# DS 2006 HW 7 Solutions

2024-03-04

**1. Create a table of counts and proportions for the cola variable. Report counts with proportions in parentheses, for example 23 (0.39). (This is what is meant by N (p).)**

```
survey_data <- read.csv('https://tgstewart.cloud/soda.csv')

table(survey_data)
```

```
##                             sugar
## cola                         Regular (full sugar) Zero sugar or diet
##   Cola (Coke, Pepsi, etc.)                    10                 10
##   Something else                              30                 16
```

| Colas | No Colas |
|---|---|
| 20 ($\frac{20}{66} = 0.\bar{30}$) | 46 ($\frac{46}{66} = 0.6\bar{9}$ ) |

**2. Create a similar table for the sugar preference variable.**

| Regular | Zero Sugar |
|---|---|
| 40 ($\frac{40}{66} = 0.6\bar{0}$) | 26 ($\frac{26}{66} = 0.\bar{39}$) |

**3. Use the data to generate a contingency table of counts and proportions.**

| | Colas | Not Colas |
|---|---|---|
| Regular | 10 ($\frac{10}{66} = 0.1\bar{5}$) | 30 ($\frac{30}{66} = 0.4\bar{5}$) |
| No Sugar | 10 ($\frac{10}{66} = 0.1\bar{5}$) | 16 ($\frac{16}{66} = 0.\bar{24}$) |

**4. Add to the table conditional column and row proportions and counts. With sugar preferences on the rows, the conditional row proportions are P(Drink | Sugar). Likewise, with drink preferences on the columns, conditional column proportions are P(Sugar | Drink).**

| | Colas | Not Colas |
|---|---|---|
| Regular | 10 ($\frac{10}{66} = 0.1\bar{5}$) | 30 ($\frac{30}{66} = 0.4\bar{5}$) |
| Row | 10 ($\frac{10}{40} = 0.25$) | 30 ($\frac{30}{40} = .75$) |
| Col | 10 ($\frac{10}{20} = 0.5$) | 30 ($\frac{30}{46} = 0.6521$) |

| | Colas | Not Colas |
|---|---|---|
| No Sugar/Diet | 10 ($\frac{10}{66} = 0.1\bar{5}$) | 16 ($\frac{16}{66} = 0.2\bar{4}$) |
| Row | 10 ($\frac{10}{26} = 0.3846$) | 16 ($\frac{16}{26} = 0.6154$) |
| Col | 10 ($\frac{10}{20} = 0.5$) | 16 ($\frac{16}{46} = 0.3478$) |

**5. Add marginal proportions and counts.**

| | Colas | Not Colas | Total |
|---|---|---|---|
| Regular | 10 ($\frac{10}{66} = 0.1\bar{5}$) | 30 ($\frac{30}{66} = 0.4\bar{5}$) | 40 ($\frac{40}{66} = 0.6\bar{0}$) |
| Row | 10 ($\frac{10}{40} = 0.25$) | 30 ($\frac{30}{40} = .75$) | |
| Col | 10 ($\frac{10}{20} = 0.5$) | 30 ($\frac{30}{46} = 0.6521$) | |
| No Sugar/Diet | 10 ($\frac{10}{66} = 0.1\bar{5}$) | 16 ($\frac{16}{66} = 0.2\bar{4}$) | 26 ($\frac{26}{66} = 0.3\bar{9}$) |
| Row | 10 ($\frac{10}{26} = 0.3846$) | 16 ($\frac{16}{26} = 0.6154$) | |
| Col | 10 ($\frac{10}{20} = 0.5$) | 16 ($\frac{16}{46} = 0.3478$) | |
| Total | 20 ($\frac{20}{66} = 0.3\bar{0}$) | 46 ($\frac{46}{66} = 0.6\bar{9}$) | 66 (1) |