Final Report on

The University of the South: Data Analysis Institute
Matthew Rudd, PhD
Grant no: NVF-PITU-The University of the South-Subgrant-012785-2020-11-11

Introduction

In alignment with our grant project proposal, Sewanee piloted a data analysis institute for undergraduate faculty and students in order to develop both groups and address community problems in a rural context.

To this end, in the spring of 2021 DataLab identified **four public interest projects** with a data component. **Student, faculty, and community partner participants** were identified to form a team dedicated to working on each project. Two data technologists trained students in a four-week coding for data analysis boot camps, using R. **Visiting alumni data technologists** provided plenary talks and led discussions of PIT and ways they themselves developed data-analysis careers with a social good focus for our interns. **Planning in the fall** has resulted in an expanded version of Sewanee DataLab, which will bring students from across the South to join Sewanee students in this project.

We offer how the project goals were achieved by outlining how we met the project objectives.

Final Report on

The University of the South: Data Analysis Institute
Matthew Rudd, PhD
Grant no: NVF-PITU-The University of the South-Subgrant-012785-2020-11-11

Grant Objectives

Train faculty and students to use data collaboratively in four projects to attack problems of significance to their communities. This hands-on training through a workshop that addressed data issues of various sorts through these projects:

Objective 1: Four Data for Social Good Projects

- **Equitable Broadband Internet Access**
  - **Community Partner:** Thrive Regional Partnership
  - DataLab collaborated with this non-profit based in Chattanooga to develop information that would help the partner achieve more equitable broadband internet access.
  - **Problem**: An insufficient understanding of broadband assets and access to it due to the lack of a broadband asset inventory restricts our ability to explore trends and patterns in broadband access.
  - **Solution**: The team created an interactive open-source map indicating the availability or more importantly the non-availability of broadband assets covering 16 counties in the Chattanooga tri-state area.
  - **Impact**: The interactive map illustrates the disparity in broadband asset and distribution at first glance on the map so that the necessary intervention to bridge the gap in broadband disparity can be made.
  - **DataLab Interns:** Pravesh Agarwal, Feza Anaise Umutoni, Zach Shunnarah (All from Sewanee)
  - **Data Consultants**: Professor of Mathematics, Matthew Rudd, Professor of Environmental Studies, Eric Keen (Sewanee faculty), Ben Brew, data analyst
- **Carbon for Cash in Haiti**
  - **Community Partner:** Zanmi Agrikol, a partnership between the University of the South and *Zanmi Agrikol*, a Haitian organization working with 50 family farmers to grow trees instead of selling material for charcoal through payments made by the Sewanee Green Fund.
  - **Problem:** The Carbon for Cash project needs a uniform process to analyze carbon uptake of trees and pay farmers fairly for the carbon sequestered.
  - **Solution:** The team created a dashboard using R programming language that transforms measurements of carbon sequestered by each Haitian household, as

Final Report on

The University of the South: Data Analysis Institute
Matthew Rudd, PhD
Grant no: NVF-PITU-The University of the South-Subgrant-012785-2020-11-11

well as the number of trees on each small farm, their carbon payments, and the amount of variation in the two datasets.
- ○ **Impact:** Through the work of the DataLab team, Zanmi Agrikol now has a more efficient way of calculating carbon sequestration and associate payments to farmers, and this system will make it possible to expand the project should political conditions allow. the project to expand
- ○ **DataLab: Interns:** Kate Baker, Nika Gorski, Caroline Willett
- ○ **Data and Project Consultants:** Deborah McGrath, Biology and Environmental Sciences, Professor Matthew Rudd, Mathematics.
- **Digitized Images Analysis**
  - ○ **Community Partners**: Haiti Institute and the Ancient Art Archive
  - ○ Sewanee DataLab photos team collaborated with two clients to apply machine learning techniques to curate photographic evidence of the impact of Sewanee's work in Haiti and preserve ancient images that are at risk of degradation.
  - ○ **Problems:** The lack of conclusions/ insights on large files of photos from Haiti and the continuous threat of loss of stories told by rock arts
  - ○ **Solution**: The team used convolutional neural networks to quantify objects and humans in the photos and to label and digitize rock art objects
  - ○ **Impact:** The first stage development of an analytical tool to measure of economic development (through images) and inform conversations about community development and analysis of ancient rock images for understanding and preservation of heritage sites
  - ○ **DataLab Interns**: Katherine Zelaya, Dirk Kayitare, Pratham Singhal, Ngan Nguyen, Sophie Smith (Sewanee students)
  - ○ **Data and Project Consultants**: Professor of Art Pradip Malde, Stephen Alvarez (photographer and founder of the Ancient Art Archive (https://www.ancientartarchive.org/our-work/), Professor of Mathematics Matthew Rudd, Joe and Ben Brew, data analysts.
- **Collection of Public Health Data and Development of analytical tools for identifying Disparities in Tennessee by County**
  - ○ **Community Partner:** South Cumberland Health Network
  - ○ The Public Health team partnered with South Cumberland Health Network to create a reliable resource to analyze those disparities.
  - ○ **Problem:** Wide range of TN health-related information is dispersed across the internet. This leads to difficulty understanding the depth and specific character and location of these disparities.
  - ○ **Solution:** The team collected and consolidated the data from various sources into a database. After that, this database was used to analyze trends and presented in an interactive dashboard online.
  - ○ **Impact:** This project sets the foundation for a centralized location to gain significant health information in TN.
  - ○ **DataLab Interns:** Martha Clark, Mehrael Ibrahim, Sherry Khan, Jeremiah Studivant (Sewanee students) and Esarrah Hopkins and Kirstyn George (Meharry Medical College graduate students)

