

**GRAD-C5: Statistics I**

Dr. Svanhildur Thorvaldsdottir

**1. General information**

Class time	Group A - Tuesdays 10-12 Group B – Tuesdays 18-20 Coffee and statistics open hour - XXX			
Course Format	The Statistics I lecture will be taught online. The lecture will also be recorded for later viewing. Small participatory labs will accompany the lecture and will be taught onsite at the Hertie School and via an online platform. For those attending online labs, the platform allows for an interactive, participatory style of teaching. The coffee and statistics open hour is intended to allow students to ask questions about course material in a less formal setting than lecture.			
Instructor	Dr. Svanhildur Thorvaldsdottir			
TAs	Lab 1	Maria Uttenthal (Lead TA)	Fri, 10-11:30h	Room 2.34
	Lab 2	Maria Uttenthal (Lead TA)	Fri, 16-17:30h	Room 2.61
	Lab 3	Nicolai Berk	Wed, 12-13:30h	Room 2.30
	Lab 4	Nicolai Berk	Wed, 16-17:30h	Room 2.30
	Lab 5	Francisca Castro G.	Wed, 18-19:30h	online only
	Lab 6	Francisca Castro G.	Fri, 14-15:30h	Room 2.32
	Lab 7	Joshua Cova	Wed, 10-11:30h	Room 3.30
	Lab 8	Joshua Cova	Wed, 16-17:30h	Room 2.61
	Lab 9	Korinna Lindemann	Wed, 08-09:30h	Room 2.30
	Lab 10	Korinna Lindemann	Fri, 14-15:30h	Room 2.30
	Lab 11	Cecilia Trasi	Fri, 10-11:30h	Room 2.30
	Lab 12	Cecilia Trasi	Fri, 12-13:30h	Room 2.30

	Lab 13	Cecilia Trasi	Fri, 16-17:30h	Room 2.30
	Lab 14	Sukayna Younger-Khan	Fri, 12-13:30h	Room 2.32
	Lab 15	Sukayna Younger-Khan	Fri, 16-17:30h	Room 2.32
	Lab 16	Carol Sobral	Wed, 10-11:30h	online only
	Lab 17	Carol Sobral	Fri, 08-09:30h	online only
Instructor's office				
Instructor's e-mail	<a href="mailto:thorvaldsdottir@hertie-school.org">thorvaldsdottir@hertie-school.org</a>			
Instructor's phone number				
Assistant	Name: Ayamba Kwoyila Email: <a href="mailto:kwoyila@hertie-school.org">kwoyila@hertie-school.org</a>			
Instructor's Office Hours	Wednesdays 15- 16. Please email Ayamba Kwoyila ( <a href="mailto:kwoyila@hertie-school.org">kwoyila@hertie-school.org</a> ) for appointments and link.			

Link to MIA and MPP [Module Handbooks](#)

Link to [Study, Examination and Admission Rules](#)

#### Instructor Information:

**Svanhildur Thorvaldsdottir** is a Senior Lecturer at the Hertie School. She has previously worked as a postdoctoral researcher at the Chair of Political Systems and European Integration at the Ludwig Maximilians University (LMU) in Munich and as a computational social scientist at the Chair of International Relations at the Hochschule für Politik at the Technical University of Munich. At the LMU, she studied the bureaucracies of international organizations (IOs) as part of the DFG Research Unit "International Public Administration" and at the TUM, she was (and remains) a Co-PI of the CoronaNet Research Project, which is mapping and coding government policies globally in response to Covid-19. She has published articles and book chapters on various aspects of budgeting and finance in international organizations. Svanhildur received her PhD in political science from the University of Rochester in New York and holds a BA degree in political science and a Master of International Affairs degree from Columbia University. Prior to her Ph.D. studies, she was a Senior Policy Analyst at the International Peace Institute, a New York-based think tank with strong ties to the United Nations, and a secretary in the Icelandic Foreign Service, serving in the Icelandic Embassy in Washington, D.C

## 2. Course Contents and Learning Objectives

#### Course contents:

This course offers an introduction to quantitative research methods for public policy and international affairs. By the end of the semester, students should have a better acquaintance with quantitative methods conducive to empirical policy research and the ability to understand and critique the techniques employed by others. No prior knowledge of statistics is assumed and we will make every effort to address concepts both quantitatively and theoretically. Additionally, every

student will attend a weekly lab session run by a teaching assistant in which concepts will be reinforced, where possible, through hands-on application using R.

