

Department of Sociology  
KING'S UNIVERSITY COLLEGE at the University of Western Ontario  
INVESTIGATING THE SOCIAL WORLD: QUANTITATIVE RESEARCH  
SOCIOLOGY 3306 SECTION 570

Winter 2018  
[Research Project](#)

*Introduction:*

Since one of the goals of this course is to teach you how to conduct quantitative research, it is important that you get some hands-on experience. As a result, you are required to carry out a small research project using either the Canadian Census (2006 Public Use Individual File), the Canadian Community Health Survey (2010 CCHS), the General Social Survey (2009 GSS: Victimization Survey), or the 1994 NLSCY (National Longitudinal Survey on Children and Youth). I have selected these datasets to make your life easier; as you can select one of these datasets and work with it all term long for all of the SPSS assignments as well as your final analytical assignment (due at the end of the term). *I've obtained these datasets from the **ODESI Data Library** which is available to senior undergraduate students at Western, as well as faculty and graduate students on the main campus.* It is possible to download off of *ODESI* literally 1000s of datasets on innumerable topics (data that has already been collected, typically via surveys by other academics, Statistics Canada, NGOs, other government agencies, international organizations, etc.). I've downloaded the four files, and formatted them to SPSS for your convenience (your datasets are now up and ready to go!)

The minimum length of this final assignment (paper) is 12 pages (*excluding* tables, references, and title page) double spaced. Feel free to give me a longer write-up, if you so choose.

**Available DATA!!**

Early on you **MUST** select the major “*DEPENDENT*” variable for your study. In other words, what do you hope to “**EXPLAIN**” in your research? My website has listed the user manuals for the four datasets available to you in this course (which includes a listing of the frequency distributions on all available variables).

<http://dkerr.kingsfaculty.ca/teach/sociology-3306/data-files/>

**Locate and take a look at these frequency distributions. What do you think would be an interesting “dependent variable” for a term paper?**

**In addition, you can also be thinking about: what other “variables” in this dataset might be associated with or relevant in helping to explain variation in this dependent variable (i.e. what might be the “INDEPENDENT variables” that could be useful in your analysis?) As merely an example from my own area of expertise (as a demographer), assume you pick childbearing (total number of children) as your “DEPENDENT” variable. What sorts of “INDEPENDENT “variables might be relevant in explaining differences across individuals**

**in this sample in terms of “no. of children”? (e.g. education of parents, labor force participation, ancestry).**

**I will have a short meeting with you on either Friday January 26th or Monday January 29th to verify that you are moving in the right direction, so please get a start on this *as soon as possible*. In deciding upon a manageable topic, also feel free to consult the SPSS consultant (Donna Maynard: her lab hours have been posted on the assignment page of my website). Hint (helps with motivation): you can try to locate a subject that you find to be particularly interesting or important. What variable might be an interesting “DEPENDENT” variable for your analysis?**

### ***The Canadian Census:***

Every 5 years Canada has a national census, whereby all Canadians (over 36 million currently) are contacted. One in five Canadians is provided the exhaustive 2B questionnaire, whereas 4 in 5 receive the less complex straight forward 2A questionnaire. The dataset available for this course is based on the 2B questionnaire, with detailed information on the family, socioeconomic and demographic characteristics of a random sample of the full population (the sample in this public use file has over 800,000 Canadians).

This dataset is rich with information, allowing for all sorts of possible analyses.

A few examples of variables that could be used as a DEPENDENT variable in your analysis include: Total individual income; Low income after-tax status (Persons); Gross rent; Total number of hours worked for pay in last week; Total years of schooling; Value of dwelling (asked of persons who own their own home); Hours spent doing unpaid housework, etc. Language use; etc.

