Ph.D. School in
Analytics for Economics and Business

COURSE: Statistics for Data Science II
Innovative Survey/Data collection Methods

Reference: Silvia Biffignandi, Maurizio Carpita

Address of the first reference:
Silvia Biffignandi
Dept. of Management, Economics and Quantitative Methods
University of Bergamo
Univ. tel: 0039 2052516
e-mail(s): silvia.biffignandi@unibg.it

Teachers: Silvia Biffignandi, Daniele Toninelli,

Course description and objectives

Nowadays, new data collection tools are the basic data source in the empirical analyses. New data collection technique encompasses relevant methodological and applicative problems. The course will enable doctoral students to get in-depth theoretical and practical knowledge of the construction and utilization of data sources based on the use of web, of new devices and of IT in general. The course aims to introduce students the frontier research issue in the field of data collection and related quality.
The course carries 4CFU and it is structured in three parts dedicated to: a) Advanced Sampling and Estimation, b) Questionnaire design and analyses; c) Text: analytics and modeling.

Outline

a) Advanced Sampling and Estimation for IT tools based data collection (web surveys, mobile surveys and other web based data sources)
   - Probability and nonprobability samples and data sources
   - Surveys in the big data context
   - New advanced sampling techniques
   - Mixing modes
   - Non response and imputation
   - Weighting and estimation: stratification, calibration, raking ratio, complex weighting techniques; propensity scores; estimation procedures -
   - Errors in different types of sources: categorization and measurement of the errors
   - Case studies in probability and nonprobability context
   - Software short review

b) Questionnaire design and analysis
   - Questionnaire design using IT; designing models; issues related to sampling
   - Questionnaire analysis: Models for latent variables
o Questionnaire analysis: Rasch analysis and item response theory: model, generalizations and applications

c) Text analytics and modeling

o Structured and unstructured data
o Categorization and classification
o Clustering
o Semantic and Sentiment analysis
o Fields of application and case studies
o Software short review

References

a) Advanced Sampling and Estimation


b) Questionnaire design and analysis

Latent models


Programming SEM with R:


Rasch analysis:


c) **Text analytics and modeling:**
