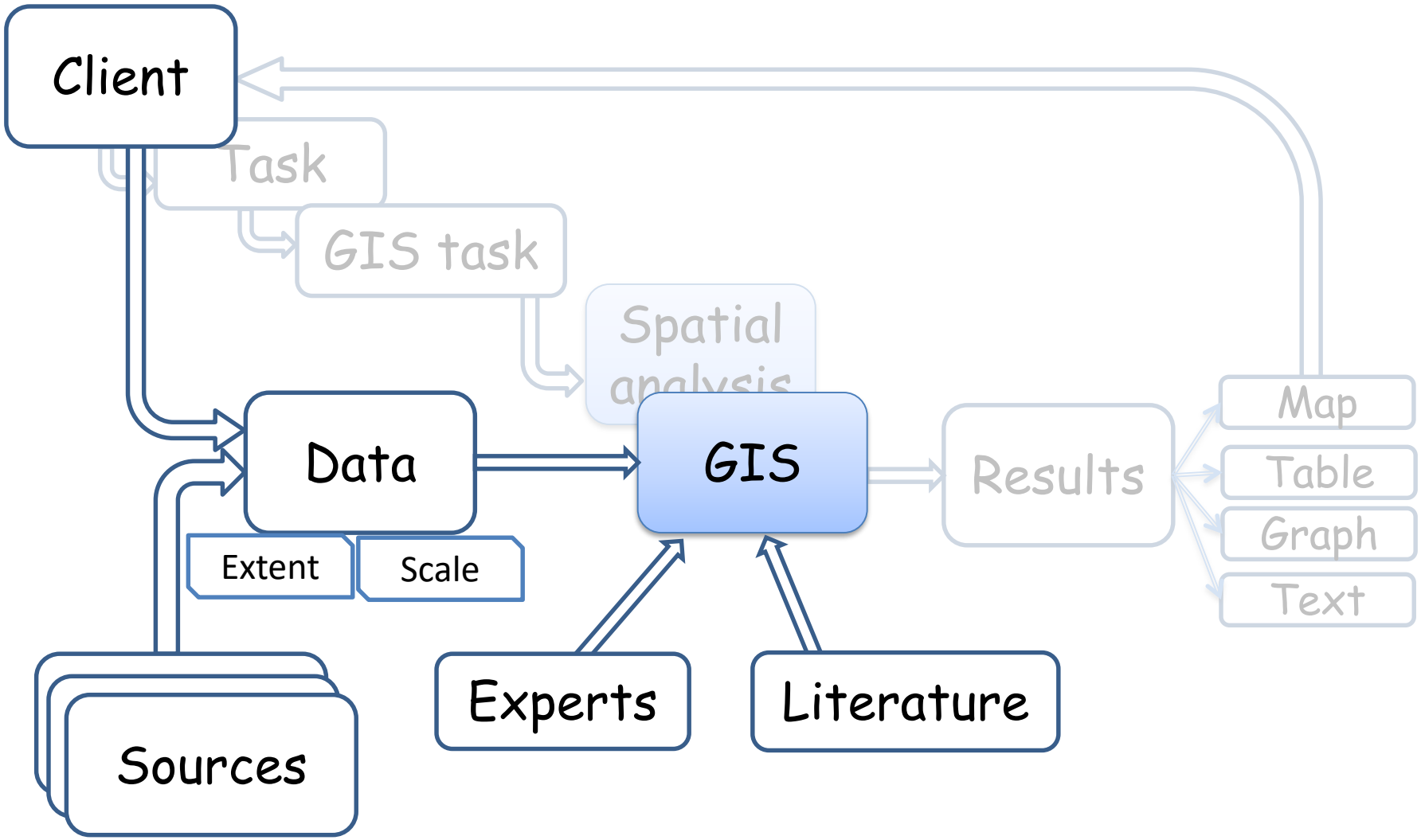
Duke
NICHOLAS SCHOOL OF THE ENVIRONMENT

Project Based GIS: Geospatial data

ENVIRON 761

Geospatial Applications for Conservation & Land Management

Geospatial data

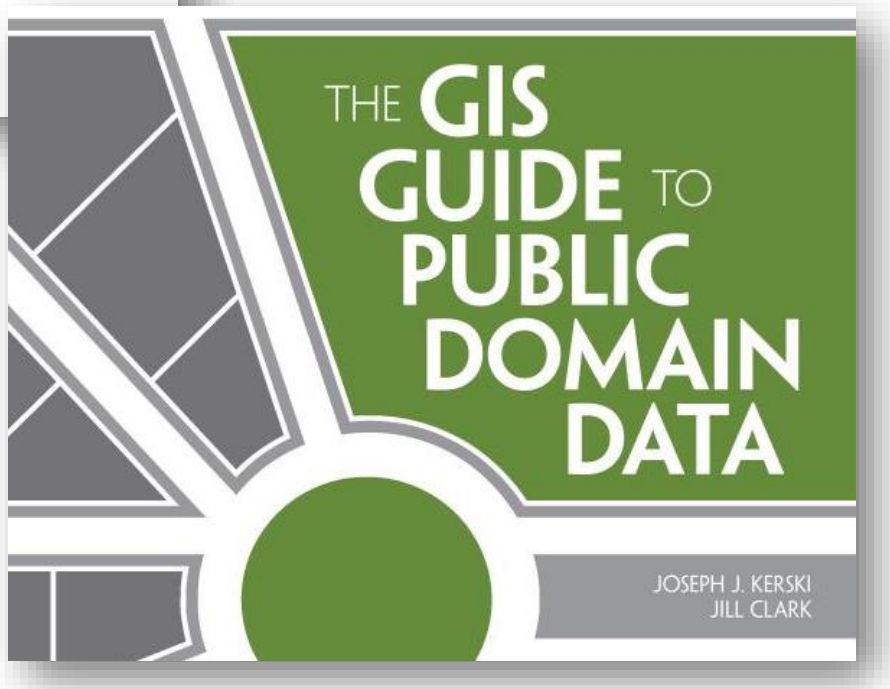


Overview

- Role of **data** in spatial analysis & GIS
- **Review**: Spatial data formats…
- **Finding** data
  - Useful data collections and on-line data portals
  - Trends data availability and access; what to expect
- **Evaluating** data
  - Considerations when selecting a dataset to use…

References

2001



2012

Data, spatial analysis & GIS

Figure 1.1 Data-to-wisdom processing pyramid.

Decker 2001. GIS Data sources.

Data, spatial analysis & GIS

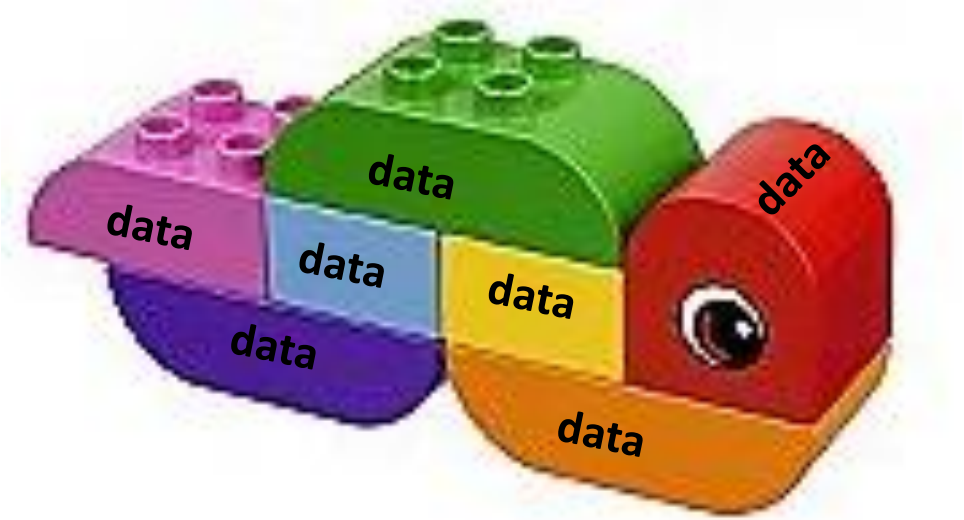
Geographic GIS stores, retrieves, displays, and manages **spatial** information

Information GIS converts **data** into **information**

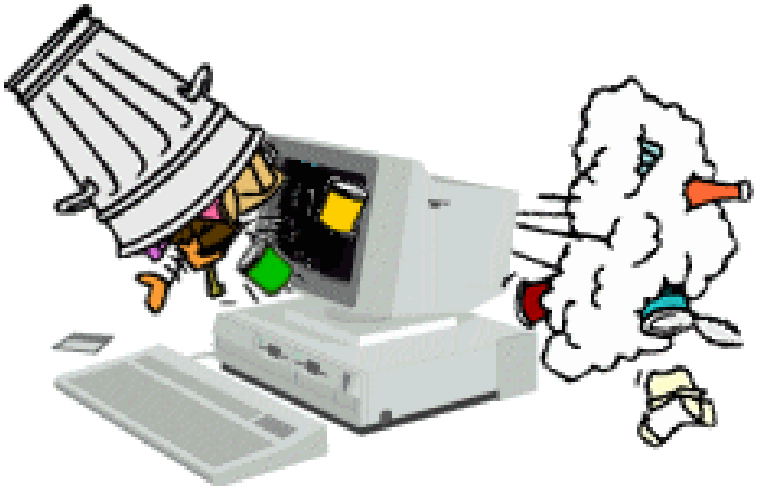
System GIS **stores**, **retrieves**, **displays**, and **manages** information

