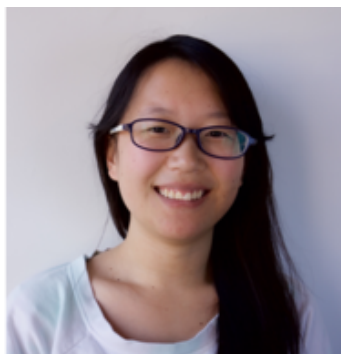




**Welcome to the Foundations of Data  
Science Course!**



Hi!

My name is Kane Li, and I'm your Student Advisor!

I'm a born and raised Los Angeles native. I recently graduated from UC Riverside with a degree in psychology, and I'm really passionate about mental health and education. Like many of us on the Springboard team, I'm also a strong believer in lifelong learning and always striving to improve yourself. My hobbies include photography, travel, and exploring new places!

As your Student Advisor, my goal is to help you succeed and support you through the workshop and beyond. So, while you'll get emails from me about weekly check ins, your weekly calls, and more, I can also help you in a variety of ways.

For example, I can:

✓ **Help you prepare for the course.** I'm happy to sit down with you to understand what kind of learner you are, what your weekly schedule might look like, what you need to do to stay on track, and more!

✓ **Check in with you!** Whether it's by email or phone, I'm happy to check in with you to see how you're doing throughout the workshop. We can talk about what frequency and medium might help you the most in our first advising call together.

✓ **Answer general questions.** If you have questions about your account, the workshop, or other general topics, I'm happy to answer them via email or schedule a call with you! I'm not a Data Scientist myself so if you have technical questions regarding the curriculum material, you can ask those questions to your mentor or in the online community!

I'm always here to chat with you, whether it's about the workshop or just to say hello. My (virtual) door is always open for you, and you can always reach me at [kane@springboard.com](mailto:kane@springboard.com).

In this packet, I have a few tips to share with you as well as the course timeline to help you complete!



## Tips for Springboard Success!

To help set you up for success, my team and I have compiled a list of ten tips we've learned from our students through advising:

1. **Set aside 'study time' and treat it as a non-negotiable commitment.** Find time in your schedule to work on the course, and treat them as you would a business meeting, interview, or in-person class: your coursework is a priority!
2. **Find an accountability buddy.** This could be a family member, co-worker, friend, or someone else in your life. Check in with your buddy each week and share your goals for the course.
3. **Understand how you're currently using your time.** Keep an activity log to track how you spend your time each day. As soon as you know where you're spending your time, you can free up spots that you would normally spend doing something else, like getting sucked into the black hole of the Internet.
4. **Minimize interruptions when possible.** Reflect on common interruptions you experience, and which ones can be reduced or removed completely? For example, when working on the curriculum, turn off your phone. You won't have the temptation to check it when the screen lights up.
5. **Overestimate how much time it will take.** By nature, humans aren't great at estimating how long it takes to complete a task. If you think it will take 1 hour, plan for 1.5 hours. That way, if you get done early, it's an added bonus. Use the estimates in the curriculum, but keep in mind that these may vary depending on your experience and learning style.
6. **Try the Pomodoro Technique.** This is one of our favorite tools for improving time estimation, focus, and productivity. A Pomodoro is a 25-minute immersive working period. After one Pomodoro, take a quick break (grab a drink, go to the bathroom), and then, complete another Pomodoro. After four Pomodoros, take a longer break.
7. **Send your mentor an agenda 24 hours before your call.** Draft a weekly agenda with action items that you'd like to cover during your call and share your agenda with your mentor the day before your call
8. **Engage in the online community.** Start conversations online or search the community for members in your area to meet up in person! If you have a question



or want feedback on a project, post it in the online community. Students, mentors, and our community manager are there to help!

- 9. Attend Office Hours.** Join our weekly mentor-hosted Office Hours to humanize this online learning experience. Chat about curriculum topics or the industry in general. If you can't attend, check out the recordings, and email your student advisor to request topics for special sessions.
- 10. Schedule calls with your student advisor.** Want to check in about your progress each week? Eager to provide feedback on the curriculum or learning experience? Schedule a call with your student advisor! We want to hear from you, and your voice matters to us.



## Course Pace Timeline

To help you complete in time, here is a snapshot of what your weekly goals should be. There is a more in-depth timeline in the next few pages.

### Week – by – Week Schedule

Week	Unit Name
1	Introduction Programming in R
2	Data Wrangling
3	Exploratory Data Analysis in R
4	Probability & Statistics
5	Data Story (Start) An Introduction to Machine Learning
6	(Finish) An Introduction to Machine Learning
7 & 8	Capstone Project

As you can see, you'll hit the ground running if you want to complete in two months. This means you should:

- ✓ **Be ready to dedicate 20-30 hours per week to the course.** Depending on your strengths and weaknesses, some weeks may take you longer or shorter than our own estimations!
- ✓ **Be driven and disciplined with managing your study schedule.** There is a lot to cover so make sure you stick to your study schedule.

You will still have one, 30-minute mentor call per week on a two-month timeline so we recommend utilizing the online community and office hours as much as possible to supplement your education!

### Remember, to complete the course, you must:

- 1) Complete at least 60% of the Curriculum,
- 2) Submit all projects, including your mini-projects and capstone project, and
- 3) Be approved by your mentor for completion

Now, let's look at the in-depth weekly schedule!



