

May 10,
2021



MATHEMATICAL ASSOCIATION OF AMERICA

MAA

PROGRAMS

StatPREP

NEWSLETTER

PREPARING FOR SUMMER AFTERNOONS

BY DONNA LALONDE

"Summer afternoon—summer afternoon; to me those have always been the two most beautiful words in the English language." I agree with Henry James and am planning some reading, watching, and listening for those anticipated afternoons. If you are also dreaming of summer afternoons, at least a few that are free of meetings, grading, and mowing, I'll share some recommendations to keep you company.

Summer afternoons are perfect for catching up on webinars that I missed or just want to revisit. At the top of my list is the "Moving to a World Beyond $p < 0.05$." My friends and colleagues, Nicole Lazar, Allen Schirm and Ron Wasserstein presented this webinar as part of the AMATYC Webinar Series. You can watch the video or just listen at (<https://www.youtube.com/watch?v=l7apbg9EgXM>). Although first presented in March 2020, it remains an important topic and Nicole, Allen, and Ron share many important insights. This is just one of the amazing webinars from this series; you can learn about upcoming webinars and past webinars at <https://amatyc.org/page/Webinars>.

On my reading list is an article from the Journal of Statistics and Data Science Education (<https://www.tandfonline.com/doi/full/10.1080/10691898.2020.1799728>), "Playing the Whole Game": A Data Collection and Analysis Exercise With Google Calendar, written by Albert Kim and Jo Hardin. "Playing the Whole Game" implies that students need the opportunity not only to practice the skills and the parts of the data science workflow but to experience the cyclic nature of the workflow. This includes collecting the data as well as completing the analysis. In this article Bert and Jo share a class assignment that allows students to "practice 'playing the whole game.'" The authors' inspiration for this assignment came from an episode of the podcast, Not so Standard Deviations, hosted by Hilary Parker and Roger Peng, so you may want to check it out - <https://nssdeviations.com/71-compromised-shoe-situation>.

...continued on page 2

WHO'S WHO:

LEADERSHIP TEAM

Mike Brilleslyper,
Air Force Academy

Jenna Carpenter,
Campbell University

Danny Kaplan,
Macalester College

Kathryn Kozak,
Coconino Community
College

Donna LaLonde,
ASA

Ambika Silva,
College of the Canyons

Deirdre Longacher Smeltzer,
MAA

HUB LEADERS

Joe Roith, St. Olaf's College,
Northfield, MN (2017-18)

Ambika Silva, College of the
Canyons, Santa Clarita, CA
(2017-18)

Helen Burn, Highline College,
Seattle, WA (2018-19)

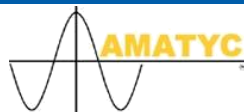
Hwayeon Ryu, Elon University,
Elon, NC (2018-19)

Carol Howald, Howard Community
College, Columbia,
MD (2019-2020)

Thomas Kinzeler, Tarrant
County College, Fort Worth,
TX (2019-2010)

Rona Axelrod, Florida SW
State College, Fort Myers,
FL (2020-2021)

Support for this MAA Program is provided by NSF DUE-1626337



SUMMER AFTERNOONS Continued...

Listening to podcasts is an essential part of summer afternoons. I have a number of episodes from Stats+Stories that I want to revisit. Sharon Lohr described the work done by the women of Hull House in this episode - <https://statsandstories.net/society1/the-women-of-hull-house>. Lohr described the work published in 1895 work, *The Hull House Maps and Papers* (<https://archive.org/details/hullhousemapsan00unkngoog>). This work contained incredible visualizations highlighting the inequitable distribution of wages in Chicago neighborhoods. Also on my playlist is to revisit the conversation with journalist and author, Angela Saini (<https://statsandstories.net/society1/the-recent-regrettable-rise-of-race-science>) because her book, *Superior: The Return of Race Science*, is on my reading list. On the podcast, *Saini and the hosts - John Bailer, Rosemary Pennington, and Richard Campbell - explore the cultural context of science.*

I'll look forward to learning what is on your summer afternoon reading, listening, or watching list, so please share on the StatPREP community!



2021 StatPREP Summer Workshops!

Register

Data Science ● R Studio ● Little Apps



MAA
CONNECT

HUB LEADER SPOTLIGHT

Thomas Kinzeler is an instructor and department co-chair of mathematics at Tarrant County College - Northeast Campus. He is currently working on strategies to improve the success of co-requisite courses which has led him to focusing on helping all developmental math students both in and out of the classroom. Lately this has led to many opportunities for collaboration with other departments and Thomas is always looking for new innovative ideas to help these students succeed in their math course.

