ANALYSIS OF THE QUESTIONNAIRE WITH DENDROGRAMS BASED ON CART MODELS

Aleš Dufek[1,2] and Miroslav Homola[1]

[1/]**Agriresearch Rapotin Ltd.**, Vyzkumniku 267, 788 13 Vikyrovice, Czech Republic

[2/]**Research Institute for Cattle Breeding, Ltd.**, Rapotin, Vyzkumniku 267, 788 13 Vikyrovice, Czech Republic

Corresponding author e-mail: **ales.dufek@vuchs.cz**

#47742

RMC 2011 RECIPROCAL MEAT CONFERENCE

June 19-22
Manhattan, Kansas
"Developing Our Future"

INTRODUCTION

Traditional Czech cuisine involves pork, chicken, fish and beef. Whereas the meal from the first three species is prepared by many ways including grilling, the majority of the Czech population does not use grilling for beef. Beef is mainly used to prepare goulash or is stewed, served with soup and rice or dumplings. The purpose of our study was to evaluate relationship between consumer characteristics (gender, age, educational attainment, frequency of purchasing of meat, the most frequently bought and the most favorite species) and preference for cooking method of beef, degree of doneness of steaks and origin of beef. The second aim of our work was to draw attention of scientists to the CART – Classification and Regression Tree models as a method of visualization of results from questionnaire, because, generally, the power of CART is in its application for demography and studies with many effects.

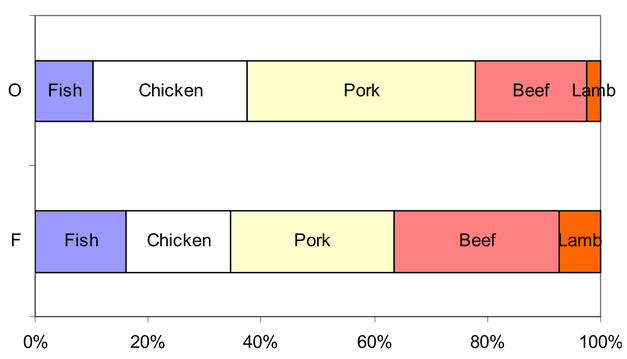
MATERIAL & METHODS

Data gained from survey forms were evaluated with CART - Classification and Regression Tree models fitted in the package rpart, version 3.1-41 in the R software, version 2.7.2. The package allows quick view and further selection of significant factors and evaluation of their ratio of explained variability (length of branches) with effective visualization of their gradual relationship by dendrogram. The total number of respondents was 132 (people qualified in agriculture) with the average age of 38.6, ranged from 11 to 73. The point is that any answer of any question from survey form can be used as dependent or independent variable in the CART model. The R language makes possible to switch it very quickly. Commands are:

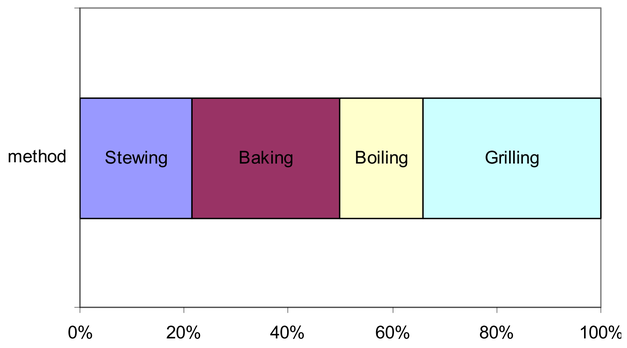
```
dendrogram1<-rpart(beefpref~.,data=RMC)
dendrogram2<-rpart(doneness~.,data=RMC)
```

The dot stands for the whole data frame (all variables are taken as effects except the selected dependent variable).

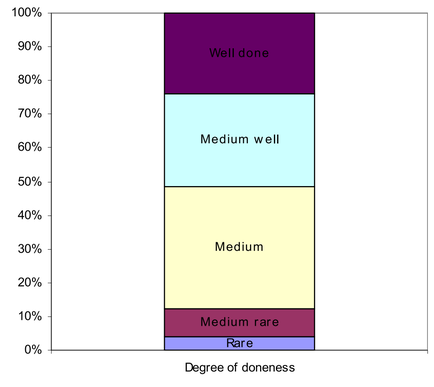
SIMPLE RESULTS



↑ Results based on the questions "What species do you consume most often?" (O) and "What species is your favorite?" (F). Difference between preferences for the most often consumed and the most favorite species was significant (Chi-sq = 18.49, df = 4, p < 0.00098).