**Note: I would have included the 2011 public use file rather than the 2006, but the Harper Government canceled the 2B Census in 2011. It has since been re-instated for 2016 but the public use data is not yet available to the public.**

### ***The Canadian Community Health Survey (CCHS):***

This cross-sectional survey collects information related to health status, health care utilization and health determinants for the Canadian population. It includes information on a wide range of topics, including alcohol consumption and dependence, chronic health conditions, fruit and vegetable consumption, general health, use of health services, height and weight, injuries, physical activity, restriction of activities and smoking. It also provides information on the socio-demographic characteristics, the income and the labor force characteristics of the population. Data were collected from over 130,000 respondents, aged 12 or older, residing in households across all provinces and territories.

As merely a few examples of some of the variables available as possible dependent variables: self perceived health; BMI classification; underweight, normal weight, overweight, obese; # of days work missed due to disability over last 2 weeks; # of consultations with family doctor; # of chronic conditions reported; Daily consumption of fruit/vegetables; index of physical activity/

energy expended; number of cigarettes smoked daily; number of drinks, self-esteem scale; satisfaction with life scale (1-10); health utilities index, # of consultations with dentist, illicit drug use scale (illicit drug interference), # of close friends, social support scale (20 categories), depression scale (10 categories), daily consumption of food indicator (different types), monthly frequency of physical activity duration of breastfeeding, etc.

**IMPORTANT POINT: WITH STATISTICS CANADA'S SURVEYS NEVER USE AS A VARIABLE IN YOUR ANALYSIS A VARIABLE WHICH INCLUDES IN ITS TITLE THE TERM "INCLUSION FLAG" (for example, with this survey: **Activities of daily living - Inclusion Flag - (F)** **Satisfaction with daily living - Inclusion Flag - (F)**; **Drug use – Inclusion Flag, etc).** These are not "Variables" and cannot be used as such in your analysis. These are in fact "filters" that have been included in the dataset to identify "subsamples" for further analysis..**

### *General Social Survey (GSS): Victimization:*

This is a particularly interesting dataset for all you aspiring "*criminologists*", and it looks very promising.. take a look.

The 2009 GSS is a survey of over 19,000 Canadians, aged 15 and older, asking them all sorts of questions relating to crime prevention, perceptions of crime, history and risk to crime, abuse, criminal incident reports, , etc. This survey also has lots of background information, relating to basic socioeconomic characteristics.

As examples of the sorts of variables available, there is information on: "Number of times one goes out in the evening", "frequency of drinking", "level of stress (6 point ordinal scale), "overall life satisfaction "(10 point scale), "frequency of drug use (5 or 6 point scale, I think), # of weeks worked, # of hours worked, # of close relatives, # of years of education, whether one has experienced a specific type of crime: vandalism; violence; theft, etc. Alternatively, you can create a scale using the information available in this dataset (see assignment 1 for details) on the "quality of one's neighborhood" (using items on vandalism, loitering, etc), the "effectiveness of police" (there are several questions on this), the "effectiveness of the courts", among many other possibilities. Merely look for interrelated questions, and it is potentially possible to create a scale (while Assignment 1 shows you how to create a scale, feel free to consult Prof. Maynard or myself, if you feel uncertain or are having difficulties in creating such a scale).

### *NLSCY*

The National Longitudinal Survey of Children and Youth (NLSCY) is a long-term survey designed to measure child development and well-being. The posted dataset is exclusively the first cycle (1994) of this longitudinal survey. The target population of the NLSCY for Cycle 1 consisted of Canadian children aged newborn to 11 years of age (and their families). The total sample size is 22,831, which is relatively large by most standards. The first cycle of the survey was conducted by Statistics Canada in 1994-1995 on behalf of Human Resources Development Canada. In this course we limit ourselves to data from this first wave (i.e. the longitudinal

dimension of this data is not available, as this would require more than 1 wave and multiple datasets).