Data, spatial analysis & GIS

1. Data are the fundamental building blocks spatial analysis…

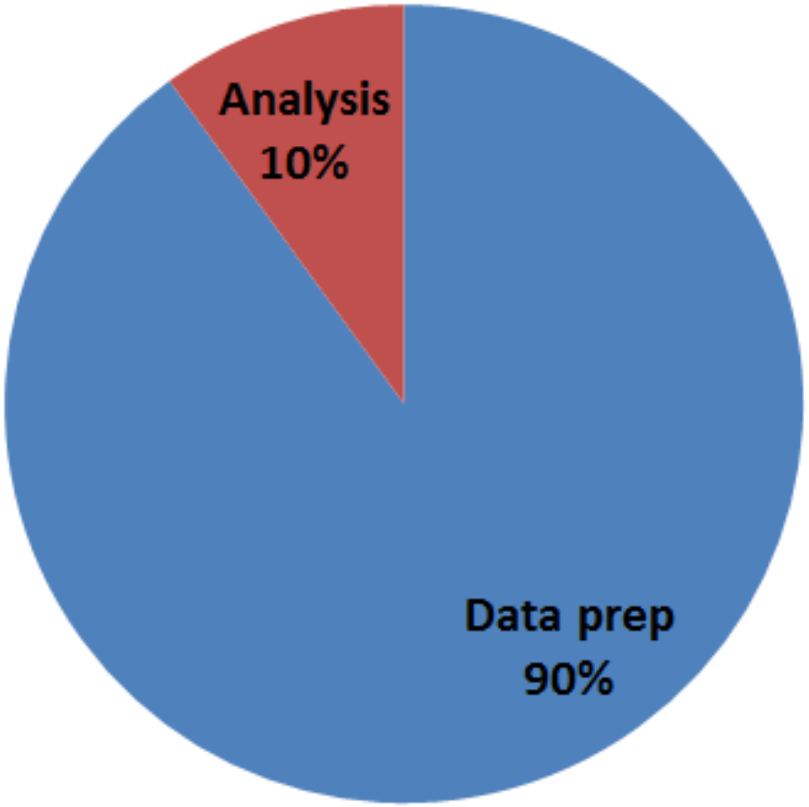


2. Garbage in…. → Garbage out…



Finding GIS Data

Ye olden days…



Today…

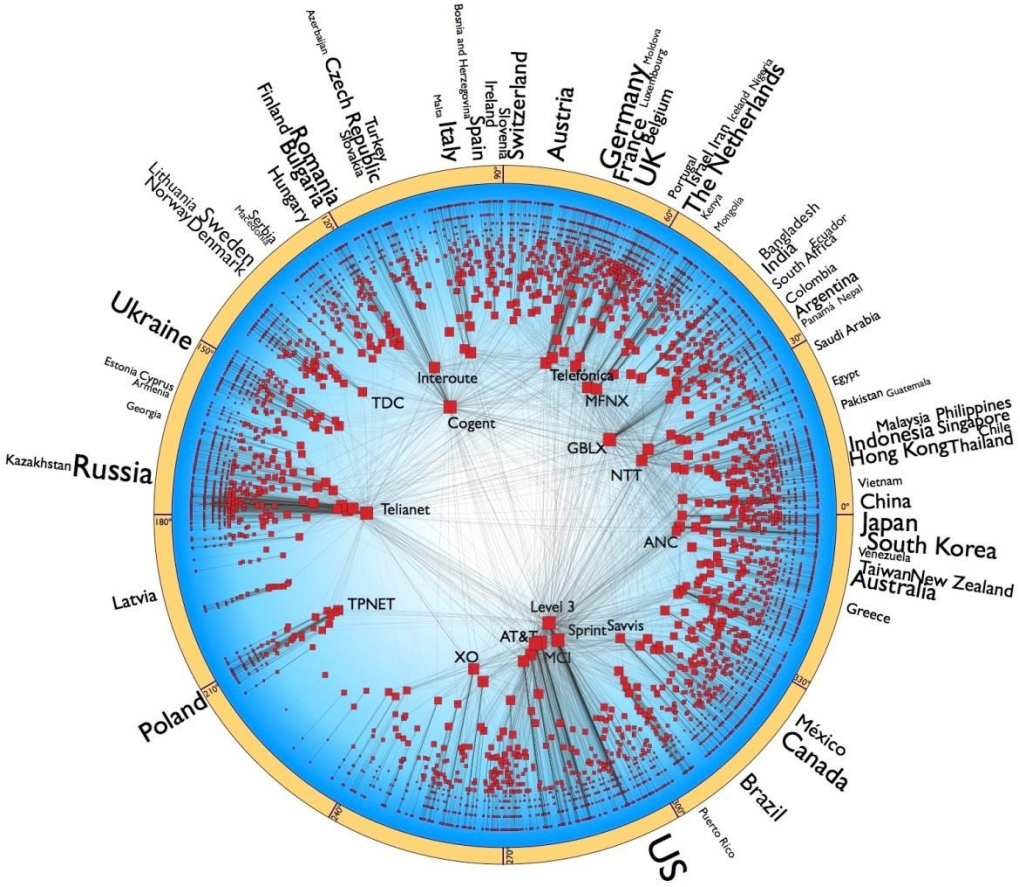
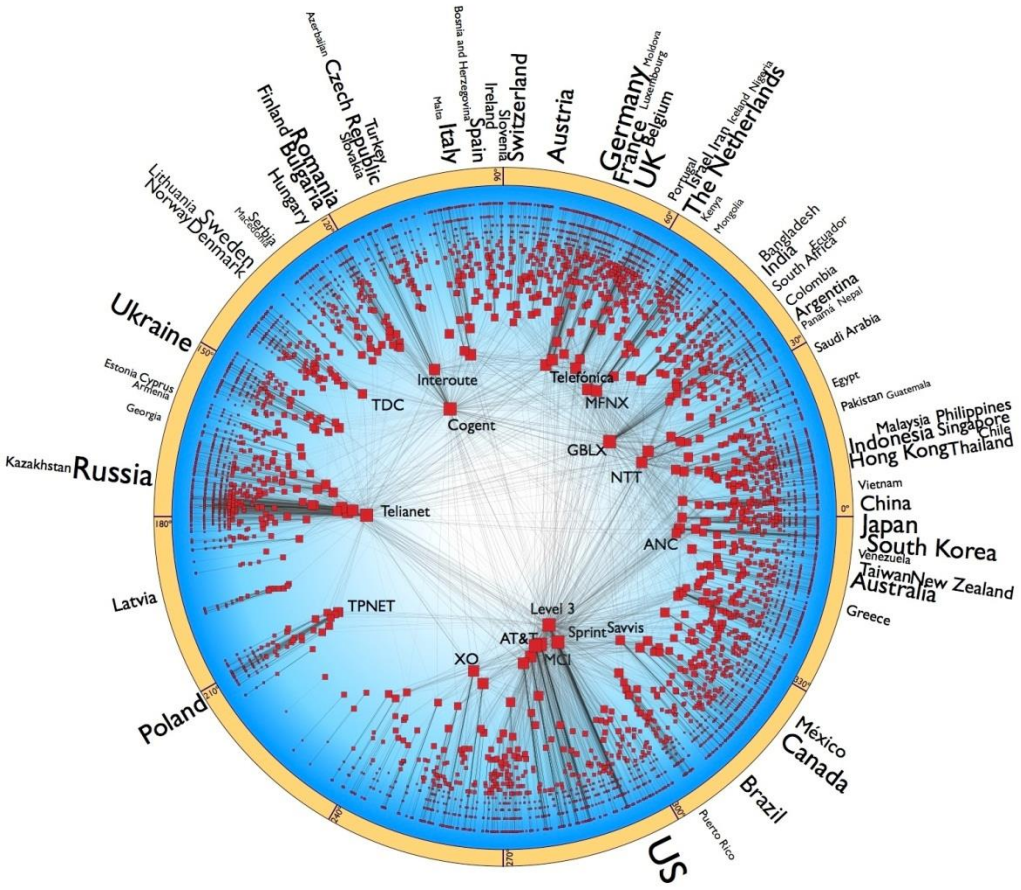


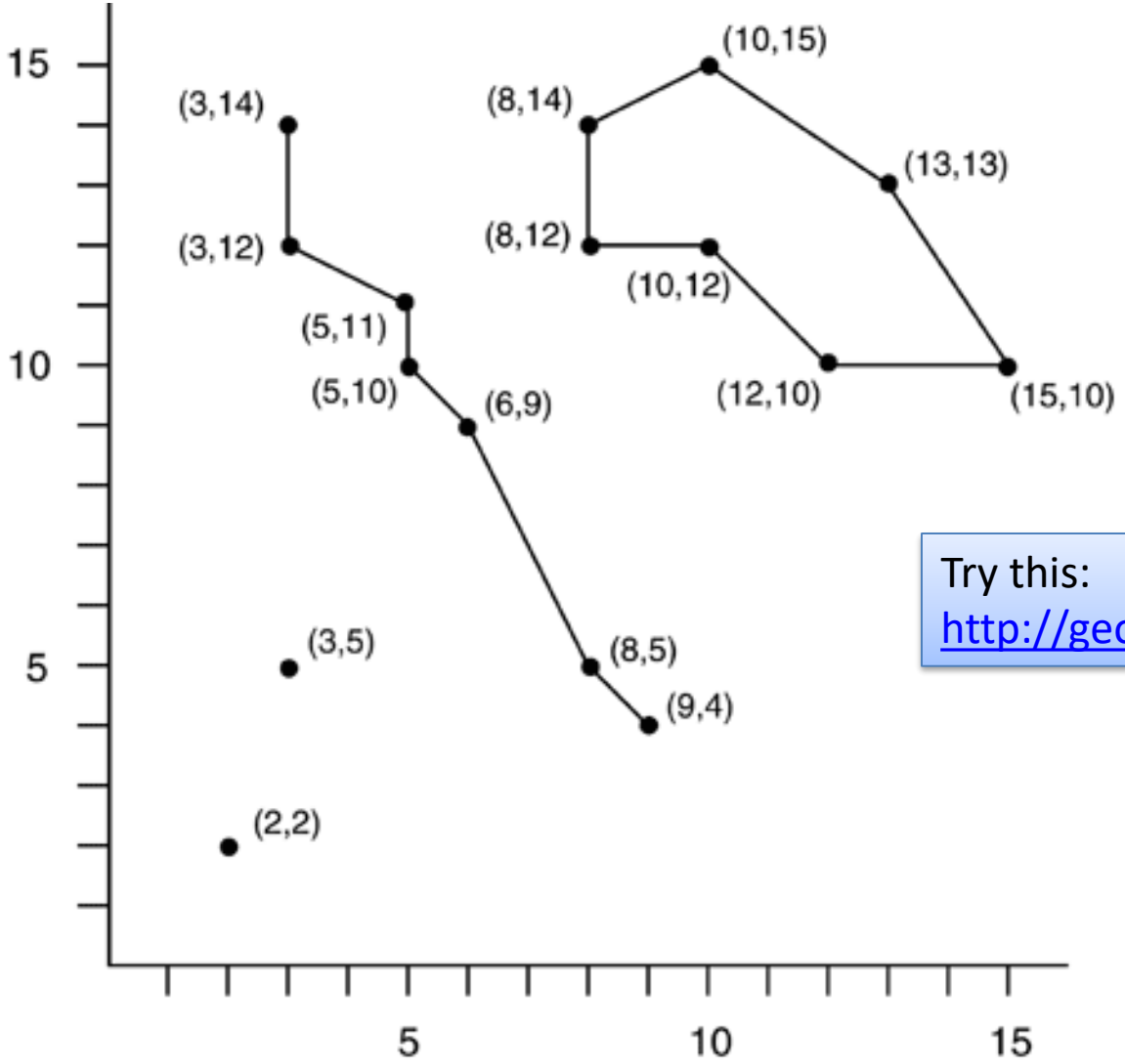
Image courtesy of Dmitri Krioukov, SDSC/CAIDA

Finding GIS Data



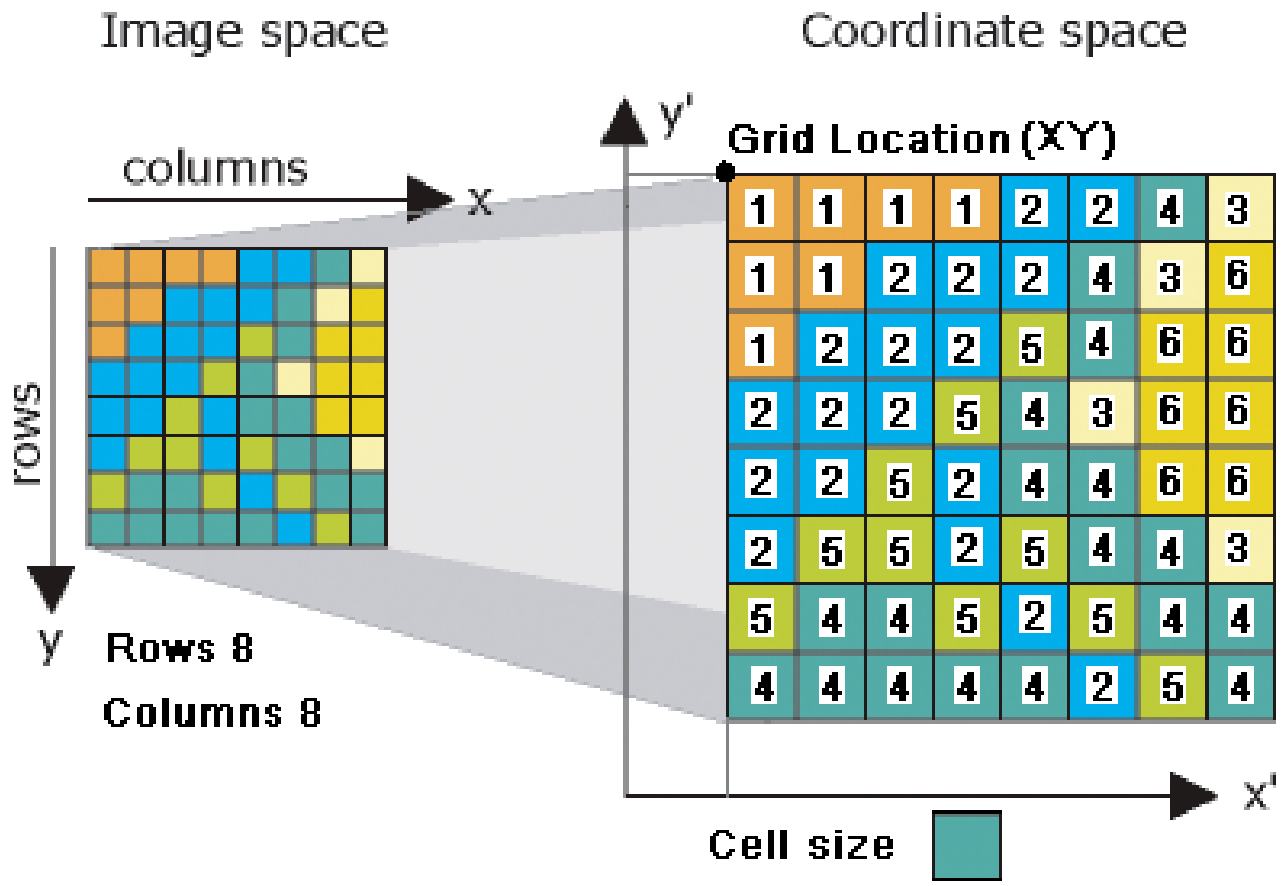
"Best Available Data"

