

**Data Management and Data Analysis
By**

**Sisira T. De Silva
B00469552
sisira@dal.ca**

**Performed at
Population Health Research Unit,
Dalhousie University,
5790 University Avenue,
Halifax, N.S. B3H 1V7
Phone:(902) 494-6456
Fax: (902) 494-1597**

**In partial fulfillment of the requirement of the Master of
Health Informatics Program, Dalhousie University**

**Report of the Internship for the period May 21, 2007 -
August 17, 2007**

Date Submitted: August 24, 2007

Acknowledgement and Endorsement:

First I want to express my deep gratitude to Mr. Mark Smith, Director of Population Health Research Unit (PHRU), for giving me this opportunity to work at Population Health Research Unit. This was a very unique experience and wonderful first step forward a career goal working in the health information field.

I want to deeply thank Ms. Yan Wang, Programmer /Data Manager, for remarkable supervision of my training. Her kindness and her availability have been a great help during the entire internship at Population Health Research Unit.

I specially would like to thank Ms. Deirdre Harvey, administrative assistance for Medical Informatics Program, and Ms. Sandra Pauls, financial administrator for Population Health Research Unit, for their help and support in all the various administrative processes prior and during my internship.

I would like to thank Dr. Grace Paterson, Clinical Informatician, for providing the information/ contacts for the internship appointments.

Finally, I want to thank and my friendliness towards all the staff from Population Health Research Unit for providing me with a really friendly and enjoyable working environment during all four months. This has been a really pleasant time working with the PHRU staff.

I couldn't complete the acknowledgments without thanking Dr. Raza Abidi, Director of Health Informatics and Dr. David Zitner, Director of Medical Informatics, at Dalhousie University, for the quality of training they provide for students during the Health Informatics program.

This report has been written by the author and has not received any previous academic credits at this or any other institution.

Due to the confidential nature of the "SAS" files I was working with, it is impossible to include them in this report.

Signed by Sisira T. De Silva

Sisira T. De Silva

Executive Summary:

Health Informatics focuses on the application of computer information systems to health care and public health. A major focus of it is the support of information systems for reasoning, decision-making, and learning to support health care delivery, education, and management. The Health Informatics internship is an opportunity for the student to acquire on-the-job training at a health organization.

The author fulfilled the internship requirement at the Population Health Research Unit (PHRU) located at Dalhousie University. Earlier course work laid the theoretical foundation for success in the health informatics field but the internship provided as setting for understanding how the theory can be applied in practice. Reinforcing the concept that health information must be constantly updated in order to maintain its relevance, the author acquired the skills necessary to perform this task through revising the databases. In order to do this the author had to be familiar with data extraction, statistical analysis, data interpenetration, and how to use relevant software applications.

Not only what was the author able to transfer concepts learned at school to the workplace setting he will be able to transfer skills gained through his internship to his future academic pursuits.

After completing the internship the author realizes that to become more confident he must familiarize himself with variety of computer operating systems and large databases.

Table of Contents

Acknowledgement and Endorsement:	1
Executive Summary:	2
Table of Contents	3
Introduction:	4
Description of the Organization:	5
Data access procedures and policies:	6
Objectives of Population Health Research Unit:	7
Description of the work and how it relates to Health Informatics:	8
Role and Responsibilities:	8
Other activities during internship:	19
Discussion of the problem that the author have analyzed and the corresponding solution:	20
Conclusions:	21
Recommendations:	22
Reference:	23
Appendices:	24
Appendix A: Web Page for Research Data Repository	24
Appendix B: Web page for Physician Billing Database	25
Appendix C: Web page for CIHI Hospital Discharge Abstract Database	26
Appendix D: Web page for Mental Health Outpatient information System Database	27
Appendix E: Web page for Senior’s Pharmacare Prescription Database	28
Appendix F: Web page for Community Services Pharmacare Prescriptions Database	29
Appendix G: Web page for Insured Patient Registry Database	30
Appendix H: Web page for Postal Code Geography Database	31
Appendix I: Web page for Licensed Provider Registry Database	32
Appendix J: Web page for Vital Statistics Deaths Database	33
Appendix K: Web page for Canada Censes Survey Database	34
Appendix L: “What’s New?” web page of Population Health Unit	35
Appendix M: Website for National Diabetes Surveillance System	36
Appendix N: VMS software interface	37
Appendix O: UltraEdit text editor	38

Introduction:

Health Informatics is an integrated research and learning program with health as its focus and information technology as the enabler [1]. It teaches students the ability to use health information related technology and to facilitate better communication between health providers and IT professionals to improve the outcome and the delivery of healthcare. Students are provided with an understanding of the health care system plus a conceptual understanding of the science and the methodology for health informatics [2].

The Population Health Research Unit (PHRU) is a Dalhousie University-based research and support group conducting systematic research into population health, health services and their inter-relationships. Through an internship at PHRU the author planned to learn about and gain practical experience in health informatics.

Objectives of the internship at PHRU:

1. To identify and understand the informatics concepts driving the host institution (PHRU).
2. To describe the challenges associated with development, maintenance, and evaluation of health information systems.
3. To recognize, discern and appraise the utility of products and services offered by the host institution (PHRU).
4. To ascertain and learn practical skills.
5. To use statistical software programs, and interfaces.
6. To have an opportunity to apply theories, ideas, principles, and skills learned.
7. To understand the complex relationships within health-care organizations.
8. To maintain and update of Databases (Research Data Repositories) of host organization (PHRU).
9. To maintain existing websites and to develop new WebPages for various institutions.
10. To perform data extractions, statistical analysis and data interpenetration.
11. To identify how the lessons learned can enhance a career in informatics.

The internship is an opportunity for the student to acquire on-the-job training at a leading health organization [1].

Description of the Organization:

The Population Health Research Unit (PHRU) is a university-based research and support group conducting systematic research into population health, health services and their inter-relationships. PHRU is committed to advancing the level of knowledge and developing innovative research methods for the betterment of the health of the general population in a cost effective manner. Its efforts support both individual researchers and research teams addressing larger projects. PHRU staff work with government bodies, other research institutions, granting agencies, public health boards, as well as the private sector [3].

PHRU was established within Dalhousie University's Department of Community Health and Epidemiology in 1993 to meet the growing need for data and research support in population health, health services utilization and their interrelationships. In a time of health care reform, spending constraints and an expanding world of therapies and technologies, the need for efficient and effective support services for population-based research has never been greater. To facilitate the growth of research in these areas, the Province of Nova Scotia has supplied PHRU with complete Medicare, Pharmacare and Hospital files suitable for research purposes. The Unit has also been supplied with Workers Compensation records and has access to a variety of other data sources including clinical databases and large-scale population surveys [3]. The individual databases have been linked to create a comprehensive data system, which provides unparalleled opportunities for research in the health and social sciences:

- Evaluation studies
- Needs assessments
- Cost-benefit and cost-effectiveness analyses
- Compliance studies
- Determinants of health studies
- Health status analyses
- Patterns of care and practice studies
- Services and technology utilization studies
- Health economics & Pharmaco-economics

Data access procedures and policies:

To ensure fair and equitable access to the databases and to preserve the integrity and confidentiality of the data, a number of policies have been adopted.

- Researchers are required to submit a written request indicating the nature of the data requested, its intended use and the extent to which it will be shared with others.
- All users are required to sign a confidentiality agreement and all access is governed by the pricing policies of the Unit. Staff within PHRU will provide expert services for data interpretation and explanation.
- Micro data in the form of customized research files may be released to researchers upon approval of the PHRU Management Committee. All such data files will be rendered anonymous and subjected to special release conditions.
- The linkage of PHRU data to external research data sets requires special protocols to ensure the security of the Department of Health procedure for encrypting patient identifiers and to protect the confidentiality of subjects.
- The PHRU Database System provides efficient and effective access to a comprehensive data set, which creates unique opportunities for population health and service utilization research in Nova Scotia. The system provides information resources that are at the leading edge of the North American population health community.
- All new projects coming into PHRU will be reviewed by a committee that will convene once a month. For more information about using PHRU data, please review the Data access procedures for obtaining access to PHRU data.