Before the NLSCY was undertaken there were few statistical studies describing a broad range of characteristics of children in Canada. Measures of health and child well-being are available, as is information on many of the characteristics of the families in which the children are being raised. Data is available on the socioeconomic status of parents, their parenting style, their marital status, the degree of marital conflict, among various other factors often associated with the well-being of children in Canadian households. The following is a partial listing of variables from the NLSCY that you could potentially use in your research: Hyperactivity/Inattention scale for child; Prosocial Behavior scale of child; Emotional disorder/ anxiety of child; Aggressive behavior of child; Indirect aggression (verbal) of child; Depression scale for parent (person most knowledgeable); Family Functioning; Social support scale for parent (person most knowledgeable); Consistent parenting style; Punitive parenting style; Hostile parenting, etc.

There are several other behavioral scales that could also be potentially useful in this dataset, as well as background information on the characteristics of the children (age, sex), parents (age, sex, SES, education, marital status, etc) and family/households (family structure, lone parenthood, family income, etc).

### **Expectations:**

Further details will be forthcoming on what sort of analysis is possible for this final research project (you shall be working with either linear or logistic regression analysis to conduct a multivariate analysis; one dependent variable and several independent variables). Yet do note: your analysis in your final paper will build directly on Assignment # 3 (due in November).

### **Structure of final paper:**

Your final analytical paper must include:

1) Title page: Include the project title, your name and student ID number, and course section number. The title should clearly state the goal of the research.

2) Brief introductory paragraph (1/2 page max) where you tell me what the paper is all about. For clarity purposes, this is sound practice for all of your papers in sociology. In writing this paper, please be concise and avoid being repetitive.

3) Literature Review (3 pages max): The literature review should discuss previous research done on the topic you have chosen. The utility of the literature review is that it situates your current research within the broader context of previous research. What dependent variable have you selected, and what variables do you think to be particularly important in explaining variation in this dependent variable. Prior to developing this literature review, it will be necessary for you to (i) familiarize yourself with the variables that are available in the available datasets, and (2) to do some library research in familiarizing yourself with relevant literature. There is obviously no point in developing an elaborate literature review if you cannot operationalize it in the GSS,

CCHS, Census or NLSCY (so obviously your literature review should not focus on variables that are unavailable in this dataset),

For the purposes of the current project, I want you to locate at least 5 relevant articles from sociological journals and/or book chapters that are directly related to sorts of relationships that you are planning on studying. In coming up with a research topic, the goal here is to relate this topic to what has been previously published. That is, why is it interesting to carry out your study? What does it contribute? In putting together a literature review, it is expected that you end it by stating the primary research question(s) that you hope to examine, and to situate this research question within the context of previous research. In other words, it should be obvious to the reader what the primary “dependent” variable of your analysis is and how this variable is likely associated with at least one other “independent” variable of interest.

#### 4) Research Questions or Hypotheses (2 pages)

Here we move from our literature review to explain exactly what it is that you are going to explore and test, i.e. again clearly state the primary research question(s) or hypothesis(es) that you will examine in this final paper. In this section, it should be made explicit as to what the dependent variable will be in your analysis (note: select only one dependent variable). Similarly it should be clear as to what sub-sample of Canadians you will be focusing on and the unit of analysis involved (the child, the family, the parent of the child). You must specify what independent variables are of particular importance in explaining this dependent variable. While I don't expect you to be completely original in developing this research questions, I do hope that you can come up with something theoretically interesting, both to you personally and also in light of the previous literature.

In terms of relevant independent variables, you will ultimately be selecting a minimum of four or five distinct variables (details forthcoming in assignment 3). In other words, if your focus is on the relationship between one independent variable and your dependent variable, you should also select an additional three or four additional independent variables that can serve as control variables in your analysis.

#### 5) Method (3 pages)

Include a short description of the GSS, CCHS, Canadian Census or the NLSCY, and a rationale for using it. Here it is important that you explicitly state which variables will be used to test your hypotheses (do not give me the “technical names” as specified in the dataset). In particular, you must clearly identify your dependent variable and independent variables. You will be expected to carry out a multivariate analysis, so it is important that you identify not only the independent variables directly related to your research questions, but also the variables that will be included as control variables. Finally, describe the questions from the survey as well as the variable names in the dataset. All of this must be done in “**paragraph format**”, i.e. your paper should not merely list the variables and categories. Be concise and brief in summarizing the relevant variables. Also briefly state the method that you will be using in testing these hypotheses (which will be either linear (OLS) regression or logit analysis: you shall learn how to do this through the required assignments for this course).