Review: Data Formats



Try this:
http://geojson.io

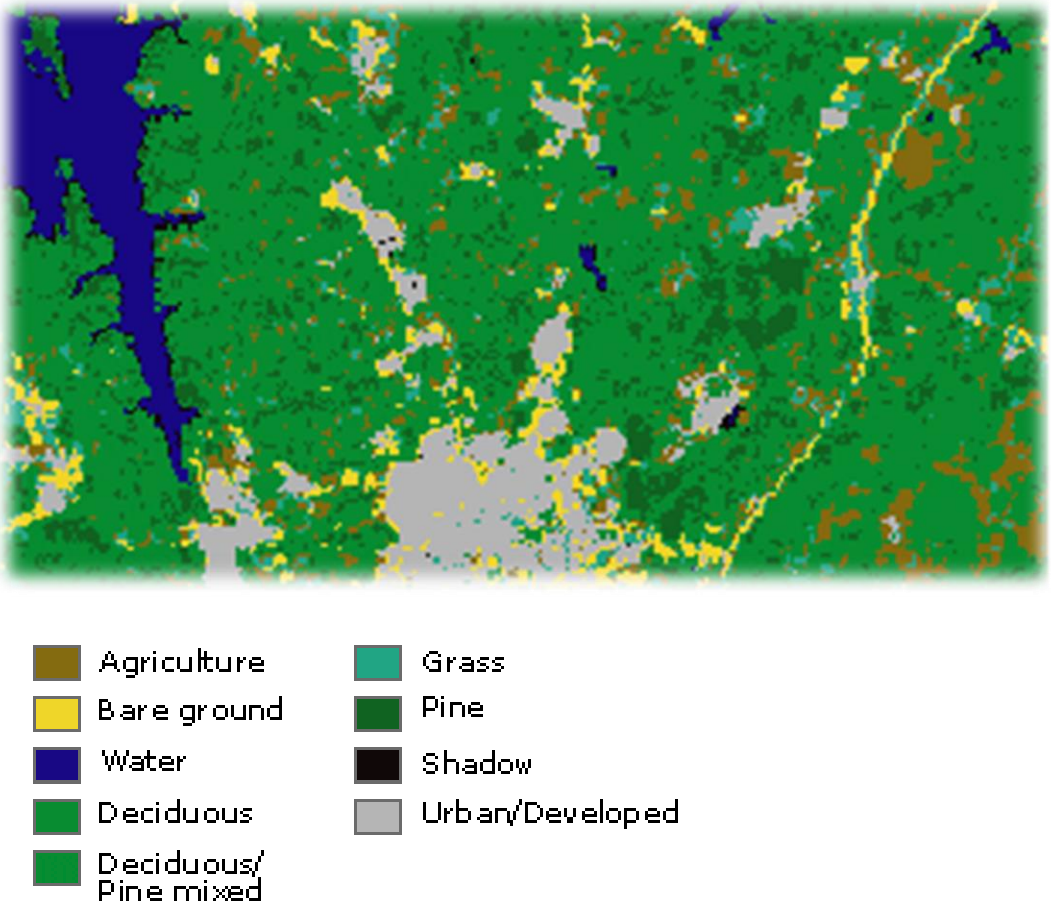
Review: Data Formats

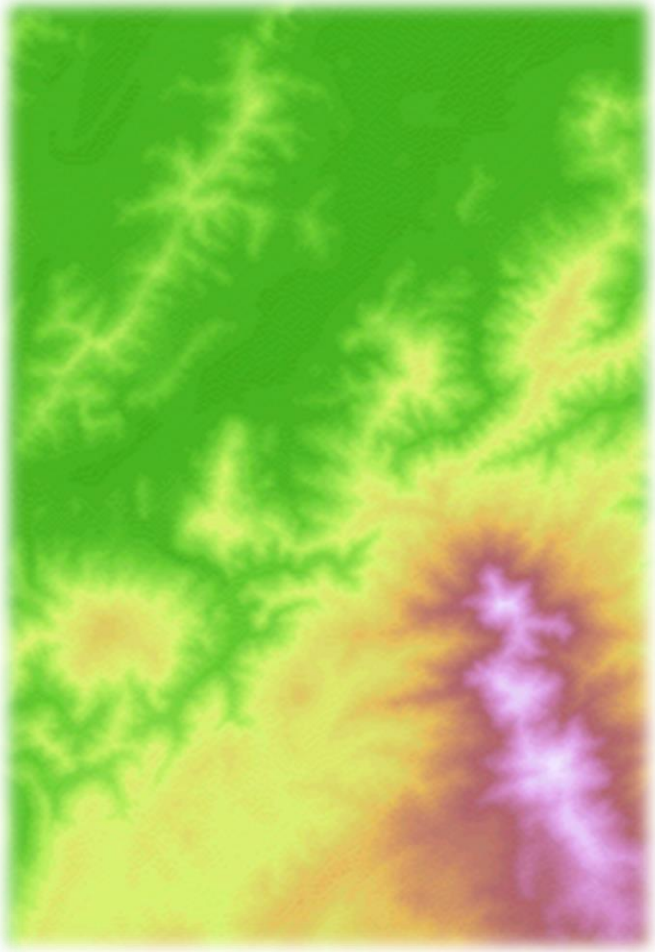


List of cell values

[1111224311222436122254662225436622524466255254435445254444444254]

Review: Data Formats





Review: Data Formats

- Annotations
- CAD
- Coverages
- Dimension features
- Domains
- Feature classes
- Feature datasets
- Geometric networks
- KML
- LAS dataset
- Locators
- Metadata
- NetCDF
- Network datasets
- Parcel fabrics
- Raster and images
- Relationships and related objects
- Shapefiles
- Subtypes
- Tables
- Terrains
- TIN
- Topologies

Why so many different formats?

- Each tied to a unique purpose
- Specific uses/applications…
- Advantages/disadvantages…
- Responses to changes in scale…

Obtaining data

You are about to begin a project looking at the biological impacts of expanding palm oil plantations in Indonesia…

You need data on:

- Land cover
- Topography
- Hydrography
- Infrastructure
- Political boundaries

How might you get these data?

Public Domain Datasets

Public Domain Data:

Publicly accessible information about a spatial theme or phenomenon, the use of which does not infringe the legal rights of an individual or organization.

Kerski (2012)

Finding Data

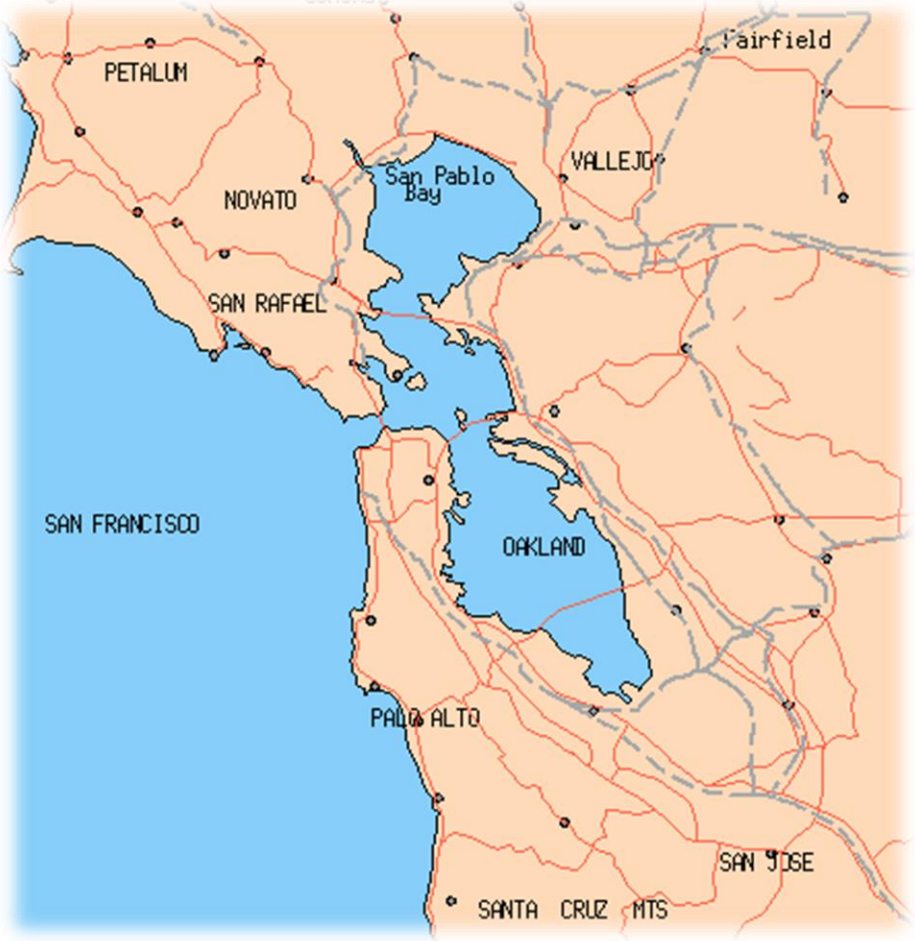
- Useful existing datasets
  - Vector
  - Raster
- Geospatial data clearinghouses & portals
- Searching for specific data

Public Domain Data Sources

Useful public domain data collections:

- Digital Chart of the World (DCW)
  http://en.wikipedia.org/wiki/Digital_Chart_of_the_World
- Global GIS Data
  http://webgis.wr.usgs.gov/globalgis/
- ESRI's Data and Maps for ArcGIS (on DVDs)
  http://library.duke.edu/data/collections/gis/esri/esri_2012/index.html

Digital Chart of the World (DCW)



- Political/ocean (country boundaries)
- Populated places (urbanized areas & points)
- Roads
- Railroads
- Aeronautical structures
- Utilities (electrical, telephone, pipelines)
- Drainage system
- Hypsographic data
- Land cover
- Ocean features
- Physiography
- Cultural landmarks
- Transportation structure
- Vegetation
- Data location

http://en.wikipedia.org/wiki/Digital_Chart_of_the_World

Digital Chart of the World (DCW)

- Source: *US Defense Mapping Agency*
  - 1:1m scale operational navigational charts
- Last updated: *1992*
- 3 resolutions:
  - VMAP0 – 1:1m *public domain*
  - VMAP1 – 1:250k *©ESRI*
  - VMAP2 – 1:50k *not publicly available*
- Distributed in *Vector Product Format*

Requires reformatting for use in ArcGIS…

http://www.mapability.com/index1.html

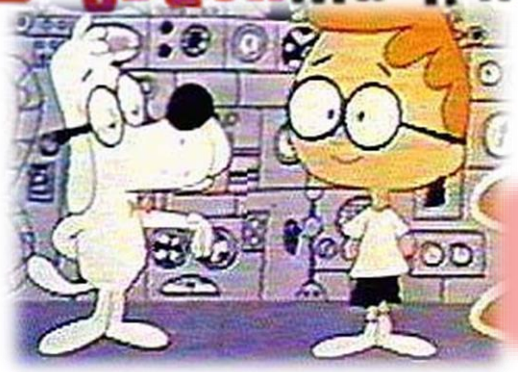
DCW Download & The **Web Archive**

Most sites point to http://www.maproom.psu.edu/dcw/ to download DCW data…

Server not found

Firefox can't find the server at www.maproom.psu.edu.

