DATA SCIENCE FOR THE PUBLIC GOOD

Communities of all types have access to data in greater quantity, detail, and variety than ever before to assess conditions, develop strategies and policies, and evaluate program impact. Rural communities, towns, cities, counties, and states face the challenge of developing a workforce that can support their need to optimally utilize these data in their planning and operations. They need individuals trained in data science and familiar with the social science and community applications of analytics to assist with community-identified issues. Examples include identifying the factors that affect obesity risk, where people live who are vulnerable and why, and policies that prevent people from advancing economically at the neighborhood level defined by the community.

We have developed a model program, **Data Science for the Public Good (DSPG)**, that examines "public good" projects at local (e.g., neighborhoods, townships, and counties, both urban and rural), state, and federal levels of government. The program builds a data science workforce pipeline interested in civic engagement by engaging Cooperative Extension Service (CES) professionals, faculty, junior scholars, and local communities in public good research and provides hands-on contributions to address social challenges, such as economic mobility. The teams examine community-identified issues by geographic areas that are actionable, e.g. school attendance areas, local political districts, or neighborhood levels.

CES professionals act as liaisons by working with community sponsors (those who have identified issues) and faculty, post-docs, and students. Through a collaborative process, the story of each community—its problems, needs, and aspirations—is told using data. These data are a mix of organizational, local, state, and federal administrative and survey and administrative data integrated with place-based opportunity (social media), and procedural data.

The DSPG program is a 10-week fully immersive experiential learning experience that may run during the summer or during the academic school year. A foundational component of the program are teams that are organized vertically across skill levels and horizontally across disciplines. Operating in this team science environment, these teams combine diverse levels of experience academic and local perspectives.

The data science skills of the selected CES professionals, faculty, students, and local sponsors are wide ranging, from none to having varying levels of statistical and data science knowledge. Selection of DPSG teams is a function of interest in both civic problems and learning how data can be used to build evidence-based insights for the public good. By the end of the program, the teams have been introduced to data science methods and tools and have become proficient in using many of them.

The DSPG program immerses the teams in the Community Learning Through Data-Driven Discovery (CLD3) process (see appendix 1). It teaches participants (e.g., local government leadership, CES professionals, faculty, and students) how to discover and sift through vast amounts of information related to topics such as public safety, employment, and the provision of services, and to integrate these data based on statistical and social science methods to develop insights on how communities can improve the quality of life of *all* its residents. The DSPG projects accelerate a facet of these community-based research projects and establish feasibility about whether to continue the research. A selection of these projects are continued to be more fully developed, driving them towards evidence-based policymaking.

As noted in the proposal, an additional byproduct of our approach is the development of a community workforce pipeline to supply new professionals with data skills and interest in applying CLD3 approaches to rural communities. These will include future local government leaders and CES professionals that are data science literate and interested in bringing data in-service of the public good to support economic mobility.