Social Science Data Library (SSDL)

What is SSDL?

- The Social Science Data Library is Temple University's repository for computerized social science data and a primary center for expertise in the analysis and presentation of such data.

What format do data come in?

- The resources below link to data in all formats. Sometimes original data files are formatted in ASCII, Excel, SAS, SPSS, or Stata. In other instances, organizations may only provide data in PDF reports. Finally, some organizations have their own, interactive online analysis platforms where the actual data cannot be accessed, but users can create tables, maps, and run other analyses.

National Data Sources

Name/Organization	1. **Centers for Disease Control and Prevention** a. **National Center for Health Statistics**
Website	http://www.cdc.gov/nchs/
Source Level	National
Data Format	ASCII + SPSS, SAS, Stata Statements, Codebooks
Description	The National Center for Health Statistics (NCHS) is a principal agency of the U.S. Federal Statistical System which provides statistical information to guide actions and policies to improve the health of the American people. NCHS is housed within the Centers for Disease Control and Prevention (CDC), which is part of the United States Department of Health and Human Services (HHS).

Topics and Health-Related Datasets	National Health and Nutrition Examination Survey (NHANES)
	National Health Care Surveys
	     National Hospital Care Survey (NHCS)
	       National Hospital Discharge Survey (NHDS)
	       National Hospital Ambulatory Medical Care Survey (NHAMCS)
	National Health Interview Survey (NHIS)
	National Immunization Survey (NIS)
	National Survey of Family Growth (NSFG)
	National Vital Statistics System (NVSS)
	Longitudinal Studies of Aging (LSOA)
	State and Local Area Integrated Telephone Survey (SLAITS)

Useful Resources	• FastStats: STATISTICS BY TOPIC - http://www.cdc.gov/nchs/fastats/ o FastStats provides quick access to statistics on topics of public health importance, including: diseases and conditions, injuries, life stages and populations, and health care and insurance.

Name/Organization	1. Centers for Disease Control and Prevention b. **CDC WONDER -- Wide-ranging Online Data for Epidemiologic Research**
Website	http://wonder.cdc.gov/
Source Level	National
Data Format	Online analysis platform
Description	CDC WONDER is an easy-to-use, menu-driven system that makes the information resources of the Centers for Disease Control and Prevention (CDC) available to public health professionals and the public at large. It provides access to a wide array of public health information. WONDER online databases utilize a rich ad-hoc query system for the analysis of public health data.

Topics and Health-Related Datasets	**WONDER Online Databases**
	
	▸AIDS Public Use Data
	▸Births
	▸Cancer Statistics
	Environment
	  ▸Daily Air Temperatures & Heat Index
	  ▸Daily Land Surface Temperatures
	  ▸Daily Fine Particulate Matter
	  ▸Daily Sunlight
	  ▸Daily Precipitation
	Mortality
	  Underlying Cause of Death
	    ▸Detailed Mortality
	    ▸Compressed Mortality
	  ▸Multiple cause of death (Detailed Mortality)
	  ▸Infant Deaths (Linked Birth/Infant Death Records)
	▸Online Tuberculosis Information System
	Population
	  ▸Bridged-Race Population (from NCHS)
	  ▸Population (from Census)
	▸Sexually Transmitted Disease Morbidity
	▸Vaccine Adverse Event Reporting