INTERNET ARCHIVE
WayBackMachine



http://archive.org/web/web.php

The Wayback Machine

http://www.maproom.psu.edu/dcw/

Digital Chart of the World Server
Penn State University Libraries

ESRI Data and Maps

Esri Data & Maps

http://www.esri.com/data/data-maps/data-and-maps-dvd

Esri Data & Maps features downloadable data layers for the following areas of interest:

- **World**: global data layers including country boundaries and cities
- **United States**: data for the USA including states, counties, and landmarks
- **North America**: data for the United States and Canada including states/provinces, highways/roads, and city areas
- **Europe**: data layers for several countries in Europe including country and province level demographics

The ArcGIS Online services published by Esri can be found in the separate Esri Maps and Data group.

http://links.esri.com/arcgisdesktop/en/esrimapsanddata

ESRI Data and Maps

Redistribution rights

Geodata » Datasets provided with ArcGIS

All data is available for internal use. Please review this information before redistributing any of this data.

Redistribution rights data

Description	Data Source	Filename	ESRIDATA Directory	Redistribution (see FAQ below)
World				
Countries (generalized)	ArcWorld Supplement	country.*	\world	Yes 1,2,3
Countries	ArcWorld Supplement	dtl_cntry.*	\world	Yes 1,2,3

- No—Internal Use Only. No redistribution rights are granted by the data vendor, and the data is for the end user's own internal use only.
- Yes 1—Redistribution rights are granted by the data vendor for hardcopy renditions or static, electronic map images (for example, .gif, .jpeg) that are plotted, printed, or publicly displayed with proper metadata and source/copyright attribution to the respective data vendor/vendors.
- Yes 2—Geodata is redistributable with a value-added software application developed by ESRI Business Partners on a royalty-free basis with proper metadata and source/copyright attribution to the respective data vendor/vendors.
- Yes 3—Geodata is redistributable without a value-added software application (that is, adding the sample data to an existing, [non]commercial dataset for redistribution) with proper metadata and source/copyright attribution to the respective data vendor/vendors.
- Yes 4—Public domain data from the U.S. government is freely redistributable with proper metadata and source attribution.

http://resources.arcgis.com/en/help/main/10.1/index.html#/Redistribution_rights/001z00000003000000/

Topical Data Sources

Hydrographic data	• Hydro1k, Digital Line Graphs*, EPA Reach files*, NHD/NHD+*
Wetlands	• GLWD, NWI*
Agriculture	• CropExplorer, FAOSTAT, NASS*
Boundaries/parcels	• GADM, Cresta Maps
Environmental	• IUCN, WWF, EPA*, CEISIN, SEDAC
Energy	• USGS Energy Data Finder
Geologic	• OneGeology, USGS*
Soils	• FAO/UNESCO, STATSGO*, SSURGO*
Geographic names	• GNDB, GNIS, European Geographics
Street data	• DeLorme, TomTom, Garmin, ESRI
Demographic	• CEISIN, US Census*

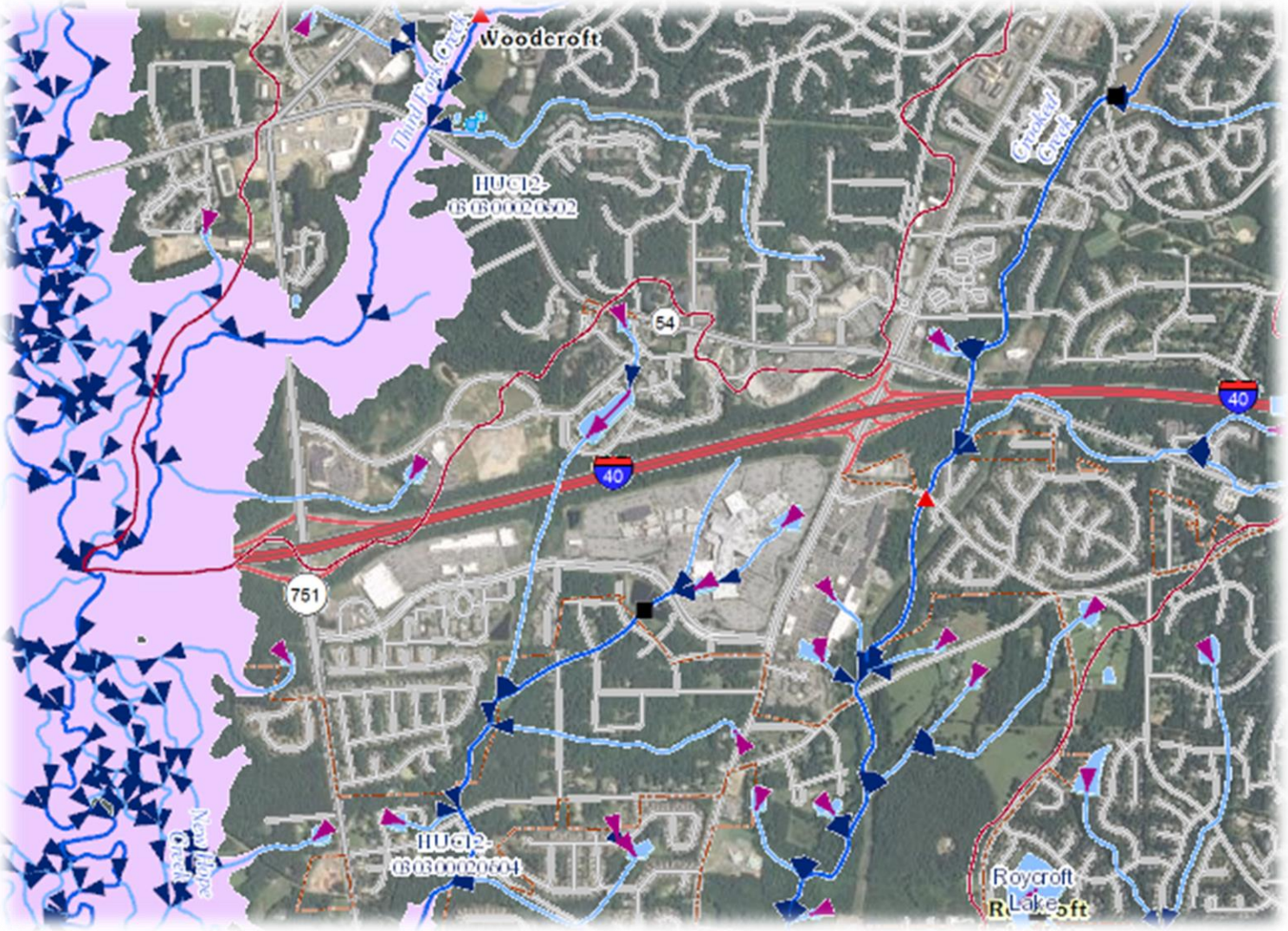
** US only*

National Hydrography Dataset (NHD/NHD+)

http://viewer.nationalmap.gov/viewer/nhd.html

- Combines EPA RF and USGS DLG data
- Includes more features than EPA RF
- 1:100,000 and 1:24,000 scale data
- Developed to work with ESRI's network analyst
- NHD+ contains "value added data"…

National Hydrography Dataset



Agricultural

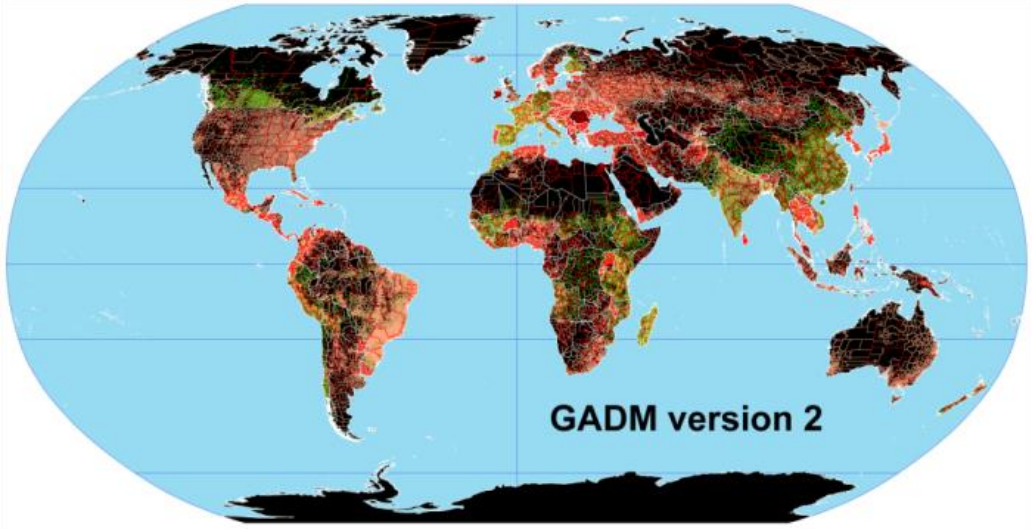
- Crop Explorer:
  http://www.pecad.fas.usda.gov/cropexplorer/index.cfm
- FAOSTAT:
  http://faostat.fao.org/
- National Agricultural Statistical Service (US):
  http://www.nass.usda.gov/
- National Atlas (US):
  http://nationalatlas.gov/mld/agcn07.html

Mostly tabular data. Cannot do very detailed analyses…

Boundaries and Land Parcel Data

- Very problematic…
- Mostly only available at small spatial extents…

- *Global Database of Administrative Areas*: http://www.gadm.org/
  - 1:1m scale
  - 2001
  - Subject to copyright restrictions



Boundaries and Land Parcel Data

- CRESTA Zones and Subzones
  - Used by insurance industry for risk zones
  - Based on postal codes, where available

https://www.cresta.org/

http://www.gfk-geomarketing.de/en/digital_maps/branchspecific_maps/cresta_zones_insurance.html

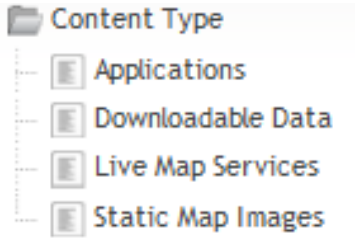
Environmental

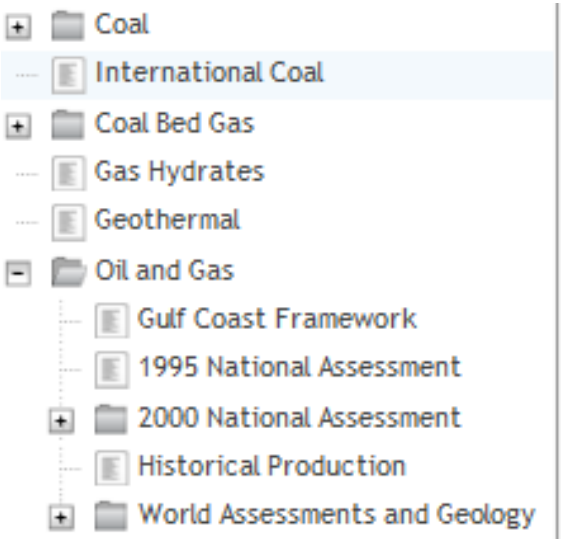
- IUCN Red List species range maps
  http://www.iucnredlist.org/
- WWF Ecoregions
  http://worldwildlife.org/biomes
- EPA toxics, water quality, etc. (US only)
- Last of the Wild
- Human footprint

