AP Computer Science Principles

Unit 5 Practice Exam: Data

AP Computer Science Principles

2026–27 · 30 Questions · 30 Points

CC BY-NC 4.0

2026–27 · 30 Questions · 30 Points

Name: _____ Period: _____ Date: _____

DIRECTIONS: Practice exam. 30 questions, 30 points. Choose the single best answer unless a question says to select two; select-two questions are worth 1 point with partial credit. Code is AP pseudocode (Exam Reference Sheet).

1. Which of the following best describes the difference between *data* and *information*?

(A) Data is always numbers; information is always text.
(B) Information is the facts and patterns that people extract from data.
(C) Data is stored on computers; information is stored on paper.
(D) There is no difference; the words mean the same thing.

2. A video file is stored on a website. Which of the following is *metadata* about the video?

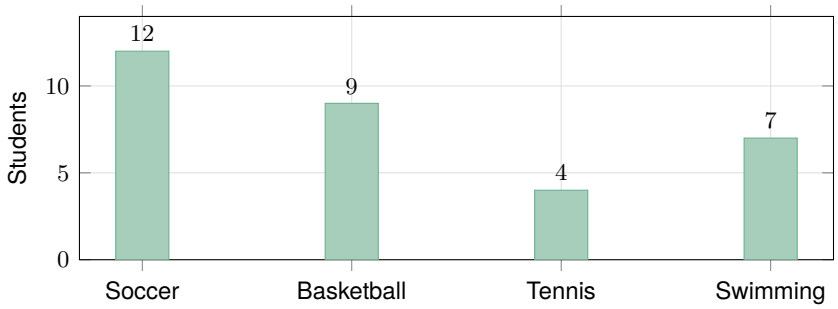
(A) The images in each frame of the video
(B) The sound of the people talking in the video
(C) The words spoken in the video
(D) The date the video was uploaded and its length

3. A student posts a photo online. The photo's metadata includes the GPS location where it was taken. Which of the following is a possible concern?

(A) Someone could use the metadata to learn where the photo was taken, such as the student's home.
(B) The metadata makes the photo blurry.
(C) The metadata prevents the photo from being shared.
(D) The metadata changes the colors in the photo.

 CC BY-NC 4.0

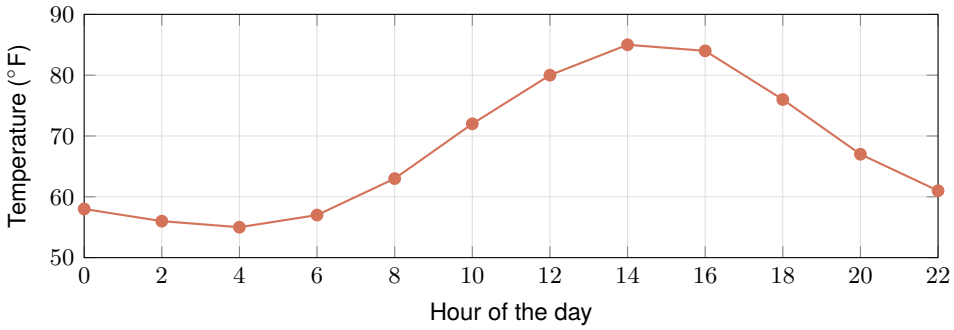
4. The bar chart shows the favorite sport of each student in a gym class.



How many more students chose soccer than tennis?

(A) 4
(B) 12
(C) 8
(D) 3

5. The line chart shows the temperature in a city every two hours during one day.

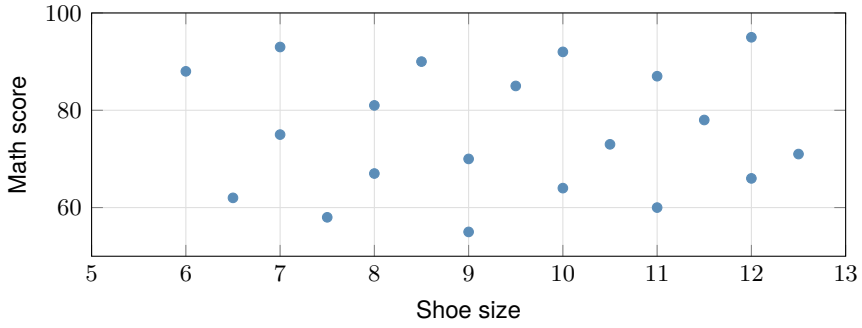


Which statement is supported by the chart?