Name/Organization	1. **Centers for Disease Control** c. National Program of Cancer Registries (NPCR)
Website	http://www.cdc.gov/cancer/npcr/
Source Level	National
Data Format	Online analysis platform
Description	Since 1994, CDC's National Program of Cancer Registries (NPCR) has funded state cancer registries to collect population-based cancer incidence data under Public Law 102-515, the Cancer Registries Amendment Act. One of NPCR's goals is to release these data to public health planners and others to monitor the burden of disease and to implement cancer prevention and control programs. Starting in 2001, NPCR began receiving data annually from funded programs with the goals of establishing the quality of the data and eventually releasing the data for use in public health planning.
Topics and Health-Related Datasets	**United States Cancer Statistics** The *United States Cancer Statistics (USCS): Incidence and Mortality* Web-based report contains official federal government cancer statistics for new cancer cases and deaths. For background information about USCS, visit About *United States Cancer Statistics.* **Interactive Cancer Atlas** CDC's Interactive Cancer Atlas (InCA) uses USCS data to create United States maps that allow you to make quick comparisons including incidence or mortality rate or count, cancer site, gender, race/ethnicity, and year. For more information about InCA, visit About Interactive Cancer Atlas. **CDC WONDER** CDC WONDER is an online query system that produces age-adjusted and crude rates in tabular, map, and chart formats. Variables include year of diagnosis, state, region or division of the United States, sex, race, ethnicity, age, primary site, and childhood cancer. Data from selected registries, covering between 86% and 96% of the United States population depending on the year of diagnosis, may be accessed through CDC WONDER. **State Cancer Facts** The State Cancer Facts application provides state-specific data on lung, colorectal, female breast, and prostate cancers, including the estimated number of new cancer cases, cancer deaths, and the age-adjusted mortality rates for cancer deaths by race for each state. **U.S. County Cancer Incidence Dataset** This dataset consists of aggregate cancer incidence rates and counts for major cancer sites for selected counties in the United States. The purpose of this dataset is to provide aggregated county-level data for cancer control planning, policymaking, and monitoring. Examples of current users include state cancer control planners, state legislators and policymakers, and the American Cancer Society. **National Center for Health Statistics Research Data Center** CDC's National Center for Health Statistics (NCHS) developed the Research Data Center (RDC) to allow researchers access to restricted data. In addition to providing access to NCHS data, the RDC also hosts restricted data from NPCR-CSS as well as a variety of other groups within the Department of Health and Human Services.

Social Science Data Library (SSDL)

Name/Organization	2. Agency for Healthcare Research and Quality (AHRQ)	
Website	http://www.ahrq.gov/	
Source Level	National	
Data Format	ASCII + SPSS, SAS, Stata Statements, Codebooks	
Description	The Agency for Healthcare Research and Quality's (AHRQ) mission is to produce evidence to make health care safer, higher quality, more accessible, equitable, and affordable, and to work within the U.S. Department of Health and Human Services and with other partners to make sure that the evidence is understood and used. AHRQ sponsors and conducts research that provides evidence-based information on health care outcomes; quality; and cost, use, and access. The information helps health care decision-makers—patients and clinicians, health system leaders, purchasers, and policymakers—make more informed decisions and improve the quality of health care services.	
Topics and Health-Related Datasets	http://www.ahrq.gov/research/data/dataresources/index.html	
	If You Need Data On…	**Use This Data Source**
	Accessibility of care	Medical Expenditure Panel Survey; MEPSnet Query Tools
	Ambulatory surgeries	Healthcare Cost and Utilization Project; HCUPnet
	Emergency department visits	Healthcare Cost and Utilization Project National Data State Data; HCUPnet; HCUP Statistical Brief
	Health care disparities	National Healthcare Disparities Report
	Health care quality	National Healthcare Quality Report State Snapshots
	Health care spending	Medical Expenditure Panel Survey; MEPSnet Query Tools
	Health care use	Medical Expenditure Panel Survey; MEPSnet Query Tools
	Health information data element definitions, values and information models	United States Health Information Knowledgebase
	Hospitalizations	Healthcare Cost and Utilization Project; HCUPnet
	By payer	Healthcare Cost and Utilization Project; HCUPnet
	By States	Healthcare Cost and Utilization Project; HCUPnet
	For children	Healthcare Cost and Utilization Project; HCUPnet
	State-specific health care quality information (strengths, weaknesses, and opportunities for improvement)	State Snapshots
	Study data extracted for systematic reviews	Systematic Review Data Repository™