Main learning objectives:

The goal of this course is not only to produce sophisticated consumers of quantitative research but to provide students with the necessary skills to conduct their own research. Thus, we spend several weeks on multivariate regression to enable students to assess situations of multiple causation. Attendance in both lecture and lab is mandatory. I urge students to keep up with the readings since the information in this class is highly cumulative.

Labs

Curricular Affairs will initially allocate you to labs and then allow you to switch – by swapping with other students – in order to avoid conflicts with your schedules. If you would like to switch labs, please use the switching partner forum on Teams to connect with your fellow students. Once you find a switching partner, please fill in the form “switching course request” on MyStudies. It is your responsibility to make sure that by switching courses you do not create time clashes with your other courses.

Teaching style:

This is a lecture course that focuses on instilling an intuitive understanding of fundamental concepts of statistics through the use of examples, visual representations and elementary mathematics.

Prerequisites:

This is an introductory statistics course that should make statistics accessible to all student willing to invest the necessary effort. No prior knowledge of statistics, experience with R or mathematics above the high school level is necessary.

### 3. Grading and Assignments

Composition of Final Grade:

<b>Homework Assignments</b>	Due: weekly	Submit via Moodle	25%
<b>Mid-term Exam</b>	Week of 18-22 Oct		20%
<b>Final Exam</b>	Week of 13-17 Dec		20%
<b>Final Data Analysis</b>	Deadline: 23:59 (Berlin time) on Monday, 20.12.21		25%
<b>Participation in Labs</b>			10%

Assignment Details

**Homework Assignments**

Weekly problem sets. There will be a total of 10 problem sets that students are required to complete. To give everyone time to get up to speed with the systems we are using and to iron out any logistical kinks, the first problem set will be ungraded. The remaining nine problem sets will all be graded; however, when calculating the final grade, the lowest score will be dropped.

**Final Data Analysis**

The two exams cover concepts and applications in statistics and research design but do not cover the use or mastery of R. To provide students with an incentive and opportunity to develop their R

skills, I will distribute a data set and questions toward the end of the semester for the students to analyze using R. This assignment, similar to a take-home exam, must be completed individually.

### **Participation grade**

The participation grade is based on the assumption that students take part not as passive consumers of knowledge but as active participants in the exchange, production, and critique of ideas—their own ideas and the ideas of others. Therefore, students should come to class not only having read and viewed the materials assigned for that day but also prepared to discuss the readings of the day and to contribute thoughtfully to the conversation. Participation is marked by its active nature, its consistency, and its quality.

**Late submission of assignments:** For each day the assignment is turned in late, the grade will be reduced by 10% (e.g. submission two days after the deadline would result in 20% grade deduction).

**Attendance:** Students are expected to be present and prepared for every class session. Active participation during lectures and seminar discussions is essential. If unavoidable circumstances arise which prevent attendance or preparation, the instructor should be advised by email with as much advance notice as possible. Please note that students cannot miss more than two out of 12 course sessions. For further information please consult the [Examination Rules](#) §10.

**Academic Integrity:** The Hertie School is committed to the standards of good academic and ethical conduct. Any violation of these standards shall be subject to disciplinary action. Plagiarism, deceitful actions as well as free-riding in group work are not tolerated. See [Examination Rules](#) §16 and the Hertie [Plagiarism Policy](#).

**Compensation for Disadvantages:** If a student furnishes evidence that he or she is not able to take an examination as required in whole or in part due to disability or permanent illness, the Examination Committee may upon written request approve learning accommodation(s). In this respect, the submission of adequate certificates may be required. See [Examination Rules](#) §14.

**Extenuating circumstances:** An extension can be granted due to extenuating circumstances (i.e., for reasons like illness, personal loss or hardship, or caring duties). In such cases, please contact the course instructors and the Examination Office *in advance* of the deadline.