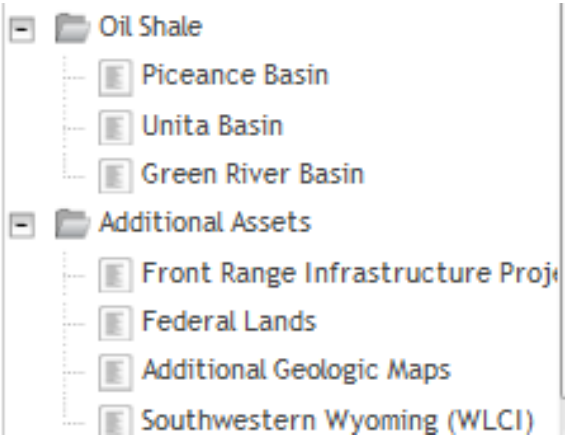
More on these later in the semester…

Energy

- Extremely difficult to obtain…
  - Proprietary, security issues
- *USGS Energy Data Finder*
  - http://energy.usgs.gov/Tools/EnergyDataFinderSplash.aspx







Geology

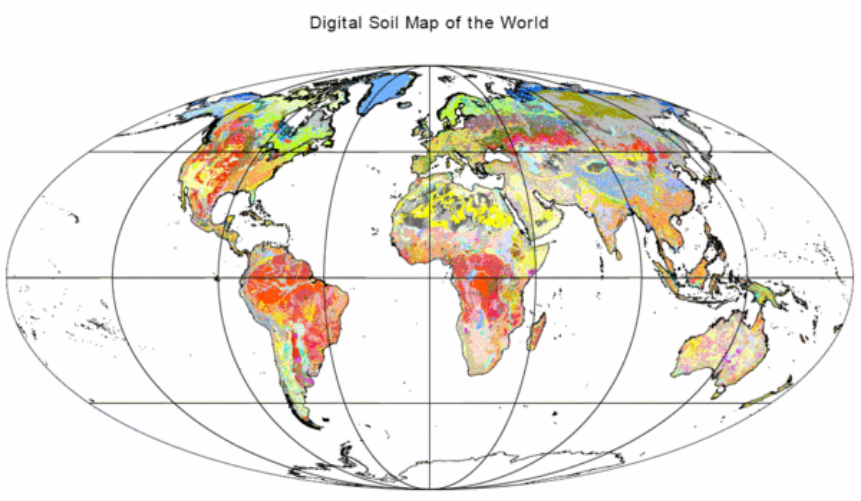
- OneGeology http://www.onegeology.org/
  - International effort “to create and distribute the best geological map”
    - 81 participants from 43 nations
  - Scale 1:1m
  - Started 2007

- USGS mineral resources on-line spatial data library:
  - http://mrdata.usgs.gov
  - Incomplete and varied format
- USGS 1:24k geology quad sheets
  - http://ngmdb.usgs.gov/ngmdb/ngmdb_home.html
  - Patchy coverage; not all are digital

Soils

- FAO/UNESCO
  - http://www.fao.org/nr/land/soils/digital-soil-map-of-the-world/en/
  - 1:5m scale

- USGS Soil Data Mart *(US only)*
  - http://soildatamart.nrcs.usda.gov/USDGSM.aspx
  - STATSGO
    - Regional, multi-state
  - SSURGO
    - For fine scale applications



Human/Demographic Data

Place and feature names databases:

These serve as official identifiers for geographic features

- National Geospatial Intelligence Agency
  - GEONet Names Server (GNS)
  - http://earth-info.nga.mil/gns/html/index.html
- European Geographics
  - http://www.eurogeographics.org/eurogeonames
- Geographic Names Information System (GNIS)
  - US only
  - http://nhd.usgs.gov/gnis.html
  - http://geonames.usgs.gov/

Human/Demographic Data

Street data (locations *and* addresses)

- Mostly proprietary at broad scales (GPS)
  - DeLorme, Garmin, TomTom, TeleAtlas

- Digital Chart of the World
- Digital Line Graph (US only)

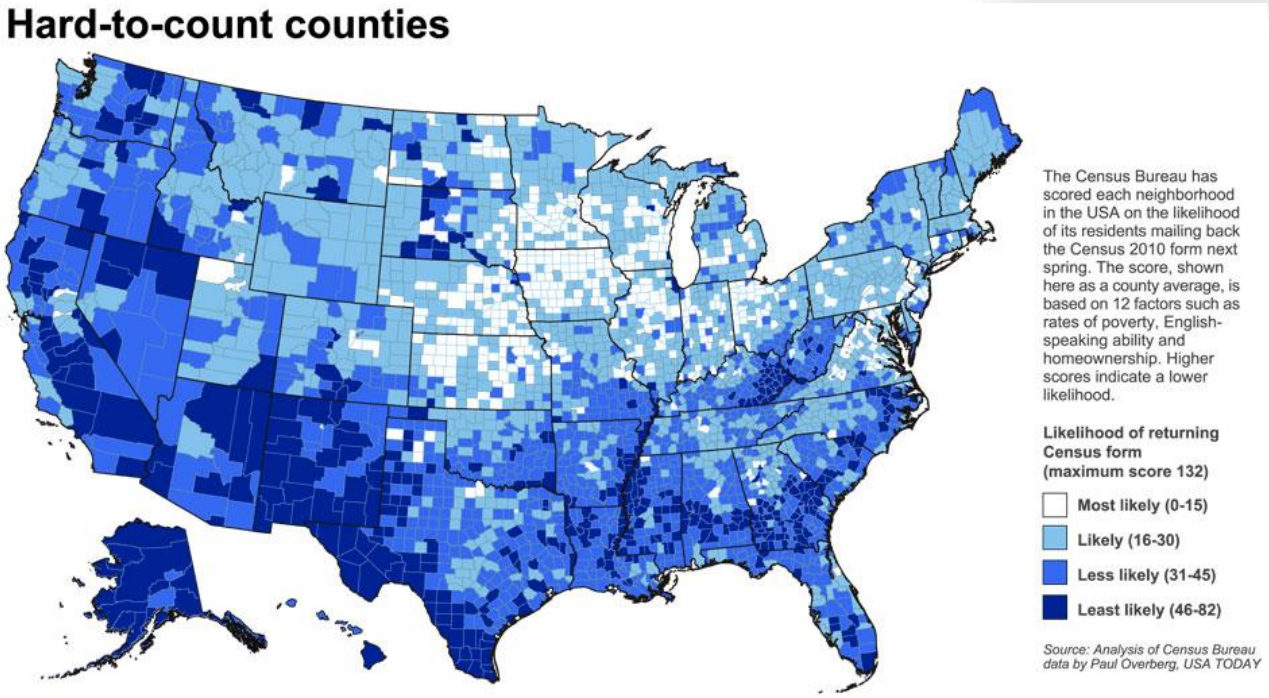
- ESRI provides US and Europe on DVDs

- OpenStreetMap project
  http://www.openstreetmap.org/

Human/Demographic Data

Population/Census

- National Census Organizations
  - http://www.census.gov/geo/maps-data/



Available data: Considerations

GIS Data Law 1 The reason all that existing GIS data often cannot help you is that data are usually created to solve a specific problem and are not designed to be applied to a wide range of applications.

Decker, p. 6

Can you use available data?

- Does the **extent** completely cover your study area?
- Are the data at an adequate **scale**?
- Are the data **accurate** enough?
- Is the **format** useable?
- Are the data **timely**?

GIS Data Law 2 If you don't know the analog data, you don't know the whole story!

Decker, p. 9

Considerations: Extent

Great data may not include your entire study area…

Considerations: Scale and Detail

1:24,000

1:100,000

1:250,000

Scale…

Larger scale maps have more detail…

1:24k	**1:100k**	**1:250k**
1/4 inch = .095 mile	1/4 inch = 1.58 mile	1/4 inch = 3.95 mile
1cm = 0.24 km	1cm = 1 km	1cm = 2.5 km

Considerations: Scale and Accuracy

Accuracy:

Distance between map coordinates and true coordinates.

http://nationalmap.gov/standards/nmas647.html

> 1:20,000 scale	< 1:20,000 scale
90% points within 1/30”	90% of points within 1/50”

Considerations: Scale and Accuracy

1:24,000


1:100,000

1:250,000

<u>1:24k</u>

90% within

- 40 ft.
- 12.2 m

<u>1:100k</u>

90% within

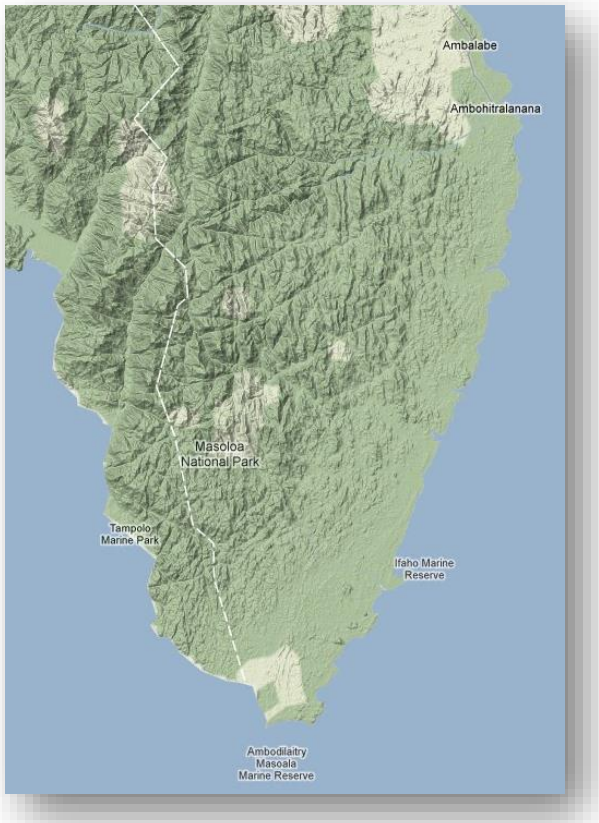
- 167 ft.
- 50.8 m

<u>1:250k</u>

90% within

- 417 ft.
- 127 m

Considerations: Format



Google

Useful *Raster* Data Sources

Imagery	• Satellite…, Airborne…, Scanned data…
Land Cover	• GLOBCOV, GLCNMO, POSTEL, NLCD*, GIRAS*
Demography	• CEISIN, GRUMP, LandScan
Climate/Environmental	• WorldClim, Univ. of East Anglia, CLIMAT, NCDC
Elevation	• GTOPO, SRTM, ASTER, NED*, LiDAR**

Imagery vs. Raster Data

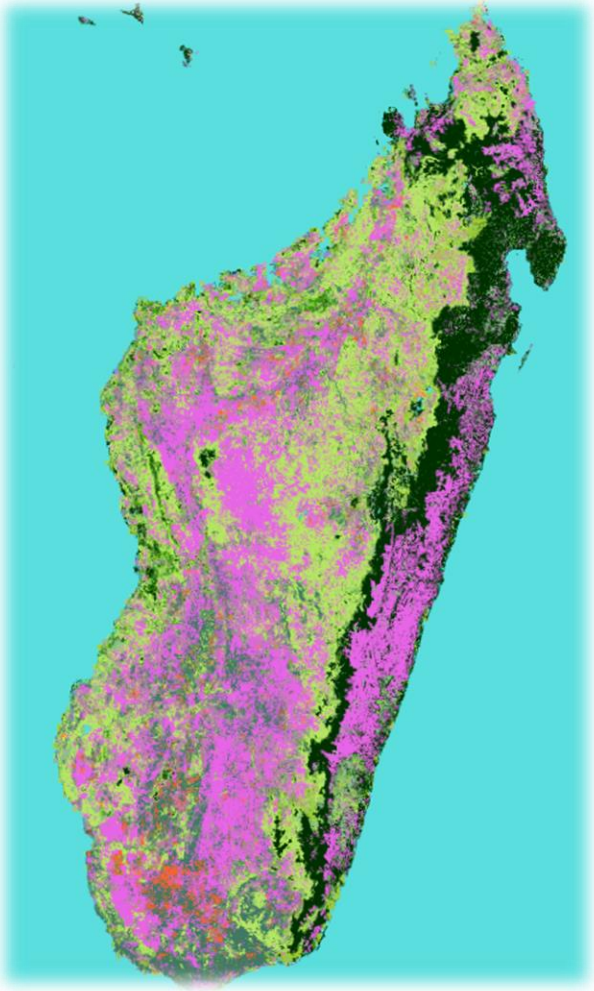
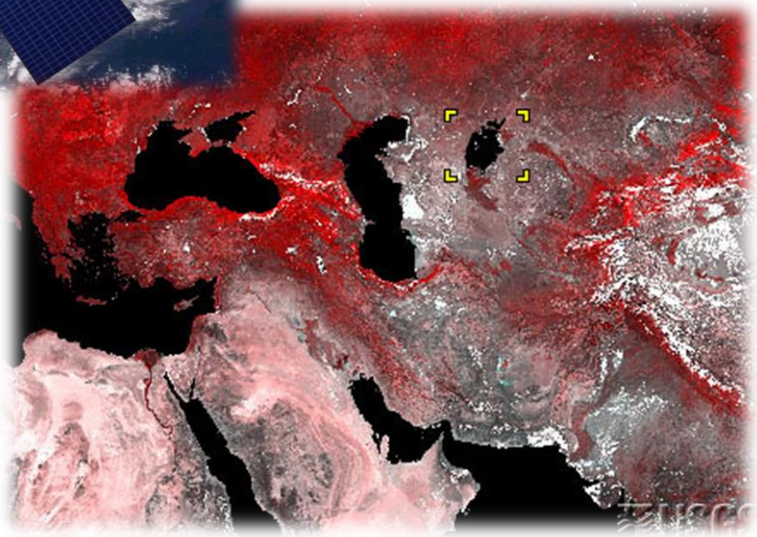