Name/Organization	2. Agency for Healthcare Research and Quality (AHRQ) a. Medical Expenditure Panel Survey (MEPS)
Website	http://meps.ahrq.gov/mepsweb/
Source Level	National
Data Format	ASCII + SPSS, SAS, Stata Statements, Codebooks
Description	The Medical Expenditure Panel Survey (MEPS) is a set of large-scale surveys of families and individuals, their medical providers, and employers across the United States. MEPS is the most complete source of data on the cost and use of health care and health insurance coverage
Topics and Health-Related Datasets	Access to Health Care Children's Health Children's Insurance Coverage Elderly Health Care Health Care Costs/Expenditures Health Care Disparities Health Insurance Medical Conditions Medicare/Medicaid/SCHIP Men's Health Mental Health Obesity Prescription Drugs Projected Data/Expenditures Quality of Health Care State and Metro Area Estimates The Uninsured Women's Health

Name/Organization	2. Agency for Healthcare Research and Quality (AHRQ) b. Healthcare Cost and Utilization Project (HCUP)
Website	http://www.ahrq.gov/research/data/hcup/index.html
Source Level	National, State
Data Format	ASCII + SPSS, SAS, Stata Statements, Codebooks Request data, Data Use Agreement
Description	The Healthcare Cost and Utilization Project (HCUP, pronounced "H-Cup") is the largest collection of nationwide and State-specific longitudinal hospital care data in the United States. It is a family of databases and related software tools and products developed through a Federal-State-Industry partnership and sponsored by AHRQ. These databases enable research on a broad range of health policy issues, including cost and quality of health services, medical practice patterns, access to health care programs, and outcomes of treatments at the national, State, and local market levels.
Topics and Health-Related Datasets	National (Nationwide) Inpatient Sample (NIS) The National (Nationwide) Inpatient Sample (NIS) is the largest publicly available all-payer hospital inpatient care database in the United States. Researchers and policymakers use NIS data to identify, track, and analyze trends in health care utilization, access, charges, quality, and outcomes.

	Kids' Inpatient Database (KID) The Kids' Inpatient Database (KID) is composed of hospital inpatient stays for children and is specifically designed to allow researchers to study a broad range of conditions and procedures related to children's health. Researchers and policymakers can use KID data to identify, track, and analyze health care utilization, access, charges, quality, and outcomes. Nationwide Emergency Department Sample (NEDS) The Nationwide Emergency Department Sample (NEDS) captures information on emergency department (ED) visits that do not result in an admission as well as ED visits that result in an admission to the same hospital. The NEDS enables analyses of ED utilization patterns and supports public health professionals, administrators, policymakers, and clinicians in their decisionmaking regarding this critical source of care. State Inpatient Databases (SID) The State Inpatient Databases (SID) are a set of hospital databases containing the universe of the inpatient discharge abstracts from participating States, translated into a uniform format to facilitate multi-State comparisons and analyses. Researchers and policymakers use the SID to investigate questions and identify trends unique to one State, to compare data from two or more States, and to conduct market area research or small area variation analyses. State Ambulatory Surgery and Services Database (SASD) The State Ambulatory Surgery and Services Database (SASD) are a set of databases that capture surgeries performed on the same day in which patients are admitted and released. Researchers and policymakers use the SASD to compare inpatient surgery data with ambulatory surgery data; conduct market area research or small area variation analyses; and identify State-specific trends in ambulatory surgery utilization, access, charges, and outcomes. State Emergency Department Databases (SEDD) The State Emergency Department Databases (SEDD) are a set of databases that capture discharge information on all emergency department visits that do not result in an admission. The SEDD combined with SID discharges that originate in the emergency department are well suited for research and policy questions that require complete enumeration of hospital-based emergency departments within market areas or States.

