



The University of Chicago
Department of Statistics

MASTER'S THESIS PRESENTATION

SAEMBYEOL PARK

Department of Statistics
The University of Chicago

**Development of Statistical Models for Prediction of
Grocery Shopping Trips and Expenditures**

THURSDAY, November 10, 2011, at 4:00 PM

110 Eckhart Hall, 5734 S. University Avenue

ABSTRACT

Most of previous research on grocery shopping trips focused on the relationship between shopping behavior and shoppers demographic characteristics. The goal of this study was to build a model that predicts when supermarket shoppers will next visit a store and how much they will spend using only the dates and grocery shopping spending by households over a one-year period. This study was performed as part of a prediction competition sponsored by Dunnhumby LLC. We develop models for the shopping trip and purchase separately. A set of variables that represent each customers shopping behavior based on the historic record are generated. Logistic regression and k Nearest Neighbor cross-validation methods are used to predict the shopping return day. Forward selection and AIC are used as model selection criteria. Given the estimated return day, kernel density estimation methods are applied to estimate the spending by the customers. This study suggests models and methods to better predict individual shopping trip patterns with limited information, and it will be applicable for marketing policy improvement of the supermarket industry.

For information about building access for persons with disabilities, please contact Matt Johnston at 773.702-0541 or send an email to mhj@galton.uchicago.edu. If you wish to subscribe to our email list, please visit the following web site: <https://lists.uchicago.edu/web/arc/statseminars>.