



School of Data Science

香港城市大學
City University of Hong Kong

Distinguished
Seminar
Series

Date: 14 January 2020 (Tuesday)
Time: 3:00pm - 4:30pm
Venue: YEUNG-G7603, Yeung Kin Man Academic Building,
City University of Hong Kong

Dr Wendy L. Martinez

**President-Elect, American Statistical Association,
Director, Mathematical Statistics Research Center, Office of Survey
Methods Research, Bureau of Labor Statistics, United States of America**

I will present two applications of analyzing text-based data, both of which employ clustering or unsupervised learning to explore and organize documents. The United Nations Economic Commission for Europe (UNECE) holds an annual workshop on Statistical Data Editing with a focus on official surveys. Organizers of these workshops proposed the creation of a taxonomy of methods presented at all previous workshops. The first example given in this presentation describes the process taken to start the taxonomy, which was done by clustering all of the workshop papers. The second example illustrates the exploration of text notes recorded by interviewers conducting official surveys. Most times the information in the notes are not examined for knowledge that could be used to help monitor and improve the interview process. These interviewer notes provide an opportunity for managers to examine the comments for information about why households might be concerned about responding to the survey. Various clustering approaches were used to explore these unstructured text notes fields and are discussed in the presentation.

Biography

Dr. Wendy Martinez is the president of the American Statistical Association for the year 2020 and directs the Mathematical Statistics Research Center of the U.S. Bureau of Labor Statistics (BLS). She is also the coordinating editor of the journal *Statistics Surveys* (jointly sponsored by four major statistical societies). Her research interests include computational statistics, exploratory data analysis, data visualization, outlier detection, statistical pattern recognition, spatial statistics, and the analysis of unstructured text.

She has published several books and journal publications in the field of computational statistics. Dr. Martinez was elected a Fellow of the American Statistical Association in 2006 and was elected as a member of the International Statistical Institute in 2007. She has received many awards including the Founders Award (2017) of the American Statistical Association for outstanding leadership and support of statistical and multidisciplinary research and BLS Commissioner's Award (2018) for management excellence.

