

University of Montana

## ScholarWorks at University of Montana

---

University of Montana Course Syllabi, 2021-2025

---

Spring 2-1-2022

### ANTY 595.50: Anthropological Statistics

Randall Royce Skelton

*University of Montana, Missoula*, [randy.skelton@umontana.edu](mailto:randy.skelton@umontana.edu)

Follow this and additional works at: <https://scholarworks.umt.edu/syllabi2021-2025>

**Let us know how access to this document benefits you.**

---

#### Recommended Citation

Skelton, Randall Royce, "ANTY 595.50: Anthropological Statistics" (2022). *University of Montana Course Syllabi, 2021-2025*. 75.

<https://scholarworks.umt.edu/syllabi2021-2025/75>

This Syllabus is brought to you for free and open access by ScholarWorks at University of Montana. It has been accepted for inclusion in University of Montana Course Syllabi, 2021-2025 by an authorized administrator of ScholarWorks at University of Montana. For more information, please contact [scholarworks@mso.umt.edu](mailto:scholarworks@mso.umt.edu).

# ANTHROPOLOGICAL STATISTICS

## Syllabus

Instructor information will be posted on Moodle.

### Course Content

This course explores introductory statistical methods and a few things normally considered more advanced methods. The focus will be on the practicalities of doing statistical analyses using a freeware statistical software program, with what I consider adequate background theory to allow students to run the analyses and interpret the results.

### Course Objectives

Students who successfully complete this course will:

- Learn the basics of how to conduct experiments and collect samples for statistical analysis.
- Learn how to describe a statistical sample and examine the basic characteristics of its cases and variables (descriptive statistics).
- Learn how to generalize from the information in a sample to ask and answer questions about the general population from which the sample was taken (inferential statistics).
- Learn enough statistical theory to be able to interpret the results of analyses.
- Develop a solid knowledge of which statistical methods or procedures to use to accomplish particular analytical goals.
- Attain basic proficiency with statistical software (PAST).
- Attain the statistical background and proficiency to understand how statistics is used in Anthropology and to be able to use this knowledge in classes and research.
- Conduct analyses related to a student-selected dataset in a supervised environment to draw conclusions relevant to a hypothesis or research question.

### ADMINISTRIVIA

#### Moodle Access:

This class is offered online through the University of Montana's Moodle System <http://umonline.umt.edu/>. Your instructor is only responsible for the content of the course -- not for the administration of Moodle and/or your access to it. For help with Moodle access or technical issues contact UMLonline Tech Support: (406) 243-4999, [umonline-help@umontana.edu](mailto:umonline-help@umontana.edu).

#### Computer Use:

This class requires students to use software that runs on either Window or Macintosh operating systems. The software won't run on Linux, Android, IOS, or ChromeOS. The University provides some computing labs. See <https://umt.teamdynamix.com/TDCClient/2032/Portal/KB/ArticleDet?ID=44372>. Students in this class should be able to use the Social Science Research Laboratory. Ask the Departmental Administrator (SS building 4<sup>th</sup> floor) about access.

#### Textbook:

There is no textbook for this class. Instead we will focus on watching videos from the Crash Course Statistics series.

## Grading:

Grades will be based on 14 weekly quizzes over the videos (1% each, total 14% of grade), 14 weekly regular assignments (1% each, total 14% of grade), 14 weekly graduate assignments (1% each, total 14% of grade), a midterm exam (18% of grade), a final exam (20% of grade), and a term paper (20% of grade). For the weekly video quizzes, regular assignments, midterm, and final exam I will give you the grade you earned on the associated quiz. I will normally grade the weekly graduate assignments as either completed (100%) or not completed (0%), but I may make comments to your assignment and in the case of severe problems I'll ask you to redo it.

The following scale will be used for assigning grades: A: 90-100%, B: 80-89%, C: 60-79%, D: 50-59%, F: under 50%. If appropriate, I may assign + and - modifiers to these grades.

## Miscellaneous

- Students with disabilities will be accommodated as much as possible upon recommendation by Disability Student Services (DSS). The Anthropology Department does not have facilities for special testing accommodations, so these will have to be provided by DSS.
- This class and the students enrolled in it are subject to the Student Conduct Code: <http://www.umt.edu/student-affairs/community-standards/default.php>.

## Some Choices I've Made For This Course

Designing a course like this one entails making a lot of choices. I think it's worth telling you what choices I've made and why I've made them.