With the pandemic of COVID-19 there is a lot of available data out there that could be used in an intro stats class to cover a wide variety of concepts. I've been thinking a bit lately though if I should or not. On one hand it is very current and something meaningful to everyone. It could be used with descriptive stats and inferential stats in some cool ways. On the other hand it can be potentially upsetting to some students who have been struggling during this time. Do our students need or want one more place that they are having to hear about COVID-19? I'm curious to hear what other people in the community think. If you have used COVID-19 data in your class, how did it go? Do you think it is a good idea to include this in our stats classes? **Share your thoughts on MAA Connect!**



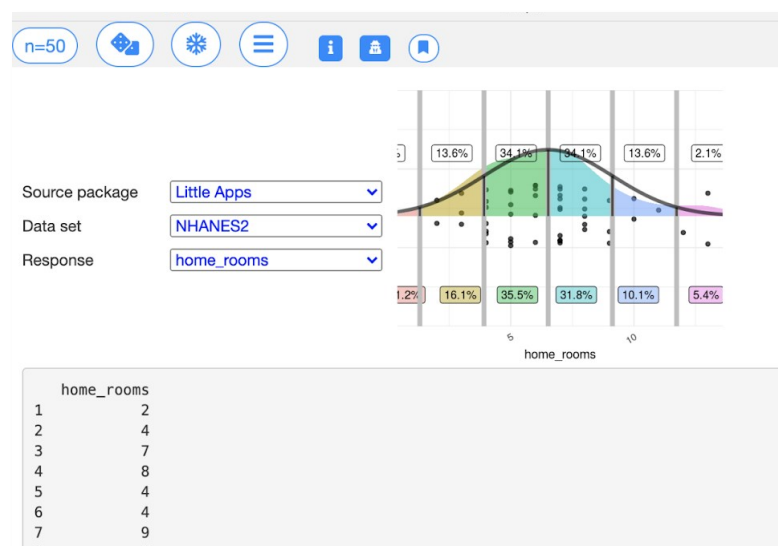
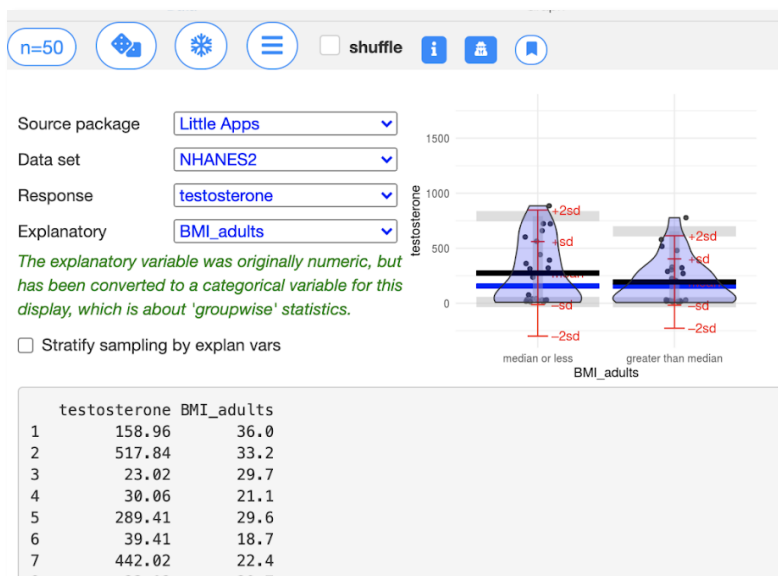
STATPREP LITTLE APPS

BY KATHRYN KOZAK

I have written about activities for Little Apps in previous newsletters. This time I want to point out that the Little Apps can be used without the activities. This article will describe what each Little App demonstrates. In each Little App, you can pick different datasets, response variables, and explanatory variables. Then you can pick different values for the sample size. There is also the ability to take multiple samples so you can compare them. There is even a way to keep one sample and then look at the graph of another sample. This way you can demonstrate how samples change when a new one is taken. There is an option to choose what descriptive statistics and inferences are displayed. You can also display the actual values of common statistics, as well as see the data itself. Lastly, all Little Apps have controls which display certain descriptive and inferential statistics. There is a code book that describes all of the variables in the chosen datasets.

