1. (a) The cases are *instances of someone coming to the concession stand*. That's a little different than two similar-sounding answers: *sales for the night*, which wouldn't account for the 8 cases where no food no drink were purchased, and *people who visited the stand*, which, in the spirit of "one row per case", would not capture the fact that perhaps there were repeat visitors.
   (b) There are a total of 397 cases.
   (c) One variable might be called "food purchased", a categorical variable with values "Pizza", "Hot Dog", "No Food"; the other is "drink purchased", again categorical with 3 possible values.
   (d) A drink was purchased in the proportion $(54+97)/158 \doteq 0.956$.
   (e) The proportion/relative frequences are $43/96 \doteq 0.448$ without food and $59/143 \doteq 0.413$ for hot dogs. So, yes, water *is* sold at a higher relative frequency when in the absence of food.
   (f) (iii)
2. (iii)
3.
```
names(houses)                                # answer to part (a)
filter(houses, sqFt > 2000)                  # answer to part (b)
gf_boxplot(~ sqFt, data=houses)              # answer to part (c)
cor(gasMileage ~ weight, data=cars)          # answer to part (d)
mean(~ weight, data=cars)                    # answer to part (e)
tally(~ numSiblings, data=survey)            # answer to part (f)
nrow(cars)                                   # answer to part (g)
lm(gasMileage ~ weight, data=cars)           # answer to part (h)
```
4. (i) and (iv)
5. (i)
6. (iii)
7. D, A, C, B, in that order
8. (a) A, B, D, C, in that order
   (b) (iii) and (iv)
   (c) (i), (ii) and (iii)
   (d) Because the distribution is right-skewed, the **mean** is larger than the median.
   (e) A
   (f) C
9. (a) You could write the mpg for each of the 110 cars on individual slips of paper, place them in a bag and mix them. You would then draw a slip, record the mpg, place it back in the bag, mix, and draw again, repeating this process until you had written 110 mpg values corresponding to the slips you drew. These mpg-values form a bootstrap sample, and from them you can calculate an average, producing a single bootstrap statistic, one dot correctly located on the dotplot we call the bootstrap distribution. The picture indicates there are 4000 dots, so we would have 3999 to go, where each is obtained similarly as the average of 110 draws with replacement from the bag.

(b) The original sample provides the point estimate $\overline{x} = 20.782$. We estimate $\mathrm{SE}_{\overline{x}}$ from the reported standard error of the bootstrap distribution, 0.425. For 95% confidence, our margin of error is twice $\mathrm{SE}_{\overline{x}}$, or $2(0.425) = 0.85$. So, lower/upper bounds of our 95% CI are

$$\text{lower} = 20.782 - 0.85 = 19.932, \qquad \text{upper} = 20.782 + 0.85 = 21.632.$$

Using interval notation, our confidence interval is $(19.932, 21.632)$.

(c) μ is the average city-mpg for all 2015 cars, a population parameter.

(d) (ii)

10. The mean:

$$\overline{x} \;=\; \frac{1}{4}(14 + 36 + 19 + 31) \;=\; 25.$$

The deviations $(x_j - \overline{x})$ from the mean:

$$14 : 14 - 25 = -11, \quad 36 : 36 - 25 = 11, \quad 19 : 19 - 25 = -6, \quad 31 : 31 - 25 = 6.$$

Sum of squared deviations from the mean:

$$\sum_{j=1}^{4}(x_j - \overline{x})^2 \;=\; (-11)^2 + 11^2 + (-6)^2 + 6^2 \;=\; 121 + 121 + 36 + 36 \;=\; 314.$$

Standard deviation:

$$\sqrt{\frac{1}{4-1} \cdot \sum_{j=1}^{4}(x_j - \overline{x})^2} \;=\; \sqrt{\frac{1}{3} \cdot 314} \;=\; 10.23.$$