Practice 8

1. The following data on marital status by gender and report of premarital sex (PMS) and extramarital sex (EMS) were obtained by the UK Marriage Research Centre.

			Marital Status	
Gender	PMS	EMS	Divorced	Still Married
Male	Yes	Yes	28	11
		No	60	42
	No	Yes	17	4
		No	68	130
Females	Yes	Yes	17	4
		No	54	25
	No	Yes	36	4
		No	214	322

The data are stored in file `Mstatus.csv`.

   (a) Fit a saturated log-linear model to the data and see what terms are significant.

   (b) Use the `step` procedure to find the "best" model. Use `scopems=list(lower=freq ~ gend*PMS*EMS, upper=freq ~ gend*PMS*EMS*MS)` to define the range of the models to be examined. Note that we may treat `gend`, `PMS` and `EMS` as the "explanatory" factors and `MS` as the response factor here. So we want to always keep the 3-factor interaction `gend:PMS:EMS` term in the model.

   (c) Based on the "best" model found in (b), test whether the effect of `gend:MS` is significant.

   (d) Regard `MS` as the response factor, fit logistic regression models to the data, find the "best" logistic model and explain it.

2. Carry out an analysis of the following 1959 data (`mobilityDenmark.csv`) on the occupational mobility of males in Denmark, and state your conclusions. (Hint: a similar analysis for the UK mobility data can be found in the lecture notes.)

Status Category of	Status Category of Sons's Occupation					
Father's Occupation	1	2	3	4	5	Total
1	18	17	16	4	2	57
2	24	105	109	59	21	318
3	23	84	289	217	95	708
4	8	49	175	348	198	778
5	6	8	69	201	246	530
Total	79	263	658	829	562	2391