Aman Patel

CSCI-B 365

February 17, 2021

Homework 2

1. Kale, Diabetes in Massachusetts
   1. cannot be computed because we do not have data for the whole population, only a sample of Massachusetts adults.
      1. The sample data can be used to estimate the population proportion, but an exact value cannot be calculated.
   2. confidence intervals
      1. Kale
      2. No Kale
   3. There is no overlap between the confidence intervals, and it can be concluded that the kale-eaters in this sample were less likely to have diabetes. Also,
   4. Causation cannot be concluded because the study did not account for confounding variables, such as exercise level, family history, etc.
   5. People who eat kale may be more likely to have a healthier diet and exercise frequently, lowering their risk of diabetes.
2. Is there evidence to conclude causation?
   1. No, the 4 zip codes selected may not be representative of the population.
   2. No, the samples should be taken from the same population, (health-conscious or not health-conscious) and then randomly assigned kale or no kale.
   3. No, this works to randomly sample people, but it does not isolate kale as the cause of lower diabetes rates.
3. See attached R file
4. See attached R file
5. Independent or Dependent events
   1. Dependent
      1. The mayor and police chief may be linked (same political party, etc.)
   2. Independent
      1. The first person’s choice is unlikely to affect the second person’s choice.
   3. Dependent
      1. If event A is false, event B must be true. The outcome of event B will always be opposite that of event A.
   4. Dependent
      1. If a person likes “The Incredibles,” they are also likely to like “The Incredibles 2.” If a person dislikes “The Incredibles,” they are also likely to dislike “The Incredibles 2.”
6. Traits in the population
   1. and are not independent because
      1. Independent events:
         1. If independent,
      2. and are not independent.
   2. See attached R file
   3. See attached R file