

STAT 330: STATISTICAL COMPUTING WITH SAS  
California Polytechnic State University, San Luis Obispo  
Fall 2017

---

<b>Instructor:</b>	Shannon Pileggi	<b>Office:</b>	25-109
<b>Email:</b>	spileggi@calpoly.edu	<b>Phone:</b>	756-2946

---

**Class meetings**

<i>Section</i>	<i>Time</i>	<i>Days</i>	<i>Location</i>
70	2:10-4:00	TR	38-123A
71	4:10-6:00	TR	38-123A

**Office hours**

M 10:10-11:10

T 10:30-11:30

W 10:10-11:10

R 12:30-1:30

By appointment (please email me three times you are available)

**Diversity statement**

This is an inclusive class that welcomes and values participation from individuals of **all** identities, which includes but is not limited to: race, ethnicity, culture, religion, gender, sexual orientation, language, national origin, age, physical/emotional/developmental ability, and socio-economic class.

**Course description**

Techniques available to the statistician for efficient use of computers to perform statistical computations and to analyze large amounts of data. Use of SAS throughout the course. Includes data preparation, report writing, and basic statistical methods.

**Learning objectives**

1. Formulate a game plan that states the objective before coding.
2. Identify multiple ways to achieve the game plan; consider pros and cons of the options before coding.
3. Apply techniques to prepare data for analysis, including: merging or transposing data, creating new variables, and identifying data errors and “cleaning” data.

4. Prepare summary statistics of data; create graphical displays of data.
5. Execute various statistical analyses and interpret results.
6. Apply arrays, loops, and SAS macros for efficient coding.
7. Discuss how SAS's program data vector operates.
8. Import data of various sources and formats into SAS.
9. Verify that code renders the desired result in the presence of missing data.

## Course materials

Required text	<i>The Little SAS Book: a primer, fifth edition</i> , by Lora D. Delwiche and Susan J. Slaughter 2012. There will be assigned readings from this text. Please note this text book is available as an e-book (for free!) through the Kennedy Library - search for STAT 330 at this site: <a href="http://lib.calpoly.edu/search-and-find/open-resources/required-textbooks/">http://lib.calpoly.edu/search-and-find/open-resources/required-textbooks/</a> .
Optional text	<i>Exercises and Projects for The Little SAS Book</i> , by Rebecca A. Ottesen, Lora D. Delwiche and Susan J. Slaughter 2015. This text may be useful to you if you want to do extra practice problems on your own.
Software	This course will require the use of SAS, SAS Studio, and Jupyter notebooks, all of which are available in the classroom 38-123A. For access outside of class time, which will be necessary for the completion of assignments, please see <i>Ways to access/install SAS, SAS Studio, and Jupyter notebooks</i> on the PolyLearn course site for more information. Difficulty accessing SAS outside of the classroom is not a valid excuse for late homework.
Flash drive	Please bring a personal flash drive to class everyday to save your work.

## Course evaluation

Item	Percentage
Pre-class homework	10%
Lab assignments	20%
Lab check	5%
Quizzes	15%
Midterm	20%
Final	30%

Letter	Percent
A	93-100
A-	90-92.9
B+	87-89.9
B	83-86.9
B-	80-82.9
C+	77-79.9
C	73-76.9
C-	70-72.9
D+	67-69.9
D	60-66.9
F	0-59.9

Your lowest homework, lab, lab check, and CFU/quiz grade will be dropped. Late assignments will be accepted with a 30% per day late penalty. The PolyLearn grade book is only a **record** of your grades and is not set up to calculate an accurate overall grade. Overall grades will be computed outside of PolyLearn.

## **Pre-class homework**

Each class period will be preceded by a pre-class reading assignment and/or a SAS script to complete. Homework will be graded as:

- 0 = no submission
- 1 = incomplete
- 2 = mostly complete
- 3 = complete submission

## **Lab assignments**

Each class period approximately one hour (some days more, some days less) will be in-class time devoted to completing a lab assignment. If you do not complete the lab assignment in the allotted time, you are responsible for completing the assignment prior to the next class period. Labs will be graded according to a more detailed rubric.

## **Lab checks**

In the last 10 minutes of each class period, a “lab check” prompt will be provided and students will submit responses PolyLearn. This is to ensure that students are making both adequate and correct progress in the allotted lab time. It also gives me an opportunity to notify you of anything that looks incorrect prior to submission of your assignment. Lab checks will be graded as:

- 0 = no submission
- 1 = in progress
- 2 = satisfactory

## **Quizzes**

You will have quizzes every 2-4 lectures throughout the quarter.

## **Exams**

Both the midterm and final exam will consist of a written portion (not using SAS) and a programming portion (using SAS). The final exam is cumulative, but will have an emphasis on material presented after the midterm.

## **Optional project**

There will be the opportunity to complete an optional project for students who want extra practice. More details will be provided in a separate document.

