1. **Data**: Explain in your own words what each of the following kinds of data is:

   (a) anecdotal data

   (b) available data

   (c) data from a sample survey

   (d) data from an experiment

   (e) census data

   **Food for thought**: sometimes, data can belong to several of the above classes. For example, census data is a special kind of data from a sample survey. Think about which of the above data categories are exclusive of one another, and which have some overlap. Record your thoughts.

2. **Experiments**: 

   (a) Define the following terms:

      i. experimental unit (subject)
ii. treatment

iii. explanatory variable (factor)

iv. response variable (outcome)

(b) Food for thought:

i. Why do we use randomization in experiments?

ii. Why do we use repetition in experiments?

iii. Give an example of a lurking variable in action.

iv. Why would ever make sense to use a “block” design?

v. In a “matched paris” design, each subject is matched with a pair. In general, what is the criterion for the match?

vi. What is the purpose of conducting a “double blind” experiment?

3. Samples

(a) Define the following terms:

i. Population
ii. Sample

iii. A simple random sample of size \( n \)

iv. A stratified random sample

v. A multistage sample

vi. A voluntary response sample

(b) **Food for thought:**

i. What is the difference between *undercoverage* and *non-response*?

ii. Why is a *voluntary response* survey generally a bad idea?

iii. Why is it important to word survey questions clearly and precisely?

4. **Inference**

(a) Define the following terms:

i. Parameter
ii. Statistic

(b) **Food for thought:** The purpose of a statistic is to estimate a parameter.
   i. If a study is well designed, will the statistic generally *equal* the parameter? Why or why not?

   ii. In practice, would you ever *know* if your statistic was equal to the parameter?

   iii. What does it mean to say that “a statistic has low bias but high variability”?

5. **Probability**
   (a) Define the following terms:
      i. Sample space

      ii. Event

      iii. Random variable

      iv. Discrete random variable

      v. Continuous random variable

      vi. Probability distribution
(b) **Food for thought**

i. Explain in words what it means to say that “the probability of flipping a head is $1/2$”.

ii. Let $X$ be a random variable that can take on value $-1$, 0, or 1. Is the mean of $X$ necessarily 0? Why or why not?

iii. Write down the 5 basic properties of any probability, and make sure you can explain in simple words what each one means.

iv. Write down the formulas for the mean, variance, and standard deviation of a discrete random variable, and make sure you have done (and understand) HW 4.72 and 4.75

v. Suppose $X$ has mean $\mu_X$ and variance $\sigma_X^2$, while $Y$ has mean $\mu_Y$ and variance $\sigma_Y^2$. If $a$ and $b$ are constants, what are the mean and variance of the random variable $Z = a + bX + Y$?