Title: Using Data Science to Understand Soil, Wildfire, & Social Disparity of Climate Change and Air Pollution
Institution: Lincoln University
City/State: Jefferson City, Missouri
PI: Sougata Bardhan
FY: **2023**

Summary: The lack of representation of minority groups in technical and leadership positions is an issue rooted in many industries. When it comes down to ethnicity, we see an extremely unbalanced picture in racial and ethnic diversity in the professional and educational fields of data science. To be more specific, data published in Forbes 2018 report and METIS KAPLAN suggest that fewer than 3% of data scientists are women of color, fewer than 5% are of Hispanic origin, fewer than 4% are African American, and fewer than 0.5% are Native American.

To maintain a competitive advantage in the world, the US needs to increase both the number and diversity of STEM-trained individuals soon. The need for young scientists in science, technology, engineering, and mathematics (STEM) fields for conducting multi and transdisciplinary research has been highlighted by several national and international reports. The under-representation of minority students, including African Americans, in undergraduate and graduate STEM fields, has continued to be a challenge. Data Science training at HBCUs is important given the lack of diversity in data sciences at U.S. government agencies. While diversity in the environmental workforce has increased with the rise of diversity programming during the 1960s, there remains a lack of diversity in data sciences at U.S. government agencies. This proposal will build upon previous efforts supported by NASA focused on earth science at Lincoln University and at other institutes. Data Science, Machine Learning (ML), and AI are going to transform the world in many positive and sustainable ways. Our project will achieve data science problem-solving, skill development, and professional development of minority and underserved students. Each of the proposed research projects is designed to utilize existing state-of-the-art machine learning methods and develop new data analytic approach to solve some of the core problems in earth science research. We also proposed machine learning approaches that can incorporate spatial variability in the prediction and management of forest fires. Another project uses modeling and prediction of soil EC and sodicity through Bayesian hierarchical approaches which can also produce uncertainty quantification. For the last project, our attempt is to measure the disparity of surgical outcomes due to climate change-induced air pollution in relation to socio-economic, racial, and language disparities. The proposed activities will strengthen statistics and data science training, education, experiential learning, and research in HBCUs and draw from the foundational work of the University of Missouri. This connection is important for the future of data science workforce development through minority-serving institutes; and will increase the marketability of minority students as future scientists and part of

an innovative data science labor force. This current proposal is relevant to NASA's 2022 Strategic Plan and is highly relevant to the Strategic Goals/Objectives. Specifically, it is aligned with Strategic Goal 1: To Expand Human Knowledge Through New Scientific Discoveries and overlaps Strategic Objective 1.1: Understand the Earth System and its Climate.