## Submission instructions

- Unless otherwise stated, students will mostly be submitting SAS program files for assignments (.sas extension).
- Please include a header such as this one on all assignments with your name, the date, and the assignment number:

```
/*-----*  
| Name: Shannon Pileggi          |  
| Date: September 14, 2017      |  
| Assignment: Homework 1        |  
*-----*/
```

- Please name your submitted assignments consistently with your last name, assignment type, and number as follows: `LastNameFirstinitial_HW#.sas`

## Tips for success

- *Attend class!* Attending class gives you the opportunity to interact with your instructor and ask meaningful questions.
- *Ask questions!* Your instructor is your best resource to answer any questions or provide any clarification. Please visit office hours early and often! Seek clarification on course content as soon as questions arise.
- *Form a study group!* It is much easier to engage in the learning process if you have peers with which to discuss challenging concepts.
- *Put in the time!* You know the 25-35 drill: you should be studying **at least 8 hours a week** for this course.

## Academic integrity

All students are expected to uphold high standards of academic integrity. The university provides [broad definitions](#) of academic misconduct, and also provides [due process](#) for students with an alleged violation. Please note that there may be serious consequences for academic misconduct, including failing an assignment or the course. If you are ever unsure as to whether or not an action is acceptable, please do not hesitate to contact the instructor.

Collaboration and study groups is strongly encouraged. However, each student is responsible for submitting his/her own work independently of others. If submitted computer code appears to be an exact (or near exact) copy of someone else's code, this is academic misconduct. Any student who does so will be reported to the Office of Students Rights and Responsibilities.

## Disability resources

If you have a disability for which you are or may be requesting an accommodation, you are encouraged to contact both your instructor and the [Disability Resource Center](#), Building 124, Room 119, at (805) 756-1395, as soon as possible.

## Communication

A productive quarter requires good communication between instructors and students, which includes both personal and course related issues. Course related questions can include clarification on course policy, but mostly arise from questions regarding course content. All course related questions should be posted to the PolyLearn discussion forum. This is because it is likely that other students may have similar questions, and then the whole class can benefit from your question. On the other hand, personal communication may entail issues like missing class for personal reasons, or requests for a meeting. Please email me any personal communication.

## Difficult conversations

Occasionally students encounter challenges in classroom dynamics with other students or even with the professor. If you ever find yourself in the situation where you aren't comfortable with something that was said or done, please consider taking one of the following actions:

- Discuss the issue with me privately during office hours.
- Discuss the issue with the class or individual of concern if you feel you can do so in a respectful and well-communicated way.
- Discuss the issue with someone that you feel that you can trust (another faculty member, a mentor, or advisor) and who can, in turn, communicate your concerns with me.

Change cannot be enacted unless there is awareness of a problem - thank you for taking action.

## Other policies

- All course materials and important announcements will be posted on PolyLearn. You are expected to check PolyLearn regularly and read your emails.
- Students are expected to attend all class meetings. However, missing class for religious holidays, school-related travel for academics or athletics, serious illness, or family emergencies is excused, and missed work due to such reasons will be allowed to be made up.
- Please notify the instructor in a timely manner of any events that may adversely impact your performance in the class.

This schedule is tentative and subject to change.

Week	Tuesday	Thursday
<b>1</b> 9/11 - 9/15	<i>No class</i>	L1: Overview of SAS
<b>2</b> 9/18 - 9/22	L2: PROC IMPORT, PROC CONTENTS, PROC PRINT, PROC MEANS	<i>No class</i>
<b>3</b> 9/25 - 9/29	L3: SAS libraries, PROC UNIVARIATE, PROC FREQ	<b>Quiz</b> ; L4: Data step basics: IF/then, DROP/KEEP, SAS functions, subsetting data
<b>4</b> 10/2 - 10/6	L5: Macros	L6: Instream data, informats, formats, labels, and PROC FORMAT
<b>5</b> 10/9 - 10/13	<b>Quiz</b> ; L7: Combining data	L8: DO loops and arrays
<b>6</b> 10/16 - 10/20	<b>Quiz</b> ; L9: Data cleaning and new variable creation; verifying correctness	<b>Midterm</b>
<b>7</b> 10/23 - 10/27	L10: How the data step works, PROC SGPLOT	L11: ODS graphics, PROC TTEST, PROC CORR
<b>8</b> 10/30 - 11/3	L12: turning output into data, PROC TRANSPOSE, PROC EXPORT	<b>Quiz</b> ; L13: PROC ANOVA, PROC REG, and PROC GLM
<b>9</b> 11/6 - 11/10	L14: retain/sum, PROC SORT, First./Last.	L15: PROC TABULATE
<b>10</b> 11/13 - 11/17	<b>Quiz</b> ; L16: PROC REPORT	L17: PROC SQL
<b>11</b> 11/20 - 11/24	<i>Thanksgiving break</i>	<i>Thanksgiving break</i>
<b>12</b> 11/27 - 12/1	L18: list input, column input, formatted input	<b>Quiz</b> ; L19: more options for reading in raw data with INFILE and INPUT

#### Final Exam:

- Sec 70 (2pm class): Tuesday, Dec 5 4:10pm-7:00pm
- Sec 71 (4pm class): Thursday, Dec 7 7:10pm-10:00pm