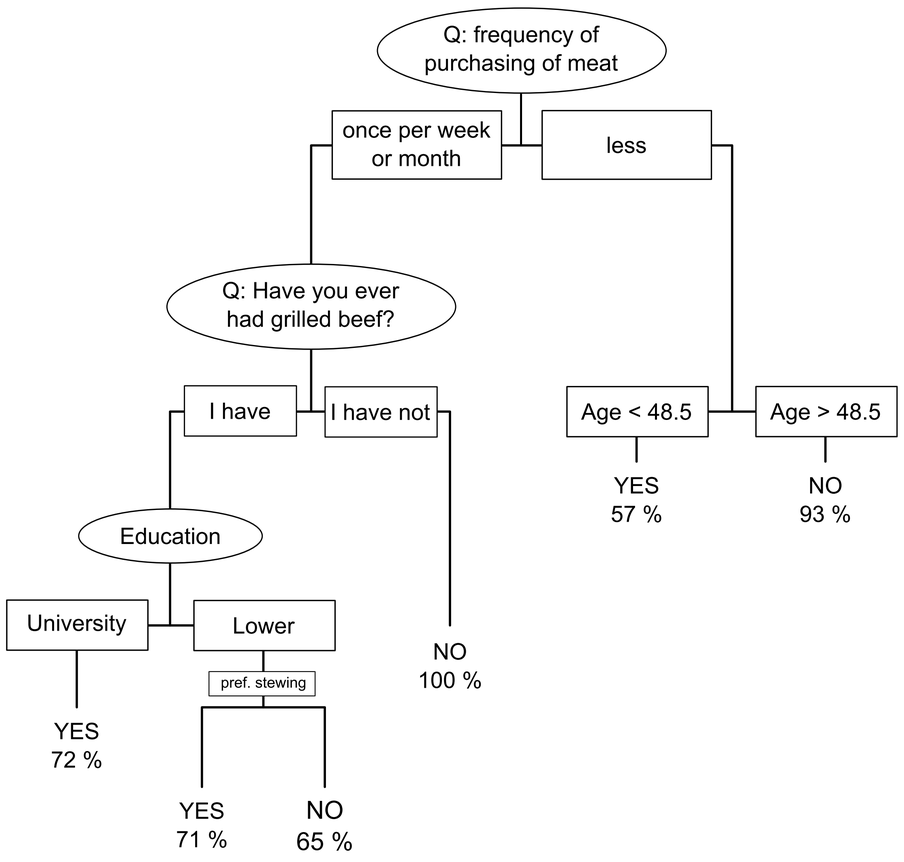


↑ Ratios of methods of cooking beef preferred by respondents.

