Professional Development Workshops to Increase Data Science Education in Delaware High Schools

James Skripchuk

May 2019

1 Executive Summary

Data science is one of the fastest growing fields in the world right now. While Delaware is doing a good job at promoting data science at the collegiate level, experts agree that foundational data science skills should be taught as early as high school. In this proposal, I address the lack of data science education in Delaware high schools and detail the possible consequences if we continue to ignore this problem. Without a giving a strong data science background, we are ill-preparing our future citizens and workforce.

To solve this problem, I propose that we start a series of professional development workshops for high school educators. I will first survey students and high school teachers on their current knowledge about data science and use these results to help decide subject matters for the workshops. Concurrently, I will be networking with Delawarian data science academics, professionals, and educators who are willing to help with the initiative. After reviewing the survey results and notes with various professionals, I will set up two preliminary meetings and invite data science professionals to discuss how we can create a strong data science curriculum in Delaware. These meeting will cover a variety of topics in data science, such as statistics, machine learning, and the current state of the industry. This project will be relatively cheap, only costing 2,000$ for website design and catering costs.

The end goal of this initiative is to create a "ground-zero" for data science education in Delaware, and hopefully serve as a launch pad for future work of a Delaware high school data science curriculum.

2 Introduction

What is Data Science?

Data science is an interdisciplinary field that applies mathematics, computer science, and domain knowledge in order to gather meaningful and relevant knowledge from unstructured data. Due to the rapidly increasing growth of technology, institutions ranging from healthcare to finance have massive amounts of data that are waiting to be converted into meaningful results. A good Data Scientist uses cutting edge techniques in order to extract insights from raw data, and can effectively report to stakeholders the implications of the results.

The current state of Data Science in Delaware.

Delaware has started to become aware of the necessity of good Data Scientists, as evident by the founding of the University of Delaware's data science Institute (DSI) this past year. While this is an amazing start, we as a state can improve further by introducing data science as early as high school. IBM states that "Ideally, students graduating from high school should already have reached a base-line data literacy that they can then apply across college and university departments" [3]. Delaware is already on track by requiring all public high schools will offer at least one computer science course by 2020 [1]. However, currently there is no work being done on introducing data driven courses into Delaware high schools.

Without data science education:

Delaware will not produce enough qualified individuals for the future workforce.

IBM states that Data Science and Analytics jobs are the hardest to fill, and in 2015 projected a 5-year growth of 15% to over 2,700,000 job listings by 2020 [3]. A recent survey by Qlik survey states that 55% of employees lack the education and resources to make sound decisions based on insight, and that only 21% of respondents classify themselves as data-literate [4].

We would be critically under-utilizing Delaware's unique resources.

The three largest industries in demand for data science professionals are Professional Services, Finance & Insurance, and Manufacturing. Due to Delaware's unique corporate laws, over 50% of publicly traded corporations in the United States and 60% of the Fortune 500 are incorporated in Delaware [7]. As a result, Delaware attracts a wealth of businesses in need of data science talent. One of Delaware's largest employers is the banking industry, with companies such as Bank of America, JPMorgan & Chase, and Deutsche Bank. Delaware also has a rich history with the chemical manufacturing industry due to it being the home of the DuPont company. In addition to industry, Delaware is also home to the University of Delaware, one of the largest research institutions in the United States. With a critically acclaimed Computer Science department and the newly formed Data Science Institute, Delaware is possibly the best place to start a high school data science initiative.

We would be ill-preparing the next generation of adults to be good citizens in a data-driven world.

Over 95% of teenagers have a smartphone, and over 45% of teenagers say that they are online "almost constantly" [2]. Almost every website or app collects and uses data to personally tailor experiences. While these personalized experiences can be a boon for most cases, poor data-literacy can lead to disastrous results. There have been a host of political scandals centering around the misuse and mistrust of data. Recent examples include the Facebook-Cambridge Analytica affair and the misinformation campaigns by Russia's Internet Research Agency. Teenagers are especially vulnerable to misinformation. To not give our students a fundamental knowledge in data-literacy would be doing a disservice to them as future citizens.

If we continue to ignore this problem, Delaware will lag behind in creating skilled citizens and a skilled workforce. I aim to solve this problem by introducing a series of professional development workshops that will hopefully serve as a starting point for future conversation in data science education.

3 Plan

In order to increase the quality of data science education in Delaware high schools, I propose **to create a series of professional development data science workshops for high school educators**. This plan will hopefully achieve the following goals:

- Educate students and educators about the need for a data science curriculum in high schools.
- Provide an infrastructure for introducing a data science curriculum in high schools
- Increase diversity in high school data science education in Delaware

This plan consists of the three following phases:

Assess State of Data Science in Delaware

Since there has been little focus on data science education in Delaware, there is a lack of statistics on the current weak spots that Delaware has. My first step will be to design two surveys: one for students and one for educators. Without a foundational knowledge in where our weaknesses lie, it will be very difficult to customize a curriculum going forward.