Final Report on

The University of the South: Data Analysis Institute
Matthew Rudd, PhD
Grant no: NVF-PITU-The University of the South-Subgrant-012785-2020-11-11

- ○ **Data and Project Consultants:** Professor James Peterman, Chair, South Cumberland Health Network, Ben and Joe Brew, data analysts, Professor of Mathematics Matthew Rudd.

Objective 2 Provide mentoring models that participants can adapt to future work with colleagues and students

DataLab brought to the summer project two successful alumni, who graduated from Rayid Ghani's Data Science for Social Good Fellowship Program (https://www.dssgfellowship.org/). They built the undergraduate version of this DSSG program, taught most of it, and were also available as mentors and role models for future data analysts. We also invited fifteen Sewanee alumni who work in data to offer stories and their professional journeys and commitment to data analysis for the social good.

Objective 3: Gain actionable insights into issues that are important to the community
As indicated in the project summations given above, the DataLab projects were result-driven and produced results. The projects were planned in consultation with the various partner organizations.

In the case of the machine learning image identification, the project only allowed for developing the first iteration of a machine learning tool, and that tool will be developed further in DataLab 2022. Moreover, even though the interactive data dashboard for public health data accomplished the goal of creating an interactive public health dashboard, the data it covers is not the data the community partner wished to analyze: hospital discharge data from the Tennessee Department of Health, which was not delivered to the DataLab in the time proposed. So we worked on related publically available data, which the community partner found quite useful.

Objective 4; Gain guidance from experienced practitioners into building a data science program tied to the public interest
The ongoing advice and assistance we received from Sewanee Alumni and our data analysis consultants, provided invaluable insight on how to build the DataLab curriculum and also how to bring a start-up planning framework to our planning for Data 2022. That has resulted in the development of a funding model, that requests payment for the cost of one intern for any organization we partner with, and for collaboration of financial support from two regional higher education consortia, Appalachian College Association and the Sullivan Foundation network.

Objective 5: Build enthusiasm for a sustainability plan
We built an elaborate media and social media marketing campaign to let interested parties and prospective students know about Sewanee DataLab, but also to ensure that the higher administration and Board of Regents were aware of this project and the way in which it exemplifies the highest ideals of what George Kuh has called

Final Report on

The University of the South: Data Analysis Institute
Matthew Rudd, PhD
Grant no: NVF-PITU-The University of the South-Subgrant-012785-2020-11-11

high impact educational practices. That has garnered ongoing conversation in the current academic strategic planning about the need to sustain DataLab and develop a cross-disciplinary data for social good program.

Benefits for PIT UN
   We have developed an exciting and—in-principle—sustainable liberal arts data science for social good co-curricular program that we believe can be sustained and brought to scale. We currently plan to enhance this program through a second iteration of Sewanee DataLab in 2022, which will
- Add to the number of community partners involved in DataLab,
- Provide a potential funding stream through community partners, and
- Develop an extension of this work to other southern schools.
In doing this, Sewanee will become a training hub for schools in our region, with much preliminary work with HBCUs, including our current partners, Fisk University and Meharry Medical College, and other HBCU's in the region.

Also, based on this year's work and our communications about it, one key community partner has taken a keen interest in developing a community-wide data-to-action program. Motivated by training offered this year from the Robert Wood Johnson Foundation-funded Data Across Sectors of Health (DASH) program, DataLab is exploring partnering with the South Cumberland Community Fund to support data-to-action programming their network of community organizations.

