In [1]:
```
# DSC530-T302
# Stephen Smitshoek
# Week08
# Exercise 11-3
```

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.prglngth>30]

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

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

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

```
Optimization terminated successfully.
         Current function value: 1.687055
         Iterations 5
```

Out[4]:

Poisson Regression Results

Dep. Variable:	numbabes	**No. Observations:**	8884
Model:	Poisson	**Df Residuals:**	8878
Method:	MLE	**Df Model:**	5
Date:	Wed, 27 Jul 2022	**Pseudo R-squ.:**	0.03109
Time:	07:15:51	**Log-Likelihood:**	-14988.
converged:	True	**LL-Null:**	-15469.
Covariance Type:	nonrobust	**LLR p-value:**	1.106e-205

	coef	std err	z	P>|z|	[0.025	0.975]
Intercept	1.0842	0.045	23.995	0.000	0.996	1.173
C(race)[T.2]	-0.1398	0.015	-9.464	0.000	-0.169	-0.111
C(race)[T.3]	-0.0914	0.025	-3.717	0.000	-0.140	-0.043
age_r	0.0208	0.001	20.474	0.000	0.019	0.023
educat	-0.0443	0.003	-15.139	0.000	-0.050	-0.039
totincr	-0.0179	0.002	-9.442	0.000	-0.022	-0.014

In [7]:
```
columns = ['age_r', 'race', 'educat', 'totincr']
new = pandas.DataFrame([[35, 1, 16, 14]], columns=columns)
```

```
y = results.predict(new)
print(f'The model predicts the woman has {float(y):.0f} children')
```

```
The model predicts the woman has 2 children
```