Image Data: Satellite data

- Landsat
- MODIS
- ASTER
- AVHRR
- VEGETATION
- SPOT
- RADARSAT



- GeoEye†
- DigitalGlobe †

† *private*

Image Data: Aerial photography

Aerial

- USGS Digital Ortho Quarter Quads (DOQQs)
  - Orthorectified aerial photos; B/W or color infrared
  - Collection began in 1965
  - http://gis.apfo.usda.gov/arcgis/services

- National Agricultural Imagery Program (NAIP)
  - 2003-2009 leaf-on images
  - 1m resolution; < 10% cloud cover
  - http://datagateway.nrcs.usda.gov/
  - http://gis.apfo.usda.gov/arcgis/services

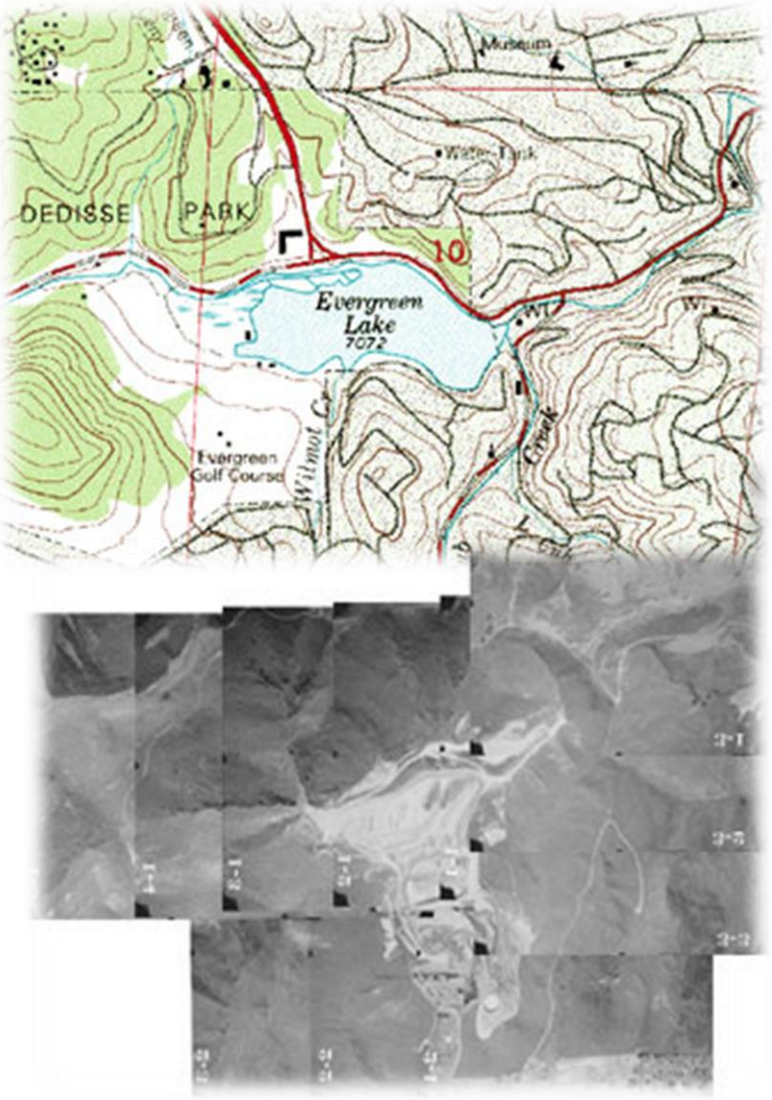
- Microsoft TerraServer
  - http://www.terraserver.com/

Image Data: Other stuff

Scanned maps

- USGS Digital Raster Graphics (DRGs)
  - Scanned USGS Quad sheets
  - http://topomaps.usgs.gov/drg/

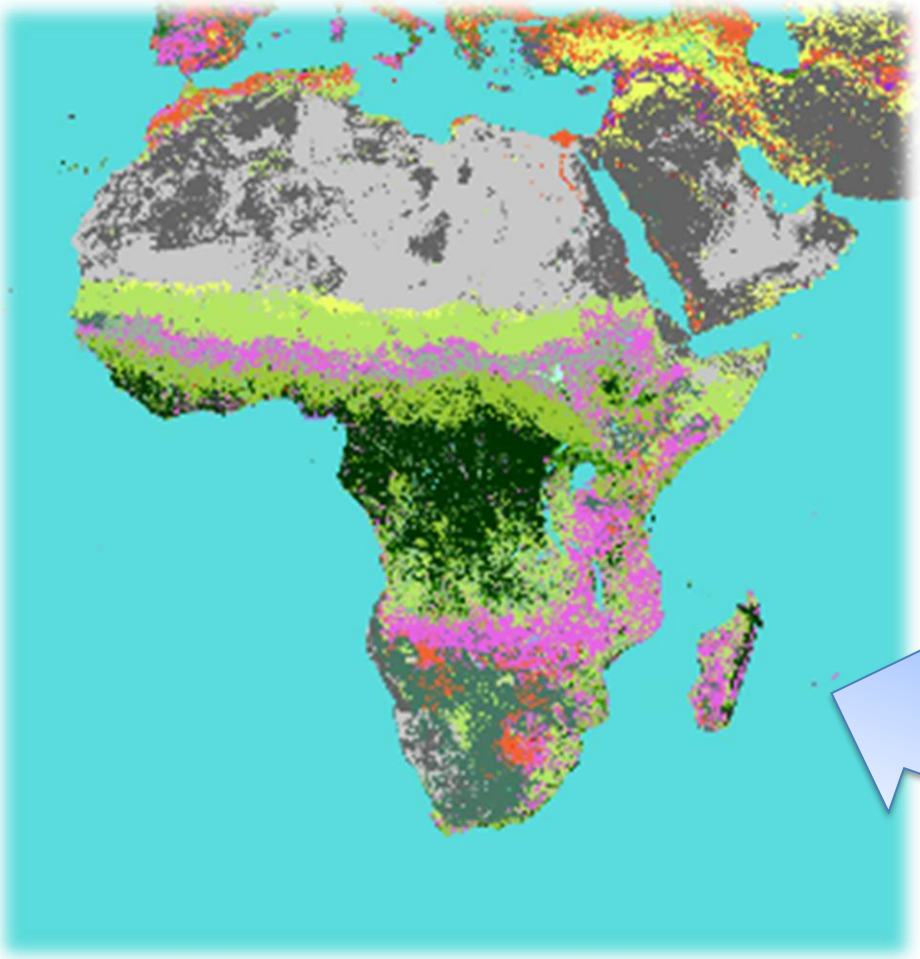
- Self-scanned maps/photos



Geo-Tagged Photos

- Flickr
- Picasa
- Instagram…

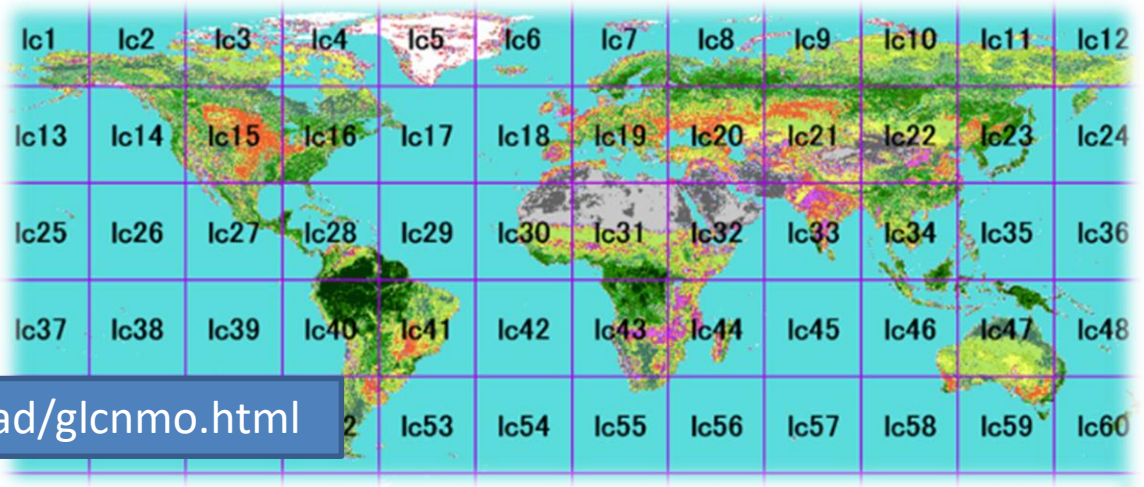
Land Use/Land Cover



- GLCNMO
- POSTEL/GlobCover
- NLCD*

GLCNMO

- ***G**lobal **L**and **C**over by **N**ational **M**apping **O**rganizations*
- Part of ***Global Map*** database
- 1km resolution (30 arc-seconds)
- Based on **MODIS** data
- 20 classes
- Time period: 2003



http://www.iscgm.org/gmd/download/glcnmo.html

LEGEND

- Broadleaf Evergreen Forest
- Broadleaf Deciduous Forest
- Needleleaf Evergreen Forest
- Needleleaf Deciduous Forest
- Mixed Forest
- Tree Open
- Shrub
- Herbaceous
- Herbaceous with Sparse Tree/Shrub
- Sparse Vegetation
- Cropland
- Paddy field
- Cropland/Other Vegetation Mosaic
- Mangrove
- Wetland
- Bare area, consolidated(gravel,rock)
- Bare area, unconsolidated(sand)
- Urban
- Snow/Ice
- Water Bodies

POSTEL/GlobCover

- European Space Agency/POSTEL
  http://toyo.mediasfrance.org/?-land-cover-68-
  http://dup.esrin.esa.int/globcover/

- Source: EnviSat MERIS data
  (European Space Agency)

- 300 m resolution

- Time period: 2005-06
  Now also 2009!

