Data Sets for Biostatistics 624 Students

To assist Bio 624 students looking for data sources, we at the Center for a Livable Future (CLF) pulled together this list of ideas. These databases would enable you to conduct analyses relevant to diet, food production, food education, environment and public health, as well as other topics. If you do use one of these databases, no obligation, but we’d love it if you shared your papers with us at the end!

- Roni Neff, PhD, Research Director rneff@jhsph.edu

(Disclaimer: This is not a comprehensive collection of all public access databases on all issues CLF works on, it’s just some resources we’ve collected across time. Also, CLF does not vouch for the quality or user-friendliness of these data sources.)

Center for a Livable Future
Community Supported Agriculture (CSA) data analysis:
In a Community Supported Agriculture (CSA) model, consumers buy directly from regional farmers. Consumers pay up front for a share in the season’s harvest. This helps cover production costs and ensures a steady market, helping smaller farmers remain in business. The CLF and Maryland’s One Straw Farm initiated the first CSA project at JHMI in 2007, connecting faculty, students and staff to fresh, certified organic produce. Based out of JHSPH, the project expanded to over 150 participants in 2009. We have unanalyzed survey data from 2009 participants, as well as earlier year survey data and a small survey of those who did not continue with the program. Surveys inquire into member satisfaction, barriers to participation, changes to dietary behaviors and other measures. Contact: Brent Kim: bkim@jhsph.edu

Community food assessment dataset:
Based on a community food assessment conducted in southwest Baltimore in 2006, new data were collected in 2009 that has not been coded or entered yet. There are 53 surveys that need cleaning, coding, entering and analyzing. The 2006 codebook can be used and will require some modification based on 5 new questions added in 2009. The survey asks respondents about their food shopping habits and nutrition knowledge. If preferred, two students could work together to enter and organize the data, and then each could address separate research questions for their analyses. Contact: Anne Palmer ampalmer@jhsph.edu

Baltimore Area Food Map:
The Center for a Livable Future is developing a geographic tool with a number of different data layers regarding different sectors of the area food system, from stores including those that accept SNAP and WIC, to farms, to emergency food programs. Some Census data is included in the database. Contact: Amanda Behrens: abehrens@jhsph.edu

US Department of Agriculture Economic Research Service (USDA- ERS)
http://www.ers.usda.gov/Data/
Contains links to data on U.S.:
- Diet, Health & Safety/Food and Nutrition Assistance: food cost, food security, food availability, government assistance, foodborne illness, and consumption patterns
- Farm Economy/Food Sector: prices, trade, income, and commodity costs
Data sources from the Center for a Livable Future, http://www.jhsph.edu/clf

- Farm Practices & Management/Natural Resources & Environment: organic wholesale prices, adoption of genetically engineered crops, productivity, environmental quality, confined animal manure nutrients, and major land uses
- Policy Topics/Rural Economy: farm employment, rural/urban definitions, and agricultural research funding
- Within USDA, see also: http://www.fas.usda.gov/data.asp for data on agricultural imports and exports.

USDA-NASS Census of Agriculture 2007. The census is conducted every 5 years; data were released in February 2009. Much data collected at national/state/county levels. http://www.agcensus.usda.gov/.

EPA STORET Databases
http://www.epa.gov/storet/dbtop.html
- STORET’s query-able data reflects millions of surface and ground water quality samples from EPA across time.

CDC Outbreak Response and Surveillance Team
http://www.cdc.gov/outbreaknet/surveillance_data.html
Foodborne outbreak surveillance data

U.S. Department of Energy, Alternative Fuels and Advanced Vehicles Data Center
http://www.eere.energy.gov/afdc/data/index.html
- Downloadable excel files on vehicles, fuels, infrastructure, biomass resources, geographic, incentives and laws, clean cities, state and alternative fuel provider fleets, federal fleets.

FAOSTAT
http://faostat.fao.org/
- “FAOSTAT provides access to over 3 million time-series and cross sectional data relating to food and agriculture. FAOSTAT contains data for 200 countries and more than 200 primary products and inputs, just in its core. The national version of FAOSTAT, CountrySTAT, is being implemented in about 20 countries and three regions. It offers a two-way bridge amongst sub national, national, regional and international statistics on food and agriculture.”
- More specifically, you can find data by country and across time on topics including: food aid; food imports and exports around the world; crops, livestock and fish produced; food consumption; food security; prices, and more.
- “As part of its response to high food prices, FAO has developed an interactive database of staple food prices on national markets in 55 developing countries from Afghanistan to Zimbabwe. http://www.fao.org/giews/pricetool/

Climate Analysis Indicators Tool (CAIT)
http://cait.wri.org/links.php
- Contains links to international data on:
  - greenhouse gas emissions
  - environmental, social, economic, and development indicators
  - energy use projections and scenarios
Food and Water Watch
http://www.factoryfarmmap.org/
  ▪ Factory farm database/map

Center for Science in the Public Interest
  ▪ Food safety outbreak database http://www.cspinet.org/foodsafety/outbreak/pathogen.php

Also...Swivel.com has all kinds of interesting data sources -- use your own judgment as to source reliability/accuracy! http://www.swivel.com/