Objectives of Population Health Research Unit:

The objectives of PHRU are to encourage research and develop health data resources for academic researchers, government and other clients. In accordance with the policies outlined herein PHRU manages (for research purposes) the following Nova Scotia databases including, but not limited to:

- MSI Physician Services
- Hospital Admission/Separation/Day Surgery
- CIHI Discharge Abstract Database
- MSI Seniors' Pharmacare
- MSI Community Services Pharmacare
- Physician Demographics
- MSI Master Registration File
- Vital Statistics
- N.S. Department of Environment Data
- Community Services Family Benefits Database
- Mental Health Outpatient Information System

Description of the work and how it relates to Health Informatics:

Title during internship: Analyst

Role and Responsibilities:

1. To maintain and update the Population Health Research Unit Databases (PHRU Research Data Repositories).
2. To maintain and Update “What’s New?” web page of Population Health Unit.
3. To contribute to research projects by performing data extractions, statistical analysis and data interpenetration; as well as acquired hands-on research experience relating directly to today’s major health problems.
4. To produce technical reports and prepare data for presentation at scientific meetings for both academic and government audiences.
5. To apply IT Project Management skills effectively to develop questions and solutions as well as to implement solutions effectively.
6. To develop websites for other institutions.
7. To apply principles of project and information management within projects on daily basis.
8. To become familiar with additional software applications

Elaborating upon these roles and responsibilities:

1. Maintained and updated the Population Health Research Unit Databases (Appendix A):
 - 1.1. Provincial Health Utilization databases:
 - 1.1.1 Physician Billings Database (MED)
 - 1.1.2 CIHI Hospital Discharge Abstract Database (ASD/CIHI)
 - 1.1.3 Mental Health Outpatient Information System Database (MHOIS)
 - 1.1.4 Senior’s Pharmacare Prescriptions Database (CSPHARM)
 - 1.1.5 Community Services Pharmacare Prescriptions Database (CSPHARM)

1.2. Population Databases

1.2.1 Insured Patient Registry Database (MASTER)

1.2.2 Postal Code Geography Database (WHERE)

1.2.3 Licensed Provider Registry Database (DOCTOR)

1.2.4 Vital Statistics Deaths Database (VITAL)

1.3 Health Status Determinants Databases

1.3.1 Canada Censuses Survey Database (CENSUS)

1.3.2 Nutrition Survey Database

1.3.3 National Population Health Survey Database

1.3.4 Nova Scotia Health Survey Database

Steps taken to update the PHRU Research Data Repositories:

- I. Created SAS data files using VMS software and saved to GEORG4.BP.DAL.CA server using UltraEdit software.
- II. Used five SAS files to create the final HTML documentation file using VMS software:
 - CONFIG.SAS file – This is online documentation configuration file. It includes information such as library where SAS file is located, name of the file is documented, name of 1st documentation file, name of 1st HTML output file, title for documentation.
 - PASS1.SAS file - Program to create an HTML formatted output file of proc contents, proc means, proc freq's, proc tabulates, and proc format FMTLIB table for documentation purposes. Footnote numbers are included in the FORMAT program. MSI and Hospital numbers are excluded from the report for confidentiality reasons.
 - PASS2.SAS file – Program to read the log file from the last step to extract IDX information for modifications to the HTML file. What this program does is identify all the objects that were output to the HTML file and assigns a

sequential IDX# to each. It builds a table (array) and then uses that to modify the HTML output.

- PASS3.SAS file – Program to read the PASS2 output, build an IDX# table for the FORMAT entries in the FMTLIB output, and then modify proc content entries to link to the FORMAT table.
 - FINAL.SAS file – Program to handle footnotes. If no footnotes are required just remove all of the text between <pre> and </pre> in footnotes.htm.
- III. Created online data documentation (Selected which users can view/edit based on their role. Identified the taxonomy term associations; specified the taxonomy terms with which the user wishes to associate the node).
- IV. Uploaded the final HTML documentation file to PHRU website.
- V. Saved copies of final HTML documentation file and printable version of final HTML documentation file to METADATA server.

Description of the PHRU Databases (Research Data Repositories):

The Physician Billing database contains administrative records for each insured health service rendered by a physician and paid for by the Nova Scotia provincial healthcare system (Appendix B). The data include:

- Patient demographic information (e.g. birth date, sex, postal code)
- Physician information (e.g. specialty, ID number for linkage to the physician registry)
- Diagnostic (only one from 1989-1996, up to three from 1997 forward) and procedure codes
- Costing information (i.e. the amount paid for each billing)

Therefore, Physicians delivering health services in Nova Scotia can be remunerated by Nova Scotia Medical Services Insurance (MSI) in either of two ways:

- MSI by Fee-for-Services
- MSI by Alternative Funding.

The CIHI Hospital Discharge Abstract Database (ASD/CIHI) has two distinct data series. Such as ASD (Admissions /Separations/ Day Surgeries) and DAD (Discharge Abstract Database) (Appendix C). The database contains diagnostic codes from ICD -9 format from 1989-2000 and diagnostic codes ICD-10 format from 2001 forward. Both ASD and DAD data series contain individual patient-level information including:

- Patient demography (age, gender, location, etc)
- Attending physicians
- Diagnoses and procedures performed

In addition, the newer, CIHI database contains information on

- Service transfers while in hospital
- Specialty services received (e.g. physiotherapy, occupational therapy)
- Case complexity (e.g. resource intensity weight)

The MHOIS, Mental Health Outpatient Information System Database, facilitates clinician access to patient data and enables case management and support staff to quickly edit client data (Appendix D). The system was implemented at all Mental Health Clinics across Nova Scotia in 1992. The system uses a form called “Mental Health Outpatient Record” or “Event Form “. The submission of the form is a clinician’s responsibility and it contains client encounter information. Such as:

- Registration and Re-registration
- Assessment
- Repeat Transaction
- Closure

PHRU maintains anonymized and de-identified copies of this information for research purposes.

The Senior's Pharmacare Prescriptions database contains information about pharmaceutical prescriptions paid for by the Nova Scotia provincial government to beneficiaries over the age of 65 (Appendix E). The database includes

- Demographic information (gender, birth date, county, postal code).
- Provider information (prescribing physician identifier, specialty, and county).
- Pharmaceutical (Drug Identification Number (DIN), Anatomical Therapeutic Classification (ATC), compound indicator, quantity of drug dispensed).
- Costing data (prescription cost, amount paid, copay amount, amount of markup).

The Community Services Pharmacare database contains information about pharmaceutical prescriptions paid for by the provincial government to beneficiaries on provincial income assistance programs (Appendix F). All individuals in the database are under the age of 65.

The Insured Patient Registry contains longitudinal information about every individual registered as a beneficiary of provincial MSI healthcare services (Appendix G).

The database includes:

- Demographic information such as gender and date of birth.
- Insured Health Benefits program eligibility start / end dates
- Patient Geography (county, postal code, Canada Post Community)

The WHERE, Postal Code Geography Database, contains additional geographic information derived from the patient's postal code. The database employs the PCCF+ (Postal Code Conversion File) programs to assign specific geography codes to records so that geography-specific (e.g. census information) can be linked to these records. In areas where postal codes do not map exactly to other geographic boundaries (in rural areas, for example), the geographic code is assigned probabilistically using the relative population weights of the surrounding areas. This WHERE geography database is used for analyzing population health data by various geographic units, including dissemination areas, census tracts, census subdivisions, and metropolitan-influenced zones (Appendix H).

The Licensed Provider Registry contains demographic information about the ‘population’ of health services providers (currently only containing physicians that are registered and licensed to practice in Nova Scotia) (Appendix I). It includes information for each physician such as

- Date of birth, gender, place of birth, geography.
- Education including year of graduation, name of school, area(s) of specialization.
- Eligibility start / end dates (period during which physician was licensed to practice in Nova Scotia).

