



United States Department of Agriculture

Opportunities for Statistics and Data Science (SDS) Internships in Agriculture. The Sustainable Agricultural Systems Lab (SASL) conducts cutting-edge research to merge science-driven evidence with new digital technologies to increase agricultural productivity and sustainability. The immense amount of data now available from myriad real-time sources (i.e. sensors, satellites, farm equipment) creates new opportunities to develop predictive tools that facilitate farmer and policy decision-making. *We are looking for talented and motivated undergraduate and graduate students interested in honing and applying their statistical and modeling skills to real-world agricultural problems.*

Benefits: Gaining valuable experience and developing skills much needed in SDS positions.
Options: Working at Beltsville Agricultural Research Center (Beltsville, MD) or working remotely, as agreed upon.

Description. Interns will work with data scientists and domain researchers, using their statistical expertise to help improve the accuracy and precision of existing predictive models and prototype new ones. Interns are expected to contribute to refining existing data pipelines by integrating datasets with different formats and levels of abstraction by dplyr or other SQL tools and performing statistical analyses of multi-location agricultural datasets under supervision.

Qualifications. Upper-level undergraduates or graduate students pursuing a degree in statistics, econometrics, data science, or computer science, or practical experience in:

- Estimating and interpreting parametric models (linear, generalized, and mixed-effects models). Familiarity with nonparametric techniques (e.g. bootstrapping)
- Applying supervised and unsupervised machine learning algorithms (decision trees, support vector machines, XG-boosting, k-means, multi-level clustering)

Technical requirements:

- Strong programming skills (R and one of Java, Python, or MATLAB) and familiarity with basic scientific computation, data analysis, and visualization.
- Obtaining and storing data inputs and products from different platforms and/or formats.
- Excellent interpersonal communication skills

Preferred skills:

- Previous experience with analysis and visualization of vector and raster datasets (e.g. Python: GeoPandas, rasterio; some R special packages, such as Shiny, sf, or gdal)
- Experience or familiarity with cloud-based computing in AWS, Microsoft Azure
- Experience or knowledge for aggregating and forecasting sparse geo-referenced datasets across time and space (**weather, soil, crop yield datasets a plus*)

If you are interested in joining a fantastic team where agriculture and data science can make a difference, please send a current resume, including references to: steven.mirsky@usda.gov. Please indicate you are applying for a Statistics and Data science (SDS) internship.