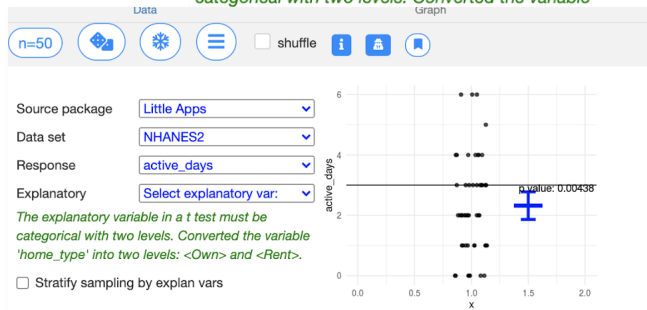
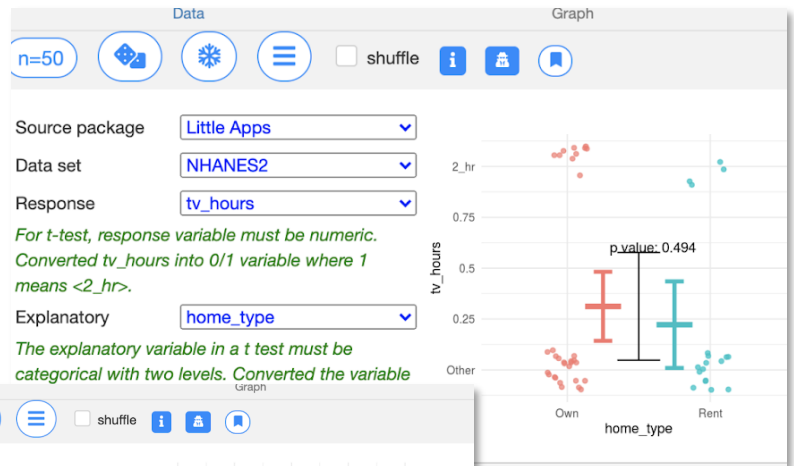
One app is Center and Spread, which calculates and displays descriptive and inferential statistics. You can display the mean, median, and a ruler that shows the standard deviation. You can also show the confidence interval and the coverage interval. All of this can be displayed in a violin plot.

Another Little App is Points and Densities. This app shows a density graph for quantitative variables. There are several options available, such as superimposing a normal curve on the density curves. Other options include displaying the center as the mean or the median. You can also see what percentage of the data falls within certain standard deviations from the mean. The bars can be moved to change at what z-score the lines are placed. This Little App can be useful when discussing the z-score and the empirical rule.



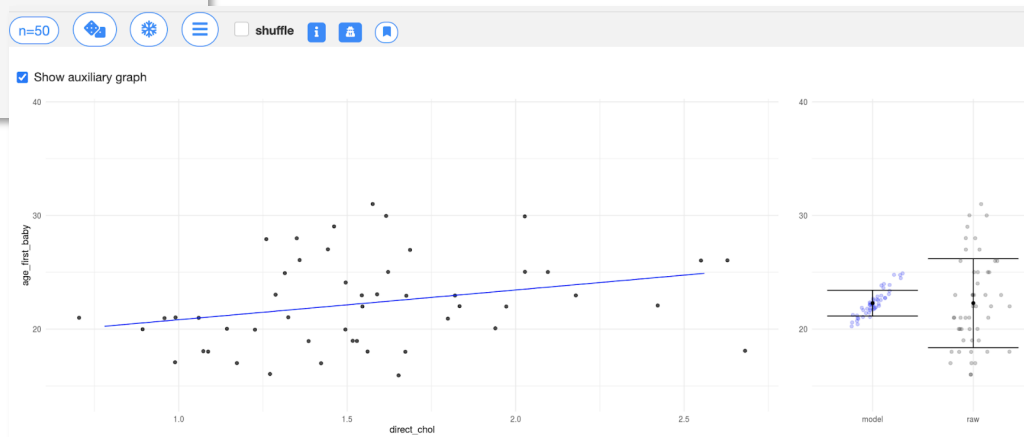
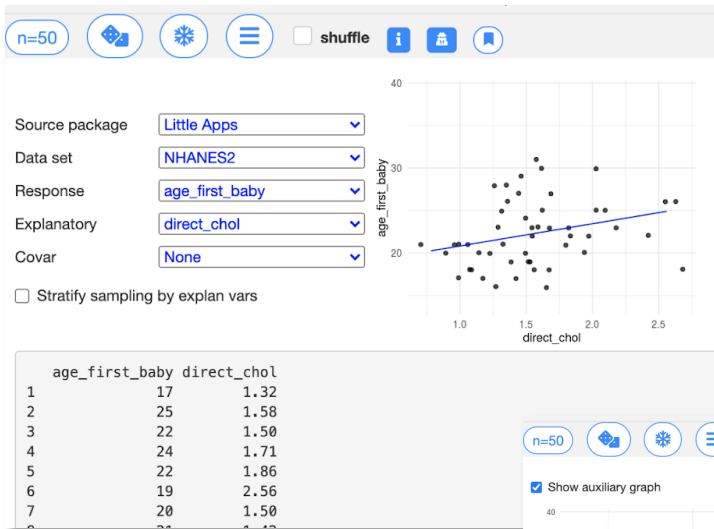
LITTLE APPS Continued...