Vital Statistics provides PHRU with data around all mortalities in Nova Scotia (Appendix J). It includes:

- Demographic information such as gender, date of birth, date of death, county of residence, hospital, birth province, province of residence, marital status.
- Information around the mortality event such as the place of injury, attendant type, and whether a medical exam or autopsy was performed.
- Diagnostic data such as the nature of injury, causes of death, and underlying cause of death

The Nova Scotia portion of the census data conducted by Statistics Canada every 4 years housed at PHRU is broken down by geography, providing a complete census profile for each dissemination area in Nova Scotia (Appendix K). The database contains information about

- Demography (age, sex, location).
- Number of households, dwellings and families.
- Socioeconomic indicators (income, unemployment, housing, occupation)
- Languages, ethnicity, religion.

Nutritional Survey provides PHRU with data:

- Population demographics
- Diet/ food composition patterns
- Knowledge / opinions related to diet
- Food preparation method

National Population Health Survey provides PHRU with data:

- Population demographics
- Health status
- Use of health services
- Determinants of health
- Socioeconomic indicators

Nova Scotia Health Survey contains information about:

- Population demographics
- Health status
- Determinants of health
- Socioeconomic indicators

2. Maintained and Updated “What’s New?” web page of Population Health Unit (Appendix L). The web page contains links to both external and internal events of other organizations such as workshops, seminars, bulletins, reports, fundraising events etc.

Therefore, it gives an opportunity for users to collect and manage health information of the important events of other institutions through PHRU website.

Adding links required modification of the index page using HTML/PHP codes and uploading various types of pictures and PDF files to the server using FTP software.

3. Contributed to research projects by performing data extractions, statistical analysis and data interpenetration; as well as acquired hands-on research experience relating directly to today's major health problems. Two examples of this are:

(i). "Growing Together", a program designed to promote healthy child development in strong families and supportive communities. This research was a joint partnership between the Dartmouth Family Resource Center in North Dartmouth, the Family Place Resource Center in Northside Victoria, the IWK Health Center and Dalhousie University. This study tested how well the program assisted parents in getting to know their newborns and promoting healthy child development. The information from this study will help policy makers to understand how best to support young families. For this the author used SPSS statistical software to analyze following data:

- Family-Centered information
- Child-Centered information
- Child literacy information
- Caregiver's childhood experiences
- Difficult life circumstances (DLC)
- Community life skills scale
- CES-D scale
- Parent protective factors
- Measurement of language skills in infancy (Reel-2 scale)
- NCAST teaching scale
- Home observation for measurement of the environment

Involvement in this project allowed the author to gain a greater appreciation for how to acquire information, how to analyze it, and how to disseminate it. The author enhanced his ability to integrate data from multiple sources, and not only combine them, but to determine its validity and quality so the data becomes useful information.

The results could be used for policy development, program planning and evaluation purposes.

(ii).Surveillance of Mood Disorders and Anxiety Disorders in Nova Scotia from 2001 to 2005. Using Provincial / Territorial administrative databases, this study was conducted by the Center for Chronic Disease Prevention and Control, which is part of the Public Health Agency of Canada.

The author was involved in preparation of the following reports in order to interpret the data on Mood and Anxiety disorders:

- Sex and Age distribution of individuals with treated disorder for each year.
- Means and Median age of individuals with treated disorder for each year.
- Prevalence of treated disorder for each year
- Prevalence of treated disorder by sex for each year
- Prevalence of treated disorder by age group for year
- Among individuals with treated disorder, proportion who saw each type of service provider for each year.
- Among individuals with treated disorder, proportion who saw each type of service provider by sex for each year.
- Among individuals with treated disorder, proportion who saw each type of service provider by Age group for each year.
- Among individuals with treated disorder from physician billings, the number of visits to a physician for each year.
- Among individuals with treated disorder from hospital separation, the number of hospitalization for each year.

The results are to be used for policy development, program planning and evaluation purposes.

4. Produced technical reports and prepared data for presentation at scientific meetings for both academic and government audiences.
5. Applied IT Project Management skills effectively to develop questions and solutions as well as to implement solutions effectively.
6. Developed website for “National Diabetes Surveillance System” (NDSS) using PHP language (Appendix M). The website contains total of 10 WebPages in both English and French. The first page is a portal page for the NDSS with links to the home pages in both English and French. The home pages have links to the sub pages. The visitor count and date functions are included.

Steps taken to design the website:

A website design is very important as it displays the potential of the body it represent. A second-rate design webpage would repel most people from the site and this would eventually effect the institution. Therefore, the author has considered the following rules “Dos and Don’ts” for the web design [4]:

- *The scope of design web site-* Do check the scope of the web site you are designing. It is very important that the target of the web site is clearly in your mind when you are designing a web page. Don’t design heavy sites with several links if it is used for personal purposes or representing a small business. Most online sites are design with the aim to be user friendly and easy to surf.
- *The layout for design websites –* Do the designing of the website according to the requirement of the websites. Heavy animations and images often imposes heavy burden on the site. Such websites consume much time to load and moist people quickly lose interest in such websites. Don’t ever try to heavily decorate a web page to impress people. It is evidently proven that more people are attracted to simpler built web site as compare to those which are heavily filled with images, sounds and animations.

- *The compatibility of design websites* – Do check that the design website is compatible with the major browsers which are being used in the world. It is also necessary that when you add animations and designing to your webpage you are aware what type of graphics the browser support and which animation won't work on the internet. Don't upload your website on server if they are not run on most common browsers such as Internet Explore and Netscape Navigator. Always check the authenticity of a server before uploading the work online.
 - *Documentation of design websites* – Do make sure that the text on your design website is relevant. It is very necessary that the documentation on the design website is easily understandable and to the point. The uses of keywords while describing the contents of the website is also important as it keeps the focus of the reader to the topics on the design website. Don't fill your design site with useless and irrelevant text as this diverts most of the traffic from the web site. The amount of information or any knowledge which is provided on the design website should be done in a professional manner.
 - *Safe and secure mechanism* – Do make sure that your website is safe from viruses and hackers. It is very necessary that latest security protocols are implemented with the website design. Don't encourage the sharing security protocols and protection mechanism with the general public.
7. Applied principles of project and information management within projects on daily basis and worked in fast paced team environment at PHRU.
8. Became familiar with additional software applications:
- **VMS Software** - VMS is a multi-user, multiprocessing virtual memory-based operating system (OS) designed for use in time sharing, batch processing, real time (process priorities can be set higher than OS kernel jobs) and transaction processing (Appendix N). It offers high system availability through clustering, or the ability to distribute the system over multiple physical machines this allows the system to be "disaster-tolerant" against natural disasters that may disable individual data-processing facilities [5].

- **UltraEdit Software** - UltraEdit is a commercial text editor for Microsoft Windows created in 1994 by IDM Computer Solutions (Appendix O). The editor contains tools for programmers, including macros, syntax highlighting, code folding, file type conversions, project management, regular expressions for search-and-replace, a column-edit mode, remote editing of files via FTP, interfaces for APIs or command lines of choice and more.
- **SPSS Software** – SPSS is a modular, tightly integrated, full-featured product line for the analytical process—planning, data collecting, data access, data management and preparation, data analysis, reporting, and deployment and widely used in research and for teaching of statistical methods.
- **SAS Software** – SAS system is an integrated statistical package that allows user to do statistical analyses--such as descriptive statistics, multivariate techniques and time series--using a consistent set of commands.

Other activities during internship:

In addition to the work, the most valuable experience during the internship was the PHRU staff meetings where we presented our progress reports and discussed about the work that we participated. The author attended three staff meetings and one seminar at PHRU. It was a great opportunity to learn how to discuss progress reports of health informatics because this was the first opportunity the author has exposed to Health Informatics work. The seminar that the author has attended has given an opportunity to familiarize with new software.

