



# 2022 Summer Statistics and Programming Bootcamp

**July 11<sup>th</sup> – September 2<sup>nd</sup> (8 weeks) on Zoom**

## **Sections & Hours**

**Section I:** Tuesdays, 10 am – 12 pm (<https://ucr.zoom.us/j/99822608667>)

**Section II:** Tuesdays, 2 – 4 pm (<https://ucr.zoom.us/j/93800832833>)

**Section III:** Wednesdays, 10 am – 12 pm (<https://ucr.zoom.us/j/96090373957>)

**Office hours (on Zoom):** Wednesdays, 3 – 5 pm (<https://ucr.zoom.us/j/4999938154>)

\*Please only attend the section you are assigned to.

## **Team**

### **Yiwang Li**

Statistics and Programming Bootcamp Instructor

Ph.D. candidate in Education Policy Analysis and Leadership

**Email:** [yli394@ucr.edu](mailto:yli394@ucr.edu)

**Appointments:** Office hours on Wednesdays, 3 – 5 pm. If you are unavailable during this time, please email me to discuss your questions.

### **Jason Chou**

GradQuant and Graduate Student Resource Center Coordinator

**Email:** [jason.chou@ucr.edu](mailto:jason.chou@ucr.edu) | <https://gradquant.ucr.edu> | <https://gsrc.ucr.edu>

*\*Jason serves in an administrative role in the bootcamp.*

## **Purpose**

This bootcamp provides an introduction to the basic concepts of descriptive and inferential statistics, as well as applications within your fields. It is intended for graduate students with little to no background in statistics and who wish to review the fundamentals before taking graduate-level statistics courses. You will also receive a brief introduction to Excel and R programming language that will help you get started on performing analyses on various kinds of research data. Because research is always a team effort and relies upon effective communication, you will interact with other students in your sections throughout this course and communicate your research ideas and understandings as we progress.

### **Key Learning Objectives**

The goals of this bootcamp are to provide a solid understanding of the principles underlying statistical inference and null hypothesis significance testing. Upon successful completion of this course, you should be able to:

1. Understand the role of statistics in your field.
2. Recognize various types of data.
3. Summarize and make sense of data through plots and descriptive statistics.
4. Choose the appropriate statistical methods to analyze data in routine experimental settings.
5. Perform statistical analyses in Excel and R software.
6. Interpret the results of statistical analyses in various software environments (i.e., Excel and R), and that appear in scientific literature.

### **Prerequisites**

None. If the topics within this syllabus look completely foreign to you, do not be scared off! You have the perfect background knowledge for this bootcamp. If you have previously learned some of this material but have not used it in a while, you should take this bootcamp to refresh your knowledge of the topics. If any of the content intimidates you or brings back bad memories (hopefully not), you can rest assured that Yiwang is an approachable, responsive instructor. Feel free to ask questions during office hours or over email, if you prefer, regarding any topic you are uncertain about.

### **Structure**

The bootcamp will consist of one workshop a week for each section. This bootcamp is not for credit and has no exams; however, there will be weekly homework. Only you will know how well you did, so you can test yourself by participating in the workshops as much as possible. Workshops will focus on major conceptual topics in statistics and modeling that are widely applied across a range of scientific fields. Each week, office hours will be available, in addition to the workshops. During office hours, you can work with the instructor to solve problems or clarify examples with which you might have struggled. You are encouraged to work on the workshop exercises individually or in small groups of two or three.

### **Homework**



There will be weekly homework complementary to what we learn in the lecture that week. All homework are optional to turn in, but we will be operating under the assumption that everyone is completing them each week. Turn in your homework by 11:59 pm on the day that it is "due". The homework solution will be posted at the same time, but please attempt to solve the problems by yourself before checking the answers.

### **Turnaround/Feedback**

During the week, I will check emails multiple times. If you have concerns and send me an email, you can expect a response within two days. Please be sure to include your name and section in the email. If you are emailing about programming problems, sending the code and error message is also helpful.

### **Computing**

You will need access to a computer with Excel, R, and RStudio. The links to download each are listed below.

- **Excel** (free for UCR students)  
[https://ucrsupport.servicenow.com/ucr\\_portal/?id=kb\\_article&sys\\_id=20eb980a1bf0685063c62fc02a4bcb77](https://ucrsupport.servicenow.com/ucr_portal/?id=kb_article&sys_id=20eb980a1bf0685063c62fc02a4bcb77)
- **R** (free - required) <https://cran.r-project.org/>
- **RStudio** (free - required) <https://www.rstudio.com/products/rstudio/download/>
  - Install RStudio after installing R
  - RStudio is a friendly platform that adds many useful features to simplify using R. All the functions used in this class can be performed without RStudio, but I will demonstrate their use within RStudio.