Name/Organization	3. Centers for Medicare and Medicaid Systems (CMS)
Website	http://www.cms.gov/
Source Level	National
Data Format	Access Request Form
Topics and Health-Related Datasets	Medicare Medicaid/CHIP Medicare-Medicaid Coordination Private Insurance Innovation Center Regulations and Guidance Research, Statistics & Data Systems National Health Expenditure Data

	Historical Projected Age and Gender State (Residence) State (Provider) NHEA Related Studies NHE Fact Sheet
	Outreach & Education

Name/Organization	4. United States Census Bureau - Health
Website	http://www.census.gov/topics/health.html
Source Level	National, State, Local
Data Format	Reports, Detailed Tables (Data); Online analysis platform (Data Base)
Description	Health statistics are increasingly important in measuring the nation's overall well-being. The Census Bureau provides accurate, detailed and up-to-date statistics—covering people and business—relating to health in America.
Health-Related Datasets	**Health Insurance Data -** These data have information on types of health coverage by income, personal characteristics, and health expenditures. **Disability Employment Data -** These tables provide over 1 billion estimates on labor force characteristics of people with disabilities with detail on occupations, education, and earning **Health Care Industries (NAICS 62) -** The Industry Statistics Portal provides access to data sources that produce economic statistics for a user-selected industry, such as Health Care Industries. **Disability Data -** Statistics about types of disabilities and the characteristics of the population of people with disabilities. **Fertility Data -** Data are available on characteristics of women age 15-50 with a recent birth, along with historical data. **Healthcare and Social Assistance Services Report -** The Service Annual Survey produces economic statistics for selected service industries including the Health Care and Social Assistance sector. **HIV/AIDS Surveillance Data Base -** Access data from various sources on HIV/AIDS prevalence and incidence for countries in Africa, Asia, Europe, Latin America and the Caribbean, and Oceania. **Longitudinal Employer-Household Dynamics -** The new LED Extraction Tool allows streamlined access to the entire set of raw Quarterly Workforce Indicators (QWI) data. **Small Area Health Insurance Estimates (SAHIE) -** The SAHIE program annually publishes the most current single year health insurance coverage estimates for selected characteristics for all counties. **Service Annual Survey -** The Service Annual Survey produces economic statistics for selected service industries including the Health Care and Social Assistance sector. **Quarterly Services Survey -** The Quarterly Services Survey produces economic statistics for selected service industries including the Health Care and Social Assistance sector. **Economic Census -** The Economic Census paints a detailed portrait of the

	economy every five years. It provides information by industry down to the local area level.

Useful Resources	• American Factfinder – http://factfinder2.census.gov/faces/nav/jsf/pages/index.xhtml
	     o Facts are organizable by geographic units
	• Current Population Survey (CPS) - http://www.census.gov/cps/
	     o Useful information about uninsurance rates
	• DataFerrett - http://dataferrett.census.gov/
	     o Data analysis and extraction tool to customize federal, state, and local data to suit your requirements. Using DataFerrett, you can develop an unlimited array of customized spreadsheets that are as versatile and complex as your usage demands then turn those spreadsheets into graphs and maps without any additional software.

Name/Organization	5. Substance Abuse and Mental Health Data Archive (SAMHDA)
Website	www.datafiles.samhsa.gov
Source Level	National
Data Format	Data Files in SAS, SPSS, Stata, R, ASCII, Delimited
Description	SAMHDA is an initiative funded under contract (#283-10-0324) with the Center for Behavioral Health Statistics and Quality (CBHSQ), Substance Abuse and Mental Health Services Administration (SAMHSA), U.S. Department of Health and Human Services (HHS). CBHSQ has primary responsibility for the collection, analysis, and dissemination of SAMHSA's behavioral health data. CBHSQ promotes the access and use of the nation's substance abuse and mental health data through the Substance Abuse and Mental Health Data Archive (SAMHDA). SAMHDA provides public-use data files, file documentation, and access to restricted-use data files to support a better understanding of this critical area of public health.
Topics and Health-Related Datasets	Drug Abuse Warning Network (DAWN) Series (8 studies) National Survey of Substance Abuse Treatment Services (N-SSATS) Series (15 studies) National Survey on Drug Use and Health (NSDUH) Series (31 studies) National Youth Survey (NYS) Series (7 studies) Treatment Episode Data Set - Admissions (TEDS-A) Series (22 studies) Treatment Episode Data Set - Discharges (TEDS-D) Series (7 studies)