Discussion of the problem that the author have analyzed and the corresponding solution:

Problem encountered	Reason for the problem
Unable to change the format for data documentation nodes in the PHRU websites.	The system uses “Drupal” software to manage the data repositories.

Proposed Solution:

At the moment the system uses “Drupal” software to communicate between PHRU website and the PHRU data repositories. This makes it impossible to changr the format for the data documentation nodes in the PHRU websites. Though this does not affect the integrety of the data it does make it more difficult to work with.

Therefore, it may be beneficial to redesign the system so that the PHRU website directly interfaces with PHRU data repositories. The proposed system could be restructured using PHP language and MySQL Server or MS SQL Sever.

Conclusions:

The past four months at the Population Health Research Unit was an excellent opportunity for author to gain some hands on experience in the health informatics field. It gave the author an opportunity to work as a health information analyst, not just a student. The author acquired many new skills and developed many already existing skills including innovativeness and creativity in web-design, data-analysis using SAS/SPSS software, working in a team based environment, dealing with management and customers, and report writing. Although these skills were developed in the work force, the author has found that they could easily be applied at school. The author has found that his ability to absorb material has drastically improved, as has his time management ability.

Recommendations:

The author's experience at Population Health Research Unit has been invaluable. Author thoroughly enjoyed his position and learned a great deal about all aspects of a major health research institution.

However, reflecting upon the internship experience the author identifies a couple of areas in which he would like to strengthen his knowledge. These include:

- To be familiar with UNIX /LINUX operating systems. The author discovered that such knowledge would be helpful in the health institution environment because many of these places rely on these systems.
- To gain experience working with large health databases. The author found that health informaticians working in the field have the ability to compile information from large databases.

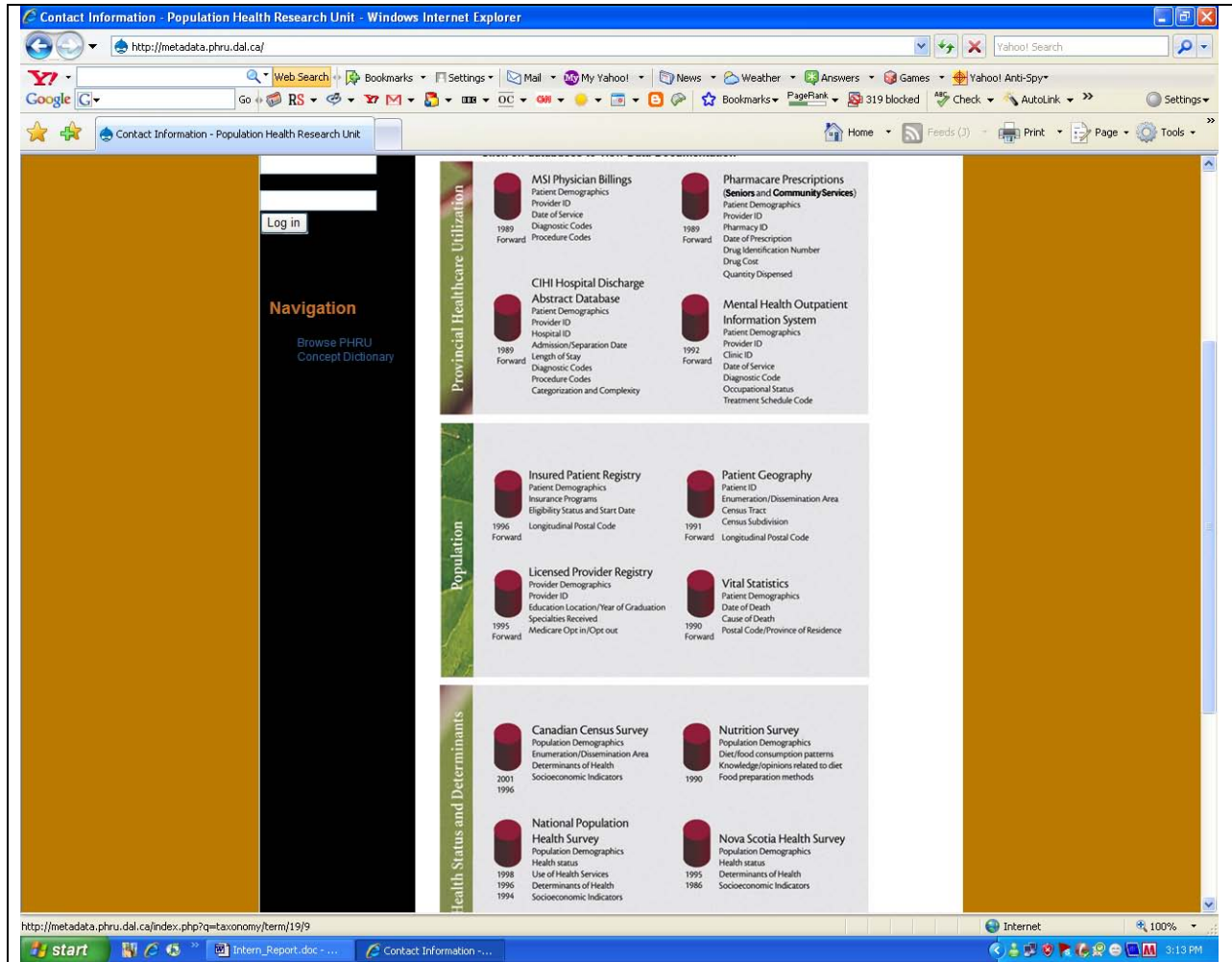
Further, the author suggests that all students do an internship whether they are pursuing master thesis or not. Work experience while in school would significantly increase the chances of employment after graduation.

Reference:

- [1] Dalhousie University (2007). Faculty of Computer Science, Health Informatics Overview. Retrieved July 20, 2007, from <http://www.healthinformatics.dal.ca/overview.php>
- [2] Northeastern University, (2007). College of Computer and Information Science. Retrieved August 03, 2007, from <http://www.healthinformatics.neu.edu/>
- [3] Dalhousie University (2007). Faculty of Medicine, Population Health Research Unit. Retrieved July 22, 2007, from <http://www.phru.dal.ca/>
- [4] Small Business Bible (2007). Dos and Don'ts for website design. Retrieved July 22, 2007, from http://www.smallbusinessbible.org/dos_donts_websitedesign.html
- [5] Wikipedia (2007). OpenVMS. Retrieved July 23, 2007, from <http://en.wikipedia.org/wiki/OpenVMS>

Appendices:

Appendix A: Web Page for Research Data Repository



Appendix B: Web page for Physician Billing Database

5790 UNIVERSITY AVENUE | HALIFAX, NOVA SCOTIA | CANADA B3H 1V7 | Phone +1 (902) 494-1785 | Fax +1 (902) 494-1697

MED05

- view
- edit

Physician Billings (MED)

MSI Physician Services Fiscal 2005

The CONTENTS Procedure

Data Set Name:	LIBRARY.MED05	Observations:	7200279
Member Type:	DATA	Variables:	35
Engine:	V8	Indexes:	0
Created:	12:41 Friday, May 26, 2006	Observation Length:	184
Last Modified:	12:55 Friday, May 26, 2006	Deleted Observations:	0
Protection:		Compressed:	NO
Data Set Type:		Sorted:	YES
Label:		Printable Version	

-----Engine/Host Dependent Information-----			
Data Set Page Size:	14848		
Number of Data Set Pages:	90005		
First Data Page:	1		
Max Obs per Page:	80		
Obs in First Data Page:	52		
Number of Data Set Repairs:	0		
Filename:	MED05.SAS7BDAT		
Release Created:	8.0202M0		
Host Created:	OpenVMS		
File Size (blocks):	2610146		

-----Alphabetic List of Variables and Attributes-----						
#	Variable	Type	Len	Pos	Format	Label
32	addate	Num	4	20		Hospital Admit Date
25	age	Num	3	178		Patient Age
28	approved	Num	4	4		Amount Approved (Cents)
11	bspecial	Char	4	50	\$DSPECF	Provider (Billing) Specialty
14	ccpcat	Char	4	62	\$CCPCATF	CCP (Procedure) Category
15	ccpcode	Char	5	66		CCP (Procedure) Code
7	ccpqual	Char	1	29		CCP (Procedure) Qualifier

