

**SOCIOLOGY 328 – SOCIAL STATISTICS I (DRAFT)**  
**SEP - DEC, 2015**

**Instructor:** Dr. Gerry Veenstra  
**Office:** ANSO 1321  
**Phone:** (604) 822-4351  
**Email:** [gerry.veenstra@ubc.ca](mailto:gerry.veenstra@ubc.ca)

**Teaching Assistant:**  
**Office:**  
**Email:**

---

## **Purpose**

This course introduces students to elementary techniques of quantitative data analysis common in sociological research. It emphasizes selection of appropriate statistical techniques, examination of assumptions associated with them and interpretation of the results provided by them. The course will not emphasize calculation and will not involve memorization of complex formulae. It might even be fun on occasion.

## **Prerequisite and Anti-requisites**

SOCI 100 or its equivalent is a prerequisite for this course. Note that UBC students cannot receive credit for any two of the following introductory statistics courses: STAT 200, 203, BIOL 300, COMM 312, 291, ECON 325, EPSE 482, 483, FRST 231, GEOG 374, KIN 371, POLI 380, PSYC 218, 366, SOCI 328.

## **Format**

We meet 12:30–13:50 pm T/TH in Frederic Lasserre 102. Classroom time will entail a mixture of lectures, exercises and tutorials.

## **Textbook (required)**

Garner, Roberta. 2010. *The Joy of Stats. Second Edition*. Toronto: The University of Toronto Press. The website for the textbook ([www.garnerjoyofstats.com](http://www.garnerjoyofstats.com)) has useful supplementary material. Please note that the lectures contain material that is not contained in the textbook and vice versa.

## **Evaluation**

Evaluation will be based on four assignments (25%), a slew of pop quizzes (10%), a midterm exam (25%) and a final exam (40%). There will be no opportunities for extra credit. Final grades may be scaled if the average final grade for the course is inordinately high or low.

## **Policy on Missed Classes**

Please note that there are no marks allocated for attendance or participation. Students are, however, expected to attend all classes. **Because of the cumulative nature of the course, misunderstanding can compound quickly and students who miss class or do not keep up with the readings will have difficulty catching up.** Students who encounter medical, emotional or personal problems that affect their attendance or academic performance should contact the Faculty of Arts Academic Advising Services, located in Buchanan D-111, phone (604) 822-4028. Please refer to the UBC Calendar for a more thorough discussion of academic concession. UBC accommodates students with disabilities who have registered with the Disability Resource Centre. The university also accommodates students whose religious obligations conflict with attendance, submitting assignments, or completing scheduled tests and examinations. A list of religious holidays involving fasting, abstention from work or study or participation in religious activities is available on the UBC website. Students should let the instructor know in advance (in the first week of the course) if they will require accommodation on these grounds. Students who plan to be absent for varsity athletics, family obligations or other commitments should not assume they will be accommodated and should discuss their commitments with the instructor early in the term.

## **Policy on Assignments**

The assignments involve implementing statistical techniques with real data and interpreting the results. Late assignments will be penalized 10% per day (a weekend counts as one day). Assignments handed in after 3:00 p.m. on the due date are deemed to be one day late. Assignments should be handed to the course instructor or T.A. or at the SOCI main office (not to someone else, not by email, etc.).

## **Policy on Examinations**

The exams will be comprised primarily of multiple choice and short answer questions. Students are expected to write examinations on the scheduled dates. A make-up exam will be scheduled only if Arts Advising formally indicates that this is appropriate.

## **Software and Data**

The best way to learn statistics is to do statistics which means messing around with real data. Students and instructor will use a statistical software package called Stata which is available in Buchanan computer labs B101, B121, B123, B125 and B126. We will use Stata and data provided by the instructor for classroom examples, exercises and assignments.

