

### 4.3 Relations in Categorical Data

When bivariate data is quantitative, we display it with a scatterplot.  
 When two or more variables are categorical (qualitative), we can display them with a **two way table** (like a spreadsheet).

#### 2002 General Social Survey

	Happiness			
Income	Not too happy	Pretty Happy	Very Happy	Total
Above Average	21	159	110	290
Average	53	372	221	646
Below Average	94	249	83	426
Total				1362

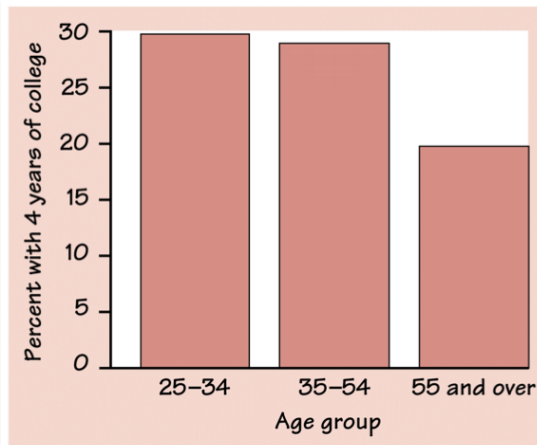
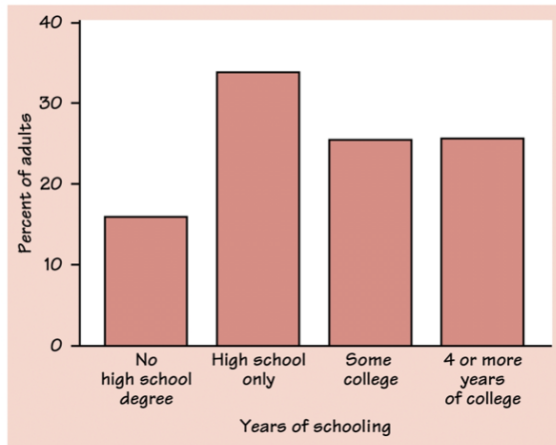
column variable ↓  
 row variable →  
 if row and column totals are missing, calculate them FIRST  
 marginal distribution of income  
 numbers in the table represent counts in each category  
 marginal distribution of happiness

Power of percent comparison:

	Happiness			
Income	Not too happy	Pretty Happy	Very Happy	Total
Above Average	7%	55%	38%	290 (100%)
Average	8%	58%	34%	646 (100%)
Below Average	22%	59%	19%	426 (100%)

**TABLE 4.6** Years of school completed, by age, 2000 (thousands of persons)

Education	Age group			Total
	25 to 34	35 to 54	55+	
Did not complete high school	4,474	9,155	14,224	27,853
Completed high school	11,546	26,481	20,060	58,087
1 to 3 years of college	10,700	22,618	11,127	44,445
4 or more years of college	11,066	23,183	10,596	44,845
Total	37,786	81,435	56,008	175,230



EDU	AGE			Total
Frequency	25-34	35-54	55 over	
NoHS	4474	9155	14224	27853
	11.84	11.24	25.40	
HSonly	11546	26481	20060	58087
	30.56	32.52	35.82	
SomeColl	10700	22618	11127	44445
	28.32	27.77	19.87	
Coll4yrs	11066	23183	10596	44845
	29.29	28.47	18.92	
Total	37786	81435	56008	175230

## Conditional Distribution