Challenges encountered/lessons learned

First Challenge
Our first challenge was planning and designing the intensive boot camp approach to DataLab projects in the midst of COVID 19 and campus restrictions. If the summer school were not open to students, the campus would be closed to special programs. As the spring planning semester evolved, and summer programs were allowed, we were able to move the program to mostly in-person Our initial planning required us to plan the boot camp online, but as the campus open up-to our great relief, we were able to move the main site of the boot camp to Sewanee's campus. Some of our interns were not able to come to the campus and so we developed a hybrid boot camp model. Some students had to join us through Zoom from domestic and some from international locations. The intensity of the boot camp was infectious enough to keep everyone engaged, and the development of collaborative teams that connected the away interns with the on-campus interns helped keep the work of the team moving forward.

Second Challenge

The University of the South: Data Analysis Institute
Matthew Rudd, PhD
Grant no: NVF-PITU-The University of the South-Subgrant-012785-2020-11-11

█████████████ Also, some faculty who were interested in data analysis did not see the unique features of working with local and regional organizations to address data problems. So we hired, through our own funds a marketing strategist to help us design a marketing program. We were able to use other internal funds to hire a student intern to implement our social media plan.

Third Challenge

The most significant design challenge was to design a boot camp for both computer-savvy students with not much background in community and civic engagement and those whose interests in the project were motivated by social justice concerns with little background in coding or data analysis. We found that by putting interns with different skill sets and interests on each team, and with concentrated training on how collaboration works in team projects, and how it was needed that interns find the ways in which their skills could contribute to the projects, we were able to resolve that issue.

██████████████████████████████
██████████████████████████████
██████████████████████████████
██████████████████████████████

Copies of any generated publications or media

1. Sewanee DataLab Social Media Instagram, Facebook, Twitter, and YouTube
2. Sewanee Features: Coding for Brighter Future
3. Sewanee DataLab Campus In-person Presentation Video shared with community partners
4. Press release

IRS Certification

All The University of the South activities conducted with the Grant funds were and are consistent with charitable purposes as set forth in Section 501(c)(3) of the Internal Revenue Code, and The University of the South complied with all provisions and restrictions contained in this Agreement, including, for example and without limitation, those provisions relating to lobbying and political activity.

Intellectual property and assets purchased or created with the Grant.

We created interactive data dashboards that are available to the public free of charge.

Final Report on

The University of the South: Data Analysis Institute
Matthew Rudd, PhD
Grant no: NVF-PITU-The University of the South-Subgrant-012785-2020-11-11

New Venture Fund Budget Template

Proposed Project Budget	NVF Grant Budget \<Grant Period\>	Total Final	Total Project \<Period\>
REVENUE			
Grants and Donations	173,400		
New Venture Fund			173,400
The University of the South			43,350
The Algernon Sydney Sullivan Foundation			16,800
Subtotal Grants & Donations	173,400		233,550
Fee for Services			
Program fees			
Program fees			
Subtotal Fee for Services	-		-
Donated Services and Supplies			
Type of donated supplies			
Administrative staff donated time and benefits			16,080
Subtotal Donated Services and Supplies	-		16,080
Fundraising Campaigns and Special Events			
Name of campaign or event			
Name of campaign or event			
Subtotal Fundraising Campaigns and Special Events	-		-
Total Revenue	**173,400**		**249,630**
EXPENSES			
Salaries	40,000		
Faculty Fellows (summer salary)		40,000	40,000
OCE staff time and effort commitment to project		24,000	12,000
Title or Function			
Subtotal Salaries	40,000	64,000	52,000
Payroll Taxes and Employee Benefits	13,600		
Fellows' Benefits at 34% pooled rate		14,000	13,600
OCE staff benefits for commitment to project		8,400	4,080
Description			
Subtotal Payroll Taxes and Employee Benefits	13,600	22,400	17,680
Subgrants to other organizations (include name if known)			
Description of services	-		
Description of services			
Description of services			
Subtotal Subgrants	-		-
Travel, Conferences and Meetings			
Insitute Materials	2,500	2,400	2,500
Institute Hospitality	5,000	7,135.26	5,000
Description			
Subtotal Travel Conferences and Meetings	7,500	9,535	7,500
Professional/Consultant Services	33,000	38,500	33,000
Honoraria for PIT consultants			
Description			
Description			
Subtotal Professional/Consultant Services	33,000	38,500	33,000
Other Direct Costs	50,400	54,230	
Student support (internship stipends)			
Description			
Description			
Subtotal Other Direct Costs	50,400	54,230	67,200
Indirect Costs	28,900	72250	72,250
Total Project/Program Budget (total should tie to total funding above)	**173,400**	**260,915**	**249,630**
NET OPERATING SURPLUS/(DEFICIT)	**0**	**11,285**	**0**

*Benefits % raised from .34 to .35