3. Five hundred randomly selected middle-aged men and five hundred randomly selected young adult men were rated on a scale from 1 to 10 on their physical flexibility, with 10 being the most flexible. Their ratings appear in the frequency table below. For example, 17 middle-aged men had a flexibility rating of 1.

Physical Flexibility Rating	Frequency of Middle-Aged Men	Frequency of Young Adult Men
1	17	4
2	31	17
3	49	29
4	71	39
5	70	54
6	87	69
7	78	83
8	54	93
9	34	73
10	9	39

GO ON TO THE NEXT PAGE.

(a) Display these data graphically so that the flexibility of middle-aged men and young adult men can be easily compared.

(b) Based on an examination of your graphical display, write a few sentences comparing the flexibility of middle-aged men with the flexibility of young adult men.

GO ON TO THE NEXT PAGE.

4. Baby walkers are seats hanging from frames that allow babies to sit upright with their legs dangling and feet touching the floor. Walkers have wheels on their legs that allow the infant to propel the walker around the house long before he or she can walk or even crawl. Typically, babies use walkers between the ages of 4 months and 11 months.

   Because most walkers have tray tables in front that block babies' views of their feet, child psychologists have begun to question whether walkers affect infants' cognitive development. One study compared mental skills of a random sample of those who used walkers with a random sample of those who never used walkers. Mental skill scores averaged 113 for 54 babies who used walkers (standard deviation of 12) and 123 for 55 babies who did not use walkers (standard deviation of 15).

   (a) Is there evidence that the mean mental skill score of babies who use walkers is different from the mean mental skill score of babies who do not use walkers? Explain your answer.
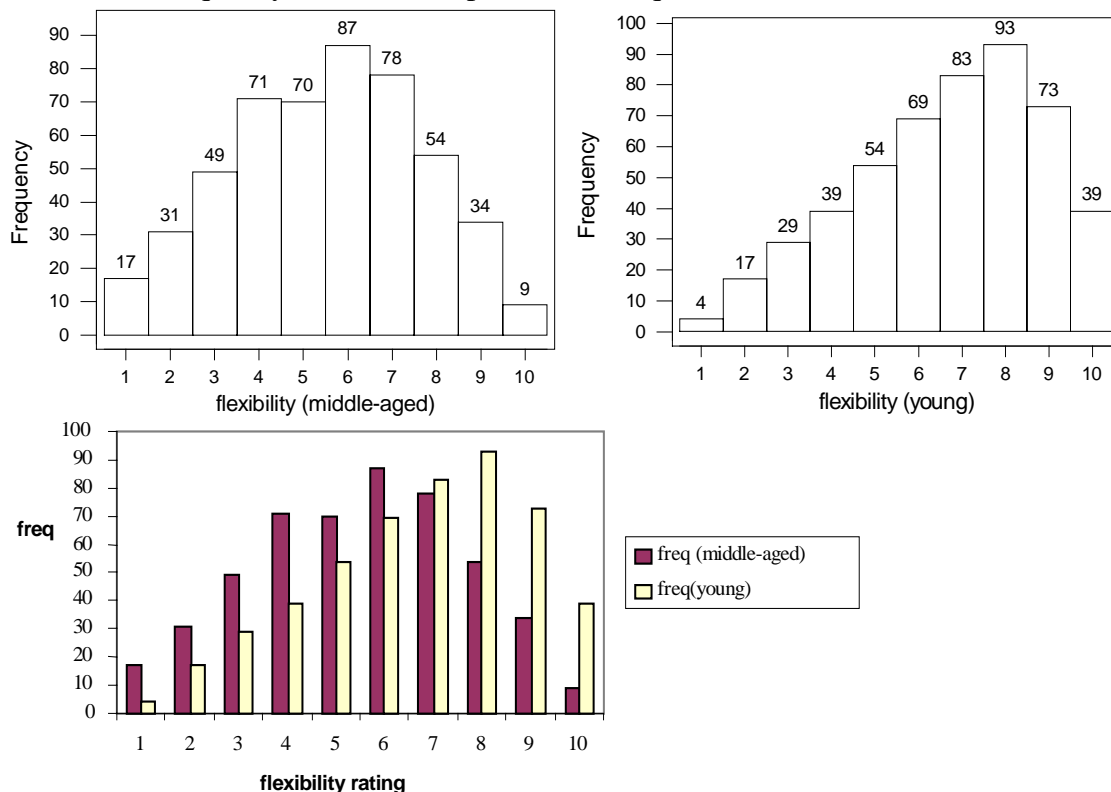
   (b) Suppose that a study using this design found a statistically significant result. Would it be reasonable to conclude that using a walker causes a change in mean mental skill score? Explain your answer.

GO ON TO THE NEXT PAGE.

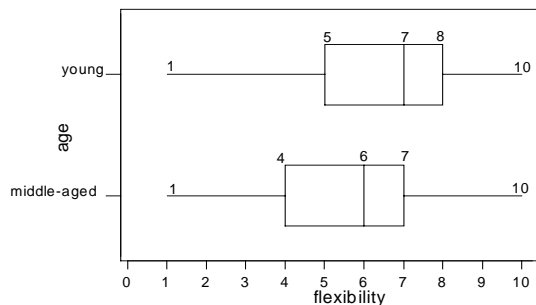# AP® Statistics 2000 – Scoring Guidelines