Appendix C: Web page for CIHI Hospital Discharge Abstract Database

DALHOUSIE UNIVERSITY
Inspiring Minds
Faculty of Medicine
Population Health Research Unit

5790 UNIVERSITY AVENUE | HALIFAX, NOVA SCOTIA | CANADA B3H 1V7 | Phone +1 (902) 494-1785 | Fax +1 (902) 494-1697

HOME
NEWS
ABOUT PHRU
USING OUR SERVICES
POPULATION HEALTH DATABASES
CONCEPT DICTIONARY
RESEARCH
CONTACT US

Search data documentation

Search

phruadmin

Browse PHRU
Concept Dictionary
Create Concept Dictionary Entry
Create Data Description
Create Data Documentation
search
my account
administer
log out

CIHI05

- view
- edit

Hospital Discharge Abstract Database (ASD / CIHI)

CIHI Fiscal 2005

The CONTENTS Procedure

Data Set Name:	LIBRARY.CIHI05	Observations:	192844
Member Type:	DATA	Variables:	601
Engine:	V8	Indexes:	0
Created:	14:43 Thursday, January 4, 2007	Observation Length:	2816
Last Modified:	14:43 Thursday, January 4, 2007	Deleted Observations:	0
Protection:		Compressed:	NO
Data Set Type:		Sorted:	YES
Label:		Printable Version	

-----Engine/Host Dependent Information-----

Data Set Page Size:	31232
Number of Data Set Pages:	17534
First Data Page:	3
Max Obs per Page:	11
Obs in First Data Page:	6
Number of Data Set Repairs:	0
Filename:	CIHI05.SAS7BDAT
Release Created:	8.0202M0
Host Created:	OpenVMS
File Size (blocks):	1069575

-----Alphabetic List of Variables and Attributes-----

# Variable	Type	Len	Pos	Format	Label
371 ACCDATE	Num	4	236		Accident Date
98 ACULOS	Num	4	24		Acute Care LOS
363 ADDATE	Num	4	204		Date of Admission
554 ADFROM	Char	52605		\$HOSPITE	Institution From Number
349 ADHOUR	Num	4	148		Hour of Admission
372 ADMITNUM	Num	52796			Admission Number
385 ADTYPE	Char	12133		\$ADTYPE2	Admission Type

http://www.phru.dal.ca/contact

Internet 100%

3:15 PM

Appendix D: Web page for Mental Health Outpatient information System Database

DALHOUSIE UNIVERSITY
Inspiring Minds
Faculty of Medicine
Population Health Research Unit

5790 UNIVERSITY AVENUE | HALIFAX, NOVA SCOTIA | CANADA B3H 1V7 | Phone +1 (902) 494-1785 | Fax +1 (902) 494-1597

HOME
NEWS
ABOUT PHRU
USING OUR SERVICES
POPULATION HEALTH DATABASES
CONCEPT DICTIONARY
RESEARCH
CONTACT US

Search data documentation

Search

phruadmin

Browse PHRU
Create Concept Dictionary Entry
Create Data Description
Create Data Documentation Search
my account
administer
log out

MHOIS04

- view
- edit

Mental Health Outpatient Information System (MHOIS)

MHOIS Fiscal 2001

The CONTENTS Procedure

Data Set Name:	LIBRARY.MHOIS04	Observations:	239839
Member Type:	DATA	Variables:	55
Engine:	V8	Indexes:	0
Created:	16:09 Tuesday, November 29, 2005	Observation Length:	316
Last Modified:	14:53 Wednesday, November 30, 2005	Deleted Observations:	0
Protection:		Compressed:	NO
Data Set Type:		Sorted:	YES
Label:		Printable Version	

-----Engine/Host Dependent Information-----

Data Set Page Size:	25600
Number of Data Set Pages:	3000
First Data Page:	1
Max Obs per Page:	80
Obs in First Data Page:	57
Number of Data Set Repairs:	0
Filename:	MHOIS04.SAS7BDAT
Release Created:	8.0202M0
Host Created:	OpenVMS
File Size (blocks):	150001

-----Alphabetic List of Variables and Attributes-----

#	Variable	Type	Len	Pos	Format	Label
29	age	Num	3	310		Patient Age
1	central	Char	4	8	\$CLINICF	Central Clinic
49	clin1	Char	24	161		Clinician 1 Identifier
50	clin2	Char	24	185		Clinician 2 Identifier
51	clin3	Char	24	209		Clinician 3 Identifier
52	clin4	Char	24	233		Clinician 4 Identifier
44	clinds1	Char	1	133	\$DISCIPLF	Clinician 1 Discipline

Appendix E: Web page for Senior's Pharmacare Prescription Database

DALHOUSIE UNIVERSITY
Inspiring Minds
Faculty of Medicine
Population Health Research Unit

5790 UNIVERSITY AVENUE | HALIFAX, NOVA SCOTIA | CANADA B3H 1V7 | Phone +1 (902) 494-1785 | Fax +1 (902) 494-1697

HOME
NEWS
ABOUT PHRU
USING OUR SERVICES
POPULATION HEALTH DATABASES
CONCEPT DICTIONARY
RESEARCH
CONTACT US

Search data documentation

Search

phruadmin

Browse PHRU
Concept Dictionary
Create Concept Dictionary Entry
Create Data Description
Create Data Documentation
search
my account
administer
log out

PHARM06

- view
- edit

Senior's Pharmacare Prescriptions (PHARM)

Senior's Pharmacare 2006

The CONTENTS Procedure

Data Set Name:	LIBRARY.PHARM06	Observations:	3184812
Member Type:	DATA	Variables:	32
Engine:	V8	Indexes:	0
Created:	15:24 Wednesday, June 27, 2007	Observation Length:	176
Last Modified:	15:24 Wednesday, June 27, 2007	Deleted Observations:	0
Protection:		Compressed:	NO
Data Set Type:		Sorted:	YES
Label:		Printable Version	

-----Engine/Host Dependent Information-----	
Data Set Page Size:	14336
Number of Data Set Pages:	39319
First Data Page:	1
Max Obs per Page:	81
Obs in First Data Page:	54
Number of Data Set Repairs:	0
Filename:	PHARM06.SAS7BDAT
Release Created:	8.0202M0
Host Created:	Open/VMS
File Size (blocks):	1100934

-----Alphabetic List of Variables and Attributes-----						
#	Variable	Type	Len	Pos	Format	Label
21	age	Num	3	168		Patient Age
22	approved	Num	4	0		Amount Approved (Cents) [1]
11	atc	Char	7	148		Anatomical Therapeutic Class
23	claimed	Num	4	4		Amount Claimed (Cents)
6	claimtyp	Char	1	92	\$CLM1TYPE	Pricing Agreement Number
1	compound	Char	1	40	\$YESNOF	Compound Indicator
25	copay	Num	4	8		Amount of Copay (Cents) [2]

Appendix F: Web page for Community Services Pharmacare Prescriptions Database

5780 UNIVERSITY AVENUE | HALIFAX, NOVA SCOTIA | CANADA B3H 1V7 | Phone +1 (902) 494-1785 | Fax +1 (902) 494-1597

CSPHARM06

- view
- edit

Community Services Pharmacare Prescriptions (CSPHARM)

Community Services Pharmacare 2006

The CONTENTS Procedure

Data Set Name:	LIBRARY.FAM06	Observations:	1031917
Member Type:	DATA	Variables:	32
Engine:	V8	Indexes:	0
Created:	15:25 Wednesday, June 27, 2007	Observation Length:	176
Last Modified:	15:25 Wednesday, June 27, 2007	Deleted Observations:	0
Protection:		Compressed:	NO
Data Set Type:		Sorted:	YES
Label:		Printable Version	