The student survey will consist of a variety of questions related to opinions and familiarity of data science. These questions will first assess how well students know how their data is used in today's world. Further questions will ask them if they are familiar with the basics of data science, and if they are comfortable in mathematics and computer science. For an ease of sample size, this survey will be distributed to Freshman and Juniors.

The educator survey will be distributed to high school teachers in the state. This survey is primarily targeted towards (but not limited to) high school STEM (Science, Technology, Engineering, and Math) teachers. The goal of the teacher survey is to assess how comfortable educators are with data science, and if they feel prepared to teach students about data-driven science and analytics.

This survey will be digital and designed using the Qualtrics survey software. Ideally, the survey would be distributed to all high schools in Delaware, but due to time and budget constraints it will first be limited to Red Clay

Consolidated School District (5 public schools, 2 charter schools). In order to incentivize responses, there will be a set of prizes worth $500 in total that will be distributed to randomly selected survey participants. The prizes will most likely be gift cards for a popular online shopping retailer.

Network with Academics, Professionals, and Educators

Concurrently, I will be spending my time finding and meeting with potential contacts who would be willing to be part of this initiative. Meetings will be discussing what makes a successful data science program for high schools.

Academics

The University of Delaware is home to the recently founded Data Science Institute (DSI). The mission statement of the DSI is "to accelerate research in data science, serving as a nucleating effort to catalyze interdisciplinary research collaborations across fields impacting our society." [5] They also state that "the Institute will further involve partnerships with industry and other institutions in the region." I will reach out to the Institute and inform them about this plan. Hopefully, this will lead to a fruitful collaboration and a wealth of professors who are willing to share their knowledge of data science pedagogy.

Industry Professionals

Another avenue that I will explore will be contacting industry professionals. Many industries have outreach programs that are willing to be put into contact with initiatives that seek to improve the local community. Some of the major companies I will contact include

- JPMorgan & Chase
- Bank of America
- Christiana Care Health System
- DowDuPont
- Chemours

Delaware also has a wealth of independent startups that reside in Wilmington and Newark, so I will reach out to them as well.

In addition to contacting companies directly, I will also network with the community organization Open Data Delaware. Open Data Delaware is home to over 500 "civic-minded" data professionals ranging from developers, teachers, and entrepreneurs [6]. The group hosts weekly meetings where data professionals gather to discuss policy and work on projects. Open Data Delaware will serve as a valuable resource to learn community event organization.

High-School Educators

I will also reach out to high school educators. Having a personal conversation with STEM educators will supplement the survey results greatly and allow a detailed look into what resources these educators need. In the interest of time, I will limit myself to networking with high school teachers in the Red Clay school district.

After this networking spree, I hope to have a sizable list of data-minded contacts who are passionate about data science education and will be available to help design and teach workshops. I will also have a compendium of notes of what would make a successful data science program based off of the conversations that were had.

Host Professional Development Workshops

Using the results of the student and teacher survey and notes compiled from conversations with professionals, I will organize a series of semi-monthly professional development workshops for high school teachers. These workshops will bring together data professionals willing to teach as well as high school educators who are willing to learn. These workshops will hopefully serve as a ground-zero for further discourse about data science education. These meetings will hopefully consist of approximately 15-30 individuals, and will be open to people willing to help data science education.

These workshops will consist of a variety of topics all centered around data science education. Topics include, but are not limited to:

- Programming Tutorials
- Machine Learning Methods
- Education for Underrepresented Minorities
- Statistics
- Data Science Pedagogy
- Industry Trends
- Professional Networking

New Castle County Libraries are open to the public for activities of "civic and educational nature", so meetings will be held at a local library. The libraries also provide AV equipment for presentations. To promote participation, we would use funding to provide catering for the first few meetings. I will also create a website and mailing list that will be used to keep interested participants involved.

Conclusion

If this plan is put into motion, the result would be a tightly knit community of individuals who are committed and trained to bring data science into the classroom. All parties involved would benefit from the exchange of ideas in this forum, and will take pride in knowing that they are part of a program that aims to further their community.