- Make the course non-calculational. Most statistics classes emphasize doing the calculations required by statistical procedures using pencil and paper, or at best a hand calculator. The rationale is that students won't understand the underlying math if they don't have to calculate it by hand. I agree with this up to a point, but I'm not convinced that understanding the underlying calculations is as important as some other things. I find that hand calculation inspires fear in students because they know that if they make a small arithmetic error the answer they get will be incorrect. This leads to students thinking of statistics as difficult, when it is actually fairly simple in most ways. Also, I would rather have my students spend their time learning the principles of statistical analysis and how to actually set up, run, and interpret a statistical analysis.
- Using software. Following from the point above, I've decided that the best way for students to learn what is really important in statistics is to use software. Seriously, no professional ever does their statistical calculations by hand. It's just too easy to make a mistake in writing down a value or in doing the calculations. All professionals use software.
- Choice of software. Professionals use a variety of software, and some swear by certain apps. I believe that by far the best software to use is commercial software such as SPSS, SAS, Stata, or the like. However, all of these have a fairly steep cost. Most will offer a 30-day free trial, but after that you have to pay for it. Some, such as SPSS will rent the software to students for a semester at a moderate cost (about \$70 in early 2020), but after that period you either have to rent it again or buy it. The cost to buy it is often in the thousands of dollars. Many anthropologists like using Microsoft Excel as their statistical software, and indeed Excel will do a lot once you learn how to use the functions and set them up. However, function programming in Excel is complex, and I didn't want this to be a class about using Excel functions. I'll leave that to a computer science or business oriented class. A huge number of professionals are now using the R statistical programming language. R is free, and it can do anything. However, it is a computer programming language and I, again, don't want to spend my class teaching students how to program. I'll leave that to Computer Science or Math classes. So, given all this I decided on PAST. PAST has many advantages, the most important being that it's free software. It does most things that a researcher would want to do,

but, being freeware, there are some things that are not up to par with commercial software. Some procedures seem broken or very flakey, as we will see. However, free is good. PAST comes in versions for Windows or Macintosh. It's also portable (at least on Windows) meaning that you don't need to install it – just download it and run it. Thus, you could keep PAST on a memory stick along with your data and run it from there.

- Using the Crash Course Statistics video series. I could have made my own videos, or some documents for you to read. However, my screen presence is not the best. I also stutter and stammer when talking. So, I will rely on the Crash Course Statistics series of videos on YouTube. This is a very good series of videos. It does a good job of discussing basic theory and methods. Some of the videos also explore additional topics such as machine learning and big data, both of which are based in statistics and very commonly used today. The host of this series is both better looking than I am and far more articulate. The videos are relatively short, with a mean length of 11.68724444 minutes and standard deviation of 1.629339188 minutes. They are far more entertaining than anything I could make. Over the 15 weeks we will watch the whole series. There are some problems with the Crash Course Statistics series. Sometimes they use what I consider to be questionable terminology. There are a few major methods that aren't discussed at all. In these cases I will use additional videos not from the Crash Course Statistics series or post a document for you to read.
- Weekly things to do to promote engagement in the course. Remaining engaged in a course over the whole semester is important for any course, but especially critical for online classes. I've been teaching online courses since the early 2000's and I've observed that if a student gets behind in an online course they almost never get caught up. Therefore, I do everything I can to encourage and enforce weekly participation. Because of this I have 3 things for you to do most weeks. One is weekly quizzes on the content of the videos and/or posted documents. The second is weekly assignments that call on you to use PAST to run some procedure using a dataset I supply. The third is weekly assignments that ask you to run the procedures covered that week with your own data.
- Encouraging learning via repetition. For all the weekly video quizzes I allow unlimited tries. For these quizzes I have Moodle set to allow you to check your answer (immediate feedback is effective) and allow you to change it if you got it wrong. Assignment quizzes restrict this some. You get to immediately see whether your answer was correct or not, but you are not allowed to change it. Instead you should figure out the correct answer and try the quiz again. I allow 3 attempts on assignment quizzes. Exams are different, in that they won't allow you to check your answers, and you are only allowed one attempt. Frankly, this is to prevent students from sharing answers with each other. However, at least 50% of the questions on exams will come from the weekly quizzes, so if you have mastered them you should do well on exams. My reason for doing all of this is to encourage students to actually learn from their mistakes, even to the point of doing it over and over to get it right.
- Choice of a class dataset. I have created a dataset that we will use throughout the semester. It consists of several demographic, economic, health, and other variables for most of the countries of the world. There are various data types. My source was mostly data supplied publicly by the United Nations. I won't claim that the variables are especially important or interesting, but I will point out that the United Nations thinks that they are interesting or important enough to report. This dataset is probably most appropriate for applied or medical anthropologists, but all anthropologists should be able to understand what the variables are about. To tell the truth, I'm more comfortable working with measurements of bones or bodies, but that sort of data has limited appeal. Occasionally I will need to provide you with datasets that are formatted differently or which have some form of processed data (for example, means for a set of groups). However, they are all based on the same original dataset.

