



Data Analysis and Modelling

1 Which city has a higher mean?

City A: 29, 31, 30, 27, 33

City B: 17, 11, 29, 20, 23

A

City B

B

City A

2 Find the mean of these numbers

8, 15, 15, 14, 3

A

9

B

13

C

11

D

16

E

15

F

5

3 What is the median of this data set?

55, 59, 63, 37, 67, 68, 65, 93, 50

A

67.5

B

52.5

C

63

D

60

4 What is the median of this data set?

87, 76, 16, 24, 75, 59, 19, 63, 79, 42

A

50

B

61

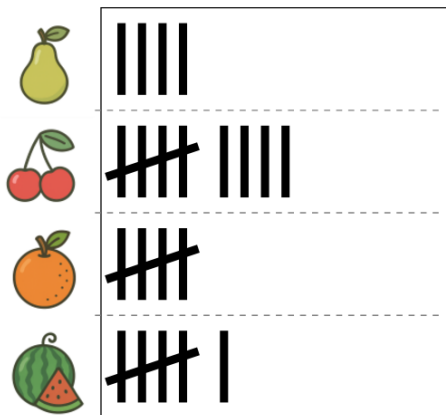
C

24

D

76

5 What is the average number of kids who like each fruit?



A

11

B

7

C

12

D

6

E

13

6 What is the interquartile range (IQR) of this data set?

29, 32, 34, 54, 56, 69, 72, 92, 95

A

115

B

49

C

82

D

66

7 What is the interquartile range (IQR) of this data set?

40, 58, 50, 90, 58, 54, 27, 39

A

18.5

B

63

C

58

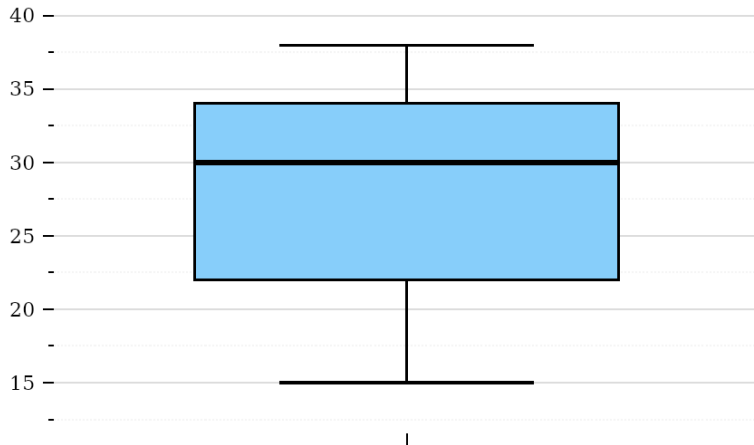
D

97.5



Data Analysis and Modelling

8 What is the interquartile range (IQR) of this box plot?



A

23

B

56

C

12

D

34

9 Find the range of these numbers

4, 8, 7, 4

A

6

B

5

C

1

D

2

E

4

F

9

10 How is this variable best classified?

Height of a randomly selected student

A

Quantitative (continuous)

B

Quantitative (discrete)

C

Qualitative

11 Which of these variables is best classified as quantitative (continuous)?

Quantitative (continuous)

A

Height of a randomly selected student

B

Number of siblings a student has

C

Number of customers entering a store in one hour

12 Why does this sampling approach introduce bias?

Goal: Student satisfaction with cafeteria food
Sampling approach: Ask students, "How bad is the cafeteria food this year?"

A

The wording of the question may encourage students to report lower satisfaction.

B

Students may remember cafeteria meals from earlier in the year more clearly than recent meals.