4 Qualifications

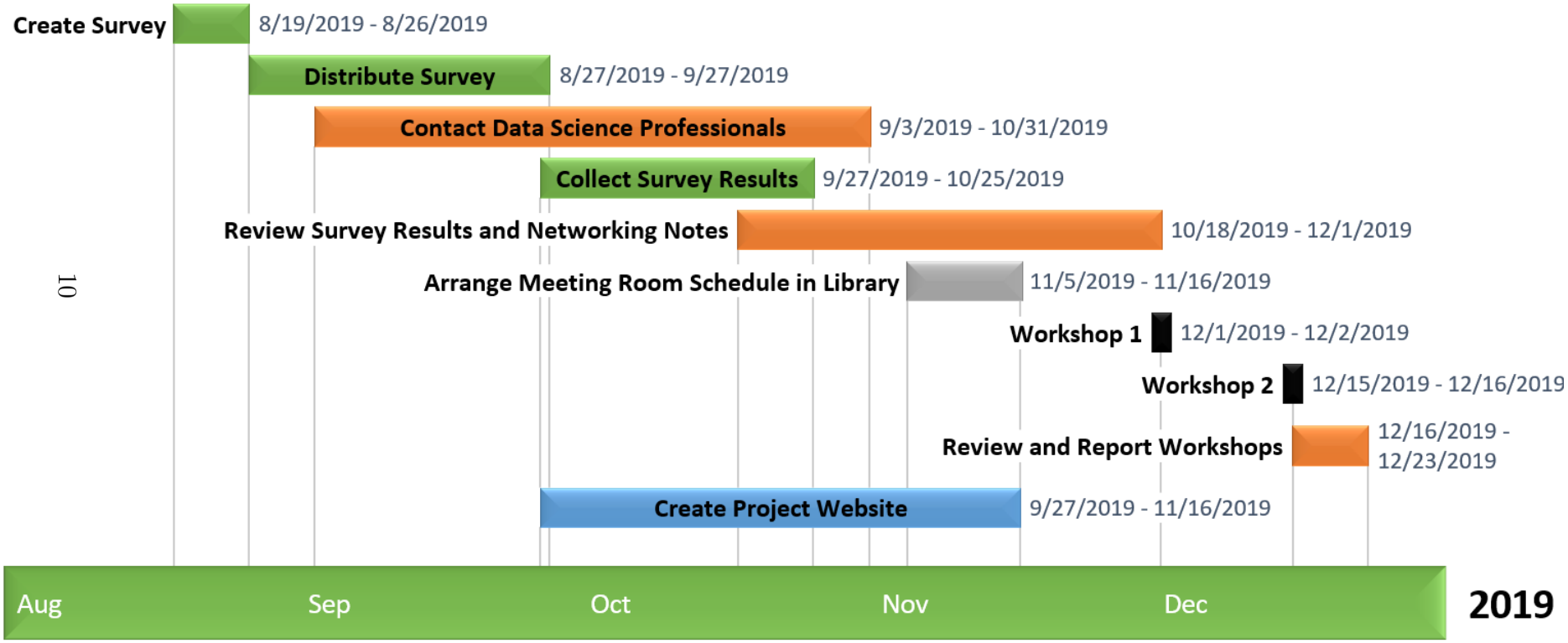
Technical

I am a Junior at the University of Delaware majoring in Computer Science with a concentration in Artificial Intelligence and Machine Learning. I have taken a variety of classes that fall under the umbrella of data science, such as Data Mining, Machine Learning, Artificial Intelligence, and Statistics. I also have minors in Physics, Electrical & Computer Engineering, and Mathematics. I have experience doing cross-disciplinary research at the University of Delaware, and have interned as a software engineer at JPMorgan & Chase for over a year. These technical skills will allow me to talk with confidence about data science concepts among academics and industry professionals.

Organizational

I have experience with designing and organizing coursework via creating a new course at the University of Delaware - MUSC106: Computational Thinking in Music. Designing this course was a year long project, and gave me experience in organizing a large-scale project from start to finish. I was a Teacher's Assistant for this course, and I collected feedback from students directly and adjusted the course accordingly. I am also the Public Relations chair for the University of Delaware's Association for Computing Machinery chapter, and I have helped organize UD's semi-annual Hackathon (an event where programmers gather together and work on projects).

Project Timeline



5 Itemized Budget

This initiative will not take a large sum of money to get started. The most limiting factor is time, which will be spent on finding and contacting willing individuals. The Qualtrics survey software is free for University of Delaware students, so the cost of that is mitigated. As stated, we will reward survey participation by providing gift cards. According to hosting companies, the cost of designing a website yourself and keeping the website up for a year will cost around 100$. Catering assumes that we will spend 25$ on a single meal for 30 people during a period of 2 weeks.

Details	Amount
Website Design and Hosting	100
Survey Prizes	500
Catering	1,500
Total	2,000

References

[1] An Act To Amend Title 14 Of The Delaware Code Relating To School Curriculum And Computer Science, HB 15, 149th General Assembly (De. 2017-2018)

[2] Anderson, M., Jiang, J., Anderson, M., & Jiang, J. (2018, November 30). Teens, Social Media & Technology 2018. Retrieved May 24, 2019, from https://www.pewinternet.org/2018/05/31/teens-social-media-technology-2018/

[3] Miller, S., Hughes, D. (2017) The Quant Crunch: How the Demand for Data Science Skills is Disrupting the Job Market, Retrieved from https://www.ibm.com/downloads/cas/3RL3VXGA

[4] Qlik. (2017, September 26). Data Literacy Campaign. Retrieved May 23, 2019, from http://www.dataliteracy.info/research

[5] University of Delaware. (n.d.). Institutional Strengths. Retrieved May 22, 2019, from https://dsi.udel.edu/

[6] Open Data Delaware. (n.d.). Open Data Delaware. Retrieved May 24, 2019, from https://www.opendatadelaware.com/

[7] State of Delaware. (n.d.). About the Division of Corporations. Retrieved May 25, 2019, from https://corp.delaware.gov/aboutagency/