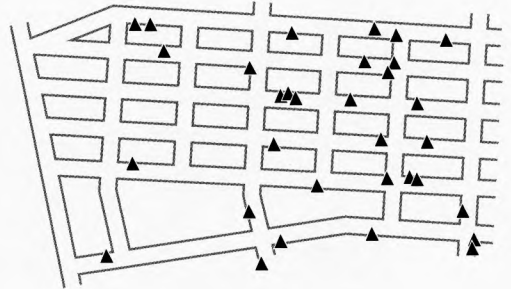


Turning Phenomena into Data

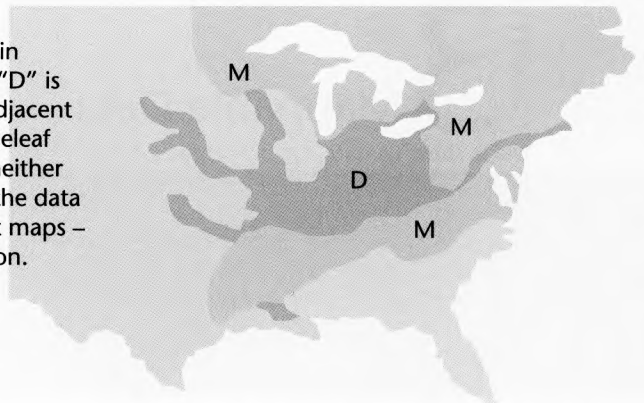
Phenomena are all the stuff in the real world. Data are records of observations of phenomena. Maps show us data, not phenomena. Carefully consider the data you are mapping, how they relate to the stuff in the world, how they are similar, how they are different, and how that may affect our understanding of the phenomena. Differentiate, for example, between individual and aggregated data:

A map of individual pin oak trees in the southern part of the Clintonville neighborhood in Columbus, Ohio. The phenomena are trees, and the data, individually mapped by members of the local "urban arboretum," retain this individuality.



An old map shows the range of the pin oak tree in the U.S. The phenomena are not trees but the potential range of a species, and the data – aggregated from other maps – embody this abstraction.

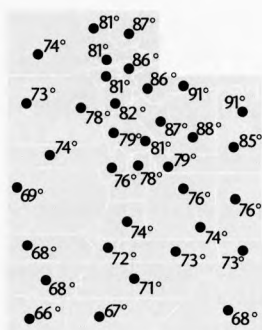
A map of major vegetation zones in the eastern U.S. The area labeled "D" is broadleaf deciduous forest. The adjacent "M" is mixed broadleaf and needleleaf forest. The phenomena here are neither trees nor species, but forest, and the data – aggregated from many different maps – embody this still greater abstraction.



Aggregating data *changes phenomena*, as here, from individual trees, to species, to forest.

Another issue is variation: how phenomena, and so data, change over space and time. Continuous phenomena vary more or less smoothly. Good examples include atmospheric pressure or temperature. Discrete phenomena exist at some places and not others (although they may move around). Good examples are people, pin oak trees, and Studebaker Champ pickup trucks. There is no necessary relationship between phenomena and data, as it's possible to have different kinds of data, as below left, for a single phenomenon.

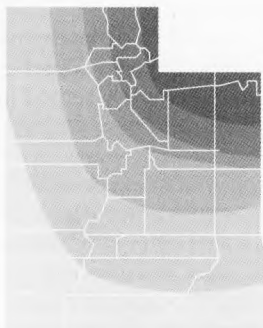
Air temperature varies *continuously* and is everywhere, but thermometers in weather stations can only record it at points. We can map temperature as collected revealing the structure of the *data*...



Humans are *discrete* phenomena. The U.S. Census counts the number of people at individual addresses during each decennial census. We could map people as collected at addresses, except we can't: the Census Bureau keeps this resolution of data secret to protect privacy.

Four people live at
100 East 3rd Avenue
Mexican Hat, UT 84531.

...or we can transform the point data into continuous data by interpolating data likely to exist between the readings. We then create a continuous surface on the map with this interpolated data to reveal the structure of the *phenomena*:



We map detailed census data into single-value areas as small as a city block, but usually counties or states. The map below implies that people are continuous throughout each county, but we know that not to be the case. Here the map shows us more about the structure of the *data*:

