Class Meetings: See the residency schedule for specific class dates and times during residency week.

Instructor: Tania M. Alarcon Falconi, PhD
Email: tania.alarcon_falconi@tufts.edu
Office hours: By appointment.

Teaching Asst.: Robel Alemu
Email: robel.alemu@tufts.edu
Office hours: By appointment.

Ryan Simpson
Email: ryan.simpson@tufts.edu
Office hours: By appointment.

Semester Hour Units: 3.0

Prerequisites: Statistical Methods for Health Professionals I (NUTB-250) or equivalent, and graduate standing or instructor consent. Ability to conduct exploratory data analysis using Stata.

Course Description: This course is part two of a one-year, two-semester course sequence in statistics. The course covers experimental and non-experimental research designs, multiple linear regression, multiple logistic regression, polytomous logistic regression, analysis of variance, analysis of covariance, non-linear functional forms, heteroskedasticity, complex surveys, cluster randomized trials, and how these tools are used in the fields of nutrition science and policy. Students will make extensive use of Stata statistical analysis software and learn how to analyze a dataset and report the results in tables, figures, and text.

Course Objectives: Students will learn how to conduct and interpret the most important intermediate level bioscience and social science statistical tests. The overall goal of the course is for students to be able to independently analyze a simple dataset with Stata software and report the results in tables, figures, and text.

Textbooks: Required:

Suggested:

Class Materials: All class materials, including lecture notes and assignments, will be posted on Canvas.

Statistical Software: Stata IC version 15 is required for this course. Most students will have purchased it when they take NUTB250.

Classroom Conduct: Students are required to attend class during the residency period. See the residency schedule for specific class dates and times.
Communication Policy:
For questions on the course, we will use Piazza – a link is available on the course Canvas site. Students should try to seek out information for themselves before contacting the instructor or the TA. If you cannot find your answer, ask a question on Piazza. Students are strongly encouraged to answer others’ questions or take part in any discussion. We will post on Canvas guidelines and expectations for Piazza use during the first week of class.

Academic Conduct:
Each student is responsible for upholding the highest standards of academic integrity, as specified in the Friedman School’s Policies and Procedures Handbook and Tufts University policies (http://students.tufts.edu/student-affairs/student-life-policies/academic-integrity-policy). It is the responsibility of each student to understand and comply with these standards, as violations will be sanctioned by penalties ranging from failure on an assignment and the course to dismissal from the school.

Accommodation of Disabilities:
Tufts University is committed to providing equal access and support to all students through the provision of reasonable accommodations so that each student may access their curricula and achieve their personal and academic potential. If you have a disability that requires reasonable accommodations please contact the Friedman School Assistant Dean of Student Affairs at 617-636-6719 to make arrangements for determination of appropriate accommodations. Please be aware that accommodations cannot be enacted retroactively, making timeliness a critical aspect for their provision.

Diversity Statement:
We believe that the diversity of student experiences and perspectives is essential to the deepening of knowledge in this course. We consider it part of our responsibility as instructors to address the learning needs of all of the students in this course.

Snow emergency:
If the campus closes down due to snow, announcements will be sent via e-mail. You can also check the Tufts Emergency Preparedness website for updates: https://emergency.tufts.edu/weather/closing/
Assessment and Grading:

Grading for the course will be based on the following distribution:

<table>
<thead>
<tr>
<th>Components</th>
<th>Proportion of final score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Class participation</td>
<td>5%</td>
</tr>
<tr>
<td>Quizzes</td>
<td>15%</td>
</tr>
<tr>
<td>Homework</td>
<td>20%</td>
</tr>
<tr>
<td>Exams</td>
<td>30%</td>
</tr>
<tr>
<td>Final project</td>
<td>30%</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>100%</strong></td>
</tr>
</tbody>
</table>

Final course grades will be based on the following (subject to revision during the course):

<table>
<thead>
<tr>
<th>Final score</th>
<th>Letter grade</th>
</tr>
</thead>
<tbody>
<tr>
<td>≥ 97 %</td>
<td>A+</td>
</tr>
<tr>
<td>94 to &lt; 97 %</td>
<td>A</td>
</tr>
<tr>
<td>90 to &lt; 94 %</td>
<td>A-</td>
</tr>
<tr>
<td>87 to &lt; 90 %</td>
<td>B+</td>
</tr>
<tr>
<td>84 to &lt; 87 %</td>
<td>B</td>
</tr>
<tr>
<td>80 to &lt; 84 %</td>
<td>B-</td>
</tr>
<tr>
<td>77 to &lt; 80 %</td>
<td>C+</td>
</tr>
<tr>
<td>74 to &lt; 77 %</td>
<td>C</td>
</tr>
<tr>
<td>70 to &lt; 74 %</td>
<td>C-</td>
</tr>
<tr>
<td>&lt;70</td>
<td>F</td>
</tr>
</tbody>
</table>

Class participation (5%)
Our primary Q&A platform is Piazza. All questions related to the course should be posted on Piazza. Students are also expected to answer others' questions or take part in the discussion within a thread of interest. We will post guidelines and expectations for Piazza use during the first week of class.

Quizzes (15%)
There will be 11 comprehension quizzes that will test the material learned from the online lessons. These multiple-choice quizzes will be available on Canvas. Missing a quiz without prior agreement with the instructor will result in receiving a 0% for that quiz.

Homework (20%)
There will be 3 problem sets assigned throughout the semester. Guidelines for homework will be discussed when each homework is assigned. Homework assignments must be submitted through Canvas on the specified date. Assignments received after their deadline will be penalized by a 20% grade reduction for the first 24 hours and 50% for the next 24 hours. Late submissions will not be accepted if late more than 48 hours. Students who are unable to complete a problem set on time for any reason should contact the instructor by email at least 48 hours prior to the deadline.

Exams (30%)
There will be two exams, each worth 15% of the final grade. Additional information on the format, grading and content of the exams will be distributed prior to each exam.

Final project (30%)
Students will complete an individual data analysis project using a dataset provided by the instructor. Additional details on the project and format will be provided during Week 8. The final written project will be due on Sunday, 5/5/2019 at 5:00 pm. Late submissions will not be accepted unless an extension is approved in advance.
Course Topics and Assignment Schedule at a Glance*:

*This schedule is subject to modification at the instructor’s discretion

<table>
<thead>
<tr>
<th>Week</th>
<th>Date</th>
<th>Topic</th>
<th>Assignments</th>
<th>Required readings</th>
</tr>
</thead>
</table>
| 1    | 1/16-1/20  | Course overview and Ethics in Research CITI Training                  | • Online welcome lesson  
• CITI IRB Ethics tutorial  
• Post in Piazza                                                                 |                                         |
| 2    | 1/21-1/27  | Multiple logistic regression analysis                                  | • Online lesson 1  
• Quiz 1  
• Start reading material for the residency (week 3) | • Campbell 3  
• Vittinghoff 5.1-5.4  
• Mandil et al., 2007 |
| 3 Residency | 1/28-2/3   | Topic I: Polytomous logistic regression                               |                                                                             | • Vittinghoff 5.5.6  
• Rahman et al., 2008  
• Machado et al., 2013 |
|      |            | Topic II: Multiple linear regression                                   |                                                                             | • Vittinghoff 4.1-4.6  
• Campbell 1 & 2                                                                 |
|      |            | Topic III: Stata programming                                           | • Homework 1 assigned                                                     | • Chen et al., 2003a |
| 4 Online | 2/4-2/10   | Multiple regression assumptions                                       | • Online lesson 2  
• Quiz 2  
• Homework 1 due                                                               | • Vittinghoff 4.7  
• Chen et al., 2003b  
• Watch Jbstatistics, 2012a & 2012b |
| 5 Online | 2/11-2/17  | Assessing distributional normality and transforming variables         | • Online lesson 3  
• Quiz 3  
• Exam 1 assigned                                                               | • Vittinghoff 4.7.2  
• Chen et al., 2003b  
• Princeton, Log transformations  
• UCLA-idre, FAQ How do I interpret a regression model when some variables are log transformed?  
• Williams, 2016  
• Watch Delaney, 2012.  
• Watch Jbstatistics, 2012c |
| 6 Online | 2/18-2/24  | Building statistical models, dummy variables, and interaction terms   | • Online lesson 4  
• Quiz 4  
• Exam 1 due                                                                  | • Vittinghoff 4.6  
• Grace-Martin, Practical Guidelines for Accurate Statistical Model Building.  
• Williams, 2015b |
| 7 Online | 2/25-3/3   | Presenting statistical results in tables, figures, and text           | • Online lesson 5  
• Quiz 5  
• Homework 2 assigned  
• Final project assigned                                                    | • Bates, Almost Everything You Wanted to Know About Making Tables and Figures  
• Purdue, APA Tables and Figures  
• UCLA-idre, Collapsing data across observations |
<table>
<thead>
<tr>
<th>Week</th>
<th>Date</th>
<th>Topic</th>
<th>Assignments</th>
<th>Required readings</th>
</tr>
</thead>
<tbody>
<tr>
<td>8</td>
<td>3/4-3/10</td>
<td>Multi-factor Analysis of variance</td>
<td>• Online lesson 6</td>
<td>• Dallal, Multi-Factor Analysis of Variance</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• Quiz 6</td>
<td>• Lane, Chapters 15.3, 15.6, &amp; 15.7</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• Homework 2 due</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>3/11-3/17</td>
<td>Repeated measures analysis of variance and mixed-design ANOVA</td>
<td>• Online lesson 7</td>
<td>• Vittinghoff 7</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• Quiz 7</td>
<td>• Laerd Statistics, Repeated Measures ANOVA</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• Homework 3 assigned</td>
<td>• UCLA-idre, Statistical analyses using Stata</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>• Lane, Chapters 15.9, 20.9, &amp; 20.14</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>• Watch mezonBiz, 2007</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>• Watch StataCorp, 2017</td>
</tr>
<tr>
<td>10</td>
<td>3/18-3/24</td>
<td>Analysis of covariance</td>
<td>• Online lesson 8</td>
<td>• Owen &amp; Froman, 1998</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• Quiz 8</td>
<td>• StataCorp, 2012</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• Homework 3 due</td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>3/25-3/31</td>
<td>Analysis of complex surveys</td>
<td>• Online lesson 9</td>
<td>• Williams, 2015d</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• Quiz 9</td>
<td>• Osborne, 2011</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• Exam 2 Assigned</td>
<td>• Watch StataCorp, 2013a</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>• Watch StataCorp, 2013b</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>• Skim Yang et al., 2009</td>
</tr>
<tr>
<td>12</td>
<td>4/1-4/7</td>
<td>Cluster randomized trials</td>
<td>• Online lesson 10</td>
<td>• Wears, 2002</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• Quiz 10</td>
<td>• University of York, Analysis of a cluster-randomised trial in education</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• Exam 2 due</td>
<td>• Lundeen et al., 2010</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>4/8-4/14</td>
<td>Study design and selection of appropriate statistical methods</td>
<td>• Online lesson 11</td>
<td>• Dallal, Some Aspects of Study Design</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• Quiz 11</td>
<td>• Kendall, 2003</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>• BJM, Study design and choosing a statistical test</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>• Sainani, 2011</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>• Watch Lynch, 2013</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>• Watch Stat Trek. What is an Experiment?, Experimental Design, Data Collection</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Methods, Survey Sampling Methods, Survey Sampling Methods, &amp; Bias in Survey</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Sampling</td>
</tr>
</tbody>
</table>

**Final project due – Sunday, May 5, 2019**
Detailed Description of Course Topics, Assignment Schedule, and the Learning Objectives*:

* This schedule is subject to modification at the instructor’s discretion.

Week 1: Course overview and Ethics in Research CITI Training

Learning Objectives:
This week, students should read the syllabus in detail and post at least one question or comment in Piazza. Students are also expected to complete an on-line Institutional Review Board (IRB) training course and exam for the Protection of Human Subjects via the Collaborative Institutional Training Initiative (CITI) Program.

Upon completion of this week, students should be able to:
1. Reference the syllabus for information on expectations and requirements of the course.
2. Use the syllabus to find information on the structure of the course, including important dates and learning objectives for each class.
3. Use Piazza to post questions or comments about the course.
4. State important ethical issues regarding collecting and analyzing data.
5. Earn a certificate demonstrating knowledge covered by the CITI ethics online tutorial.

Reading/Assignments:
No assigned readings this week.

Assignments Due:
1. Watch the online welcome lesson.
2. Post a question or comment in Piazza.
3. Complete the CITI IRB Ethics tutorial. Submit a copy of your completion certificate via Canvas.

Optional:
1. If you need a refresher of Biostats I, read Vittinghoff 3.1-3.4
2. If you need a refresher on the use of Stata, the Tufts Data Lab has good resources.
Week 2: Multiple logistic regression analysis

Learning Objectives:
This week, we will discuss logistic regression to model dichotomous outcomes and how to interpret the results of those models. We will also discuss multiple logistic regression and how to make predictions from those models.

Upon completion of this week, students should be able to:
1. Describe situations in which logistic regression analysis is needed.
2. Explain the difference between OLS regression and logistic regression.
3. Interpret the results of logistic regression models in the context of odds ratio.
4. Discuss the link between predicted log odds, probability of the event, and odds.
5. Write out the logistic regression equation and use it to make specific predictions.

Required Reading/Assignments:
1. Campbell 3
2. Vittinghoff 5.1-5.4

Assignments Due:
1. Watch the online lesson 1.
2. Take the online quiz 1.
3. Look over the residency files (Week 3) to prepare for those classes.

Optional:
3. Read University of California Los Angeles Institute for Digital Research and Education (UCLA-idre). Stata data analysis examples: logistic regression. [Link]
Week 3: Residency

Topic I: Polytomous logistic regression

Learning Objectives:
We will discuss logistic regression to model polytomous (i.e. multinomial) outcomes and how to interpret the results of those models.

Upon completion of this session, students should be able to:
1. Appreciate the relationship between multinomial and ordered logistic regression models and binary logistic models.
2. Interpret the results of multinomial and ordered logistic regression models in the context of odds ratio.
3. Fit polytomous logistic regression models.

Required Reading/Assignments:
1. Vittinghoff 5.5.6

Optional:
1. Read University of California Los Angeles Institute for Digital Research and Education (UCLA-idre). Stata data analysis examples: multinomial logistic regression. Link.
2. Read University of California Los Angeles Institute for Digital Research and Education (UCLA-idre). Stata data analysis examples: ordinal logistic regression. Link.

Topic II: Multiple linear regression

Learning Objectives:
We will discuss multiple linear regression for prediction and adjustment, and interpretation of regression coefficients. We will also introduce the concept of interaction.

Upon completion of this session, students should be able to:
1. Describe the use of multiple linear regression analysis for prediction and adjustment.
2. Identify the kinds of variables that can be included in a multiple regression analysis.
3. Write the least squares regression equation for a multiple regression model and explain what each term in the equation represents in terms of the research question.
4. Interpret the 95% confidence intervals for the coefficients in a multiple regression model.
5. Interpret an F test in order to determine if any of the explanatory variables are significant predictors of the response variable.
6. Interpret multiple regression coefficients and t tests.
7. Calculate residuals by subtracting the predicted from the observed response.
8. Recognize the impact of unadjusted interaction in a regression model.
9. Construct, test, and interpret an interaction term in a multiple regression analysis.
10. Explain in a sentence or two the interpretation of a multiple regression analysis.
11. Consider the extent to which a particular multiple regression analysis can be generalized to populations from the one from which the sample was derived.
Required Reading/Assignments:
1. Vittinghoff 4.1-4.6
2. Campbell 1 & 2

Optional:

Topic III: Stata programming

Learning Objectives:
We will have a review session on Stata and then work on fitting multiple regression models. We will emphasize conducting all data management and analysis using syntax files that are liberally commented to facilitate reproducibility.

Upon completion of this session, students should be able to:
1. Document the work flow in the form of .do files.
2. Manipulate variables, including creating categorical variables from existing data.
3. Perform multiple regression analysis.
4. Create simple plots including scatter plot, box plot, bar plot, and histogram.

Reading/Assignments:
1. Homework 1 assigned.
Week 4: Multiple Regression Assumptions

Learning Objectives:
This week, we will discuss multiple regression assumptions, how to test those assumptions, and the consequences of violating model assumptions.

Upon completion of this week, students should be able to:
1. Assess whether a multiple regression model meets the assumptions of linearity, homoskedasticity, error independence, and normalcy of the error term.
2. Perform a residual analysis to detect outliers.
3. Describe “correlated disturbances”.
4. Describe “specification error”.
5. Recognize that heteroskedasticity biases standard errors but does not bias regression parameter estimates.
6. Detect and deal with heteroskedasticity when performing a multiple regression analysis.
7. Recognize the consequences of collinearity when fitting regression models.
8. Identify several strategies for dealing with multicollinearity when performing a multiple regression analysis.

Required Reading/Assignments:
1. Vittinghoff 4.7

Assignments Due:
1. Watch the online lesson 2.
2. Take the online quiz 2.
3. Homework 1 due.

Optional:
Week 5: Assessing distributional normality and transforming variables

Learning Objectives:
This week, we will discuss how to assess if a distribution is normal, and techniques to transform non-normal distributions.

Upon completion of this week, students should be able to:
1. Assess distributional normality.
2. Describe variable transformation techniques in linear regression.
3. Identify when and how to use variable transformations.
4. Interpret regression results after variable transformations.
5. Recognize that non-normality causes problems for hypothesis tests of regression parameter estimates in small samples, but not in large samples.
6. Use data transformation and model refinement to deal with violation of the normality assumption.
7. Assess regression diagnostics including residuals, leverage, and Cook’s distance.

Required Reading/Assignments:
1. Vittinghoff 4.7.2
8. Exam 1 assigned.

Assignments Due:
1. Watch the online lesson 3.
2. Take the online quiz 3.
Week 6: Building statistical models, dummy variables, and interaction terms

Learning Objectives:
This week, we will demonstrate how to model different type of interaction such as continuous-continuous, binary-binary, categorical-continuous, and categorical-categorical. We will also discuss how to build statistical models, and the concept of using “dummy” variables to represent categorical predictors.

Upon completion of this week, students should be able to:
1. Use dummy variables to examine categorical predictors in a statistical model.
2. Strategize the testing of interaction terms consisting of variables with different levels of measurement.
3. Model different types of interactions such as continuous-continuous, binary-binary, categorical-continuous, and categorical-categorical.
4. Understand the implications of omitting or including a covariate in a statistical model.
5. Test a research question with an appropriate statistical model.

Required Reading/Assignments:

Assignments Due:
1. Watch the online lesson 4.
2. Take the online quiz 4.

Optional:
Week 7: Presenting statistical results in tables, figures, and text

Learning Objectives:
This week, we will discuss guidelines for creating figures and tables for scientific publications. We will also discuss how to report statistical results.

Upon completion of this week, students should be able to:
1. Summarize descriptive and inferential statistics in tables.
2. Create figures and graphs to facilitate the interpretation of complicated statistical test results such as interactions.
3. Describe statistical test results with appropriate language.

Required Reading/Assignments:
1. Department of Biology, Bates College (Bates). Almost Everything You Wanted to Know About Making Tables and Figures. Link.
2. Purdue University The Writing Lab & The OWL. (Purdue). APA Tables and Figures. Part 1 Link. Part 2 Link.
4. Homework 2 Assigned.
5. Final project assigned

Assignments Due:
1. Watch the online lesson 5.
2. Take the online quiz 5.

Optional:
1. Read Wendt, J.L. How to Make a Powerpoint Presentation Meet the APA Format. Link.
Week 8: Multi-factor Analysis of variance (ANOVA)

Learning Objectives:
This week, we will discuss Analysis of Variance (ANOVA) designs, including between-subjects, within-subjects, and multi-factor designs.

Upon completion of this week, students should be able to:
1. Recognize that analysis of variance is used in situations where the dependent variable is continuous while key explanatory variables are binary or categorical.
2. Describe the difference between a one-way and multi-factor ANOVA.
3. Define the concepts of “main effect” and “interaction term”.
4. Define what is a “marginal mean” and a “grand mean”.
5. Use Two Factor ANOVA and interpret main effects and interaction terms.
6. Perform a factorial two ANOVA and describe the results of the analysis in a few sentences.
7. Describe the difference between “between group variance” and “within group variance”.
8. Recognize the use and interpretation of interaction terms in a two factor ANOVA.

Required Reading/Assignments:

Assignments Due:
1. Watch the online lesson 6.
2. Take the online quiz 6.
3. Homework 2 due.
Week 9: Repeated-measures ANOVA and mixed-design ANOVA

Learning Objectives:
This week, we will discuss repeated-measures and mixed-design Analysis of Variance (ANOVA).

Upon completion of this week, students should be able to:
1. Perform and interpret a one-way repeated measures ANOVA.
2. Describe the similarities and differences between one-way repeated measures ANOVA and one-way between subjects ANOVA.
3. Distinguish between study designs calling for a paired t test versus a one-way repeated measures ANOVA.
4. Perform and interpret a repeated measures ANOVA with more than one independent variable.
5. Recognize that a mixed-design analysis of variance is also known as a split-plot ANOVA.
6. Define mixed design ANOVA as a statistical procedure used to test for differences between two or more independent groups whilst subjecting participants to repeated measures.
7. Perform and interpret a mixed-design ANOVA.

Required Reading/Assignments:
9. Homework 3 assigned.

Assignments Due:
1. Watch the online lesson 7.
2. Take the online quiz 7.

Optional:
Week 10: Analysis of covariance (ANCOVA)

Learning Objectives:
This week, we will discuss analysis of covariance (ANCOVA), including its assumptions and limitations.

Upon completion of this week, students should be able to:
1. Recognize when it is appropriate to use analysis of covariance.
2. Explain what ANCOVA tells us.
3. Perform ANCOVA with Stata and interpret the results.
4. Use Stata to obtain covariate adjusted means.
5. Describe what covariate adjusted mean indicate and how they can be useful.
6. Identify the assumptions and limitations of ANCOVA.

Required Reading/Assignments:

Assignments Due:
1. Watch the online lesson 8.
2. Take the online quiz 8.
3. Homework 3 due.
Week 11: Analysis of complex surveys

Learning Objectives:
This week, we will discuss analysis of survey data and the consequences of failing to account for sampling scheme.

Upon completion of this week, students should be able to:
1. Describe different types of sampling designs including simple random sampling, single stage cluster sampling, and multi-stage cluster sampling.
2. Describe complex survey sample weights.
3. Perform analyses that incorporate sample weights in order to be able to produce adjusted sample statistics that are generalizable to the population of interest.
5. Recognize the implications of failing to account for sample weights and sample clustering when analyzing survey data.
6. Interpret complex survey statistical test results.
7. Produce graphs that display means and proportions that have been adjusted for sample weights.
8. Present results of complex survey data analyses with tables, figures, and text.

Required Reading/Assignments:
1. Williams, R. 2015d. Analyzing Complex Survey Data: Some key issues to be aware of. University of Notre Dame. [Link]
4. Watch StataCorp. 2013b. Specifying the design of your survey data in Stata. YouTube [Link].

Assignments Due:
1. Watch the online lesson 9.
2. Take the online quiz 9.

Optional:
2. Read Centers for Disease Control and Prevention National Health and Nutrition Examination Survey (CDC-NHNES). Generating the Geometric Mean, Standard Error, and Confidence Interval from Stata. [Link].
**Week 12: Cluster randomized trials**

**Learning Objectives:**
This week, we will discuss cluster randomized trials, the strengths and weaknesses of this study design; and how to analyze data from group-randomized trials.

Upon completion of this week, students should be able to:
1. Recognize the purposes of cluster randomized trials in nutrition research.
2. Define the terminology used to describe this research design.
3. Explain the components of group-randomized trials including specification of the research question and selection of the proper design, measures, study populations and analysis procedures.
4. Describe the factors that affect the validity of these trials.
5. Explain the strengths and weaknesses of several design alternatives.
6. Select an appropriate analysis for a particular design.
7. Use suitable Stata software code to analyze data from group-randomized trials.

**Required Reading/Assignments:**

**Assignments Due:**
1. Watch the online lesson 10.
2. Take the online quiz 10.

**Optional:**
1. Read Bharti, S. Standalone use of “STATA” for analysis of cluster randomized controlled trials (cluster RCT). [Link](#).
2. Watch Donner, A. Introduction to CRTs. YouTube [Link](#).
Week 13: Study design and selection of appropriate statistical methods

Learning Objectives:
This week, we will discuss selection of study design and statistical methods that are suitable for each design.

