Analyzing Complex Survey Data Using Python

Mamadou S. Diallo[[1]](#footnote-1), PhD, PStat®

**Abstract of the Talk:**

Survey samples are often selected using predefined probabilistic methods from finite populations. In order to facilitate fieldwork and keep costs under control, complex sampling designs (e.g. stratification, clustering, stage sampling, etc.) are used resulting in samples with unequal probabilities of selection. Selecting, estimating, and analyzing data from such complex designs require advanced sampling techniques. I developed a Python package named *Samplics* which implements sample selection, parameter estimation and small area prediction to allow Python users to more easily work with survey data. The Python package *Samplics* is in beta version and not yet ready for production, but it is in active development. In this talk, I will show the functionalities of *Samplics*.

**About the Author**:

Mamadou S. Diallo has more than 14 years of experience in survey sampling, small area estimation and other topics in statistics. He currently works for UNICEF as the immunization data team lead at the global level. Prior to UNICEF, Mamadou worked as a senior survey statistician at Westat, Rockville MD where he supported numerous national and international surveys. He also worked for Statistics Canada where he contributed to the 2006 census under-coverage survey and to the Canadian Community Health Survey (CCHS). Mamadou holds a PhD in Statistics from Carleton University, Ottawa, Canada where he conducted research on small area estimation under the direction of Pr. J.N.K. Rao. He also holds a MSc in Statistics from Laval University, Quebec, Canada and a BSc (equivalent) in Mathematics from Claude Bernard University, Lyon, France.

Keywords: python, sampling methods, samplics

1. The United Nations Children’s Emergency Fund (UNICEF), New York, USA [↑](#footnote-ref-1)