

I designed my degree with two purposes in mind, being able to get a job when I graduate and being able to jump into a Master's degree when I so choose. I recently married someone who is planning on attending medical school in a little over a year, and I wanted to be prepared to go just about anywhere and either find a job or continue my education.

The job I would like to have after graduation involves working on a team of programmers handling large datasets for big companies. I want to transform the data, analyze it, find trends, and then use visualization to present the results. This is in the realm of being a Data Scientist. However, if such jobs aren't available, my degree will have trained me to work in many jobs that require programming, statistics, and visualization. It is a very versatile field, and my degree will give me a good foundation to start a career.

If I have the opportunity to start working on a Master's degree, my degree will put me in a good place to get started. I researched Data Science degrees at other schools across the country to help me choose the right classes. In particular, the classes I have chosen include a good number of the prerequisites for the Data Science Master's Certificate here at the University of Utah, as well as the Master's degree in Econometrics, which is what I would choose to do if I stay here.

When I first arrived at the University of Utah I wanted a degree in Statistics. I was disappointed to find that the Math Department only offers an *Emphasis* in Statistics degree, which contains many classes that I did not find useful and I was unwilling to take. It was around that time I began to discover Big Data, and was realizing my great interest in computer programming. I did a lot of research on the topic, and tried to find a major that would mix both of my interests. A Computer Science degree did not offer the advanced statistics skills that I really wanted, neither did the degree in Information Systems. There was no minor in statistics available, and the minor in Computer Science would not give me the “Data Mining” and “Visualization” classes that I knew would be important if I wanted to be a Data Scientist. So I decided to come to the BUS program so I could pick and choose the classes that would help me most.

The classes I chose will give me a wide range of skills that will be valuable to employers. The first cluster of classes I chose are in the Math Department. They include all the important Statistics classes that I could take with their prerequisites. Being able to analyze data and use R (which I have been learning in my Stats classes) is very beneficial and give me an edge when it come to Data Science, because most graduates with the same degree from other schools will not have as much math background as I do, and a good math background it valuable to employers. It will also help me start on a Master's degree in Statistics when I so choose.

The second group of classes I chose were for the Computer Science department. I knew that it was very difficult to get a job with only a bachelors degree in Statistics, because you will be competing with Master's graduates. I also knew that programming is a valuable skill, and these classes will teach me languages and algorithms that will make it much easier for me to get a job right out of college. I chose my classes based around how I wanted to use them with Statistics, and I made sure that I will take a class that teaches SQL, a language that is very valuable if I want to get a job. These classes also fulfill many of the prerequisites of the Data Science Master's Certificate.

The last group of classes I chose were a small number of economics classes. I chose to take these because not only did they interest me, but I knew that I was much more likely to get a job if I knew the lingo about certain subjects. Economics and Statistics are intertwined pretty often, and I thought that it could give me an edge at a bigger company if I knew how to navigate economic terms and problems while doing my job.

My thesis project has not been decided on yet, but I plan on working on some form of research where that data being collected requires me to use programming to place it in a format that I can work with, use statistics to analyze it and make predictions, and then more programming to create visualizations of the data before I present it to the program director and interested parties.

I am very excited about my chance to work on this degree, and that such an opportunity was proffered by the University. Thank you!

To whom this may concern:

I am writing in support of Sierra Allred's request for specialized degree under the Bachelors of University Studies; under the theme of "Data Science." It is a blend of Computer Science, Mathematics, and Economics, focusing on the data analysis courses in each area. I have agreed to be her faculty advisor towards this degree, and am completely supportive of Sierra.

Sierra first approach me about pursuing such a BUS degree about a year ago, and had a very clear vision of her interests. At the time I was also supervising another student who followed a similar BUS degree (Sitaram Gautum). Guided by my experience with Sitaram, I believe the proposed degree for Sierra is a better fit for the career as a data scientist and more rigorous.

### **Why is there a need for a specialized Data Science degree?**

Many computer science students (my home department) do not focus strongly enough on the mathematical skills needed for a career as a data scientists. There are a set of elective classes (Machine Learning, Data Mining, Databases, Visualization, Algorithms) offered for advanced undergraduate students. But typically in order to take these classes, a student would have already taken so many classes that they could graduate, and it usually makes sense to do an BS/MS degree. However, many of the other required CS classes, those focusing on hardware and software engineering training, are only tangentially useful for a career in data science.

On the other hand, the Economics department and Mathematics department offers many other classes in probability and statistical analysis. These go well-beyond what is covered for undergraduates in the CS department. And are very useful for more advanced statistical modeling that is covered in CS.

In short, Sierra's proposed BUS degree omits some hardware and software engineering requirements from the CS side, and replaces a single CS probability and statistics classes with 5 such classes from the Math department, and some basic mathematical Economics classes.