## Provisional\* Schedule

Week	Events	Videos (By Crash Course Statistics Unless Otherwise Noted)	Things to Read	Assignments Regular assignments are listed here. <b>Unless noted, the graduate assignments will ask you to do similar things with your own dataset.</b>
1	First Day of Class	<ul style="list-style-type: none"> <li>Statistics Preview</li> <li>What is Statistics</li> </ul>		Assignment 1: Getting Started with PAST <b>Graduate Assignment 1: Find a dataset and a research question.</b>
2	MLK Holiday	<ul style="list-style-type: none"> <li>Mathematical Thinking</li> <li>Mean, Median, and Mode</li> <li>Measures of Spread</li> </ul>	Types of Data	Assignment 2: Descriptive Statistics <b>Grad Assignment 2</b>
3		<ul style="list-style-type: none"> <li>Charts Are Like Pasta</li> <li>Plots, Outliers, and Justin Timberlake</li> <li>The Shape of Data</li> </ul>		Assignment 3: Plotting in PAST <b>Grad Assignment 3</b>
4		<ul style="list-style-type: none"> <li>Correlation Doesn't Equal Causation</li> <li>Controlled Experiments</li> <li>Sampling Methods and Bias with Surveys</li> <li>Science Journalism</li> </ul>		Assignment 4: Scatterplots and Correlation <b>Grad Assignment 4</b>
5		<ul style="list-style-type: none"> <li>Henrietta Lacks, the Tuskegee Experiment, &amp; Ethical Data Collection</li> <li>Probability Part 1</li> <li>Probability Part 2</li> <li>The Binomial Distribution</li> </ul>	<ul style="list-style-type: none"> <li>Some Comments on Week 5 Videos</li> <li>Null Hypotheses and Hypothesis Testing</li> </ul>	Assignment 5: Introduction to Hypothesis Testing <b>Grad Assignment 5</b>
6	President's Day	<ul style="list-style-type: none"> <li>Geometric Distributions &amp; The Birthday Paradox</li> <li>Randomness</li> <li>Z-scores and Percentiles</li> <li>The Normal Distribution</li> </ul>		Assignment 6: Normality Statistics and Tests <b>Grad Assignment 6</b>
7		<ul style="list-style-type: none"> <li>Confidence Intervals</li> <li>How p-values help us test hypotheses</li> <li>P-Value Problems</li> <li>Playing with Power: P-Values Pt 3</li> </ul>		Assignment 7: Single Sample Tests <b>Grad Assignment 7</b>
8	<b>Midterm Exam</b>	<ul style="list-style-type: none"> <li>You know I'm all about that Bayes</li> <li>Bayes in science and everyday life</li> </ul>		<b>Midterm Exam</b> NO assignment 8
9		<ul style="list-style-type: none"> <li>Test Statistics</li> <li>T-Tests</li> <li>Degrees of Freedom &amp; Effect Sizes</li> </ul>		Assignment 9: Two-Sample Tests <b>Grad Assignment 9</b>
Spring Break				
10		<ul style="list-style-type: none"> <li>Chi-Square Tests</li> <li>P-Hacking</li> <li>The Replication Crisis</li> </ul>		Assignment 10: Chi-Square Tests <b>Grad Assignment 10</b>

		<ul style="list-style-type: none"> <li>• Regression</li> <li>• Transforming nonlinear data (by Khan Academy)</li> </ul>	<ul style="list-style-type: none"> <li>• From Regression to Multiple Regression to Principal Components to Canonical Variates</li> </ul>	Assignment 11: Regression and Non-Linear Curve Fitting <b>Grad Assignment 11</b>
12		<ul style="list-style-type: none"> <li>• ANOVA</li> <li>• ANOVA Part 2</li> <li>• Fitting Models Is like Tetris</li> </ul>		Assignment 12: ANOVA and MANOVA <b>Grad Assignment 12</b>
13		<ul style="list-style-type: none"> <li>• Supervised Machine Learning</li> <li>• Unsupervised Machine Learning</li> <li>• Intro to Big Data</li> <li>• Big Data Problems</li> </ul>	<ul style="list-style-type: none"> <li>• Some Comments on Week 13 Videos</li> </ul>	Assignment 13: Canonical Variates and Discriminant Functions <b>Grad Assignment 13</b>
14		<ul style="list-style-type: none"> <li>• Statistics in the Courts</li> <li>• Neural Networks</li> <li>• War</li> <li>• Euclidean Distance and Manhattan Distance (by Krish Naic)</li> </ul>	<ul style="list-style-type: none"> <li>• Distance Methods</li> </ul>	Assignment 14: Distance and Clustering <b>Grad Assignment 14</b>
15	Last Day of Regular Instruction	<ul style="list-style-type: none"> <li>• When Predictions Fail</li> <li>• When Predictions Succeed</li> <li>• Visual Explanation of Principal Component Analysis, Covariance, SVD (Freedman)</li> <li>• StatQuest: MDS and PCoA</li> </ul>		Assignment 15: Principal Components and Multi-Dimensional Scaling <b>Grad Assignment 15</b>
16	<b>Final Exam</b>			<b>Final Exam Term Paper Due</b>

\*This schedule is tentative and provisional. Things may change, and if so will be updated on Moodle.