- Uses UN Land Cover classification
  - 22 classes
  - http://www.fao.org/docrep/003/x0596e/x0596e00.htm

POSTEL/GlobCover

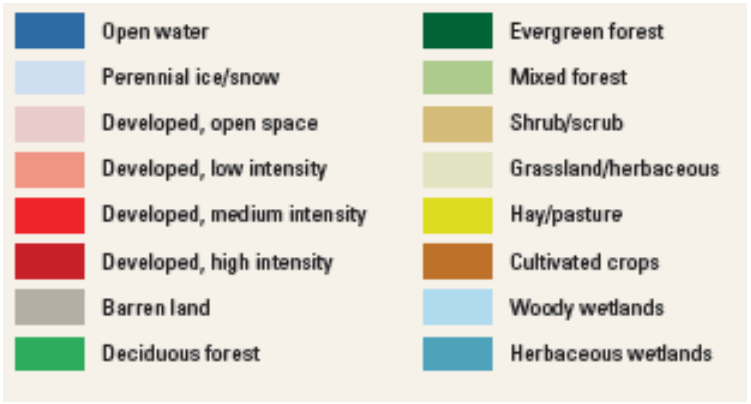
http://spatial-analyst.net/worldmaps/globcov.htm

- POST-FLOODING OR IRRIGATED CROPLANDS (OR AQUATIC)
- RAINFED CROPLANDS
- MOSAIC CROPLAND (50-70%) / VEGETATION (GRASSLAND/SHRUBLAND/FOREST) (20-50%)
- MOSAIC VEGETATION (GRASSLAND/SHRUBLAND/FOREST) (50-70%) / CROPLAND (20-50%)
- CLOSED TO OPEN (>15%) BROADLEAVED EVERGREEN OR SEMI-DECIDUOUS FOREST (>5M)
- CLOSED (>40%) BROADLEAVED DECIDUOUS FOREST (>5M)
- OPEN (15-40%) BROADLEAVED DECIDUOUS FOREST/WOODLAND (>5M)
- CLOSED (>40%) NEEDLELEAVED EVERGREEN FOREST (>5M)
- OPEN (15-40%) NEEDLELEAVED DECIDUOUS OR EVERGREEN FOREST (>5M)
- CLOSED TO OPEN (>15%) MIXED BROADLEAVED AND NEEDLELEAVED FOREST (>5M)
- MOSAIC FOREST OR SHRUBLAND (50-70%) / GRASSLAND (20-50%)
- MOSAIC GRASSLAND (50-70%) / FOREST OR SHRUBLAND (20-50%)
- CLOSED TO OPEN (>15%) (BROADLEAVED OR NEEDLELEAVED, EVERGREEN OR DECIDUOUS) SHRUBLAND (
- CLOSED TO OPEN (>15%) HERBACEOUS VEGETATION (GRASSLAND, SAVANNAS OR LICHENS/MOSSES)
- SPARSE (<15%) VEGETATION
- CLOSED TO OPEN (>15%) BROADLEAVED FOREST REGULARLY FLOODED (SEMI-PERMANENTLY OR TEMPORARILY) - FRESH OR BRACKISH WATER
- CLOSED (>40%) BROADLEAVED FOREST OR SHRUBLAND PERMANENTLY FLOODED - SALINE OR BRACKISH WATER
- CLOSED TO OPEN (>15%) GRASSLAND OR WOODY VEGETATION ON REGULARLY FLOODED OR WATERLOGGED SOIL - FRESH, BRACKISH OR SALINE WATER
- ARTIFICIAL SURFACES AND ASSOCIATED AREAS (URBAN AREAS >50%)
- BARE AREAS
- WATER BODIES
- PERMANENT SNOW AND ICE
- NO DATA (BURNT AREAS, CLOUDS,...)

National Land Cover Data (NLCD)

http://www.mrlc.gov/ - US only

- Source: Landsat
- Dates: 1976, 1992, 2001, 2006, 2011
- 30 m resolution

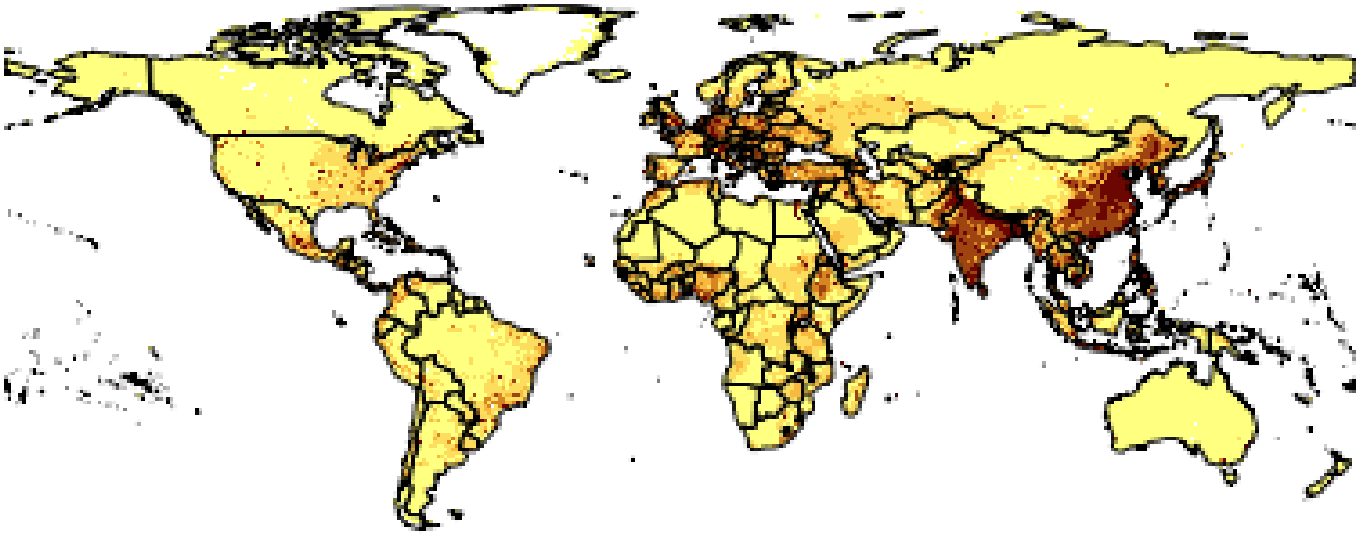


Factsheet:

http://pubs.usgs.gov/fs/2012/3020/fs2012-3020.pdf

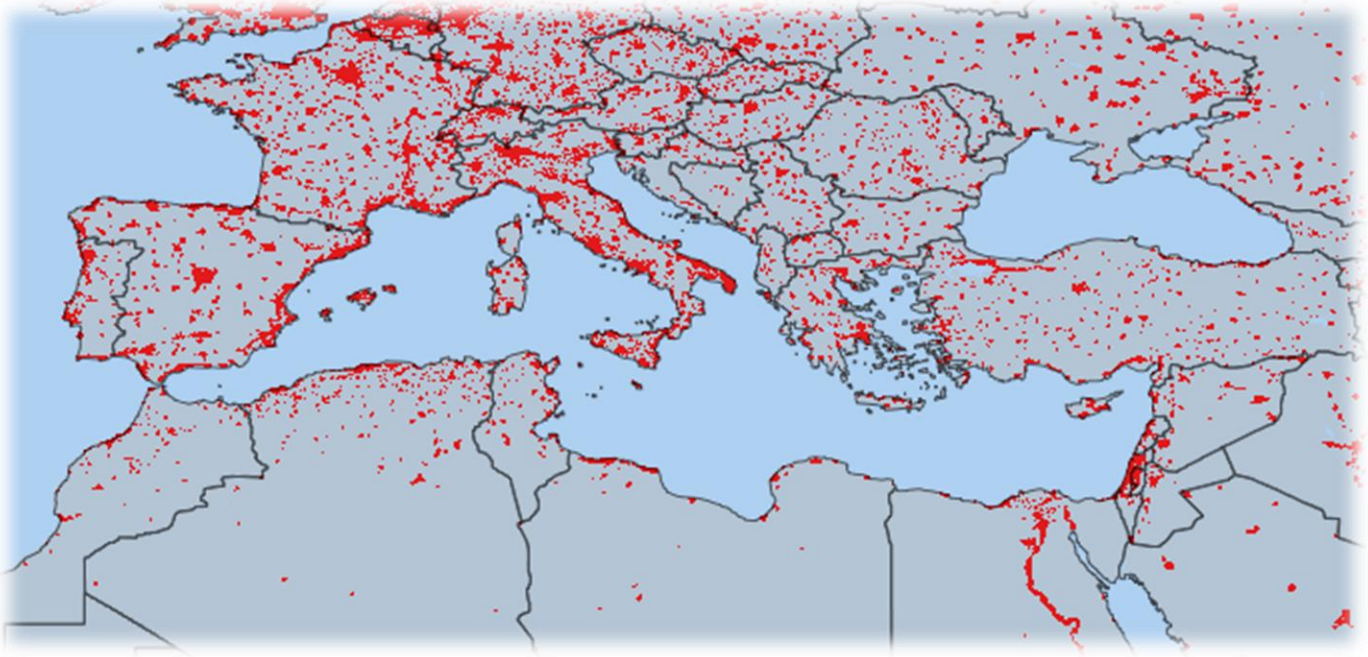
Demographics: GPW

- Gridded Population of the World, V.3 *(GPWv3)*
  http://sedac.ciesin.columbia.edu/data/collection/gpw-v3/sets/browse
  - 1990, 1995, 2000 estimates; 2005, 2010, 2015 projections
  - 2.5 arc-minute (~5 km) resolution
  - Population and Population Density products
  - Constructed from national or subnational input units (usually administrative units) of varying resolutions



Demographics: GRUMP

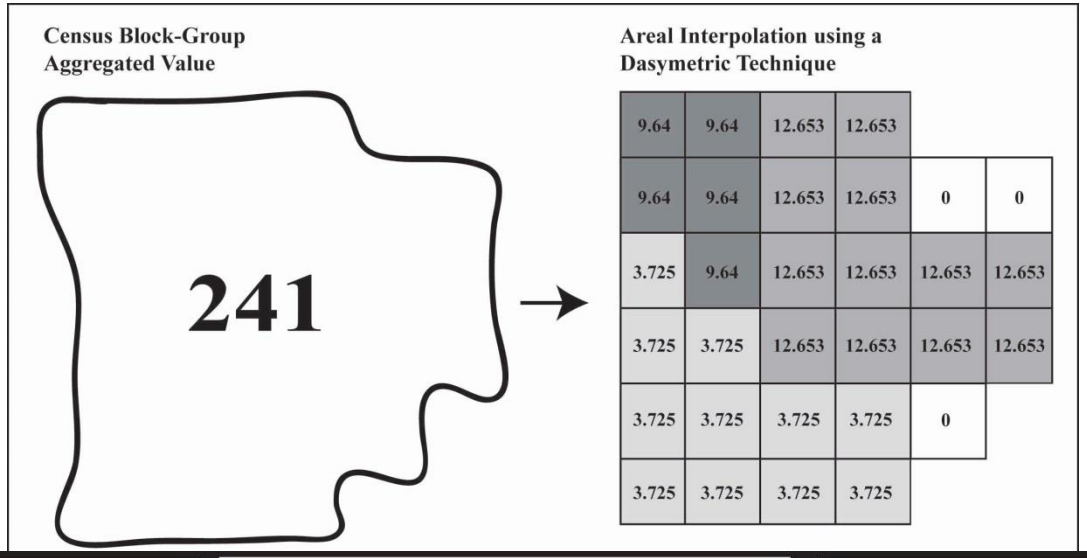
- Global Rural-Urban Mapping Project, V.1 (GRUMPv1)
  http://sedac.ciesin.columbia.edu/maps/gallery/collection/grump-v1
  - Dates: 1990, 1995, 2000
  - 30 arc-second (~1 km) resolution
  - Modeled from night-time lights and buffered settlement points