-----Engine/Host Dependent Information-----

Data Set Page Size:	14336
Number of Data Set Pages:	12741
First Data Page:	1
Max Obs per Page:	81
Obs in First Data Page:	54
Number of Data Set Repairs:	0
Filename:	FAM06.SAS7BDAT
Release Created:	8.0202M0
Host Created:	OpenVMS
File Size (blocks):	356749

-----Alphabetic List of Variables and Attributes-----

# Variable	Type	Len	Pos	Format	Label
21 age	Num	3	168		Patient Age
22 approved	Num	4	0		Amount Approved (Cents) [1]
11 atc	Char	7	148		Anatomical Therapeutic Class
23 claimed	Num	4	4		Amount Claimed (Cents)
6 claimtp	Char	1	92	\$CLMTPFF	Pricing Agreement Number
1 compound	Char	1	40	\$YESNOF	Compound Indicator
25 copay	Num	4	8		Amount of Copay (Cents) [2]

Appendix G: Web page for Insured Patient Registry Database

DALHOUSIE UNIVERSITY
Inspiring Minds
Faculty of Medicine
Population Health Research Unit

5790 UNIVERSITY AVENUE | HALIFAX, NOVA SCOTIA | CANADA B3H 1V7 | Phone +1 (902) 494-1785 | Fax +1 (902) 494-1697

ELIGIB06

- view
- edit

Insured Patient Registry (MASTER)

Patient Eligibility 2006

The CONTENTS Procedure

Data Set Name:	LIBRARY.ELIGIB06	Observations:	248662
Member Type:	DATA	Variables:	7
Engine:	V8	Indexes:	0
Created:	10:56 Thursday, July 26, 2007	Observation Length:	48
Last Modified:	11:28 Thursday, July 26, 2007	Deleted Observations:	0
Protection:		Compressed:	NO
Data Set Type:		Sorted:	YES
Label:		Printable Version	

-----Engine/Host Dependent Information-----

Data Set Page Size:	8192
Number of Data Set Pages:	14715
First Data Page:	1
Max Obs per Page:	169
Obs in First Data Page:	131
Number of Data Set Repairs:	0
Filename:	ELIGIB06.SAS7BDAT
Release Created:	8.0202M0
Host Created:	OpenVMS
File Size (blocks):	235441

-----Alphabetic List of Variables and Attributes-----

#	Variable	Type	Len	Pos	Format	Informat	Label
4	ESTAT	Char	3	13	\$ELIGIB.		Eligibility Status
2	FROMDATE	Num	4	0			Eligibility Start Date
1	PROGRAM	Char	5	8	\$PROGRAM.\$5.		Program
6	SOURCE	Char	2	19			Year of Data
3	TODATE	Num	4	4			Eligibility End Date
5	TSTAT	Char	3	16	\$TERM.		Termination Status
7	msi	Char	24	21			Patient MSI Number

Appendix H: Web page for Postal Code Geography Database

DALHOUSIE UNIVERSITY
Inspiring Minds
Faculty of Medicine
Population Health Research Unit

5790 UNIVERSITY AVENUE | HALIFAX, NOVA SCOTIA | CANADA B3H 1V7 | Phone +1 (902) 494-1785 | Fax +1 (902) 494-1997

WHERE07

- view
- edit

Postal Code Geography (WHERE)

GeoCodes/PCCF+ Version 4H (2007)

The CONTENTS Procedure

Data Set Name:	LIBRARY.WHERE07	Observations:	2718506
Member Type:	DATA	Variables:	46
Engine:	V8	Indexes:	0
Created:	15:48 Thursday, April 12, 2007	Observation Length:	144
Last Modified:	11:53 Tuesday, May 8, 2007	Deleted Observations:	0
Protection:		Compressed:	NO
Data Set Type:		Sorted:	YES
Label:		Printable Version	

-----Engine/Host Dependent Information-----			
Data Set Page Size:	11776		
Number of Data Set Pages:	33564		
First Data Page:	1		
Max Obs per Page:	81		
Obs in First Data Page:	37		
Number of Data Set Repairs:	0		
Filename:	WHERE07.SAS7BDAT		
Release Created:	8.0202M0		
Host Created:	OpenVMS		
File Size (blocks):	771973		

-----Alphabetic List of Variables and Attributes-----						
#	Variable	Type	Len	Pos	Format	Label
19	AR	Char	2	68	\$ARF.	Census Agricultural Region (Crop Dist)
12	BLK	Char	2	56		Census Block
14	BLKURB	Char	1	59	\$BLKURBF.	Urban(1)/Rural(0) Indicator
16	CCS	Char	3	62		Census Consolidated Subdivision
7	CD	Char	2	43	\$CDF.	Census Division
8	CMA	Char	3	45	\$CMAF.	Census Metropolitan/Agglomeration Area
33	CSDTYPE	Char	3	99	\$CSDTPF.	Census Subdivision Type

phruadmin

- Browse PHRU
- Concept Dictionary
- Create Concept Dictionary Entry
- Create Data Description
- Create Data Documentation
- search
- my account
- administer
- log out

Appendix I: Web page for Licensed Provider Registry Database

DALHOUSIE UNIVERSITY
Inspiring Minds
Faculty of Medicine
Population Health Research Unit

5790 UNIVERSITY AVENUE | HALIFAX, NOVA SCOTIA | CANADA B3H 1V7 | Phone +1 (902) 494-1785 | Fax +1 (902) 494-1587

HOME
NEWS
ABOUT PHRU
USING OUR SERVICES
POPULATION HEALTH DATABASES
CONCEPT DICTIONARY
RESEARCH
CONTACT US

Search data documentation

Search

phruadmin

- Browse PHRU
- Concept Dictionary
- Create Concept Dictionary Entry
- Create Data Description
- Create Data Documentation
- search
- my account
- administer
- log out

OPTIN03

- view
- edit

Licensed Provider Registry (DOCTOR)

2003 Physician Opt In/Out

The CONTENTS Procedure

Data Set Name:	LIBRARY.OPTIN03	Observations:	6081
Member Type:	DATA	Variables:	5
Engine:	V8	Indexes:	0
Created:	12:25 Monday, February 2, 2004	Observation Length:	16
Last Modified:	16:40 Friday, March 18, 2005	Deleted Observations:	0
Protection:		Compressed:	NO
Data Set Type:		Sorted:	NO
Label:		Printable Version	

-----Engine/Host Dependent Information-----

Data Set Page Size:	8192
Number of Data Set Pages:	14
First Data Page:	1
Max Obs per Page:	506
Obs in First Data Page:	413
Number of Data Set Repairs:	0
Filename:	OPTIN03.SAS7BDAT
Release Created:	8.0101M0
Host Created:	OpenVMS
File Size (blocks):	225

-----Alphabetic List of Variables and Attributes-----

#	Variable	Type	Len	Pos	Format	Label
3	doctor	Num	4	0		New Provider ID (6 Digit)[1]
1	doctype	Char	2	12	\$DTYPEF	Provider Type
4	optin	Num	4	4		Optin Date
5	optout	Num	4	8		Optout Date
2	program	Char	2	14		Program Code

http://metadata.phru.dal.ca/index.php?q=taxonomy/term/19

start | intern_report.doc - ... | Contact Information - ... | Internet | 100% | 3:19 PM

Appendix J: Web page for Vital Statistics Deaths Database

The screenshot shows a web browser window displaying the Vital Statistics Deaths Database page for VITAL05. The browser is Windows Internet Explorer, and the URL is <http://metadata.phru.dal.ca/index.php?q=node/280>. The page header features the Dalhousie University logo and the Faculty of Medicine Population Health Research Unit. The main content area displays the title "VITAL05" and a list of actions: "view" and "edit". Below this, the text "Vital Statistics Deaths (VITAL)" is followed by "Vital Statistics 2005" and "The CONTENTS Procedure". A table provides detailed metadata for the data set, including the number of observations (8307), variables (42), and creation/modification dates. A second table, titled "Engine/Host Dependent Information", lists technical details like page size (19968) and filename (VITAL05.SAS7BDAT). A third table, "Alphabetic List of Variables and Attributes", lists variables such as UCAUSE9, age, ageday, agemon, attype, autopsy, and bprov with their respective data types and labels.

