

## 1. Name (first and last)

**Text Response**

Emily Sheen

**Statistic**

Total Responses	Value
1	

## 2. Email

**Text Response**

sheen@psu.edu

**Statistic**

Total Responses	Value
1	

## 3. Contact Phone

**Text Response**

(623)332-5558

**Statistic**

Total Responses	Value
1	

#### 4. In Fall 2018 you will consider yourself to be a:

#	Answer	Bar	Response	%
1	VT Freshmen Undergraduate Student		0	0%
2	VT Sophomore Undergraduate Student		0	0%
3	VT Junior Undergraduate Student		0	0%
4	VT Senior Undergraduate Student		0	0%
5	Between Undergrad and Graduate school		0	0%
6	Graduate school		1	100%
	Total		1	

Statistic	Value
Min Value	6
Max Value	6
Mean	6.00
Variance	0.00
Standard Deviation	0.00
Total Responses	1

#### 5. What degree(s) and major(s) are you pursuing along with institution?

Text Response
Penn State PhD Statistics

  

Statistic	Value
Total Responses	1

**6. List any Minors you are pursuing.**

Text Response

Statistic

Total Responses

Value

0

**7. List any honors and/or awards received:**

Text Response

Graduate Teaching Assistantship

Statistic

Total Responses

Value

1

**8. Currently we have 1 program accepting applications which includes:**

#	Answer	Total Responses
15	Data Science for the Public Good	0
	Total	0

  

Statistic	Data Science for the Public Good
Min Value	-
Max Value	-
Mean	0.00
Variance	0.00
Standard Deviation	0.00
Total Responses	-

**9. Essay (up to 500 words): "What do you want to get out of this experience?"**

Text Response

As a DSPG Fellow, I want to gain research experience on real, pressing issues affecting governments and communities in the US right now. I think my strong background both in economics and statistics makes me a versatile applicant, eager to engage both with policy leaders and technically driven researchers in statistics. The reason I chose to study statistics was the opportunity to collaborate across disciplines, and use statistics to solve actual problems in our economy and communities. This opportunity at DSPG is perfect for me, because it offers the chance to build connections with policy and government leaders, while also expanding my technical background and preparing me to write and publish papers in the research areas that excite me. During my undergraduate studies in economics, I developed a strong thirst for understanding social issues and how government policies can both help and greatly hinder economic growth, access to healthcare and education, and employment. I developed an economist's intuition for understanding the complexities underlying millions of people interacting in our society, and learned the difficulty of producing rigorous research that can generalize complex human behavior into meaningful statistics and actionable results. At DSPG, I hope to continue grow my economic intuition for a greater variety of social issues, and expand my technical toolkit and programming skills to produce high quality research in the social sciences. I want to collaborate and write research reports with the insight of experienced policy leaders and social scientists in the heart of our nation's capital, and grow both technically and professionally as a researcher and collaborator. Social sciences research involves both hard and soft skills. The social scientist needs the statistical tools to deal with huge and imperfect datasets, and also the soft communication and professional skills to engage with people from different backgrounds and gain insights from collaborators. This research opportunity at DSPG is perfect for expanding both my hard and soft skills to make me a stronger researcher and professional as I start my career as a statistician.

Statistic

Total Responses

Value

1

**10. Essay (up to 500 words): "Please describe any previous research experience and/or work experience you may have."**

Text Response

During my time at the University of Arizona, I conducted my honors thesis research on high school AP calculus enrollment using data from Civil Rights Data Collection and the National Center for Education Statistics. Using SAS, I merged and cleaned education data from over 18,000 schools across the United States, to study the relationship between calculus enrollment and factors like race/ethnicity, school size, percentage of low-income students, and other demographic variables. This research familiarized me with data cleaning and dealing with imperfect government data, and helped me develop skills in programming, statistical modeling, multiple linear regression, and descriptive and conditional statistics. The hands-on research amplified my excitement for government and social sciences research and convinced me to pursue my PhD. After graduating from the University of Arizona in May 2016, I worked as a financial analyst at a utility company, Arizona Public Service (APS), for one year before I started graduate school at Penn State. Although my role at APS was not related to statistics or the social sciences, it helped me develop strong communication and professional skills to deal with different personalities and work in a corporate setting. I frequently engaged with top executives at APS, and met with the Vice President of Customer Service monthly to discuss financial results, budgets, and spending plans for that department. This experience gave me confidence to present my work to powerful people, and aid in executive decision making. In August 2017, I started my first year in Penn State's Statistics PhD program. The first year of coursework at Penn State is rigorous and challenging, and does not leave extra time for starting research. However, I did one applied statistics project for the regression course, STAT 511. With the help of Dr. Ephraim Hanks and two other students, we examined elk movement data from a herd in the Yellowstone region, and used R to perform spatial multiple linear regression using velocity vectors as the response and gradient vector covariates. We examined how covariates like elevation and terrain roughness correlated with elk movement over time, and tracked the elk's spring and fall migrations to lower and higher elevations. Although the project was short and limited in scope, it provided a valuable introduction to spatial regression and the challenges in that field. In my first semester, I also completed a literature review and presentation on Bayesian decision theory topics and Value of Information analysis for the colloquium course, STAT 590. Together with two other graduate students, we read three papers on Bayesian decision theory topics, and wrote an in-depth review of one of the papers: "Estimating the Expected Value of Sample Information..." by Mark Strong, Jeremy Oakley, Alan Brennan, and Penny Breeze. In our review, we summarized the motivation, methods and results of the paper, and future research that could be done in this field. We concluded the course with a presentation on Value of Information Analysis, where we taught other graduate students what this analysis entails and why it is valuable in decision theory.