The Confidence and T Little App can be used to demonstrate hypothesis testing and confidence intervals. You can display both one and two sample hypothesis tests, as well as means and confidence intervals. The t-interval shows the difference in means. If you don't pick an explanatory variable, you can look at a one sample test and compare the sample mean to a known mean.



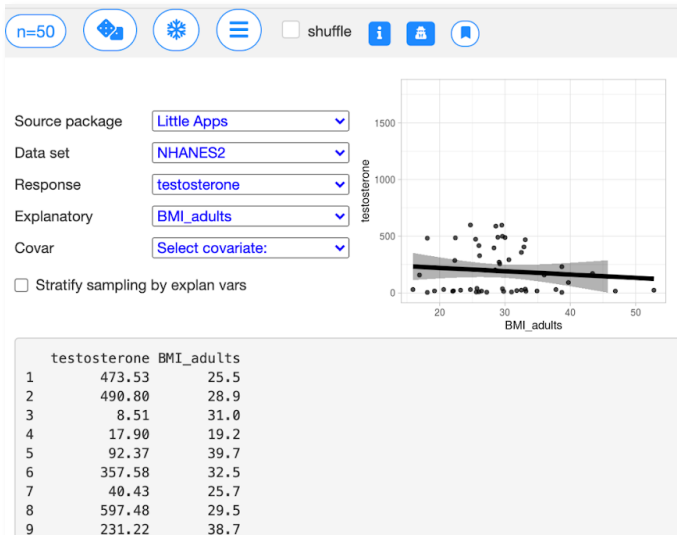
Regression is one of the more important concepts in statistics. This graph can be used to determine if there is a relationship between the variables. A line is displayed so you and your students have a reference to determine if there is a linear relationship.

In addition, a graph can be displayed to show how much variability is in the data and how much is explained by the model. This will give a visualization of the coefficient of determination.



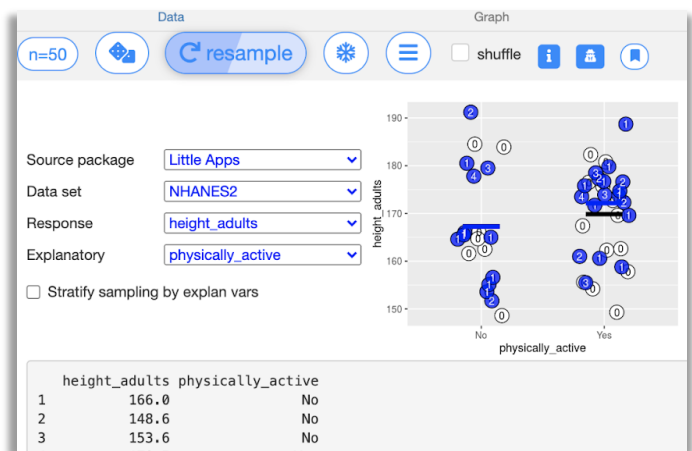
...continued on page 6

LITTLE APPS Continued...



The Stratification and Confounding Little App is useful for looking at stratification or “layering”. This app divides the covariate into layers. That way, we can show the relationship between the response and the primary explanatory variable separately for each layer of the covariate. These are the most popular Little Apps and while there are activities that complement each of these, the apps can also be used as group work activities or class discussions.

