

Relations in Categorical Data

- After a large fund drive to help the Boston City Library, the following information was obtained from a random sample of contributors to the library fund.

Number of People Making Contribution

Ethnic Group	\$1 – 50	\$51 – 100	\$101 – 150	\$151 – 200	Over \$200	Row Total
A	310	715	201	105	42	1373
B	619	511	312	97	22	1561
C	402	624	217	88	35	1366
D	544	571	309	79	29	1532
Column Total	1875	2421	1039	369	128	5832

- a. Make a percent bar graph of the ethnic group marginal distribution.
 - b. Make a percent bar graph for the conditional distribution for ethnic group C.
 - c. Why do we use bar graphs instead of histograms?
 - d. What percent of the contributors are from ethnic group A?
 - e. What percent of the contributors donated over \$200?
 - f. What percent of the contributors in ethnic group D donated \$100 or less?
 - g. What percent of the contributors who donated \$50 or less are from ethnic group B
2. The following table shows the Myers-Briggs personality preference and professions for a random sample of 2408 people in the listed professions. T refers to thinking and F refers to feeling.

Personality Preference Type

Occupation	T	F	Row Total
Clergy (all types)	114	420	
M.D.	785	818	
Lawyer	176	95	
Column Total			

- a. Complete the table.

b. Give the marginal distribution of personality types. Draw a bar graph to display this distribution.

c. What percent of lawyers are feeling?

d. What percent of professionals classified as feeling are lawyers?

3. Suppose that 2400 women are categorized according to whether or not they use oral contraceptives and whether or not they have high blood pressure. The results are shown below.

Age 18-34			Age 35-49		
	<i>High B.P.</i>	<i>Normal</i>		<i>High B.P.</i>	<i>Normal</i>
<i>Use</i>	36	564	<i>Use</i>	28	172
<i>Don't</i>	16	384	<i>Don't</i>	120	1080

a. Compute by age group the percents of women who have high blood pressure for those who use and don't use oral contraceptives. What do you conclude?

b. Make a two way table of use of contraceptives by blood pressure status for both age groups.

c. Compute the percents of women who have high blood pressure for those who use and don't use oral contraceptives. What do you conclude?

d. Explain Simpson's paradox.