Tutorial Session on

Transportation Planning Data Analysis and Data Collection

Part A: Data analysis and predictive modeling in R

December 20, 2017 (14.00-16.00 hrs.)

Instructor: Dr. Sabya Mishra, University of Memphis, USA

Abstract: In the last decade emerging technologies (sensors, smart phones, GPS traces, location based services, social media, drones) have resulted in generating plenty of data sources opening the door for researchers to utilize the data for modeling purposes. This session will introduce students to open source software R, and demonstrate how to program in R and how to use R for effective data analysis. In the class we will use a transportation planning data set and illustrate how several techniques (hierarchical clustering, k-means clustering etc.) are applied before formal predictive modeling commences. A set of multivariate predictive modeling techniques (regression, count data, time series, and choice) using different data structures (continuous, integer, categorical) will be demonstrated using R. Components of modeling such as model development, interpretation of results, inference, elasticity, and cross-validation will be presented.

Expected Outcome: Participants will be able to gain familiarity in analyzing data and building predictive models in R. The skills obtained in the workshop will introduce the participants the power of R for future research and professional practice endeavors.

Focus areas: Transportation planning, traffic safety, and open to any discipline where model building is required.

Preparation before Workshop: No prior knowledge in R is necessary for the workshop though introductory knowledge is a plus but again not necessary. Just one requirement:

Please bring your own laptop with R-studio installed. Click HERE to download.

Questions: If there are any questions before or after the workshop, please feel free to email me: smishra3@memphis.edu

Part B: Developing a Stated Preference Survey: Going through the Three Stages

December 20, 2017 (16.00-17.00 hrs.)

Instructor: Neeraj Saxena, UNSW, Australia

Module I, Designing the SP Tasks: Brief description of key concepts and definitions; Sampling strategies like fractional factorial, D-optimal, D-efficient designs; Case study on SP design; Hands-on exercise on developing an SP experiment - Setting the attributes, levels, number of choice tasks, etc.; Data analysis techniques; Generating the set of choice tasks; Implementing the generated choice tasks in an online survey interface.

Software Requirement: Experimental design softwares: Ngene (preferred) OR SPSS

Module II, Designing the Online Survey: Brief overview of the technologies involved in online webpage design like HTML, Javascript, PHP and host server; Discussing a case study on online survey instrument; A hands-on activity where participants will design a webpage that collects and stores information from respondents.

Software Requirement: Wamp Server (a freeware)

Module III, Analysing the Obtained Dataset: Introduction to statistical modeling in Biogeme (a discrete choice modeling package); A demo on coding and executing a statistical model in Biogeme; A hands-on exercise where participants will code a simple statistical model (multinomial logit) on a dummy dataset provided to the participants; A quick overview on the fundamentals of statistical models and their estimation routine (optional)

Software Requirement: Biogeme (a freeware)

Learning outcomes: Attendees will get a first-hand experience on the stages like design, data collection and analysis of Stated Preference (SP) data. They will also become aware of the relevant software packages that can be useful in an SP study.