There is one more Little App that can be used to learn more modern techniques. It is a great way to learn more about resampling and display it to your students. To demonstrate the process, just click on the resampling button to watch the process take place.

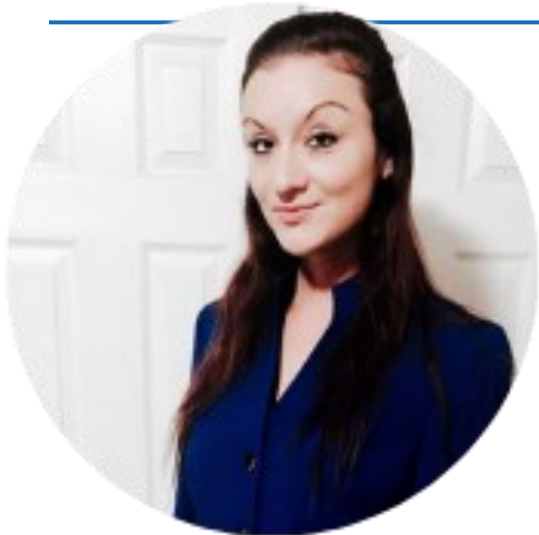


The Stratification and Confounding Little App and the Resampling Little Apps do not have activities, but they can still be useful in your classes. Whether you use activities or determine your own ways to use the Little Apps in your classroom, you will find that your students are able to visualize statistics better than they have in the past. I hope you find these tools useful in your classes.

SUMMER DATA SCIENCE WORKSHOPS: ALL ABOUT THE PANELISTS

We hope that you will join us on June 4 or July 31 from 12-3pm ET!

[Register Now!](#)



Crystal Wiggins is a Professor of Mathematics and STEM Department Chair at Northwestern CT Community College. Crystal has taught in higher education for over a decade, teaching courses ranging from basic math and advanced calculus to statistics and data science. She spent two years researching the field of data science in order to develop an associate degree program and a certificate in Data Science. The first cohort for these programs started in Fall 2019 and the first graduates are set to walk the stage May 2021. She has attended StatPREP as a participant two years in a row and is now excited to share her experiences and resources as a panelist.

Professor **Kelly Fitzpatrick** is an Associate Professor of Mathematics and President of MATYCNJ from the County College of Morris (CCM) located in Randolph, NJ. She is also the principal investor on "Expanding Pathways to a Data Science Career by Developing a Certification in Data Science and Analytics," NSF-ATE Award 2000887. CCM has developed a 16-credit certificate in Data Analytics. The certificate includes three new classes: Introduction to Data Science (R), Data Science Programming (Python) and Machine Learning. The certificate also includes Probability and Statistics, Database Programming, and a Fundamentals of Programming class. Professor Fitzpatrick looks forward to sharing the development and ideas of expanding Data Analytics initiatives at a two-year college.



PANELIST BIOS Continued...



Rachel Saidi is the Math, Statistics, and Data Science department at Montgomery College in Rockville, MD. She is currently also the Data Science Certificate Program Coordinator. For this program, students must complete five courses, or 16 credits, to earn the certificate: one statistics class of their choice; Introduction to Data Science (DATA 101), which focuses primarily on data ingestion and cleaning; Data Visualization and Communication (DATA 110); Statistical Methods in Data Science (DATA 201), which includes advanced statistical methods, machine learning, clustering and regression; and the Capstone Experience in Data Science (DATA 205) where they work with an industry or governmental partner. Students are able to complete this certificate in three semesters or even two semesters if they have already taken a statistics course, though some will take more. She will talk more about the program and challenges faced along the way.

Jim Ham has been teaching at Delta College, a Michigan two-year college, since 1994. In recent years, Jim has taught classes mostly in the pathways: introductory statistics and the math for liberal arts courses. Like most of his colleagues, his courses have been delivered remotely during the pandemic. In the past couple of years, Jim led a team of novices in the creation of a new data science certificate and degree at Delta College. Jim has served in several leadership positions in AMATYC, including his current role as AMATYC Past President and Chair of the Foundation Board. Jim has also served as an AMATYC affiliate (MichMATYC) president and Vice Chair of the Michigan Section of the MAA. Jim currently serves on TPSE-MAG and is involved in several other national organizations, initiatives, grants, and committees.