As far as I know, there are no specified programs for undergraduates which provides this breadth in training. The College of Engineering has recently put forward an undergraduate certificate in "Data Center Engineering" that can be received while getting a BS in Computer Science, Mechanical Engineering, or Electrical Engineering. But this is about the operations in a data center, and much less about the analysis of that data.

### **Is this a worthwhile program?**

Many other students are taking a similar career to Sierra, only hers is more streamlined. There is a specific MS (and PhD) "track" in CS where students can take many of the same classes (Machine

Learning, Data Mining, Databases, Visualization, Algorithms) as the basis of an MS degree. Moreover, due to their popularity, we have recently (started Fall 2014) a "Graduate Certificate in Big Data," which requires all of these classes. It already has about 10 students in this program, they are mainly students who got a regular CS degree but realized that what they really needed were the skills in these classes and are going back to the U to now take these precise classes. There are forms of each of these 5 classes at the intro grad and advanced undergraduate level, and are all part of Sierra's BUS degree plan.

In fact, many CS undergraduate students this year have asked me about taking a similar set of the data CS classes that Sierra is planning to take. Unfortunately, because of the many other CS requirements, they often cannot fit them in. And sometimes, they take the probability/statistics prerequisites for them too late. Sierra is in a fortunate position to have planned this degree out far enough ahead of time to be able to fit these classes in her schedule.

Moreover, I am involved in discussions across several departments (CS, Math, BioEngineering) towards building the University's emphasis and curriculum in these areas. There are plans under way towards creating a Statistics center (and if things go well, perhaps soon-after a department). These discussions have brought up the idea (which many are in favor of) of making this very Data Science focused. If it takes this direction, which I hope it well, then Sitaram and Sierra are the pioneers in this program for us.

I myself wish such a degree had been available when I was a student. I took many of the similar classes as Sierra is proposing, but had to get two separate degrees (Computer Science and Mathematics) to do so.

#### **Will there be jobs available after such a degree?**

Yes, yes and yes. The position of data scientist is one of the fastest growing (and lucrative and rewarding) new jobs of this decade, and, in my opinion, will continue to be so into the future. I get contacted by potential employers all of the time asking for recommendations of students with these skill sets. For instance, Goldman Sachs has recently been saying they are planning to hire many analysts in this area in Salt Lake City in the coming years. They want to hire students with only undergraduate degrees, and mainly do so from places other than the U. Most of our students who have such skills are already at the masters level.

I approve Sierra's proposed degree program. Some of the upper level CS data classes have prerequisites that Sierra will not plan to take. I have personally talked with the usual instructors of all of those classes, and they have no problem with it. These requirements are either to ensure the student is mature enough academically (as proxy for major status) or to ensure they have the statistics and probability background which Sierra has from her extensive math classes.

I will also supervise Sierra towards a senior project. I will plan to help guide Sierra towards a project topic so she can start pursuing it this summer. I have experience with this including with Sitaram last year, and two other undergraduate thesis students I am supervising this year. There are a tremendous number of great projects related to Sierra's interest, so I have no worries about this materializing.

In summary, I am very supportive of the "Data Science" focused degree program that Sierra has designed. My experience with supervising her will be of great help towards the university in potentially designing a formal degree program in this area in the future.

Sincerely,

A handwritten signature in black ink that reads "Jeff Phillips". The signature is written in a cursive style with a large, stylized "J" and "P".

Jeff M. Phillips  
Assistant Professor  
Director, Data Management and Analysis Track  
School of Computing  
University of Utah

### Major Emphasis Courses

Code	Course #	Course Title	Department	Hours
+	1210	Calculus 1	Mathematics	4
+	1220	Calculus 2	Mathematics	4
+	2210	Calculus 3	Mathematics	3
ip	2270	Linear Algebra	Mathematics	4
+	3070	Applied Statistics 1	Mathematics	4
ip	3080	Applied Statistics 2	Mathematics	3
	5010	Intro to Probability	Mathematics	3
	5080	Statistical Inference 1	Mathematics	3
	5090	Statistical Inference 2	Mathematics	3
+	1410	Intro to Object-Oriented Programming	Computer Science	4
+	2100	Discrete Structures	Computer Science	3
ip	2420	Intro to Algorithms & Data Structures	Computer Science	4
	3500	Software Practice	Computer Science	4
	4150	Algorithms	Computer Science	3
	5140	Data Mining	Computer Science	3
	5350	Machine learning	Computer Science	3
	5530	Database Systems	Computer Science	3
	5630	Visualization	Computer Science	3
+	2010	Princ Of Microeconomics	Economics	3
+	2020	Princ of Macroeconomics	Economics	3
ip	3620	Mathematics for Economists	Economics	3
<b>Total Major Emphasis Hours:</b>				<b>70</b>
<b>4000 level and above:</b>				<b>24</b>