Navigation Menu:

- HOME
- NEWS
- ABOUT PHRU
- USING OUR SERVICES
- POPULATION HEALTH DATABASES
- CONCEPT DICTIONARY
- RESEARCH
- CONTACT US

Search data documentation

Search

phruadmin

- Browse PHRU
- Concept Dictionary
- Create Concept Dictionary Entry
- Create Data Description
- Create Data Documentation
- search
- my account
- administer
- log out

Vital Statistics Deaths (VITAL)

Vital Statistics 2005

The CONTENTS Procedure

Data Set Name:	LIBRARY.VITAL05	Observations:	8307
Member Type:	DATA	Variables:	42
Engine:	V8	Indexes:	0
Created:	13:45 Tuesday, May 1, 2007	Observation Length:	248
Last Modified:	13:45 Tuesday, May 1, 2007	Deleted Observations:	0
Protection:		Compressed:	NO
Data Set Type:		Sorted:	YES
Label:		Printable Version	

-----Engine/Host Dependent Information-----			
Data Set Page Size:	19968		
Number of Data Set Pages:	105		
First Data Page:	1		
Max Obs per Page:	80		
Obs in First Data Page:	56		
Number of Data Set Repairs:	0		
Filename:	VITAL05.SAS7BDAT		
Release Created:	8.0202M0		
Host Created:	OpenVMS		
File Size (blocks):	4096		

-----Alphabetic List of Variables and Attributes-----						
#	Variable	Type	Len	Pos	Format	Label
41	UCAUSE9	Char	5	220		ICD-9 Underlying Cause of Death
17	age	Num	8	16		Age at Death
11	ageday	Num	3	239		Age - Days
10	agemon	Num	3	236		Age - Months
12	attype	Char	1	94	SATTYPEF.	Last Attendant Type
13	autopsy	Char	1	95	YESNO.	Autopsy Performed
18	bprov	Char	2	97	SPROVF.	Birth Province

Appendix K: Web page for Canada Censuses Survey Database

DALHOUSIE UNIVERSITY
Inspiring Minds
Faculty of Medicine
Population Health Research Unit

5790 UNIVERSITY AVENUE | HALIFAX, NOVA SCOTIA | CANADA B3H 1V7 | Phone +1 (902) 494-1785 | Fax +1 (902) 494-1597

HOME
NEWS
ABOUT PHRU
USING OUR SERVICES
POPULATION HEALTH DATABASES
CONCEPT DICTIONARY
RESEARCH
CONTACT US

Search data documentation

Search

phruadmin

- Browse PHRU
- Concept Dictionary
- Create Concept Dictionary Entry
- Create Data Description
- Create Data Documentation
- search
- my account
- administer
- log out

CENSUS01

- view
- edit

Canada Census (CENSUS)

Nova Scotia Census Profile by Dissemination Area (2001)

The CONTENTS Procedure

Data Set Name:	LIBRARY.NS_DAPROFILE	Observations:	1304
Member Type:	DATA	Variables:	1495
Engine:	V8	Indexes:	0
Created:	9:23 Wednesday, August 24, 2005	Observation Length:	11960
Last Modified:	9:23 Wednesday, August 24, 2005	Deleted Observations:	0
Protection:		Compressed:	NO
Data Set Type:		Sorted:	NO
Label:			

-----Engine/Host Dependent Information-----

Data Set Page Size:	24064
Number of Data Set Pages:	660
First Data Page:	9
Max Obs per Page:	2
Obs in First Data Page:	2
Number of Data Set Repairs:	0
Filename:	_DKB1000.[KDRUHAN.CENSUS01] NS_DAPROFILE.SAS7BDAT
Release Created:	8.0202M0
Host Created:	OpenVMS
File Size (blocks):	31022

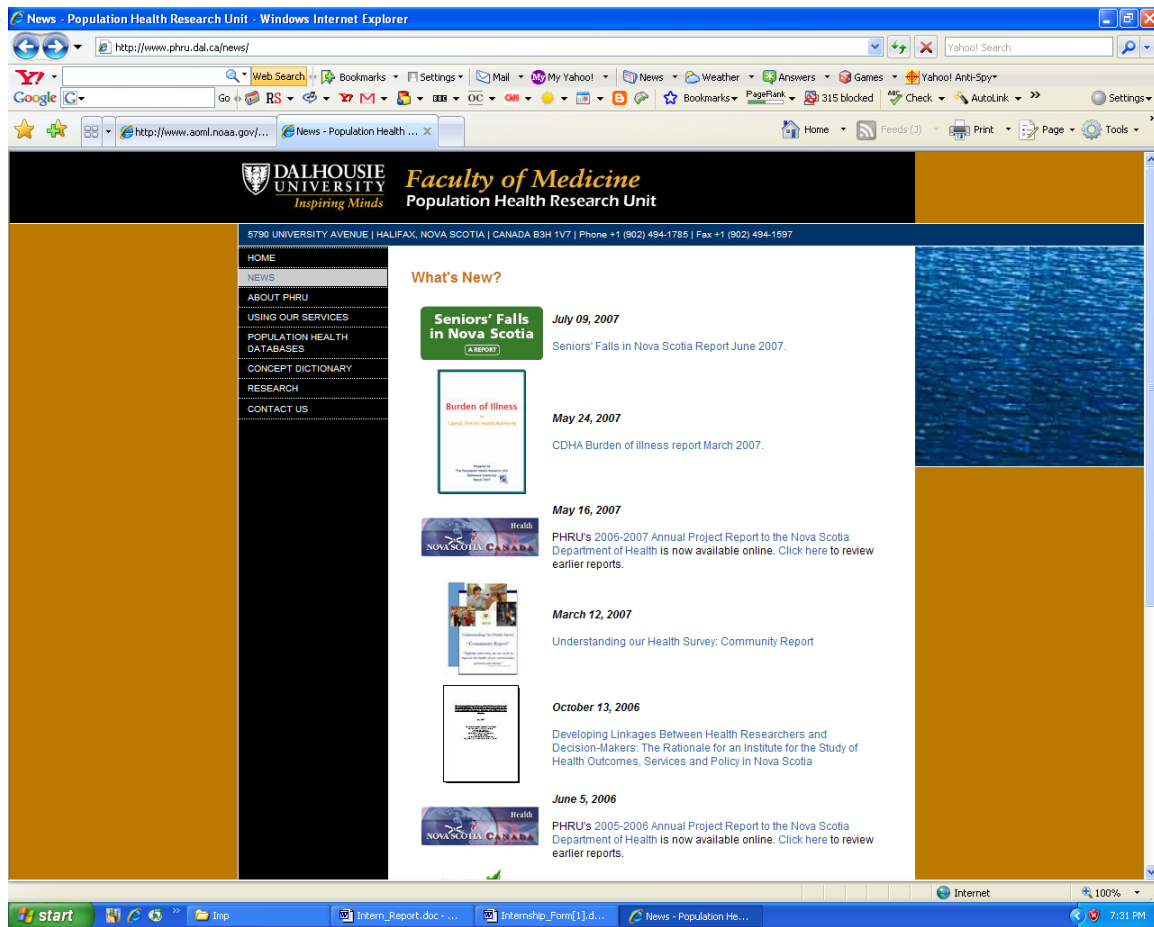
-----Alphabetic List of Variables and Attributes-----

# Variable	Type	Len	Pos	Label
1 COL0	Num	8	0	Dissemination Area
2 COL1	Num	8	8	Province Code
3 COL2	Num	8	16	CD Code
4 COL3	Num	8	24	Province Abbrev.
5 COL4	Num	8	32	CD Name