Statistic

Total Responses

Value

1

**11. Essay (up to 500 words): "Please describe your background (courses taken, research projects, etc.) in statistics and mathematics."**

## Text Response

Please see the previous essay for details on my research experiences in statistics and mathematics. During my first semester of the PhD program at Penn State, I took two theoretical statistics courses: Theory I and Theory II. Theory I focused on probability and distribution theory, conditional expectation, real analysis, moment generating functions, series and sequence convergence, and likelihood ratios. Theory II focused on point estimators, complete and sufficient statistics, UMVUE, UMP, hypothesis testing, Bayesian prior and posterior distributions, and confidence intervals. I also took one applied regression course using R, which covered multiple linear regression, generalized linear regression, Ridge and LASSO regression, regularization techniques, and exploratory data analysis and reporting. This semester I am currently taking one Asymptotic Theory course, one Markov Chain and Stochastic Processes course, and one experimental design course. The Asymptotic Theory course will cover weak and strong convergence of random variables in both the univariate and multivariate settings, Slutsky's theorem(s), delta method, the Lindeberg-Feller central limit theorem, likelihood-based estimation and testing, and other topics such as sample quantiles. The Markov Chain and Stochastic Processes course will cover conditional probability, discrete time/state Markov chains, Poisson processes, continuous time Markov chains, classic Monte Carlo Methods, and Markov Chain Monte Carlo. The Design of Experiments course will cover linear algebra, the linear model, principles of design such as causal inference, randomization, and replication, one-way ANOVA, factorial and nested designs, multivariate normal theory, random effects model, multiple comparisons (Tukey, Sheffé, false discovery rate), causal inference from observational data, survey sampling, and optimal design for computer experiments. During my undergraduate studies, I took mathematics courses on probability theory, statistical theory, linear algebra, real analysis, calculus, differential equations, and economic statistics.

## Statistic

Total Responses

## Value

1

**12. Essay (up to 500 words): "Please describe your background (courses taken, research projects, etc.) in social and behavioral sciences."**

## Text Response

Please see the previous essay for details on my economics honors thesis research. During the pursuit of my BSBA in economics, I took many courses in the social and behavioral sciences, including game theory, econometrics, world economic history, microeconomics, macroeconomics, economic writing, data analysis and economic modelling, economic statistics, ethical issues in business, business statistics, business communication, quantitative finance, organizational behavior and management, and other business classes.

## Statistic

Total Responses

## Value

1

**13. Essay (up to 500 words): "Please describe your background in programming."**

## Text Response

I used SAS for three semesters during my undergraduate studies, while conducting my yearlong honors thesis project and while taking the data analysis and economic modelling course. I used STATA for one semester for my econometrics class, and took two programming courses in Java during my undergraduate studies. Since starting the PhD program, I have worked for 7 months in R and completed the spatial regression project using R. My experimental design and Monte Carlo and Stochastic Processes courses rely heavily on using R.

## Statistic

Total Responses

## Value

1

**14. Essay (up to 500 words): "Please provide information about other significant courses you have taken within your field of study."**

## Text Response

All of my essential courses have been listed above.

## Statistic

Total Responses

## Value

1

**15. Please list the name and contact information for 2 references (teachers, mentors, or employers) that we will contact for a letter of reference/brief survey. Please make sure you list the correct email and they know we will be contacting them soon (within the next week). Only 2 references will be contacted; do not list more than 2.**

## Text Response

Dr. Aleksandra Slavkovic, sesa@psu.edu; AND Dr. Ephraim Hanks, hanks@psu.edu

## Statistic

Total Responses

## Value

1