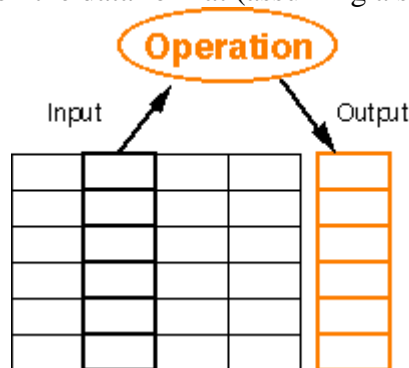
Although vector data geometries vary more than raster data geometries, the implementation of attribute-only operations are much the same when we hold the geometry constant. Thus, attribute operations on raster and vector data are much the same, assuming you are not operating on coordinates.

With attribute operations only we can assume that location stays the same for the features under study in the problem, as we are not operating on the locational data. The questions to be addressed in your analysis are thus simpler than questions wherein space and attribute or space, time, and attribute are involved in the analysis.

7.4 What are some operations that work on pairs of data values?
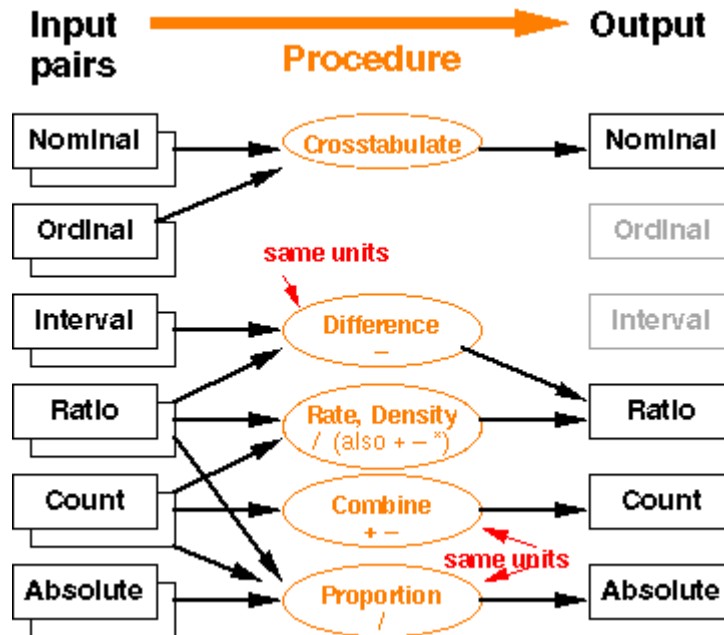
Cross-tabulate, Difference, Rate, Density, Add, Subtract, Proportion, etc. operations

If you have pairs of data values for the same set of objects (representation), you can generate new data values that depend on those **PAIRS** of data values.

Attribute operations through LOOKUP Tables – check table to get data value
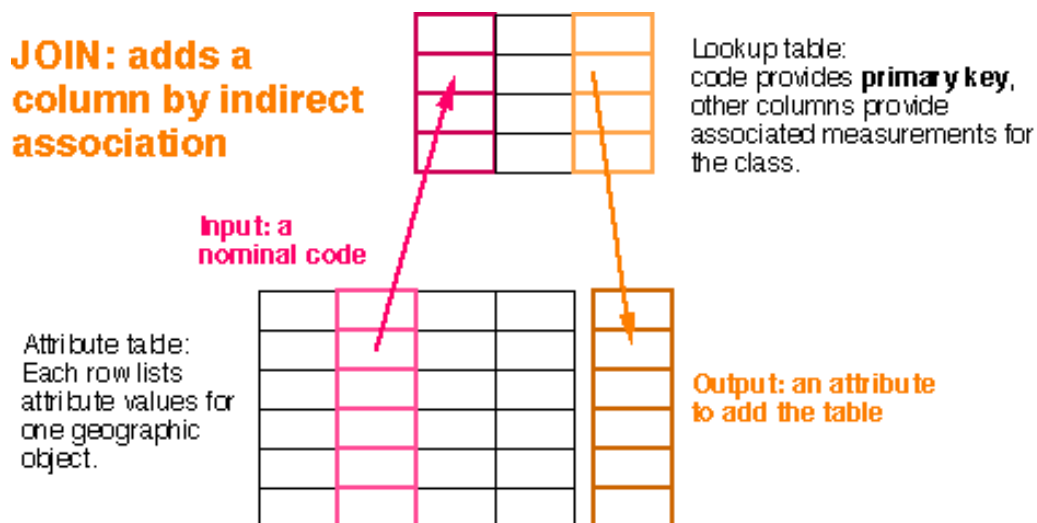
The 'attribute' provides a **foreign key** to a lookup table of attribute values. This process is called **JOIN** in a relational database package.

7.5 What are some coastal management issues that can be addressed by attribute operations?

Much of this material comes from Beatley et al. Chap 3

Remember the attribute operations DO NOT change the geometry of features (coordinate data). We are not working on coordinates, but on the descriptive aspect of data as "attribute".

How might we use attribute operations to address data issues?

Some of the most significant pressures as characterized by attribute data.
- Population pressures – people continue to move to coasts around the world
    o Population change (growth and decline in particular places)
- Land use and cover categories patterns (with population growth)
    o For land use - codes like residential, commercial, industrial
    o For land cover - codes like urban, vegetated, rock
- Ownership of the land and water interface varies
    o How far (distance) upland and how far (distance) marine-ward
    o Various zones of influence

Some critical coastal management issues identified by Beatley et al. Chap 3 are:
- Storm flood mitigation - storm surge often causes considerable impacts
- Coastal shore erosion – loss of beach or bluff, slumping into water
- Sea level rise – or lowering in some areas due to change in rise of land.
- Protecting coastal waters – from pollution such as oil spill or other hazards
- Protecting coastal wetlands and resource lands – from land development encroachment
- Energy development – wave and wind energy are renewable sources of energy
- Biodiversity and habitat conservation – greater diversity encourages stability
- Coastal recreation – is it increasing or decreasing for a given area
- Coastal community development – green development or non-sustainable development
- Public interest versus private interests – in all of the above

How might we use GIS attribute data and operations to address those issues as part of a GIS application? Remember we are not dealing directly with "space" (that is coordinate data values), but simply the attribute data values associated with spatial objects.

What is the concern about the particular issue? For example, what is the concern about flood mitigation in regards to reducing or increasing information when analyzing flood mitigation? What attribute data might we be dealing with in regards to flood mitigation? For example, create more open space as part of land use activity next to rivers to allow flooding.