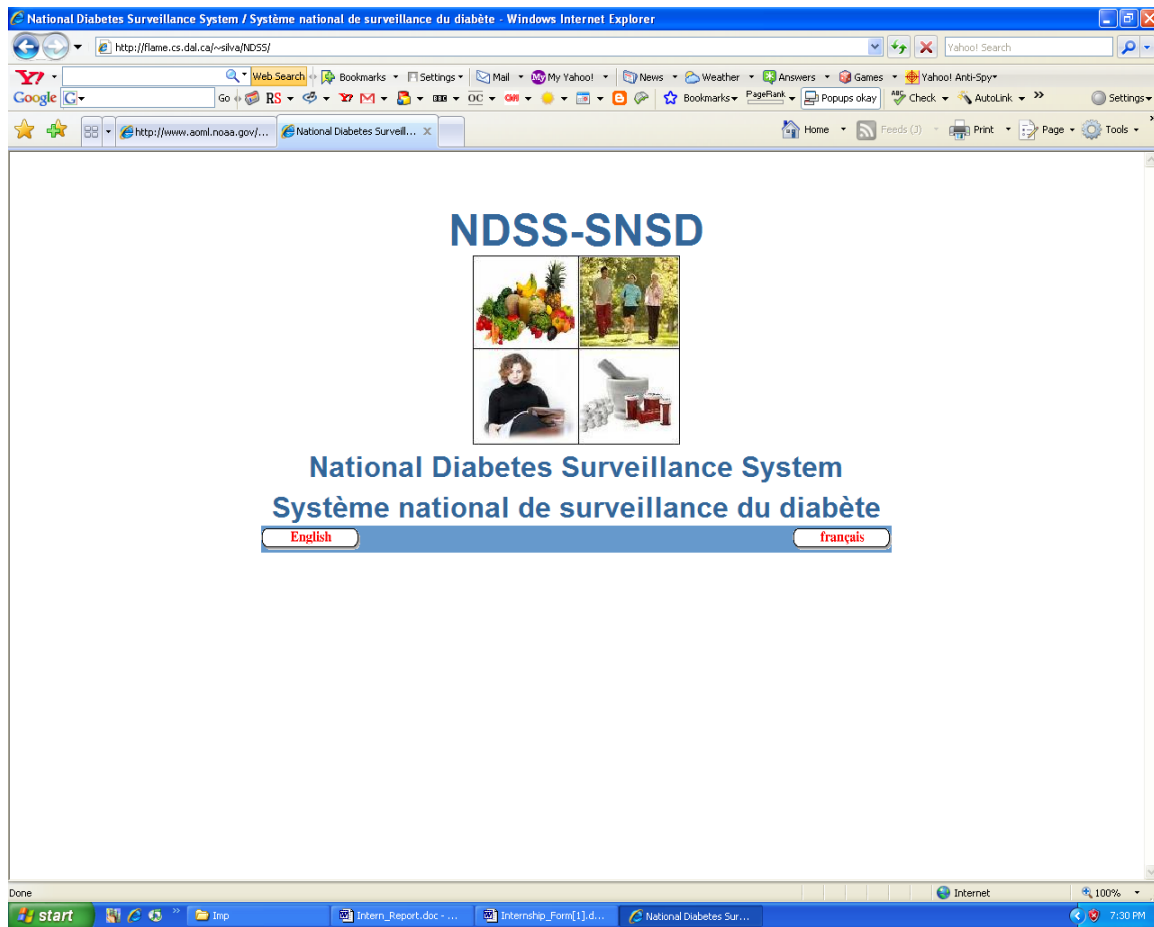
Done

start | Internet | 100% | 3:20 PM

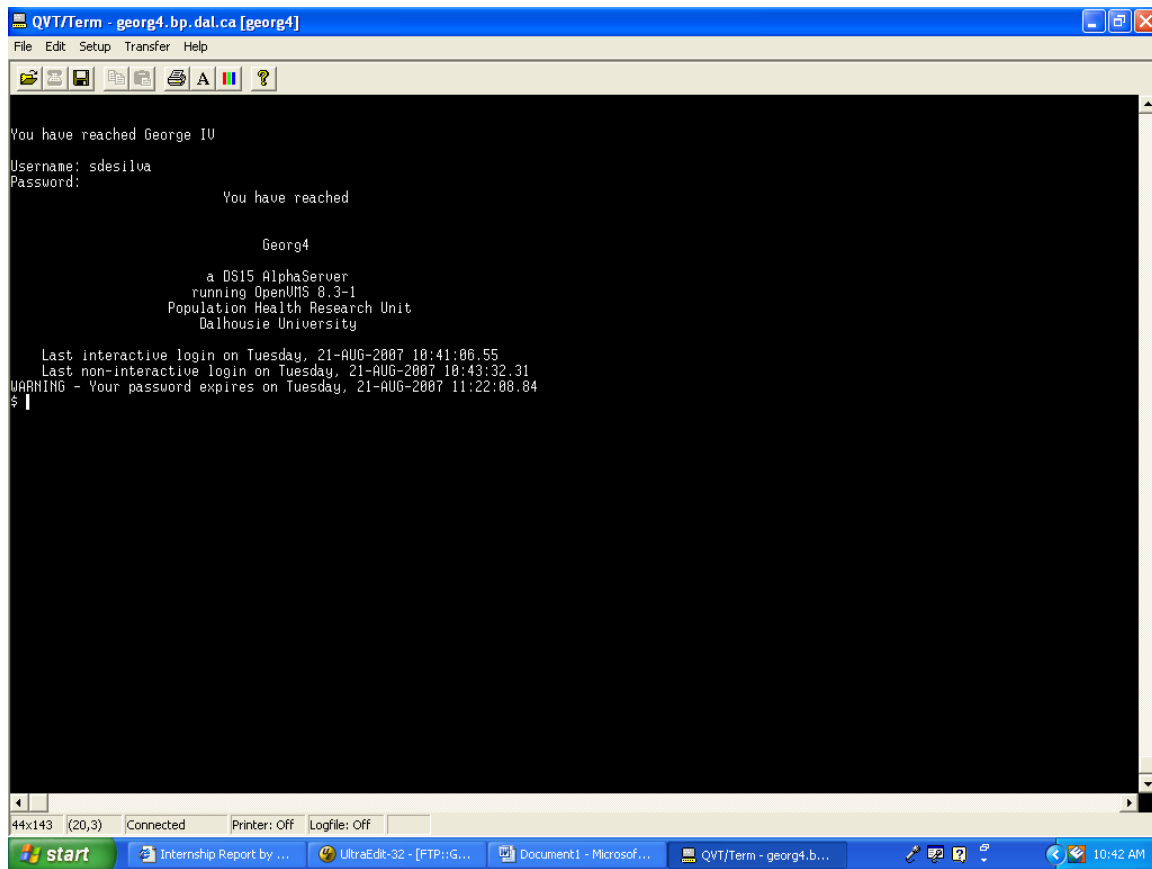
Appendix L: "What's New?" web page of Population Health Unit



Appendix M: Website for National Diabetes Surveillance System



Appendix N: VMS software interface



The screenshot shows a terminal window titled "QVT/Term - georg4.bp.dal.ca [georg4]". The window contains the following text:

```
You have reached George IV
Username: sdesilva
Password:
          You have reached
          Georg4
          a DS15 AlphaServer
          running OpenVMS 8.3-1
          Population Health Research Unit
          Dalhousie University
Last interactive login on Tuesday, 21-AUG-2007 10:41:06.55
Last non-interactive login on Tuesday, 21-AUG-2007 10:43:32.31
WARNING - Your password expires on Tuesday, 21-AUG-2007 11:22:08.84
$ |
```

The window also shows a status bar at the bottom with the following information: 44x143 (20,3) Connected Printer: Off Logfile: Off. The taskbar at the bottom of the screen shows the Start button and several open applications: Internship Report by ..., UltraEdit-32 - [FTP:G..., Document1 - Microsof..., and QVT/Term - georg4.b.... The system clock shows 10:42 AM.

Appendix O: UltraEdit text editor

