

```
In [1]: # DSC530-T302
        # Stephen Smitshoek
        # Week08
        # Exercise 11-4
```

```
In [2]: import thinkstats2
import pandas
import numpy as np
import statsmodels.formula.api as smf
import first
import nsfg
```

```
In [3]: live, firsts, others = first.MakeFrames()
live = live[live.prglnth>30]

resp = nsfg.ReadFemResp()
resp.index = resp.caseid

join_live_resp = live.join(resp, on='caseid', rsuffix='_r')
```

```
In [4]: model = smf.mnlogit('rmarital ~ age_r + C(race) + educat + totincr', data=join_live_re
results = model.fit()
results.summary()
```

```
Optimization terminated successfully.
      Current function value: 1.087603
      Iterations 8
```

Out[4]:

MNLogit Regression Results

Dep. Variable:			rmarital	No. Observations:		8884
Model:			MNLogit	Df Residuals:		8854
Method:			MLE	Df Model:		25
Date:			Wed, 27 Jul 2022	Pseudo R-squ.:		0.1655
Time:			07:45:12	Log-Likelihood:		-9662.3
converged:			True	LL-Null:		-11579.
Covariance Type:			nonrobust	LLR p-value:		0.000
rmarital=2	coef	std err	z	P> z	[0.025	0.975]
Intercept	4.4532	0.279	15.977	0.000	3.907	5.000
C(race)[T.2]	-0.9219	0.089	-10.409	0.000	-1.095	-0.748
C(race)[T.3]	-0.6334	0.136	-4.674	0.000	-0.899	-0.368
age_r	-0.0570	0.006	-9.754	0.000	-0.068	-0.046
educat	-0.2051	0.019	-11.017	0.000	-0.242	-0.169
totincr	-0.1302	0.012	-11.298	0.000	-0.153	-0.108
rmarital=3	coef	std err	z	P> z	[0.025	0.975]
Intercept	-4.5432	0.916	-4.960	0.000	-6.338	-2.748
C(race)[T.2]	-0.4405	0.236	-1.865	0.062	-0.904	0.023
C(race)[T.3]	0.0329	0.335	0.098	0.922	-0.623	0.689
age_r	0.1301	0.019	6.833	0.000	0.093	0.167
educat	-0.1010	0.048	-2.100	0.036	-0.195	-0.007
totincr	-0.3251	0.032	-10.182	0.000	-0.388	-0.263
rmarital=4	coef	std err	z	P> z	[0.025	0.975]
Intercept	-2.6485	0.302	-8.783	0.000	-3.239	-2.057
C(race)[T.2]	-0.3188	0.093	-3.417	0.001	-0.502	-0.136
C(race)[T.3]	-0.7664	0.171	-4.485	0.000	-1.101	-0.431
age_r	0.0641	0.007	9.817	0.000	0.051	0.077
educat	0.0673	0.017	4.036	0.000	0.035	0.100
totincr	-0.2270	0.012	-19.588	0.000	-0.250	-0.204
rmarital=5	coef	std err	z	P> z	[0.025	0.975]
Intercept	0.8231	0.345	2.385	0.017	0.147	1.500
C(race)[T.2]	-1.0491	0.104	-10.125	0.000	-1.252	-0.846
C(race)[T.3]	-0.5644	0.156	-3.624	0.000	-0.870	-0.259
age_r	0.0063	0.007	0.872	0.383	-0.008	0.020
educat	-0.0136	0.021	-0.639	0.523	-0.056	0.028

totincr	-0.2923	0.015	-20.149	0.000	-0.321	-0.264
rmarital=6	coef	std err	z	P> z 	[0.025	0.975]
Intercept	6.2370	0.275	22.654	0.000	5.697	6.777
C(race)[T.2]	-2.1831	0.080	-27.200	0.000	-2.340	-2.026
C(race)[T.3]	-1.9645	0.138	-14.246	0.000	-2.235	-1.694
age_r	-0.0925	0.006	-16.121	0.000	-0.104	-0.081
educat	-0.0777	0.018	-4.386	0.000	-0.112	-0.043
totincr	-0.2944	0.012	-25.343	0.000	-0.317	-0.272

```
In [5]: columns = ['age_r', 'race', 'educat', 'totincr']
new = pandas.DataFrame([[25, 2, 12, 11]], columns=columns)
y = results.predict(new)
y
```

```
Out[5]:
```

	0	1	2	3	4	5
0	0.748384	0.125474	0.001103	0.035295	0.023813	0.065931

```
In [6]: print('With the above results the model predicts there is a 75% chance the woman is ma
'\na 13% chance that she is not married but living with an opposite sex partner')
```

With the above results the model predicts there is a 75% chance the woman is married
and
a 13% chance that she is not married but living with an opposite sex partner