Name/Organization	6. County Health Rankings
Website	http://www.countyhealthrankings.org/
Source Level	National
Data Format	Data File in Excel, Summary Reports, Maps
Description	The County Health Rankings & Roadmaps program is a collaboration between the Robert Wood Johnson Foundation and the University of Wisconsin Population Health Institute. The annual County Health Rankings measure vital health factors, including high school graduation rates, obesity, smoking, unemployment, access to healthy foods, the quality of air and water, income, and teen births in nearly every county in America. The annual Rankings provide a revealing snapshot of

	how health is influenced by where we live, learn, work and play.
Topics and Health-Related Datasets	Health Outcomes (Quality of Life, Length of Life) Health Factors (Physical Environment, Social and Economic Factors, Clinical Care, Health Behaviors) Policies and Programs
Useful Resources	The data: http://www.countyhealthrankings.org/rankings/data

Name/Organization	7. Community Commons
Website	http://www.communitycommons.org/maps-data/
Data Format	Online analysis platform, Reports, Mapping platform
Source Level	National, Local
Description	Utilizes mapping to determine areas of health disparity and identify vulnerable populations.
Topics and Health-Related Datasets	Health Map Gallery Uninsured Population Diabetes Prevalence Mammogram Screenings

Name/Organization	8. LawAtlas
Website	http://www.lawatlas.org/welcome#.VBhT3vldV8F
Source Level	National, State
Data Format	Data in Excel, Codebook, Protocol, Reports
Description	In our federal system, states have broad authority to set policy through legislation and regulation. These policy differences are important to health. Understanding how state policies vary across states and how they develop over time is just as important to promoting population health as understanding the distribution and spread of disease or health risks. Knowing which states have specific policies is required for building an evidence base about the relative effects of different policy choices, and for identifying when that evidence base is not being translated back into policy. LawAtlas[sm] is a vital tool for exploring variation in policies over time and across states.
Topics and Health-Related Datasets	Alcohol, Tobacco and Other Drugs Lesbian, Gay, Bisexeual and Transgender Health Oral Health Social Determinants of Health and Health Disparities Chronic Disease Health Communication and Information Technology Injury and Violence Prevention Environmental Health Infectious Disease Prevention and Control Food Safety Health Services Maternal, Infant and Child Health Public Health Infrastructure

Pennsylvania Data Sources (State)

Name/Organization	1. Pennsylvania Health Care Cost Containment Council (PA HC4)
Website	http://www.phc4.org/
Source Level	State
Data Format	Submit a Data Request; Reports
Description	An independent state agency responsible for addressing the problem of escalating health costs, ensuring the quality of health care, and increasing access for all citizens regardless of ability to pay.
Topics and Health-Related Datasets	Hospital Performance Data http://www.phc4.org/idb/HPR/ Coronary Bypass Data: http://www.phc4.org/idb/Cabg/default.cfm Hospital Financial Data: http://www.phc4submit.org/phc4/financial/impatient.aspx

Name/Organization	2. Pennsylvania Department of Health - Statistical Registries
Website	http://www.portal.state.pa.us/portal/server.pt?open=514&objID=607222&mode=2
Source Level	State
Data Format	PDF Reports
Description	The Division of Statistical Registries collects data that can be analyzed to help solve our state's public health problems. Registries include: Vital Events (birth, death, fetal death), Cancer and Immunization. Data comes from birth, death and fetal death records, all cases of treated or diagnosed cancer and immunization records.
Topics and Health-Related Datasets	Birth, Death and Other Vital Statistics Cancer Healthy People Behavioral Risk and Injury Statistics Minority Health and Health Equity Health Facilities Communicable Diseases School Health Other Health Statistics EpiQMS