## 6) Results (2 pages)

Very briefly, interpret the results from your regression indicating whether or not your hypothesized relationships are supported empirically. Details on how to do this follow from your assignments. This section will basically present the results from a multiple regression. In other words, you will present tables and/or figures and discuss the major findings seen within them. The goal in this section is not to speculate why you have found a certain result, but simply to report them. It is good practice to refer back to the research questions or hypotheses you stated earlier. Make sure that you label and title your tables and/or figures and that within the text you clearly refer to the labels.

## 7) Discussion and Conclusions

In this section you should interpret the results. Are your findings noteworthy in any way? Do they differ from what others have found? Make sure to refer to similarities and differences compared to previous research, citing the appropriate studies. You will also want to discuss limitations (possible biases) of your study, and ways in which future research could expand on what you have done (1-2 pages).

## 8) References:

Include only sources that you cite in the assignment. Reference your citations using the American Sociological Review (ASR) guidelines, both within the text and in the reference list. These guidelines can be found by looking at any issue of the ASR or at:

Citation style

<http://www.lib.uwo.ca/files/styleguides/ASA.pdf>

Citations should be referred to in the text as (White, 2000) and citations should be listed in your bibliography at the end of your paper (do not use footnotes for references).

**NOTE: We must be very careful in getting material off the internet if it has not been academically reviewed.** i.e. is the source a reliable source? To be safe, try not to rely too heavily on the internet (i.e. resources that do not come to you via the UWO library or through legitimate academic or government agencies, Stats Can, the UN, etc.). In other words, you cannot be guaranteed that what you read has been properly refereed as are academic journals, and may include seriously biased, opinionated or inaccurate information. We should also be very skeptical of reports summarized by journalists in newspapers, among other media sources.

## 5) Format:

The analytical paper must be typed using 12-point font (times roman) and double spaced. I expect the assignments to be very well written, well organized and clearly thought out. I suggest that you do not put the assignment together hastily—such action will likely result in a very poor outcome. Not including references and the title page, you must write 12 pages minimum). Again,

this does not include the reference list, the title page, tables and figures.

It is also important to proof read your work carefully. For example, there is no excuse for having a large number of spelling errors. Although it can be helpful, do not blindly rely on the spell checker to correct spelling errors. **I will be marking this “final assignment” myself, and will be providing constructive feedback.**

This final paper should take the form of what sociologists often refer to as a “research note” (i.e., it is a small research paper). As an example, take a look at the paper written by “Beaujot and Anderson” in your course package “Stress and Adult Health: Time Spent in Paid and Unpaid Work, and its Division in Families”. Also, as mentioned in class, students who have taken this course in the past have won the departmental “essay award” at Kings (take a look at the papers prepared by either: Bosworth, Brown, Waite or Wiersma). All 4 of these papers are of exceptional quality and they provide you with a very good sense as to style and format in preparing this sort of document. Do note that all four of these papers used the NLSCY for the sole reason that this was the only dataset that I had made available in past years. These papers are linked at:

<http://www.kings.uwo.ca/academics/sociology/resources-and-information/sociology-department-academic-awards/>

In addition, you will find many similarly formatted research notes in sociological journals.

The project will require computer analysis using SPSS. In many ways, this paper will build on what you learned in Sociology 2205/2206 and the SPSS assignments that you are required to complete throughout this term. In this paper, you will set out to test the hypotheses or research questions that follow your literature review. Again, this is an independent project—although of course this does not mean you cannot ask others for advice—this is an important part of conducting good research—but it means you must do the work yourself.

These papers are due the final week of classes (I am always impressed by the work shared by Kings students!)

GOOD LUCK!!