Homework 7  
(from David Durian's study)

Data from 120 Caucasian, native-English speaking Columbus, OH store clerks (40 store clerks working in a variety of stores at each of 3 Columbus malls [Easton, Polaris, and City Center]) were elicited via the Rapid Anonymous Survey technique. The entire set of store clerks were also stratified by the following social factors: age (15–30, 35–50, 55–70, 40 speakers in each age group); social class (working class – WC, lower middle class – LMC, and upper middle class – UMC, 40 speakers of each class); and gender (60 males/60 females).

From each of the 120 speakers, two tokens of words containing word-initial “str” clusters were obtained. The first, in a less emphatic speech environment; the second in a more emphatic one. This leads to a total of 240 tokens (120 more emphatic; 120 less emphatic). The two variant realizations of “str” were rated impressionistically by the researcher as closer to [str] or [ʃtʃ].

All of the data were elicited by the researcher asking for directions to a known street in the area of the mall that was misidentified as a “road.” The speaker then corrected the researcher by saying “you mean X street,” and the first “str” token was obtained. Following this first utterance, the researcher said, “excuse me, what did you say” and the speaker would more emphatically repeat the phrase “x street” producing the second “str” token. In the case of Polaris Mall no widely identified street is located within proximity to the mall (everything there is a Road, Parkway, Place, etc.), and so the researcher asked for directions to a part of the store that would yield the word “straight” in place of “street.”

Data were written down on a sheet of paper just after leaving eyesight of the informant. No audio recordings were made. The only judgment made on the sounds was the researcher’s own impressionistic rating.

Variables

Dependent variable:
  Pronunciation: sh=[ʃ], [s]

Independent variables:
1 Emphasis (repetition): less emphatic, more emphatic (second repetition)
2 Gender: man, woman
3 Age: young, middle-aged, old
4 Mall: Polaris, Easton, City Center
5 Class: working class, lower middle class, upper middle class
Note that the malls are not stratified according to social class as in Labov. The researcher chose class-related stores in each mall to determine class variable.

The data are here.

Do a logistic regression using all the variables. The pronunciation is the dependent variable. Choose the 's' pronunciation as the reference value. Ask for the pairwise contrasts between the values of each variable. Set the sorting order for categorical predictors to descending.

1 What percent correct does the model predict?
2 What variables are insignificant?
3 What two malls can you not compare in the table? What two classes?
4 Look at the estimated means tables and interpret the influence of age, class, and mall.