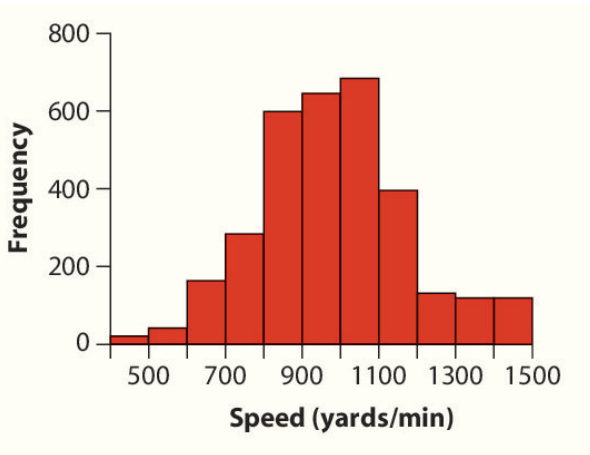


## Assignment 2

1. Francis Galton (1894) presented the following data on the flight speeds of 3207 “old” homing pigeons traveling at least 90 miles.



- What type of graph is this?
- Examine the graph and visually determine the approximate value of the mean (to the nearest 100 yards per minute). Explain how you obtained your estimate.
- Examine the graph and visually determine the approximate value of the median (to the nearest 100 yards per minute). Explain how you obtained your estimate.
- Examine the graph and visually determine the approximate value of the mode (to the nearest 100 yards per minute). Explain how you obtained your estimate.
- Examine the graph and visually determine the approximate value of the standard deviation (to the nearest 50 yards per minute). Explain how you obtained your estimate.

2. Measurements of lifetime reproductive success (LRS) of individual wild animals reveal the disparate contributions they make to the next generation. Jensen et al. (2004) estimated LRS of male and female house sparrows in an island population in Norway. They measured LRS of an individual as the total number of “recruits” produced in its lifetime, where a recruit is an offspring that survives to breed one year after birth. Parentage of recruits was determined from blood samples using DNA techniques. Their results are tabulated as follows:

Lifetime reproductive success	Frequency	
	Females	Males
0	30	38
1	25	17
2	3	7
3	6	6
4	8	4
5	4	10
6	0	2
7	4	0
8	1	0
> 8	0	0
Total	81	84

- Which sex has the higher mean lifetime reproductive success?
- Every recruit must have both a father and a mother, so it is not easy to see why male and female LRS should differ. Can you think of a biological explanation?
- Which sex has the higher variance in reproductive success?

#### R exercise

- read the data file on sparrow reproductive success and calculate the mean reproductive success for each sex.
  - Calculate the variance for each sex.
- Do the same but for males and females with lifetime reproductive success  $\geq 4$  and  $> 4$ .

**3.** A massive survey of sexual attitudes and behavior in Britain between 1999 and 2001 contacted 16,998 households and interviewed 11,161 respondents aged 16–44 years (one per responding household). The frequency distributions of ages of men and women respondents were the same. The following results were reported on the number of heterosexual partners individuals had had over the previous five-year period (Johnson et al. 2001).

	Sample size, $n$	Mean	Standard deviation
<b>Men</b>	4620	3.8	6.7
<b>Women</b>	6228	2.4	4.6

- What is the standard error of the mean in men? What is it in women? Assume that the sampling was random.
- Which is a better descriptor of the variation among men in the number of sexual partners, the standard deviation or the standard error? Why?
- Which is a better descriptor of uncertainty in the estimated mean number of partners in women, the standard deviation or the standard error? Why?
- A mysterious result of the study is the discrepancy between the mean number of partners of heterosexual men and women. If each sex obtains its partners from the other sex, then the true

mean number of heterosexual partners should be identical. Considering aspects of the study design, suggest an explanation for the discrepancy.