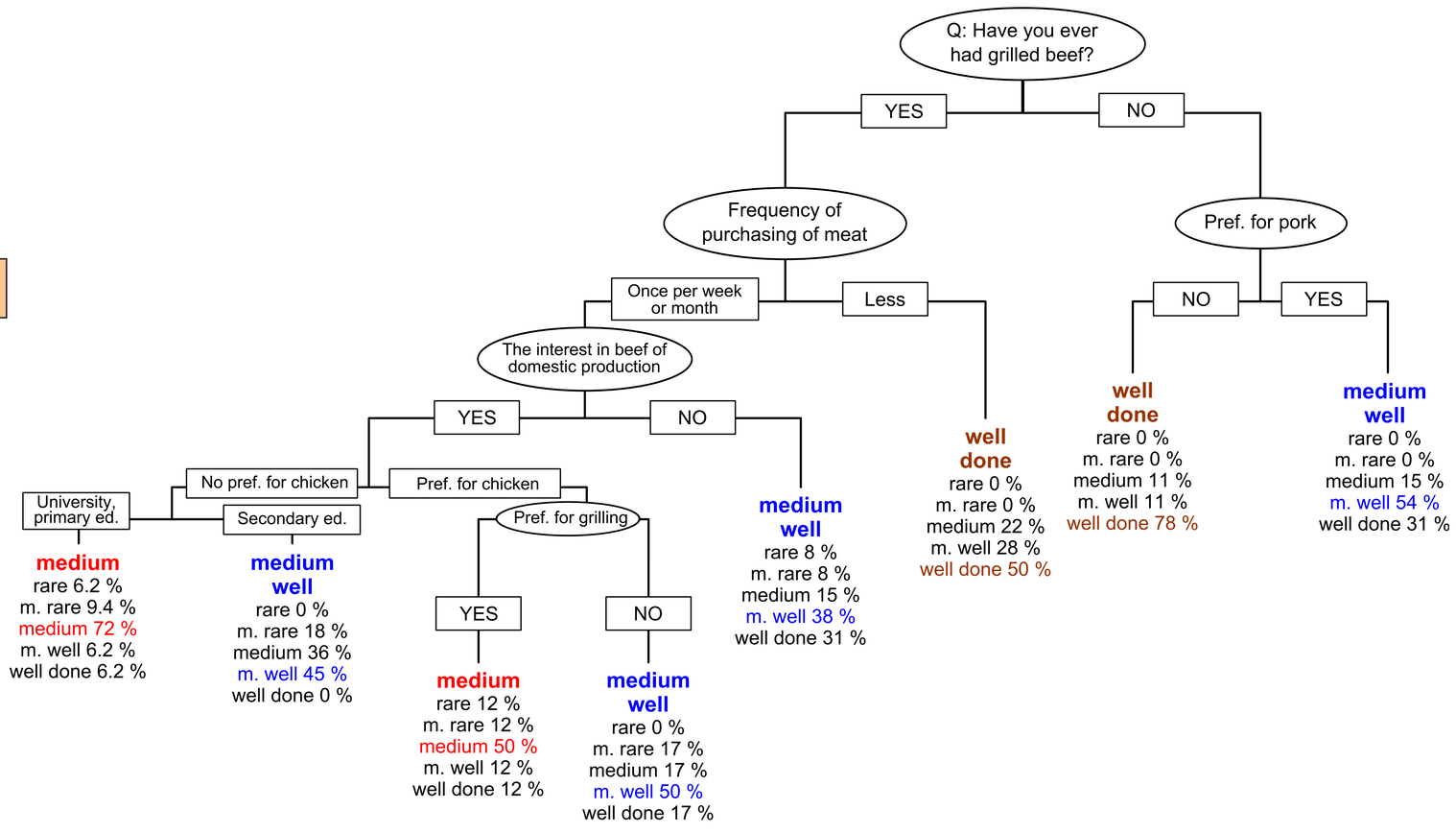


↑ Ratios of preferences for degrees of doneness.

RESULTS BASED ON THE CART DENDROGRAMS



← Dendrogram with significant characteristics of consumers who preferred beef (answer YES). Beef is preferred by respondents who purchase meat once per week or month AND had grilled beef yet AND attained to university education (LEFT BRANCH of the diagram) OR purchase meat less than once per month AND were younger than 48.5 years (RIGHT BRANCH of the diagram). Percent below the YES/NO answer is prediction of the answer. Length of branches correlate with portion of explained variability by the predictor marked above the dichotomy. The root is the top.



↑ Dendrogram with significant characteristics of consumers and their preferences for degrees of doneness. Degree medium was preferred by consumers who had had a grilled steak AND usually buy beef 1 times a week or month AND are interested in origin of beef AND do not prefer chicken (this cascade of significant criterions appeared on the very left branch of dendrogram) OR prefer chicken AND prefer grilling, whereas the respondents who do not grill beef prefer medium well. On the contrary, the very right branch of the dendrogram shows that consumers who had had not a steak marked well done. Percents by the degree of doneness are predictions of answers. Length of branches correlate with portion of explained variability by the predictor marked above the dichotomy. The root is the top.

CONCLUSION

CART method seems very effective for the first analysis of the data gained from questionnaires and selection of important effects from many variables generated from survey forms. Our results indicate differences in consumer habits in the Czech Republic and describe their variability. Consumers who have not had grilled beef yet never marked the degree rare or medium rare and potentially preferred medium well or well done.

CITATION

Therneau, T. M. & Atkinson, B. R. port by Ripley, B. (2008). rpart: Recursive Partitioning. R package version 3.1-41. URL: http://mayoresearch.mayo.edu/mayo/research/biostat/splusfunctions.cfm

Development Core Team (2008). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL: http://www.R-project.org.

ACKNOWLEDGMENTS

This study was funded by the Ministry of Education, Youth and Sports of the Czech Republic (Projects LA08035, 2B08037 and MSM2678846201) .