## Week 1 (25-28 Hours)

1 Introduction	(hh:mm)
Demystifying Data Science	0:40
Start Thinking About Your Capstone Project	1:30
<i>(Optional) Additional Resources</i>	
2 Programming in R	
2.1 Core Resources for Learning R	
What makes R different?	0:05
Create your DataCamp Springboard Account	0:10
Introduction to R (DataCamp)	8:00
Intermediate R (DataCamp)	12:00
Getting Started with R and RStudio	0:45
Version Control with Git, Github and RStudio	2:00
Getting Started with R Markdown	2:00
Submit your Capstone Project Proposal	1:00
2.2 Additional R Resources	
RSeek: A Search Engine for R	
Quick-R	
Style Guide for R	

## Week 2 (15 – 20 Hours)

3 Data Wrangling	
3.1 Core Resources for Data Wrangling	
Data Wrangling with R and RStudio	2:00
Data Manipulation in R using Dplyr	2:00
Going Deeper with Dplyr	1:00 – 2:00
Data Wrangling Exercise 1: Basic Data Manipulation	2:00 – 3:00
Data Wrangling Exercise 2: Dealing with missing values	2:00 – 3:00
<i>(Optional) Data Wrangling Exercise 3: Human Activity Recognition</i>	4:00 – 12:00
Apply data wrangling to your capstone project	2:00 – 4:00
3.2 Additional Resources for Data Wrangling	
<i>(Optional) Cheat Sheet for Data Wrangling commands</i>	
<i>(Optional) Regular Expressions</i>	
<i>(Optional) Getting Data into R</i>	
<i>(Optional) Tidy Data</i>	



## Week 3 (23 – 27 Hours)

### 4 Exploratory Data Analysis in R

Data Visualization with ggplot2   Part 1	10:00
Submit your code for the Titanic exercise	1:00 – 3:00
Data Visualization with ggplot2   Part 2	10:00
Submit your code for the CHIS exercise	2:00 – 4:00

## Week 4 (27 – 28 Hours)

### 5 Probability and Statistics

(hh:mm)

#### 5.1 Probability and Descriptive Statistics

An introductory overview of probability (Khan Academy)	5:00
Displaying and describing data (Khan Academy)	3:30
Describing relationships in quantitative data (Khan Academy)	2:00
Modeling distributions of data (Khan Academy)	3:00

#### 5.2 An Overview of Inferential Statistics

Random Variables (Khan Academy)	6:00
One Sample Confidence Intervals (Khan Academy)	2:00
Significance Tests and Confidence Intervals (Khan Academy)	3:00
Inference for categorical data (Khan Academy)	2:00

#### 5.3 Apply Statistics to your Capstone Project

Start with some statistical analysis on your capstone data set	
Submit a short report on your statistical analysis	1:00 – 2:00

#### 5.4 Additional Resources for Probability and Statistics

*(Optional) Common Probability Distributions: The Data Scientist's Crib Sheet*  
*(Optional) Statistics in a Nutshell*

## Week 5 (27 – 32 Hours)

### 6 Data Story

#### 6.1 Core Resources for Data Story

Watch DJ's talk about data storytelling and products, among other things	0:17
A follow-up chat reiterating the importance of storytelling	0:03
Submit Your Data Story - Your Capstone Project Milestone Report	4:00 – 8:00

#### 6.2 Additional Resources for Data Story

*(Optional) Hans Rosling's TED Talk*



## Week 5 (contd)

*(Optional) Data Storytelling: Turning insights into action*

*(Optional) FiveThirtyEight*

### 7 An Introduction to Machine Learning

*(hh:mm)*

#### 7.1 Linear and Logistic Regression

Feature Engineering	1:00
Linear Regression (Week 2 of edX Analytics Edge Course)	10:00
Mini-Project: Linear Regression	1:00 – 2:00
Logistic Regression (Week 3 of edX Analytics Edge)	10:00
Logistic Regression Mini-Project	1:00 – 2:00

## Week 6 (34 – 37 Hours)

### 7 An Introduction to Machine Learning

*(hh:mm)*

#### 7.2 Model Selection and Evaluation

Model Evaluation: Concepts	2:00
Cross-validation in R	0:15

#### 7.3 Clustering and K-means

Clustering (Week 6 of edX Analytics Edge course)	10:00
K-Means Clustering Mini-Project	1:00 – 3:00

#### 7.4 Apply Machine Learning

Apply Machine Learning to your Capstone Project	1:00 – 2:00
---	-------------

#### 7.5 Electives for Data Analysis in Depth: Trees and Text Analysis

*(Optional) Trees (Week 4 of edX Analytics Edge course)*

*(Optional) Text Analytics (Week 5 of edX Analytics Edge course)*

#### 7.6 Additional References

*(Optional) Introduction to Statistical Learning (With Applications in R)*

*(Optional) The Elements of Statistical Learning*

*(Optional) Data Mining in R*

### 8 ELECTIVE: Data Visualization

*(hh:mm)*

<i>(Optional) Introduction to Infographics &amp; Data Visualization</i>	2:30 – 3:00
<i>(Optional) A Survey of Advanced Visualization Techniques</i>	1:00
<i>(Optional) Books by Nathan Yau</i>	1:00





## **Weeks 7 & 8 (10 – 15 Hours)**

8 Capstone Project

Complete and Submit Capstone Project

10:00 - 15:00