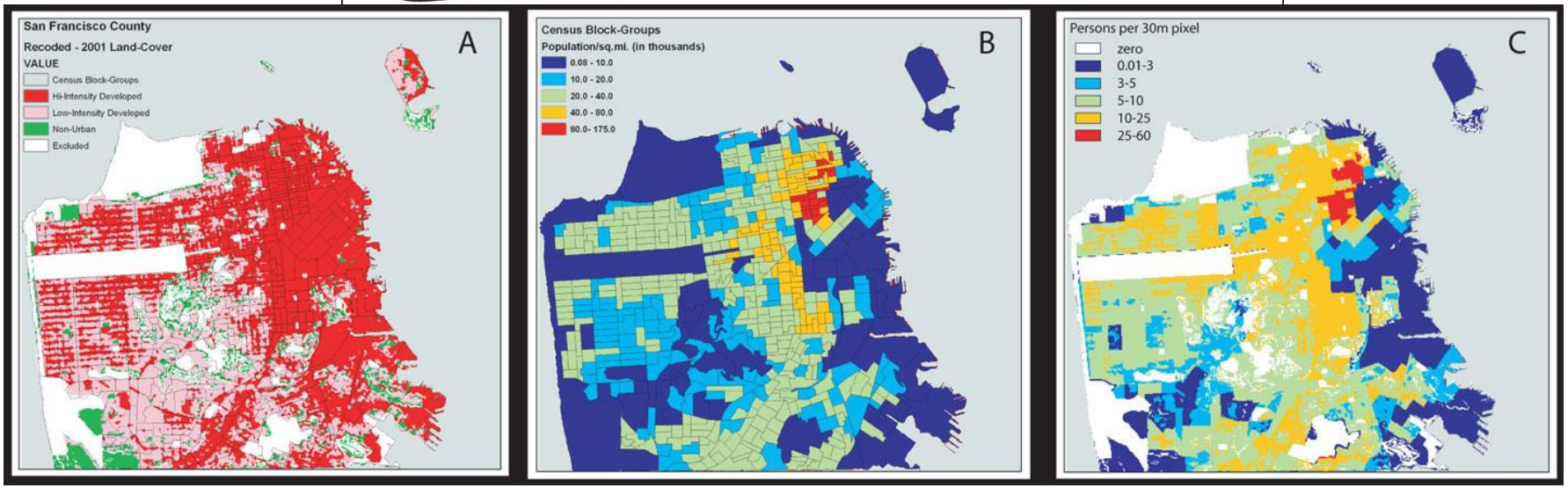


Demographics: LandScan

- LandScan (Oak Ridge National Laboratory)
  http://www.ornl.gov/sci/landscan/ *(Free, but permission required)*
  - Dates: 1990, 1995, 2000
  - 30 arc-seconds (~1km) resolution
  - Cell values are average ("ambient") population distribution; diurnal movements and collective travel habits into a single measure
  - "Smart interpolation" or "dasymetric mapping":
    Census + land cover + roads + slope + … → likely pop'n count

Dasymetric mapping





http://geography.wr.usgs.gov/science/dasymetric/methods.htm

Climate Data:

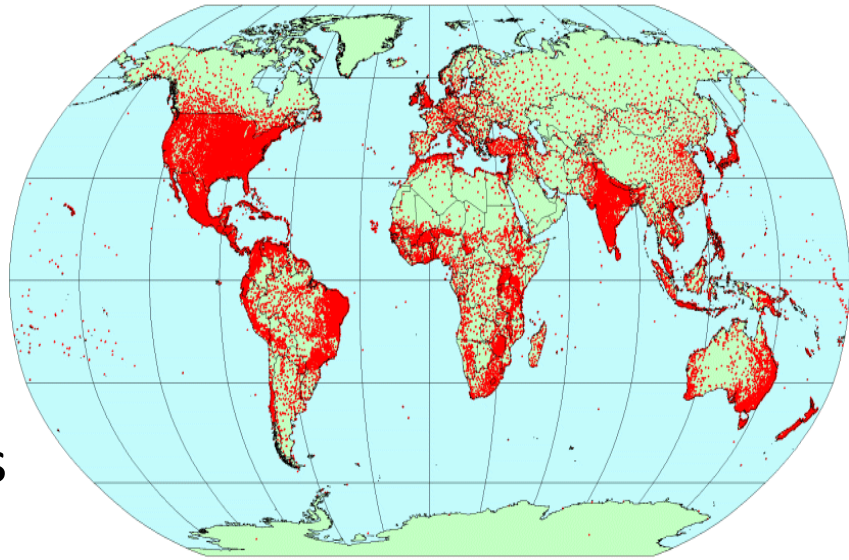
WorldClim - Global Climate Data

Free climate data for ecological modeling and GIS

http://www.worldclim.org

"Very high resolution interpolated climate surfaces"

- 47,554 weather station ; data from 1950-2000
- 30 arc-seconds (~1 km) resolution
  - Temperature:
    mean, min, and max monthly average
  - Precipitation:
    monthly total
  - Altitude
  - ***BIOCLIM*** Bioclimatic derivatives



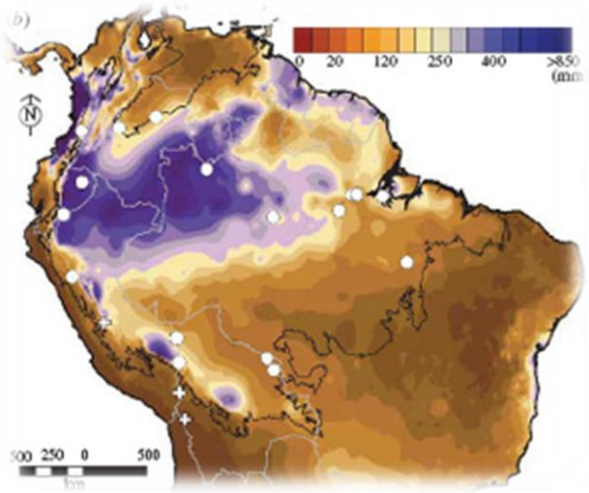
Climate Data:

WorldClim - Global Climate Data

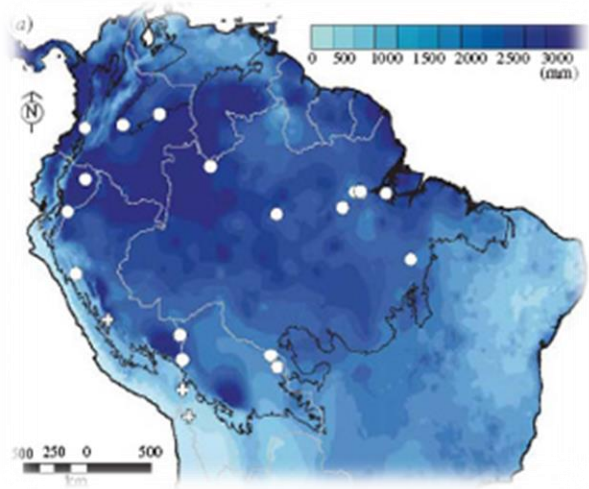
Free climate data for ecological modeling and GIS

BIOCLIM Bioclimatic derivatives

- BIO1 = Annual Mean Temperature
- BIO2 = Mean Diurnal Range (Mean of monthly (max temp - min temp))
- BIO3 = Isothermality (BIO2/BIO7) (* 100)
- BIO4 = Temperature Seasonality (standard deviation *100)
- BIO5 = Max Temperature of Warmest Month
- BIO6 = Min Temperature of Coldest Month
- BIO7 = Temperature Annual Range (BIO5-BIO6)
- BIO8 = Mean Temperature of Wettest Quarter
- BIO9 = Mean Temperature of Driest Quarter
- BIO10 = Mean Temperature of Warmest Quarter
- BIO11 = Mean Temperature of Coldest Quarter
- BIO12 = Annual Precipitation
- BIO13 = Precipitation of Wettest Month
- BIO14 = Precipitation of Driest Month
- BIO15 = Precipitation Seasonality (Coefficient of Variation)
- BIO16 = Precipitation of Wettest Quarter
- BIO17 = Precipitation of Driest Quarter
- BIO18 = Precipitation of Warmest Quarter
- BIO19 = Precipitation of Coldest Quarter



precipitation of the driest quarter



annual precipitation

Climate Data: CGIAR

Global Aridity and Potential Evapotranspiration (PET) Database

http://www.cgiar-csi.org/data/global-aridity-and-pet-database

Consultative Group on Int'l Agricultural Research Consortium

- Modeled from WorldClim data *(among other sources)*
- ~~5 x 5° cell resolution~~, downscaled to 30 arc-seconds
- Annual and monthly values

http://www.prima-klima-weltweit.de/dokumente/zomer.pdf

Climate: GIMMS

Global Inventory Modeling and Mapping Studies

http://glcf.umiacs.umd.edu/data/gimms/

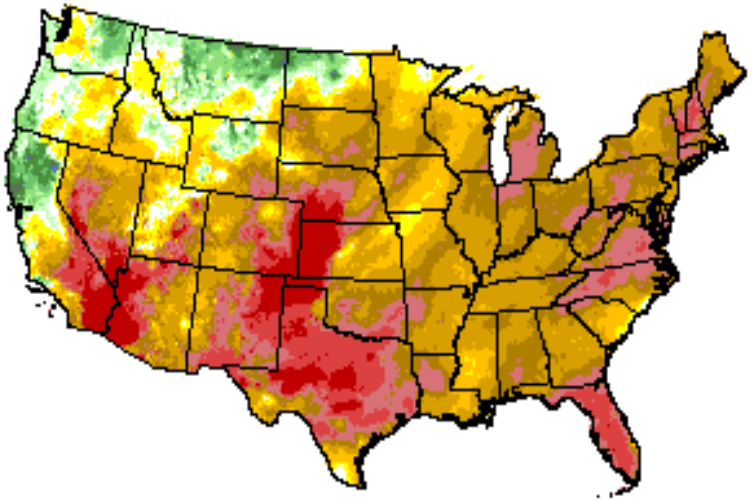
- Biophysical/vegetation change 22-year period
  - AVHRR NDVI from 1982-2004
  - Used to calculate changes in photosynthesis, CO_2 exchange, and energy flux between land and atmosphere…

Climate: PRISM

Parameter-Elevation Regressions on Independent Slopes Model

http://www.prism.oregonstate.edu/

- Monthly, yearly, and event-based climatic parameters
  - Precipitation and temperature
- Interpolated from point measurements & expert knowledge
- 30 arc-second/ 2.5 arc-minute; US only

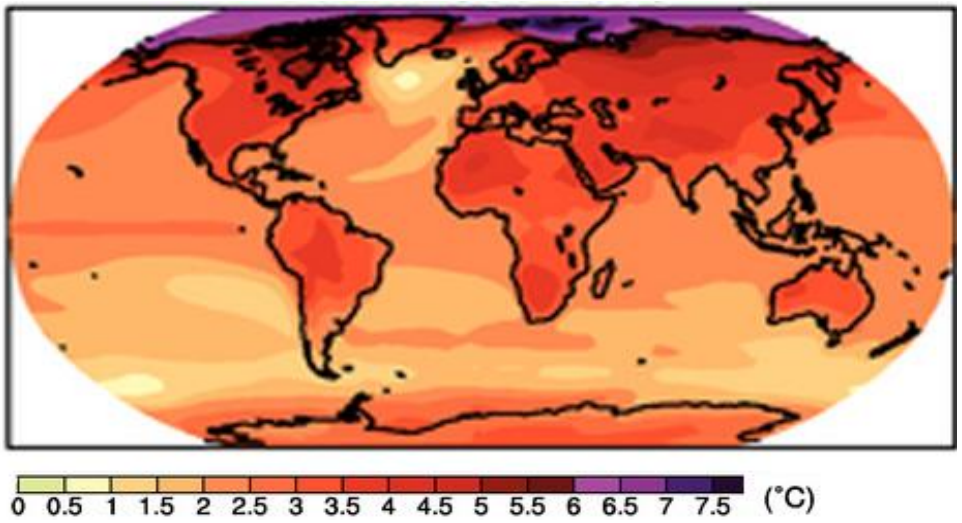


Climate: GCMs

Global climate & general circulation models

- Predictions of future climate scenarios (surface T°)
  http://regclim.coas.oregonstate.edu/dynamical-downscaling/overview-of-gcms/index.html

- Intergovernmental Panel on Climate Change (IPCC)
- National Center for Atmospheric Research (NCAR)



Later…

- An in-depth look at digital elevation data
  - Raster data considerations
- Data portals and clearinghouses
- Searching for specialized data