

Geographic information systems organize information pertaining to geographic features and provide various kinds of access to the information. A geographic feature may possess many attributes (see below). In particular, a geographic feature has a specific location.

There are a number of ways to specify location. For this project, we will use latitude and longitude, which will allow us to deal with geographic features at any location on earth. A reasonably detailed tutorial on latitude and longitude can be found in the Wikipedia at [en.wikipedia.org/wiki/Latitude](http://en.wikipedia.org/wiki/Latitude) and [en.wikipedia.org/wiki/Longitude](http://en.wikipedia.org/wiki/Longitude).

We will employ public data obtained from the Geographic Names Information System ([geonames.usgs.gov](http://geonames.usgs.gov)). Each data file consists of a list of ASCII records, each describing a single geographic feature with many attributes.

More precisely, each GIS record may contain the following fields in the indicated order (all are mandatory unless indicated otherwise):

**Figure 1: Geographic Data Fields**

Significance	Type/Format	Comments
Feature ID number (FID)	non-negative integer	unique identifier for this geographic feature
State alphabetic code	two-characters	US postal code abbreviation
Feature name	string	standard name of feature
Feature type	string	descriptive classification of feature
County name	string	county in which feature occurs
State number code	non-negative integer	numeric code for state
County number code	non-negative integer	numeric code for county
Primary latitude (DMS)	DDMMSS['N'   'S']	feature latitude in DMS format
Primary longitude (DMS)	DDMMSS['E'   'W']	feature longitude in DMS format
Primary latitude (dec deg)	decimal number	feature latitude in decimal format
Primary longitude (dec deg)	decimal number	feature longitude in decimal format
Source latitude (DMS)	DDMMSS['N'   'S']	latitude of feature source in DMS format, optional
Source longitude (DMS)	DDMMSS['E'   'W']	longitude of feature source in DMS format, optional
Source latitude (dec deg)	decimal number	latitude of feature source in decimal format, optional
Source longitude (dec deg)	decimal number	longitude of feature source in decimal format, optional
Feature elevation	integer	altitude above/below sea level, optional
Est. feature population	non-negative integer	estimated population of feature, optional
Federal status	string	???, optional
Cell	string	???

In the GIS record file, each record will occur on a single line, and the fields will be separated by pipe ( ` | ` ) symbols. Some sample records are shown below. Note that it was necessary to wrap the lines in order to fit the data between the margins of this page. In the actual data files, each record would appear on a single line by itself.

Some records will contain the string "UNKNOWN" instead of the specified longitude and latitude fields. Aside from that, GIS record files are guaranteed to conform to this syntax, so there is no explicit requirement that you validate the files. On the other hand, some error-checking during parsing may help you detect errors in your parsing logic.