

ISI WSC 2019 Short Course Programme

COURSE TITLE :	SC6 - Imputation methods for the treatment of item nonresponse in surveys
DURATION :	1.5 Days
DATE :	16 – 17 August 2019
VENUE :	Sasana Kijang
REGISTRATION FEES :	Developed Country MYR 1,760 (Approximately EUR 370) Developing Country / Student* MYR 1,140 (Approximately EUR 240) <i>* For student, proof of enrolment is required</i>

INSTRUCTOR 1

Prof David Haziza

Université de Montréal
Canada

David Haziza is a Full Professor in the department of mathematics and statistics at Université de Montréal. His research interests include the treatment of nonresponse in surveys, estimation in the presence of influential units and resampling methods.

COURSE DESCRIPTION

The most common way for treating item nonresponse in surveys is to construct one or more replacement values to fill in for a missing value. This process is known as imputation. We distinguish single from multiple imputation. Single imputation consists of replacing a missing value by a single replacement value, whereas multiple imputation uses two or more replacement values. During the course, we will review various imputation procedures used in National Statistical Offices as well as the properties of point and variance estimators in the presence of imputed survey data. It also provides the participants with newer developments in the field. Several examples and simulation studies will be presented throughout the course.

Chapter 1 **Introduction**

- Finite population and finite population parameters: total/mean, domain total/mean, distribution function, quantiles, census estimating equations
- Sampling designs
- Point estimation on the complete data case
- Editing
- Estimation after imputation: principles

Chapter 2: **Deterministic imputation procedures**

- Semi-parametric imputation: deterministic linear regression imputation, ratio imputation, mean imputation and historical imputation.
Non-parametric imputation: nearest-neighbour imputation, predictive mean matching, mean imputation within classes, kernel procedure, curse of dimensionality
- Composite imputation
- Properties of point estimators: total/mean, domain total/mean distribution functions, quantiles
- Choice of the imputation method: nature of the survey variable (continuous, semi-continuous, categorical), donor imputation vs. predicted value imputation, distortion of distributions

Chapter 3: **Random imputation procedures**

- Definition, random hot-deck imputation within classes
- Continuous/categorical variables
- Reducing the imputation variance: fractional imputation and balanced imputation
- Properties of point estimators: total/mean, domain total/mean distribution functions, quantiles

Chapter 4: **Variance estimation in the presence of singly imputed data**

- Frameworks: two-phase framework and reverse framework
- Variance estimation methods: method of Särndal (1992) and the method of Shao and Steel (1999), link with resampling methods (jackknife, bootstrap)

Chapter 5: **Multiple imputation**

- Definition, point and variance imputation
- Properties of multiple imputation: frequentist framework and Bayesian framework
- Proper imputation
- Self-efficiency and congeniality

TARGET AUDIENCE

Survey statisticians in National Statistical Offices, Researchers and graduate students. This course is an intermediate level course.