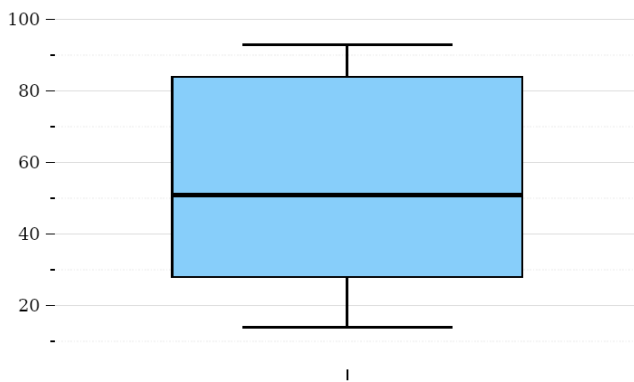


Data Analysis and Modelling

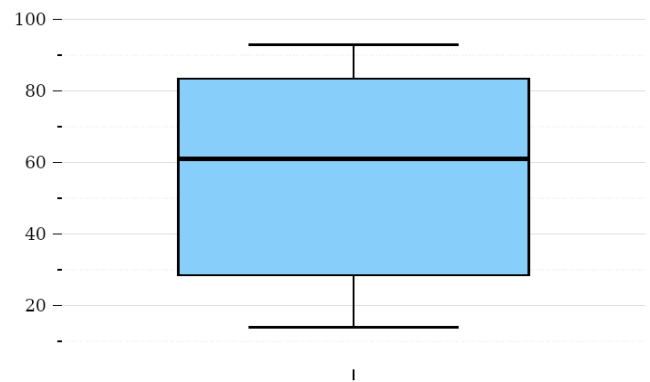
13 Which box plot represents this data set?

63, 61, 24, 77, 14, 59, 90, 93, 33

A



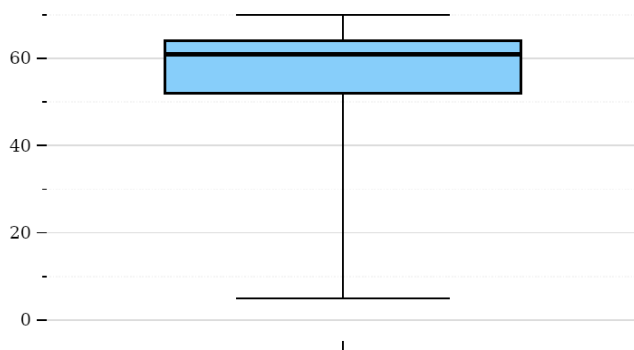
B



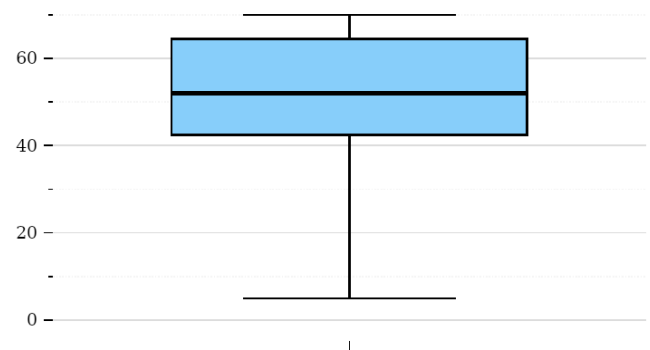
14 Which box plot represents this data set?

5, 40, 45, 48, 52, 54, 59, 70, 70

A



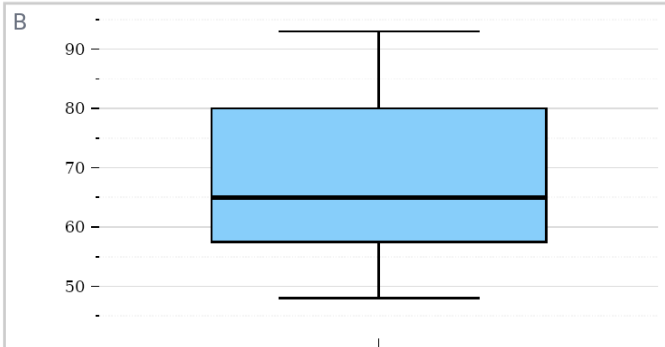
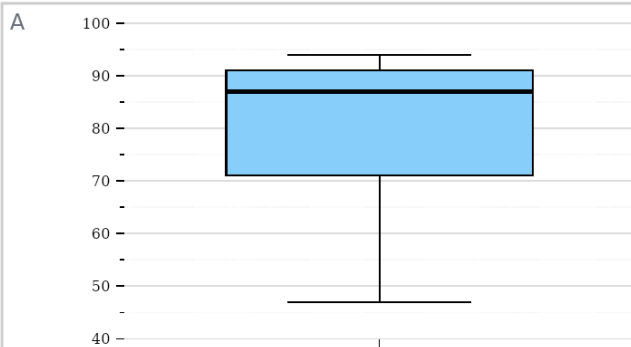
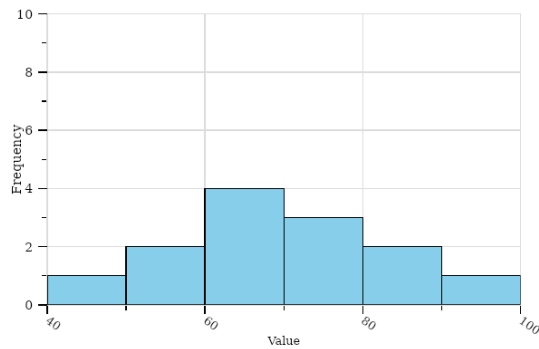
B



