

AP Computer Science Principles | Unit 2

2.1 Data Representation

Solutions

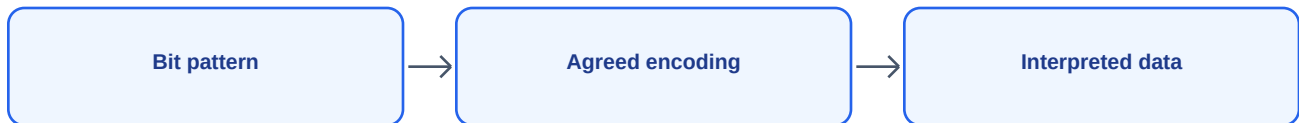
Reference scenario

Interpret a short binary value and an uncompressed image representation.

Binary: 10110

Place values: 16 8 4 2 1

Image: 20 by 10 pixels, 8 bits per pixel



2.1.1

Convert binary 11001 to decimal and state the number of distinct values representable with five bits.

11001 equals $16 + 8 + 1 = 25$. Five bits provide $2^5 = 32$ distinct bit patterns.

2.1.2

How many bytes of uncompressed pixel data are needed for a 16 by 16 image with 24 bits per pixel, excluding headers?

$16 * 16 * 24 = 6144$ bits. Dividing by 8 gives 768 bytes.

AP Computer Science Principles | Unit 2

2.2 Cleaning and Transforming Data

Solutions

Reference scenario

Prepare a survey data set for analysis without silently inventing values.

Row 1: grade 10, minutes 30
Row 2: grade "10 ", minutes 45
Row 3: grade 10, minutes missing
Row 4: duplicate of row 1

2.2.1

For observed values 10, 20, and 90, calculate the mean and median. What does the comparison show?

The mean is $120/3 = 40$ and the median is 20. The large value pulls the mean upward; the two statistics describe the data differently.

2.2.2

Why should a cleaning log record removed rows and the removal rule?

It makes the transformation reproducible and lets another analyst assess whether records were excluded unfairly or whether the cleaning changed the result.

AP Computer Science Principles | Unit 2

2.3 Visualizing and Interpreting Data

Solutions

Reference scenario

Read a small time series and limit the claim to what was measured.

Week 1: 20 visits

Week 2: 25 visits

Week 3: 40 visits

New poster displayed during week 3

2.3.1

Which visualization fits comparing total visits across four clubs, and which fits visits over ten weeks?

A bar chart fits comparing club categories. A line chart fits the ordered time trend. Label units and use an appropriate scale.

2.3.2

An app chart shows that more study time is associated with higher scores. State a defensible claim and an unsupported causal claim.

Defensible: the measured data show an association between study time and score. Unsupported: changing study time alone necessarily caused the score difference. Other factors and study design affect causal conclusions.

AP Computer Science Principles | Unit 2

2.4 Privacy and Metadata

Solutions

Reference scenario

Evaluate a supposedly anonymous location data set.

Removed: student names

Retained: exact timestamps, precise location, grade

Publication: individual rows

2.4.1

Name one direct identifier and two potential indirect identifiers in a school survey.

A name or student number is a direct identifier. A precise home location and a distinctive combination of grade, activity, and timestamp can indirectly identify a person.

2.4.2

The goal is to count total bus use by weekday. Which is the better public output: full individual trip records or weekday totals, and why?

Weekday totals better match the stated goal while exposing less individual travel behavior. Review small groups and other release risks before publication.