Begin your response to **QUESTION 2** on this page.

2. Researchers will conduct a year-long investigation of walking and cholesterol levels in adults. They will select a random sample of 100 adults from the target population to participate as subjects in the study.

   (a) One aspect of the study is to record the number of miles each subject walks per day. The researchers are deciding whether to have subjects wear an activity tracker to record the data or to have subjects keep a daily journal of the miles they walk each day. Describe what bias could be introduced by keeping the daily journal instead of wearing the activity tracker.

GO ON TO THE NEXT PAGE.

Use a pencil or pen with black or dark blue ink only. Do NOT write your name. Do NOT write outside the box.

© 2021 College Board.
Visit College Board on the web: collegeboard.org.

6

Continue your response to **QUESTION 2** on this page.

During the course of the study, the subjects will have their cholesterol levels measured each month by a doctor. The researchers will perform a significance test at the end of the study to determine whether the average cholesterol level for subjects who walk fewer miles each day is greater than for those who walk more miles each day.

(b) Selecting a random sample creates a reasonable representative sample of the target population. Explain the benefit of using a representative sample from the population.

(c) Suppose the researchers conduct the test and find a statistically significant result. Would it be valid to claim that increased walking causes a decrease in average cholesterol levels for adults in the target population? Explain your reasoning.

GO ON TO THE NEXT PAGE.

Begin your response to **QUESTION 3** on this page.

3. To increase morale among employees, a company began a program in which one employee is randomly selected each week to receive a gift card. Each of the company's 200 employees is equally likely to be selected each week, and the same employee could be selected more than once. Each week's selection is independent from every other week.

(a) Consider the probability that a particular employee receives at least one gift card in a 52-week year.

   (i) Define the random variable of interest and state how the random variable is distributed.

   (ii) Determine the probability that a particular employee receives at least one gift card in a 52-week year. Show your work.

GO ON TO THE NEXT PAGE.

Use a pencil or pen with black or dark blue ink only. Do NOT write your name. Do NOT write outside the box.

© 2021 College Board.
Visit College Board on the web: collegeboard.org.

8

Question 2: Focus on Collecting Data — 4 points

General Scoring Notes

- Each part of the question (indicated by a letter) is initially scored by determining if it meets the criteria for essentially correct (E), partially correct (P), or incorrect (I). The response is then categorized based on the scores assigned to each letter part and awarded an integer score between 0 and 4 (see the table at the end of the question).
- The model solution represents an ideal response to each part of the question, and the scoring criteria identify the specific components of the model solution that are used to determine the score.

Model Solution	Scoring
(a) Keeping daily journals could introduce response bias due to the self-reporting by subjects who may have a poor or incomplete memory of the amount of walking that was done. If most subjects who keep daily journals underreport the number of miles walked per day because they cannot remember all of their walking at the end of the day, then the estimate of mean daily miles walked for the target population will be biased too low. Wearing activity trackers would likely provide a more accurate record of daily miles walked by each subject in the study.	**Essentially correct (E)** if the response satisfies the following two components: 1. Indicates that keeping a daily journal could result in a bias that would be avoided by using activity trackers AND provides a reasonable explanation 2. Provides a description of a bias that refers to at least one of the following: • The use of a daily journal may result in a systematic/consistent underreporting, or systematic/consistent overreporting of daily miles walked • The use of a daily journal may result in a biased estimation (underestimation or overestimation) of a population parameter (e.g., mean daily miles walked for the members of the target population) **Partially correct (P)** if the response satisfies only one of the two components. **Incorrect (I)** if the response does not meet the criteria for E or P.

Additional Notes:

- A response does not need to specifically name a type of bias (e.g., response bias).
- The response may refer to the explanatory variable as "activity level."
- The direction of the bias need not be specified in order to satisfy component 1.
- Examples of reasonable explanations for indicating that keeping a daily journal may result in a bias include:
  - "Because the subjects are self-reporting their daily miles walked."
  - "Because the subjects may not accurately recall their daily miles walked."
  - "Because the subjects may forget to complete an entry in their journal."
- The direction of the bias must be specified in order to satisfy component 2.

- The response must indicate the underreporting or overreporting is systematic across the subjects (or there is a tendency to underreport or overreport) in order to satisfy component 2. Examples of responses that satisfy component 2 include:
  - "The subjects in the study may consistently underreport their daily miles walked."
  - "Subjects are likely to underreport their daily miles walked."
  - "Most subjects may overreport their daily miles walked."
  - "The bias may result in an estimate of the mean daily miles walked by members of the target population that is lower than the target population mean."
- A response that indicates the underreporting or overreporting for only some people does not satisfy component 2 (e.g., "Some people might record higher miles than they actually walk.").

Model Solution	Scoring
(b) It is necessary to have a representative sample of subjects from the population in order to make an unbiased inference about the difference between the mean cholesterol levels for all adult members of the target population who walk fewer miles per day and the mean cholesterol levels for all adult members of the target population who walk more miles per day.	**Essentially correct (E)** if the response satisfies the following two components: 1. Provides an explanation that the use of a representative sample is necessary in order to make a valid generalization about the target population 2. Refers to estimation, or inference, for cholesterol levels in the target population OR an association between cholesterol level and amount of walking in the target population **Partially correct (P)** if the response satisfies only one of the two components. **Incorrect (I)** if the response does not meet the criteria for E or P.

Additional Notes:
- A response that discusses the accuracy or validity of a significance test does not satisfy component 1 unless the response makes it clear that the inference is being generalized to the target population.
- In order to satisfy component 2, the response need not state a specific population parameter(s).
- If a parameter is specified, it must be relevant to cholesterol level or the association between cholesterol level and amount of walking. Some examples include:
  o Individual population mean cholesterol level
  o One or more differences between population mean cholesterol levels
  o Individual population median cholesterol level
  o One or more differences between population median cholesterol levels
  o A population correlation between cholesterol level and amount of walking
  o A population regression model for cholesterol level and amount of walking