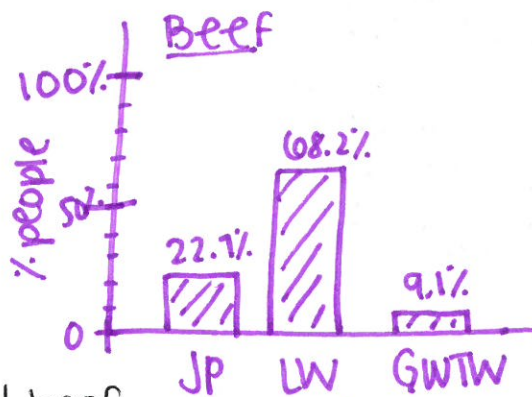


1. A researcher suspected a relationship between people's preferences in movies and preference in pizza. A random sample of 100 people produced the following two-way table:

| Favorite Movie | Pepperoni | Beef | Mushrooms | Total | % |
|--------------------|-----------|-----------|-----------|------------|------------|
| Jurassic Park | 20 | 5 | 10 | 35 | 35 |
| Lethal Weapon (I) | 8 | 15 | 12 | 35 | 35 |
| Gone with the Wind | 15 | 2 | 13 | 30 | 30 |
| Total | 43 | 22 | 35 | 100 | 100 |
| % | 43 | 22 | 35 | 100 | 100 |

- a) Enter the overall (marginal) distributions on the table.
- b) Compute (in percentages) the conditional distribution of favorite movie among those who prefer ground beef topping. Make a bar graph of the distribution below.

| | Beef |
|--------------|-----------------------------------|
| JP | $5/22 = 22.7\%$ |
| LW | $15/22 = 68.2\%$ |
| GWTW | $2/22 = 9.1\%$ |
| total | $22/22 = 100\%$ |

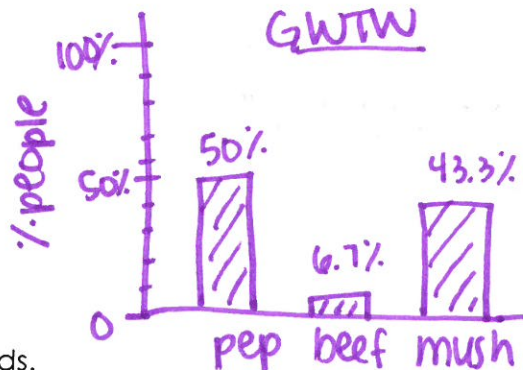


c) Briefly describe your finding in words.

Of people who prefer ground beef topping on pizza, the majority (68.2%) prefer the movie Lethal Weapon, while 22.7% prefer Jurassic Park, and 9.1% prefer gone with the wind.

d) Compute (in percentages) the conditional distribution of favorite pizza topping among those who chose Gone with the Wind as their favorite movie. Make a bar graph of the distribution below.

| | GWTW |
|--------------|-----------------------------------|
| pep | $15/30 = 50\%$ |
| beef | $2/30 = 6.7\%$ |
| mush | $13/30 = 43.3\%$ |
| total | $30/30 = 100\%$ |



e) Briefly describe your finding in words.

Of people who prefer Gone with the wind, 50% of people prefer pepperoni as a pizza topping, 43.3% prefer mushrooms, and only 6.7% prefer beef.