### **Textbook**

***OpenIntro Statistics***, 4<sup>rd</sup> edition by Diez, Barr, and Çetinkaya-Rundel. This open-source textbook is available as a free PDF: <https://www.openintro.org/book/os/>. To download the book, follow these instructions:

1. Visit the link above
2. Click on the "FREE -- OpenIntro Statistics PDF" box under Getting Started
3. Use the sliding bar to select a donation amount, including \$0 (Donation is completely voluntary)
4. Click on "Add eBook to Cart"



5. Continue to Checkout and follow the instructions to download the book

We recommend reading the sections of the textbooks, as listed on the next page, before the associated workshop to maximize your learning of the material. You will further reinforce your understanding by participating in the discussion/examples and doing the recommended practice exercises after the workshops.

Other recommendations:

- Wackerly, D., Mendenhall, W., & Scheaffer, R. L. (2014). Mathematical statistics with applications. Cengage Learning.
- Ott, R. L., & Longnecker, M. T. (2015). An introduction to statistical methods and data analysis. Cengage Learning.

### **Zoom Links**

The Zoom link for each session will be shared in an email sent before bootcamp begins. ***You will use the same Zoom link weekly to attend your session.***

Before joining the Zoom meeting for your assigned session, go to <https://ucr.zoom.us> and log in to your UCR Zoom account with your UCR credentials (netID and password).

### **Attendance**

**GradEdge/JumpStart (GE/JS) students:** Attendance is mandatory for GE/JS students. Your stipend will be reduced if you miss a bootcamp session without getting a prior excused absence request approved. To request an excused absence, email the GE/JS Coordinator, George Santellano ([gradededge@ucr.edu](mailto:gradededge@ucr.edu)), as early as you can (***but at least 48 hours prior***), and copy ("cc") the GradQuant Coordinator, Jason Chou ([jason.chou@ucr.edu](mailto:jason.chou@ucr.edu)). In your email, you must include:

- Your name and SID#,
- Day/session from which you wish to be excused, and
- Details about why you need to be excused.

We will respond to your email to confirm whether you are excused.

If you miss a bootcamp session due to an emergency, let us know as soon as it is safely possible after you deal with the emergency. Even if you miss a bootcamp session for any reason (excused absence or an emergency). We recommend that you come to the next Office Hours to ask any questions you may have about the material you missed.



**Current UCR graduate students:** Attendance is not mandatory for you, but you are still expected to attend all sessions in order to get the best out of the bootcamp. Please let us know if you will be missing a workshop.

### **Workshop Schedule**

<b>Week</b>	<b>Topics</b>
<b>Week 1</b> <i>July 11<sup>th</sup>-15<sup>th</sup></i>	<b>Types of Data; Descriptive Statistics</b> <i>Suggested Reading: §2.1-2.2</i>
<b>Week 2</b> <i>July 18<sup>th</sup>-22<sup>nd</sup></i>	<b>Probability</b> <i>Suggested Reading: §3.1, 3.2, 3.4, 3.5</i>
<b>Week 3</b> <i>July 25<sup>th</sup>-29<sup>th</sup></i>	<b>Probability Distributions</b> <i>Suggested Reading: §4.1, 4.3</i>
<b>Week 4</b> <i>August 1<sup>st</sup>-5<sup>th</sup></i>	<b>Central Limit Theorem; Confidence Intervals</b> <i>Suggested Reading: §5.1-5.3</i>
<b>Week 5</b> <i>August 8<sup>th</sup>-12<sup>th</sup></i>	<b>Introduction to Inference</b> <i>Suggested Reading: §5.3, 6.1, 7.1-7.3</i>
<b>Week 6</b> <i>August 15<sup>th</sup>-19<sup>th</sup></i>	<b>Analysis of Variance (ANOVA)</b> <i>Suggested Reading: §6.3</i>
<b>Week 7</b> <i>August 22<sup>nd</sup>-26<sup>th</sup></i>	<b>Linear Regression and Correlation</b> <i>Suggested Reading: §8.1-8.4</i>
<b>Week 8</b> <i>August 29<sup>th</sup> -September 2<sup>nd</sup></i>	<b>Specialized Topics</b> <i>Suggested Reading: Handout</i>

Each week will also feature a tutorial using the software. This tutorial will either be on features of the software or how to implement and calculate statistics and tests learned during the workshop.