(A) The temperature rose every two hours all day.
(B) The coldest temperature recorded was at hour 0.
(C) The temperature was the same at hour 8 and hour 20.
(D) The highest temperature recorded was at hour 14.

 CC BY-NC 4.0

6. Each point in the scatter plot is one student. It shows the student's shoe size and score on a math test.



Which conclusion is best supported by the scatter plot?

(A) Students with bigger feet score higher on math tests.
(B) Students with bigger feet score lower on math tests.
(C) There is no clear relationship between shoe size and math score.
(D) Shoe size causes differences in math scores.

7. A survey asked students for their grade and favorite season. The cross tab below counts the responses.

	Fall	**Winter**	**Spring**	**Summer**
Grade 9	6	3	5	16
Grade 10	8	7	4	11
Grade 11	10	2	6	12

Which statement is supported by the table?

(A) Winter was the most popular season in every grade.
(B) Summer was the most popular season in every grade.
(C) More 11th graders chose fall than summer.
(D) The same number of students in each grade chose spring.

8. A student wants to compare how many people voted for each of five candidates in a class election. Which visualization is most appropriate?

(A) A bar chart with one bar per candidate
(B) A scatter plot of votes versus candidate name length
(C) A histogram of the students' ages
(D) A line chart of votes in alphabetical order

 CC BY-NC 4.0

9. A bar chart compares two phone batteries. Battery X lasts 98 hours and Battery Y lasts 100 hours. The chart's vertical axis starts at 97 instead of 0, so Battery Y's bar looks three times as tall as Battery X's. Which is the main problem with this chart?

(A) The data is wrong.
(B) Bar charts cannot show battery life.
(C) The chart should be a scatter plot.
(D) It makes a small difference look much larger than it is.

10. An online survey accidentally recorded some responses twice when students double-clicked the submit button. What should be done before analyzing the results?

(A) Sort the responses alphabetically.
(B) Remove the duplicate responses.
(C) Double every response so that all responses count equally.
(D) Make a chart before checking the data.

11. A spreadsheet of 500 students is missing the age for 40 of them. A student calculates the average age by treating the missing ages as 0. What is the most likely effect?

(A) The average will be exactly correct.
(B) The average will be too high.
(C) The average will be too low.
(D) The spreadsheet will not be able to calculate an average.

12. A theater keeps the following table of shows.

Show	Day	Tickets Left	Price ($)
Comedy Night	Friday	0	15
Jazz Trio	Friday	40	25
Magic Show	Saturday	12	18
Puppet Play	Saturday	30	10
Dance Recital	Sunday	5	20
Film Classic	Sunday	0	8

How many shows still have tickets left *and* cost less than $20?

(A) `3`
(B) `4`
(C) `1`
(D) `2`

13. For a row of the table in the previous question, `left` contains the number of tickets left and `price` contains the price. Which expression is `true` exactly for the shows counted in the previous question?

(A) `(NOT (left = 0)) OR (price < 20)`
(B) `(NOT (left = 0)) AND (price < 20)`
(C) `(left = 0) AND (price < 20)`
(D) `(NOT (left = 0)) AND (price ≤ 20)`

 CC BY-NC 4.0

14. The table in Question 12 is sorted by Price from greatest to least. Which show appears in the first row?

(A) Jazz Trio
(B) Film Classic
(C) Comedy Night
(D) Dance Recital

15. A school has one spreadsheet with each student's ID, name, and grade level, and another with each student's ID and the clubs they belong to. What makes it possible to combine the two spreadsheets correctly?

(A) Both spreadsheets have the same number of rows.
(B) Both spreadsheets are sorted by name.
(C) Both spreadsheets include the student ID, which is unique to each student.
(D) Both spreadsheets were made by the same teacher.

16. A gym records the following for each visit: member ID, date, check-in time, and check-out time. Which question can be answered using only this data?

(A) Which exercise machine is used the most?
(B) Do members who visit more often lose more weight?
(C) On which day of the week do members stay the longest, on average?
(D) How satisfied are members with the gym?