2. The 2003 Census Bureau collected data describing the age and gender of college students (thousands of persons):

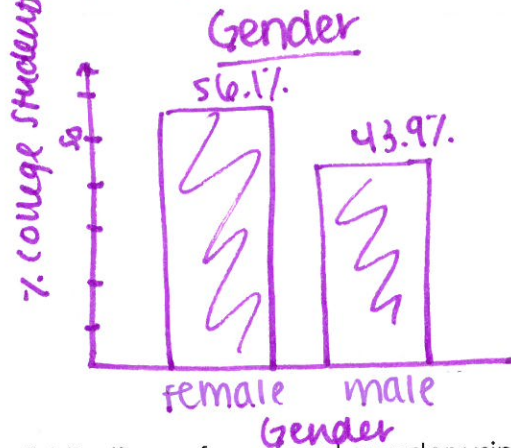
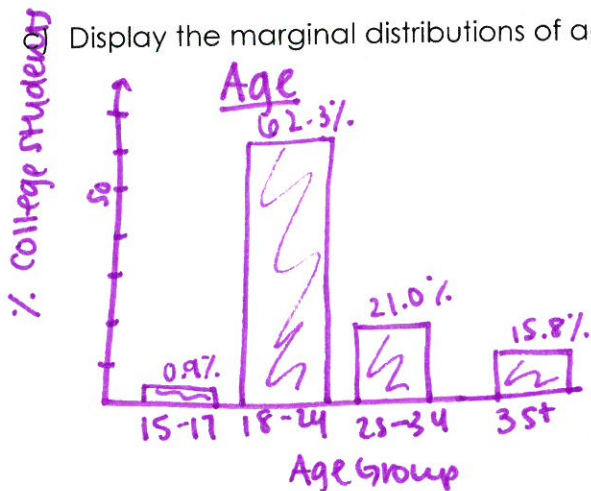
| Age Group | Gender | | Total | |
|-------------------|--------------|--------------|--------------|-------------|
| | Female | Male | | |
| 15 to 17 years | 89 | 61 | 150 | 0.9% |
| 18 to 24 years | 5,688 | 4,697 | 10385 | 62.3% |
| 25 to 34 years | 1,904 | 1,589 | 3493 | 21.0% |
| 35 years or older | 1,660 | 970 | 2630 | 15.8% |
| Total | 9341 | 7317 | 16658 | 100% |
| | 56.1% | 43.9% | 100% | |

- a) Enter the overall (marginal) distributions as totals in the table.
- b) Construct the marginal distribution of age group (in percents) and display in a table below. Then construct the marginal distribution of gender (in percents) and display in a separate table below.

| Age Group | |
|--------------|------------------------|
| 15-17 | $150/16658 = 0.9\%$ |
| 18-24 | $10385/16658 = 62.3\%$ |
| 25-34 | $3493/16658 = 21.0\%$ |
| 35+ | $2630/16658 = 15.8\%$ |
| Total | $16658/16658 = 100\%$ |

| Gender | |
|--------------|-----------------------|
| female | $9341/16658 = 56.1\%$ |
| male | $7317/16658 = 43.9\%$ |
| Total | $16658/16658 = 100\%$ |

Display the marginal distributions of age and gender as separate bar graphs below.



- d) Briefly describe, compare, and contrast the distributions of age and gender using the marginal distributions.

- e) Construct the conditional distributions of gender for students in the 18-24 age group and display in a table below. Then construct the conditional distribution of gender for students in the 35 or older age group and display in a table below.

| | 18-24 |
|--------|-----------------------|
| female | $5688/10385 = 54.8\%$ |
| male | $4697/10385 = 45.2\%$ |
| total | $10385/10385 = 100\%$ |

| | 35+ |
|--------|----------------------|
| female | $1660/2630 = 63.1\%$ |
| male | $970/2630 = 36.9\%$ |
| total | $2630/2630 = 100\%$ |

- f) Compare and contrast these conditional distributions. Do they suggest anything about the relationship between age and gender in college students?

a higher percentage of older students (35+) are female.

- g) Construct the conditional distribution of age among female students and display in a table below. Then construct the conditional distribution of age among male students and display in a table below.

| | female |
|-------|----------------------|
| 15-17 | $89/9341 = 0.95\%$ |
| 18-24 | $5688/9341 = 60.9\%$ |
| 25-34 | $1904/9341 = 20.4\%$ |
| 35+ | $1660/9341 = 17.8\%$ |
| total | $9341/9341 = 100\%$ |

| | male |
|-------|----------------------|
| 15-17 | $61/7317 = 0.8\%$ |
| 18-24 | $4697/7317 = 64.2\%$ |
| 25-34 | $1589/7317 = 21.7\%$ |
| 35+ | $970/7317 = 13.3\%$ |
| total | $7317/7317 = 100\%$ |

- h) Compare and construct these conditional distributions. Do they suggest anything about the relationship between age and gender in college students? Are they the same or different from the results you saw in part (f)?

the age distributions seem even among gender.

