Dummy variables

STAT 615 (Baron)

```
DATA C;
INFILE pulse;
INPUT row pulse1 pulse2 ran smokes gender height weight activity;
if activity=0 then delete;
if activity=1 then z1=1; else z1=0;
if activity=2 then z2=1; else z2=0;
```

/* Defining dummy variables manually */

(2 dummy variables for 3 levels of activity)

same result

```
PROC REG;
MODEL pulse1=height weight z1 z2;

PROC GLM;
CLASS activity;
MODEL pulse1 = height weight activity;
```

← Treating activity as a categorical variable

/* Automatic definition of dummy variables */

```
PROC GLM;
CLASS smokes gender activity;
MODEL pulse1 = weight activity weight*smokes smokes|gender|height;

run;
```

/* A model with interactions */

interaction (different slopes for smokers and non-smokers)

including all interactions

The REG Procedure

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	4	996.36360	249.09090	2.28	0.0675
Error	86	9406.77926	109.38115		
Corrected Total	90	10403			

Parameter Estimates

Variable	DF	Parameter Estimate	Standard Error	t Value	Pr > |t|
Intercept	1	109.52800	26.36944	4.15	<.0001
height	1	-0.45978	0.48930	-0.94	0.3500
weight	1	-0.04053	0.07495	-0.54	0.5901
z1	1	7.84646	4.18090	1.88	0.0639
z2	1	0.47400	2.67773	0.18	0.8599

The GLM Procedure