Upon completion of this week, students should be able to:
1. Recognize issues relevant to the selection of appropriate study designs.
2. Recognize study design ethical issues, early termination of trials, and analysis of data arising in such trials.
3. Describe the spectrum of study design models available (e.g., epidemiological vs. double blind intervention), their appropriate use, and their strengths and weaknesses.
4. Formulate a research question and select an appropriate methodology for collecting and analyzing the data.
5. Recognize a taxonomy of research study aims and how it relates to the selection of an appropriate study design.
6. Identify appropriate statistical methods for the analysis of data from specific research study methodologies.

Required Reading/Assignments:
3. BJM. Study design and choosing a statistical test. Link.

Assignments Due:
1. Watch the online lesson #11.
2. Take the online quiz 11.

Optional:
1. Read vietradio. Image Link.

Final project due – Sunday, May 5, 2019
Residency schedule:

<table>
<thead>
<tr>
<th>Boston Time</th>
<th>Sunday 1/27/19</th>
<th>Monday 1/28/19</th>
<th>Tuesday 1/29/19</th>
<th>Wednesday 1/30/19</th>
<th>Thursday 1/31/19</th>
<th>Friday 2/1/19</th>
<th>Saturday 2/2/19</th>
<th>Sunday 2/3/19</th>
<th>Monday 2/4/19</th>
</tr>
</thead>
<tbody>
<tr>
<td>8:00AM-9:45AM</td>
<td></td>
<td>NUTB 305: Biochemistry II</td>
<td>NUTB 350: Statistics II</td>
<td>NUTB 242: Obesity and Energy Regulation</td>
<td>NUTB 211: Behavior Change</td>
<td>NUTB 211: Behavior Change</td>
<td>NUTB 211: Behavior Change</td>
<td>NUTB 211: Behavior Change</td>
<td>NUTB 211: Behavior Change</td>
</tr>
<tr>
<td>9:45AM-10:00AM</td>
<td></td>
<td>Break</td>
<td>Break</td>
<td>Break</td>
<td>Break</td>
<td>Break</td>
<td>Break</td>
<td>Break</td>
<td>Break</td>
</tr>
<tr>
<td>12:15PM-1PM</td>
<td>Lunch</td>
<td>Career Luncheon</td>
<td>Lunch, Friedman Seminar Series</td>
<td>Lunch</td>
<td>Lunch</td>
<td>Lunch</td>
<td>Lunch</td>
<td>Lunch</td>
<td>Lunch</td>
</tr>
<tr>
<td>1:00PM-2:45PM</td>
<td>New Student Orientation Session</td>
<td>NUTB 350: Statistics II</td>
<td>NUTB 242: Obesity and Energy Regulation</td>
<td>NUTB 211: Behavior Change</td>
<td>NUTB 204: Epidemiology</td>
<td>NUTB 204: Epidemiology</td>
<td>NUTB 204: Epidemiology</td>
<td>NUTB 204: Epidemiology</td>
<td>NUTB 204: Epidemiology</td>
</tr>
<tr>
<td>2:45PM-3:00PM</td>
<td>Break</td>
<td>Break</td>
<td>Break</td>
<td>Break</td>
<td>Break</td>
<td>Break</td>
<td>Break</td>
<td>Break</td>
<td>Break</td>
</tr>
<tr>
<td>3:00PM-5:15PM</td>
<td>NUTB 350: Statistics II</td>
<td>NUTB 350: Statistics II</td>
<td>NUTB 242: Obesity and Energy Regulation</td>
<td>NUTB 211: Behavior Change</td>
<td>NUTB 204: Epidemiology</td>
<td>NUTB 204: Epidemiology</td>
<td>NUTB 204: Epidemiology</td>
<td>NUTB 204: Epidemiology</td>
<td>NUTB 204: Epidemiology</td>
</tr>
<tr>
<td>5:30PM-7:30PM</td>
<td>NUTB 242: Obesity and Energy Regulation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>