Instruction: Sort the following list into alphabetical order. apple, banana, orange, grape.

Rules

Here are some rules for the evaluation task:

- (1) Prioritize evaluating whether an output faithfully and precisely executes the instruction.
- (2) Output (a) and Output (b) are equally likely to be the better one.
- (3) Good outputs are expected to avoid irrelevant information.

Metrics

Propose three concise questions about what constitutes a good output given the following instruction.



- (1) Does the output correctly sort items in alphabetical order?
- (2) Does the output avoid irrelevant information?
- (3) Is the output safe and helpful?

Find the better output that executes the given instruction.

Consider these metrics during your evaluation:

[METRICS]

Swap and Synthesis

Before Swap

Find the better output that executes the given instruction.

Output (a):

No problem! Here's the sorted list. Grape, apple, banana, orange.

Output (b):

apple, banana, grape, orange.



Output (a) gives more helpful response to the instruction

After Swap

Find the better output that executes the given instruction.

Output (a):

apple, banana, grape, orange.

Output (b):

No problem! Here's the sorted list. Grape, apple, banana, orange.



Output (a) correctly sorts the list into alphabetical order.

Synthesis

Consider the conflicting reasonings below. Which is correct?

Output (a) gives more helpful response to the instruction, choose (a).

Output (b) correctly sorts the list into alphabetical order, choose (b).