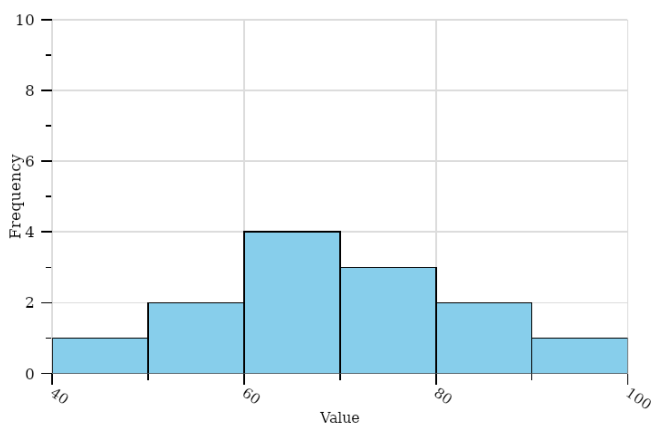


Data Analysis and Modelling

15 Which box plot best matches this histogram?



16 What is the distribution type of this histogram?



A
right-skewed

B
left-skewed

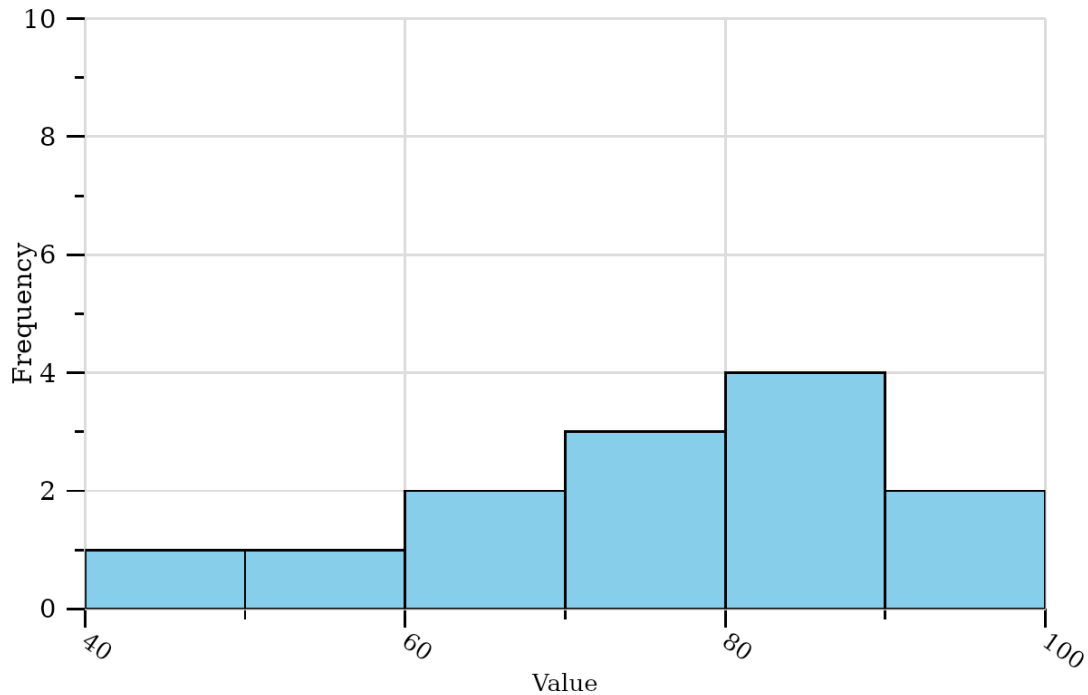
C
approximately symmetric

D
uniform



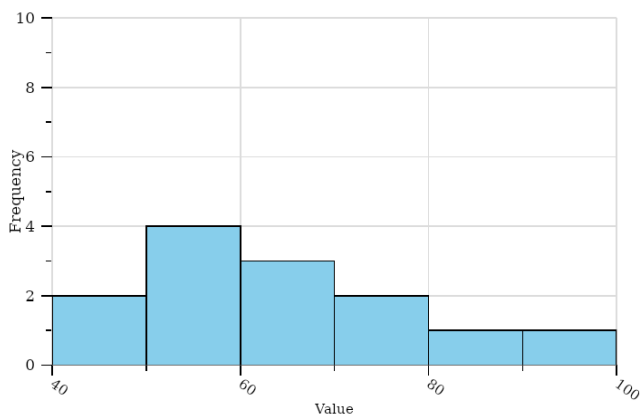
Data Analysis and Modelling

17 Which frequency table matches this histogram?

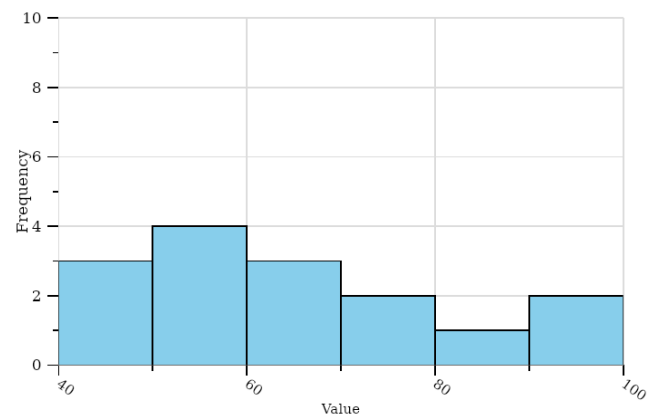