#### 4. Session Overview

Session	Session Date	Session Title
1	07.09.2021	Introduction
2	14.09.2021	Research design
3	21.09.2021	Descriptive statistics
4	28.09.2021	Probability
5	05.10.2021	Statistical inference I
6	12.10.2021	Statistical inference II
<b>Mid-term Exam Week: 18 – 22.10.2021 – no class</b>		
7	26.10.2021	Categorical variable analysis
8	02.11.2021	Linear regression
9	09.11.2021	Multivariate modeling I
10	16.11.2021	Multivariate modeling II
11	23.11.2021	Multivariate modeling III
12	30.11.2021	Overview and review
<b>Final Exam Week: 13 – 17.12.2021 – no class</b>		

#### 5. Course Sessions and Readings

Required readings are to be read and analyzed thoroughly. Optional readings to give more in-depth information about specific methods may be added over the course of the semester. Such readings are helpful for better understanding of course material, but not required. In the case that there is a change in readings, students will be notified by email.

Required readings:

- Agresti, Alan. 2018. *Statistical Methods for the Social Sciences*, 5<sup>th</sup> edition. Pearson. Earlier editions are fine, as the book changes very little between editions. Note that section numbers and pages may be different across editions.
- Paul Kellstedt and Guy Whitten. 2018. *The Fundamentals of Political Science Research*. Cambridge University Press. 3<sup>rd</sup> edition.

<b>Session 1: Introduction</b>	
<b>Learning Objective</b>	Introduction: motivation and overview. Sampling, randomness and measurement
<b>Required Readings</b>	Agresti Ch 1-2

### Session 2: Research design

<b>Learning Objective</b>	Causality and spuriousness, experimental vs observational designs, forming hypotheses, internal and external validity.
<b>Required Readings</b>	Kellstedt and Whitten, Ch 3 (sections 3.1/3.2 only), Ch 4

### Session 3: Descriptive statistics

<b>Learning Objective</b>	Levels of measurement, measures of central tendency; descriptive statistics: measures of dispersion, skew; frequency distributions
<b>Required Readings</b>	Agresti Ch 3

### Session 4: Probability

<b>Learning Objective</b>	Probability and introduction to inference. Central Limit Theorem, normal distribution, z-scores
<b>Required Readings</b>	Agresti Ch 4

### Session 5: Statistical inference I

<b>Learning Objective</b>	Statistical inference: estimation. The t-distribution, Confidence Intervals, Sample Proportions & Polling
<b>Required Readings</b>	Agresti Ch 5

### Session 6: Statistical inference II

<b>Learning Objective</b>	Inference Review; Hypothesis testing, testing the difference between two groups
<b>Required Readings</b>	Agresti Ch 6 and 7

**Mid-term Exam Week: 18 – 22.10.2021 – no class**

### Session 7: Catagorical variable analysis

<b>Learning Objective</b>	Categorical Variable Analysis: Cross-tabs, Chi-sq, Categorical Variable Analysis: Controlling for a third variable
<b>Required Readings</b>	Agresti Ch 8; Meier et al Ch 17

### Session 8: Linear regression

<b>Learning Objective</b>	Regression: Correlation, Bivariate Regression; Statistical inference and OLS Assumptions
<b>Required Readings</b>	Agresti Ch 9

### Session 9: Multivariate modeling I

<b>Learning Objective</b>	Multiple Regression: the basics. Interpretation, OVB, dummy variables, regression tables
<b>Required Readings</b>	Kellstedt and Whitten Ch 10, ch 11.1-11.2.

### Session 10: Multivariate modeling II

<b>Learning Objective</b>	Multiple Regression: specification. Interactions, time-series and panel data, polynomials, log-transformations.
<b>Required Readings</b>	Agresti Ch 14.5-14.6, KW 11.3, 12.3

### Session 11: Multivariate modeling III

<b>Learning Objective</b>	Multiple Regression: diagnostics. Outliers and influential cases, multicollinearity, heteroscedasticity and robust standard errors, partial regression plots. Limited Dependent Variables: sneak peek at logits.
<b>Required Readings</b>	Agresti Ch 14.2, KW 11.4-11.5

### Session 12: Overview and review

<b>Learning Objective</b>	Overview, and review Course evaluations will take place during the last session and will require 15 minutes.
<b>Required Readings</b>	

**Final Exam Week: 13 – 17.12.2021 – no class**

**Final Data Analysis: Due at 23:59 (Berlin time) on the Monday following the final exam**