

```

1. 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)
   resample(c(1, 0), prob = c(71, 135)/206, size = 80) # answer to part (g)
   lm(gasMileage ~ weight, data = cars) # answer to part (h)

```

2. (iii)

3. (i)

4. (a) This is an experiment. The treatments, light sources, are imposed on the subjects by the researchers.

(b) The explanatory variable is the light source.

(c) The response variable is the exam performance/score.

(d) This is an example of *blocking*.

5. D, A, C, B, in that order

6. (8 pts) From top to bottom, it goes V, II, IV, VI

7. (a) It is unimodal and right-skewed.

(b) Because of the positive skew, the median is smaller.

(c) There are three in the list which are sensitive to outliers: the range, the mean, and the standard deviation.

8. (iii)

9. (a) A, B, D, C, in that order

(b) A

(c) C

10. The mean:

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

The deviations $(x_j - \bar{x})$ 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:

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

Standard deviation:

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

11. (a) $222/434 \doteq 0.512$

- (b) $125/434 + 222/434 - 92/434 = 255/434 \doteq 0.588$
- (c) It is "not A and not B ", which is really "male and Republican".
- (d) $92/125 \doteq 0.736$
- (e) They are not disjoint, since there are female Democrats.
- (f) They are not independent, since $\Pr[B \mid A] \doteq 0.736$ is different from $\Pr[B] \doteq 0.512$.
- (g) There is an association. In particular, if you know party affiliation, this makes you better informed about sex.
- (h) side-by-side bar graphs

12. (a) $\Pr[X > 9] = \Pr[X = 10] + \Pr[X = 11] + \Pr[X = 12] = \frac{3}{36} + \frac{2}{36} + \frac{1}{36} = \frac{6}{36} = \frac{1}{6}$.

- (b) One indicator that Y is not binomial is the lack of a fixed number of trials. In binomial settings, the number of trials is known before you begin. But, here, it isn't known if James will win in the third trial, the tenth, or whenever.

Another indicator is in what Y keeps track of: it's value is the number of the trials/rolls it takes for James to win. A binomial r.v. counts the number of successes in n trials.

- (c) The value $Y = 4$ requires James to fail three times in a row, and finally succeed, in getting a higher roll than Beth. That may be visualized as a single path along a probability tree with the sequence FFFS. Since the chance of success is $5/18$, the multiplication rule for probabilities says you get this result with probability

$$\left(\frac{5}{6}\right)^3 \left(\frac{1}{6}\right) \doteq 0.0965.$$

- (d) We have

$$\Pr[Y \geq 3] = 1 - \Pr[Y \leq 2] = 1 - \Pr[Y = 1] - \Pr[Y = 2] = 1 - \frac{1}{6} - \left(\frac{5}{6}\right)\left(\frac{1}{6}\right) \doteq 0.6944.$$