Question 3

Solution

a. Two histograms, drawn to the same scale or on the same axes. Can use either frequency or relative frequency since the sample sizes are equal (500).
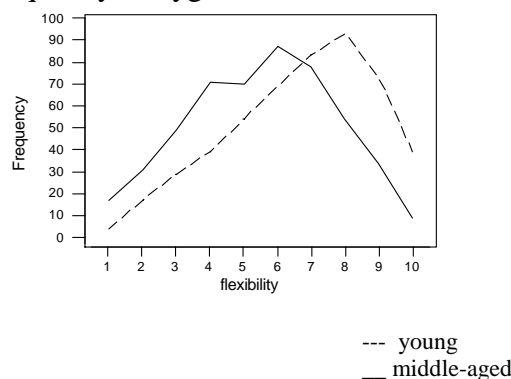
Other possibilities include boxplots or frequency polygons, if done correctly. For example,

Boxplots: Frequency Polygons:
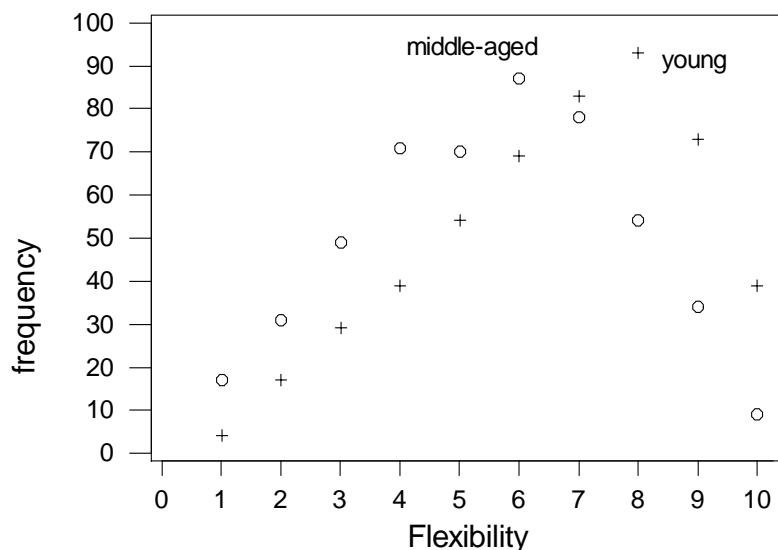
--- young
__ middle-aged

Papers that use cumulative frequency plots (ogives) may also be acceptable. These papers should be referred to a table leader.

Ideally graphs include scales, labels, title and legend.

AP® Statistics 2000 ─ Scoring Guidelines

If a student supplies the following graph, you should grade this problem holistically.

Judge from the answer to (b) whether the student is interpreting this as a frequency polygon or a scatterplot.

- With correct interpretation as a frequency polygon, the paper could be a 4.
- With weak or unclear interpretation as a frequency polygon, could be a 3.
- With mixed interpretations (e.g. association and shape), could be a 2.
- With interpretation as scatterplot (e.g. positive association/correlation), could be a 1.

b.   The distribution of flexibility rating for middle-aged men is approximately symmetric, centered around 5.5, whereas the distribution for young adult men is skewed to the left (negatively skewed) centered around 6.5, higher than the middle aged men. There is quite a bit of variability in both distributions. In general there were more young men with flexibility ratings at the high end of the scale and fewer at the low end of the scale than for middle aged men.

Note: A clear description of the relative concentration of the two distributions (e.g. more flexibility ratings for young men at high values than middle aged men but more middle aged men at lower flexibility ratings than young men) is considered equivalent to a description of shape.

Scoring

Part a is

Essentially correct if	correct graph(s) are drawn, using the same scale so that comparisons can be made easily. Either frequency or relative frequency can be used. Missing labels and legends can be recovered if the description in (b) is clear and complete. Missing scaling cannot be recovered.
Partially correct if	there are errors in the construction of the graphs or the graphs are drawn using very different scales or the graphs are incomplete (but started correctly). Example: no scaling on graph
Incorrect if	any graphical displays that treat the frequencies as data are used (scatterplot, boxplots, dot plots or stem and leaf displays of the frequencies).

Part b is

Essentially correct if	the graphical displays from part (a) are interpreted in context, with comments on the differences and similarities in at least two of center, shape and spread, and the response shows good communication of ideas. (Discussion needs to be clearly linked to graphs and comparison between the two groups must be explicit.)

Note: It is not essentially correct to say the distribution of flexibility ratings for middle-aged men "**is** normal", some qualification must be given, i.e. "approximately Normal".

Partially correct if	interpretation is correct but not in context **or** correct comparison of the two groups is made only on the basis of one of center, shape or spread **or** correct comparison of the two groups is made on at least two of center, shape and spread but communication is weak. **or** at least two of the same individual descriptions in both groups (e.g. center and shape) but no direct comparison between the two groups.

Incorrect if   response fails to compare the two groups on any of center, shape or spread (e.g. only compares the two groups for one value of the flexibility ratings.)

If a display that uses the frequencies as data is done in part (a), no credit is given in part (a) (e.g. frequency boxplots, or frequency vs. frequency/ frequency vs. flexibility scatterplots). However, if a student attempts an interpretation, and they do a credible job with the interpretation, part (b) can be scored as partially correct (resulting in a score of 1 for the problem). For example if a student says that there is a "positive linear relationship" or that center, shape and spread are similar because stem and leaf displays of the frequencies look similar, this is a credible interpretation.

4   Complete Response

Essentially correct on both parts.

3   Substantial Response

Essentially correct on one part and partially correct on the other.

2   Developing Response

Essentially correct on one part and incorrect on the other
OR
Partially correct on both parts

1   Minimal Response

Partially correct on one part

A paper using the frequencies as data can receive at most a 1.

0   No credit