4. *Amorphophallus johnsonii* is a plant growing in West Africa, and it is better known as a “corpseflower.” Its common name comes from the fact that when it flowers, it gives off a “powerful aroma of rotting fish and faeces” (Baeth 1996). The flowers smell this way because their principle pollinators are carrion beetles, who are attracted to such a smell. Baeth (1996) observed the number of carrion beetles (*Phaeochrous amplus*) who arrive per night to flowers of this species. The data are given below:

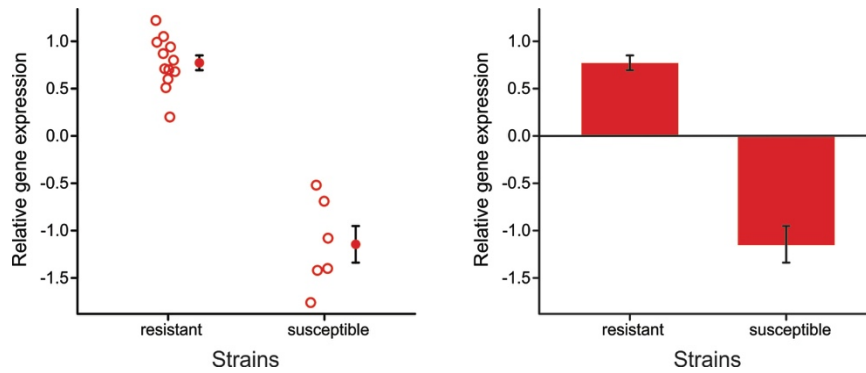
51 45 61 76 11 117 7 132 52 149

- What is the mean and standard deviation of beetles per flower?
- What is the standard error of this estimate of the mean?
- Give an approximate 95% confidence interval of the mean. Provide lower and upper limits.
- If you had been given 25 data points instead of 10, would you expect the mean to be greater, less than, or about the same as the mean of this sample?
- If you had been given 25 data points instead of 10, would you have expected the standard deviation to be greater, less than, or about the same as this sample?
- If you had been given 25 data points instead of 10, would you have expected the standard error of the mean to be greater, less than, or about the same as this sample?

#### R exercise

- Load the data into R, visualize the data and check your answer of a-c using R.
- Assume for now that the data are normally distributed (we will take about this later this week) and use the mean and standard deviation to simulate a new data set of  $n = 25$ . Check your answers for e and f. Use the function `rnorm()` and round to whole numbers using `round()`.
- Redo h for  $n = 50$  and  $n = 100$  as well to check.

5. The accompanying figure shows two alternative ways of presenting means with standard error bars in a graph. The data are from Dabord et al. (2002), who showed that elevated expression of the gene *Cyp6g1* in *Drosophila* causes resistance to DDT. Expression levels of the gene (relative to a standard) were measured in 12 resistant strains of *Drosophila* and 6 susceptible strains. Which graphical method is superior? Explain.



6. How long do you hug somebody? Nagy (2011) measured the duration of spontaneous embraces at the 2008 Summer Olympic Games in Beijing, China. The data are the durations of hugs, in seconds, of athletes immediately after competing in the finals of an event. Hugs were either with their coach, a supporter (e.g., a team member), or a competitor. Descriptive statistics calculated from the data are in the accompanying table.  $n$  refers to the sample size.

Relationship	Mean	Standard deviation	$n$
Coach	3.77	3.96	77
Supporter	3.16	2.76	75
Competitor	1.81	1.13	33

- According to the values in the table, which relationship group gets the longest hugs, on average, and which the briefest hugs? Do the values shown represent parameters or sample estimates? Explain.
- Using the numbers in the table, calculate the standard error of the mean hug duration for each relationship group. What do these values measure?
- What assumption(s) about the samples are you making in (b)?
- Using the numbers in the table, calculate an approximate 95% confidence interval for the mean hug duration when athletes embrace competitors. Provide the lower and upper limits of the confidence interval.
- In light of your results in (d), consider the most-plausible values for the mean duration of hugs-with-competitors in the population of athletes. Is 2 seconds among the most plausible values for the population mean hug duration?
- For which of the relationship groups is the possibility of a 3-second mean hug duration in the population plausible?