## Topics

### Unit 1 – Introduction

- *Topics*: thinking statistically; conditions for causality; multivariate causal scenarios
- *Readings*: Preface pp. 17–27 and Chapter Four pp. 159–165

### Unit 2 – Variables

- *Topics*: levels of measurement; characteristics of variables
- *Readings*: Chapter One pp. 29–45

### Unit 3 – Describing a categorical variable

- *Topics*: frequency distributions; pie charts and bar charts
- *Readings*: Chapter Two pp. 55–57 and pp. 70–73

### Unit 4 – Describing a continuous variable

- *Topics*: measures of central tendency (mean, median, mode); measures of dispersion (range, inter-quartile range, standard deviation); shapes (histograms, stem-plots, box-plots)
- *Readings*: Chapter Two pp. 58–67 and pp. 73–79

### Unit 5 – Relationship between two categorical variables

- *Topics*: cross-tabulations; % difference; Cramer's V
- *Readings*: Chapter Four pp. 153-155 and pp. 191–203

### Unit 6 – Relationship between two continuous variables

- *Topics*: scatterplots; Pearson's  $r$  and Spearman's rho; regression equation
- *Readings*: Chapter Four pp. 166–186

### Unit 7 – Relationship between a categorical variable and a continuous variable

- *Topics*: means- and medians-based analyses; regression equation
- *Readings*: none

### Unit 8 – Multiple regression

- *Topics*: investigating multivariate causal scenarios with multiple regression
- *Readings*: Chapter Four pp. 186–190

### Unit 9 – Sampling

- *Topics*: descriptive and inferential statistics; types of samples
- *Readings*: Chapter One pp. 46–52

### Unit 10 – Probability

- *Topics*: randomness; probability models
- *Readings*: none

**Unit 11 – Statistical inference: Tests of significance**

- *Topics*: hypothesis testing; Chi-square test of significance; significance in regression
- *Readings*: Chapter Three pp. 128–129 and pp. 135–142

**Unit 12 – Statistical inference: Confidence intervals**

- *Topics*: confidence intervals for means and proportions; confidence intervals in regression
- *Readings*: Chapter Three pp. 130–135

*<Final exam in formal exam period covering entire course>*

### **Assignment 1**

Select a categorical (nominal or ordinal) variable from the Country dataset or GSS Cycle 24 dataset. What is the variable measuring and what is its level of measurement, do you think? Next, using statistical techniques covered in the lectures and textbook, describe and summarize the distribution of values for the variable using Stata. Feel free to copy and paste output tables from Stata into your assignment. Can you provide any insights regarding the distributions of values for this variable? Now repeat the process with a continuous (interval or ratio) variable in place of the categorical one. Do not examine variables that were covered in detail in class. Type your double-spaced assignment using Times Roman 12-point font.

### **Assignment 2**

In this assignment you will use Stata to investigate a bivariate relationship between two categorical variables in the Country dataset or GSS Cycle 24 dataset.

1. Provide a theoretical rationale to guide your analysis. Why would you expect the two variables to be related? How would you expect them to be related? Are they likely to be strongly or weakly related, do you think? Is one likely to cause or influence the other, i.e., does it make sense to designate one as the independent variable and the other as the dependent variable?
2. As far as you can tell, what, exactly, do the variables measure or assess? What are their levels of measurement? What do their distributions look like in this dataset?
3. Using statistical techniques covered in the lectures and textbook, describe and summarize the relationship between the two variables. Did you have to recode or transform a variable first? How are the variables related to one another (if at all)? If they *are* related, is the relationship strong or weak (or something in between)?
4. Provide some interpretive insights regarding the relationship between the variables. Were your theoretical expectations met?

Present your analyses and insights in sentence and paragraph form (accompanied by attractive graphs and tables) as if you were writing a formal report for public consumption. Do not examine a relationship that was covered in detail in class. Include an appendix that lists the full set of commands used to execute your analysis. Type your double-spaced assignment using Times Roman 12-point font.

### **Assignment 3**

In this assignment you will investigate a bivariate relationship between two continuous variables from the Country dataset or GSS Cycle 24 dataset. Do not examine a relationship that was covered in detail in class. Follow the instructions provided for Assignment 2.

### **Assignment 4**

In this assignment you will investigate a bivariate relationship between a categorical variable and a continuous variable in the GSS Cycle 24 dataset. Do not examine a relationship that was covered in detail in class. Follow the instructions provided for Assignment 2.

**Note:** You are permitted to work in teams of two or three people and hand in one version of an assignment on behalf of both or all of you. (You can also work alone if you so desire.) Please ensure that co-authored assignments are truly co-authored.