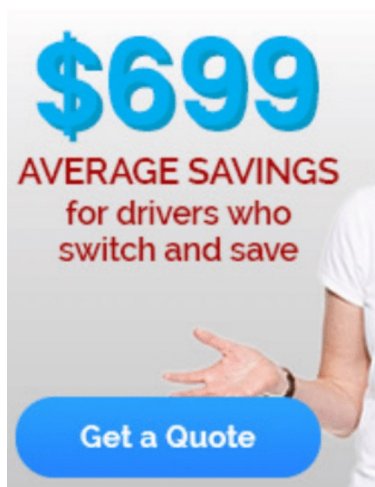
BUILDING COMMUNITY THROUGH DISCUSSION BOARDS: POPULATIONS

BY AMIBKA SILVA

In my statistics class, we talk about populations versus samples. A population is a whole, it's every member of a group. A sample is a fraction or percentage of a group. It got me thinking about a few commercials that talk about averages and what sample and population they were utilizing and if there is bias in the advertising.

Sample prompt

You may see commercials or advertising that claim the average savings for drivers who switch is high. Or that drivers who save, average certain savings a year. I want you to think about who that sample is, and who is excluded from these numbers.



Answer the following questions in complete sentences, then respond to at least two classmates. You must have a minimum of 10 sentences total for your first post.

1. Let's say that 100 people are intrigued by the commercial and decide to get a quote. About 30 of them decide to switch. What are possible reasons that the 30 people switched? Why do you think the other 70 did not switch?
2. For the 30 that switched, could they be part of the statistics offered by the commercial? Can we be sure they would be part of the statistic? Why or why not?
3. For the 70 that did not switch, could they be part of the statistics offered by the commercial? Why or why not?
4. What are the characteristics of the population that both insurance companies above are using? Meaning, who has a chance to be part of their given statistics.
5. The 100 people were hypothetical. Do we know how many people switched based on what they are telling us? Does that mean anything to us as consumers? Why or why not?

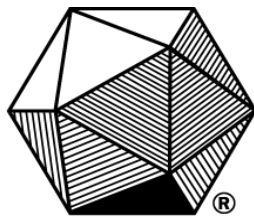
BUILDING COMMUNITY Continued...

What am I hoping they discuss or understand afterward?

For this discussion, I want students to think about what is given may not tell the whole story. To me this is a critical thinking discussion where I want them to use their vocabulary of populations and samples. In this case, the averages are only for those who did switch. What if out of 100 people only 4 people switch because they got savings? Does that impact whether I choose to get a quote from that company? Is it misleading or honest for companies to tell us that this is from only people that chose to switch?

Join the Discussion!

Do you have thoughts on using this as a discussion post? Join the conversation online at MAA Connect, <https://connect.maa.org/home>.



MAA
CONNECT

**Need Help
Joining?**

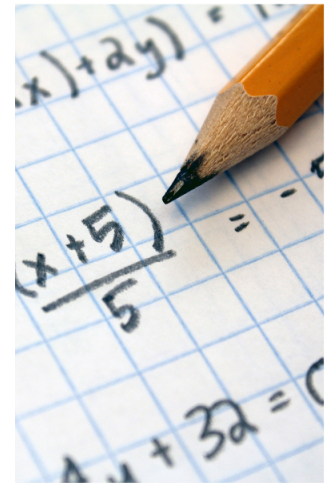
[Getting Started Guide](#) [Getting Started Video](#)

[StatPREP October Webinar](#)

WORKSHOP CALENDARS

JUNE

SUN	MON	TUE	WED	THU	FRI	SAT
		1	2	3	4 Data Science Workshop 12-3 ET	5
6	7	8	9	10 Little Apps Workshop 12-3 ET	11 Little Apps Workshop 12-3 ET	12
13	14	15	16	17	18 R Intro Workshop 12-3 ET	19 R Intro Workshop 12-3 ET
20	21	22	23	24 R Advanced Workshop 12-3 ET	25 R Advanced Workshop 12-3 ET	26
27	28	29	30			



JULY



SUN	MON	TUE	WED	THU	FRI	SAT
				1	2	3
4	5	6	7	8	9 Little Apps Workshop 12-3 ET	10 Little Apps Workshop 12-3 ET
11	12	13	14	15 R Workshop 12-3 ET	16 R Workshop 12-3 ET	17
18	19	20	21	22	23 R Advanced Workshop 12-3 ET	24 R Advanced Workshop 12-3 ET
25	26	27	28	29	30	31 Data Science Workshop 12-3 ET