17. Data shows that towns with more fire trucks tend to have more fires each year. Which conclusion is best supported?

(A) Fire trucks cause fires.
(B) Towns should get rid of fire trucks to prevent fires.
(C) There is no relationship between fire trucks and fires.
(D) The two are correlated; a third factor, such as town size, may explain both.

18. A student surveys 15 friends at lunch to learn what all 1,200 students at the school think about the dress code. Which changes would make the results more representative of the whole school?

(Select two answers.)

(A) Survey students from every grade, not only the student's friends.
(B) Ask the same 15 friends again next week.
(C) Collect responses from a much larger group chosen from across the school.
(D) Use a longer survey with more questions.

CC BY-NC 4.0

19. Which of the following is the best example of *big data*?

(A) A list of the 25 students in a class
(B) Billions of search queries collected by a search engine each day
(C) A shopping list for one week
(D) The scores from one basketball game

20. A city publishes open data about calls to its non-emergency hotline, including the street address of each call. Which is a possible risk of publishing this data?

(A) Open data cannot be downloaded.
(B) Publishing data makes the hotline stop working.
(C) Addresses could reveal private information about the people who called.
(D) Open data can only be used by the city.

21. A transit agency asks riders to report broken elevators at stations through an app. Which is the main benefit of this crowdsourced approach?

(A) Every report is guaranteed to be accurate.
(B) Problems can be reported quickly from many stations, more than staff could check on their own.
(C) Riders no longer need elevators.
(D) The agency no longer needs to store data.

22. In a citizen science project, thousands of volunteers count birds in their backyards. Which is a challenge the researchers are most likely to face?

(A) The data cannot be stored on a computer.
(B) There will be too few observations to analyze.
(C) Volunteers will all count exactly the same birds.
(D) Some volunteers may misidentify birds, so the data may contain errors.

23. A team wants to build a machine learning model that tells apart photos of cats and dogs. What does the team need first?

(A) A list of every cat and dog in the world
(B) Many example photos that are labeled “cat” or “dog”
(C) A rule for each pixel that a programmer writes by hand
(D) A faster Internet connection

24. After training a model, a team tests it using photos that were *not* part of its training data. Why is this important?

(A) It checks whether the model works on new examples it has not seen before.
(B) It makes the model run faster.
(C) It removes all bias from the model.
(D) It is required for the model to be saved.

 CC BY-NC 4.0

25. Researchers found that some facial analysis programs were much less accurate for women with darker skin than for men with lighter skin. Which is the most likely reason?

(A) Computers cannot process images of women.
(B) The training images included far fewer women with darker skin.
(C) The programs were written in the wrong programming language.
(D) The cameras were broken.

26. A loan-approval program does not use applicants' race, but it does use their ZIP code. Why might the program still produce biased results?

(A) ZIP codes are too long for the program to read.
(B) Programs cannot use numbers as inputs.
(C) ZIP codes can be closely related to race and income, so they can stand in for that information.
(D) Using ZIP codes guarantees the program is fair.

27. Which of the following can introduce bias into a computing innovation?

(Select two answers.)

(A) Storing the data on a computer instead of on paper
(B) Training data that does not represent all of the people the innovation affects
(C) Using a data set with more than 1,000 rows
(D) Training on past decisions that were unfair to some groups

28. Which order best describes the data analysis process?

(A) Visualize, collect, clean, draw conclusions
(B) Draw conclusions, collect, visualize, clean
(C) Collect, clean, visualize, find patterns and draw conclusions (then repeat with new questions)
(D) Clean, draw conclusions, collect, visualize

29. A chart shows that a store's online sales grew every year from 2019 to 2025. Which statement is best supported by the chart alone?

(A) Online sales will keep growing forever.
(B) Customers prefer shopping online to shopping in the store.
(C) The store should close its physical locations.
(D) The store's online sales were higher in 2025 than in 2019.

CC BY-NC 4.0

30. A teacher records each student's name, the number of homework assignments turned in, and the score on each test. Which questions can be answered using only this data?

(Select two answers.)

(A) Do students who turn in more homework tend to score higher on tests?
(B) Do students who sleep more turn in more homework?
(C) What was the average score on the first test?
(D) Which students have after-school jobs?

CC BY-NC 4.0