Name/Organization	3. Pennsylvania Department of Public Welfare
Website	http://www.dpw.state.pa.us/provider/healthcaremedicalassistance/managedcareinformation/
Source Level	State
Description	This section of the DPW website contains information for partners and providers on managed health care.
Data Format	PDF Reports
Topics and Health-Related Datasets	Physical Health Databooks Managed Care Information

Philadelphia Data Sources (Local)

Name/Organization	1. ***Philadelphia Health Management Corporation's (PHMC) Community Health Data Base (CHDB)
Website	Information: http://www.phmc.org/site/27-research/chdb ***Data Accessed through BlackBoard – must be added as student to the course
Source Level	Local – Philadelphia Region (5 counties)
Data Format	Data in SAS, SPSS, Stata, Excel Files, Codebooks
Description	The survey includes information about the health status, personal health behaviors, access to and utilization of area health services, and other issues for some 13,000 residents in the five county Pennsylvania side of the Philadelphia Metropolitan Area. Geographically-related data are available at the census tract, zip code, county, and regional level. SSDL has the data for 1998, 2000, and 2002. Special samples are available for children and the elderly.

Name/Organization	2. ***Philadelphia Educational Longitudinal Study (PELS)
Website	Information: http://www.philaedfund.org/research/research-studies/philadelphia-educational-longitudinal-study-pels ***Data Accessed through BlackBoard – must be added as student to the course
Source Level	Philadelphia
Data Format	Data in SPSS, Stata Files, Codebook, Data Dictionary
Description	The Philadelphia Educational Longitudinal Study (PELS) is a comprehensive database following a cohort of 2000 students in the School District of Philadelphia from the close of their 8th grade year in 1996 to three years past on-time graduation, 2003.
Topics and Health-Related Datasets	Crime Data Birth Data Census Data Parents Data Student Data

Name/Organization	3. Philadelphia Department of Public Health
Website	http://www.phila.gov/health/commissioner/DataResearch.html
Source Level	Local
Data Format	Data in CSV Files, Reports
Description	The Philadelphia Department of Public Health collects, analyzes, and reports on a wide variety of public health data, including data related to births and deaths, HIV, STDs, immunizations, children's health, smoking, obesity, air quality, and food safety. These data inform the Department's programs and policies. In addition, they enable the public and our partners to understand the health challenges faced by Philadelphians.
Topics and Health-Related Datasets	Births and Deaths HIV, STD and Communicable Diseases Smoking, Obesity, and Related Chronic Diseases Environmental Health
Useful Resources	**Data Requests**

Other Data Archives

Name/Organization	Inter-university Consortium for Political and Social Research (ICPSR)
Website	http://www.icpsr.umich.edu/icpsrweb/landing.jsp http://www.icpsr.umich.edu/icpsrweb/ICPSR/studies?classification=ICPSR.IX.
Source Level	International, National, State
Description	ICPSR advances and expands social and behavioral research, acting as a global leader in data stewardship and providing rich data resources and responsive educational opportunities for present and future generations. An international consortium of more than 700 academic institutions and research organizations, ICPSR provides leadership and training in data access, curation, and methods of analysis for the social science research community. ICPSR maintains a data archive of more than 500,000 files of research in the social sciences. It hosts 16 specialized collections of data in education, aging, criminal justice, substance abuse, terrorism, and other fields.
Topics and Health-Related Datasets	ICPSR houses thousands of datasets related to "Health Care and Facilities" including: • Health care services • Illness • Health care • Health • Medical care
Useful Resources	• good for all kinds of data • searchable • data available in multiple formats (SAS, SPSS, STATA, R, etc) • downloadable from computers with Temple IP addresses

Contact Information

Nina Hoe, PhD
Postdoctoral Fellow
Institute for Survey Research
1509 Cecil B. Moore Avenue
Suite 401 – Entertainment and Community Education Center (ECEC)
Philadelphia, PA 19121
215-204-4441 (phone)
nina@temple.edu