18 Which histogram best matches this frequency table?

A



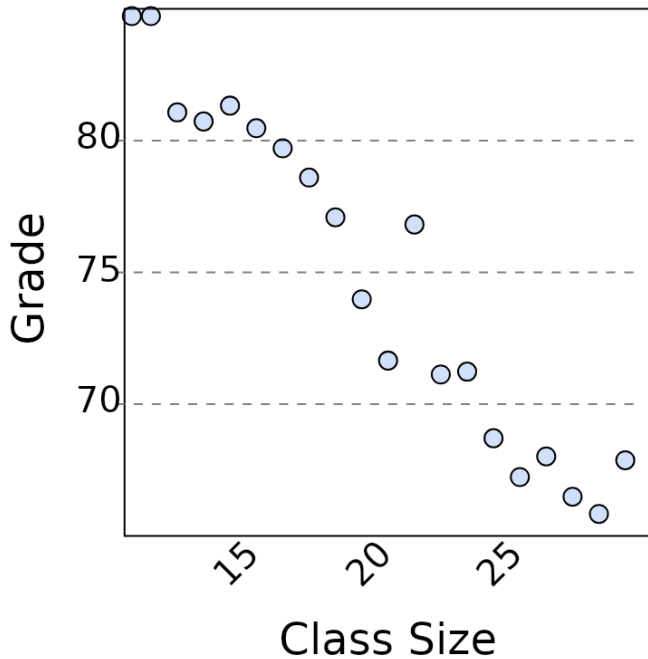
B





Data Analysis and Modelling

19 What is the correlation direction and strength on this scatter plot?



A

strong positive correlation

B

weak negative correlation

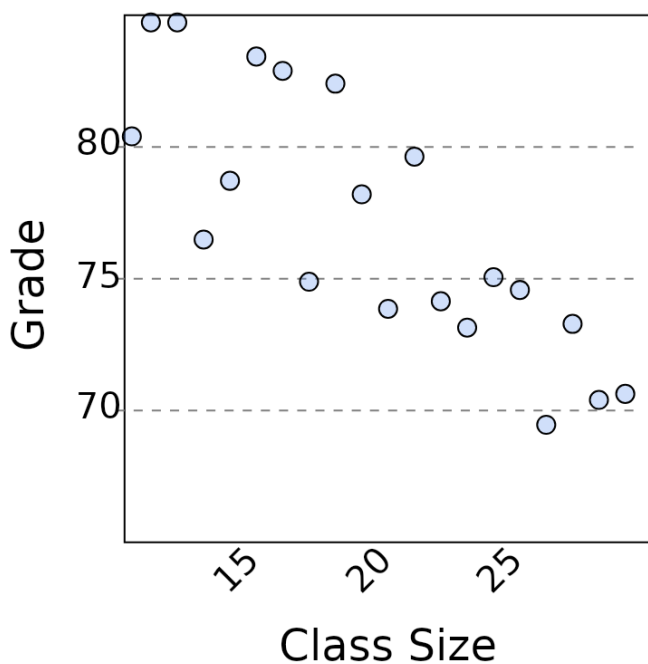
C

weak positive correlation

D

strong negative correlation

20 What would be a reasonable R value for this scatter plot?



A

$r = -0.3$

B

$r = 0.8$

C

$r = -0.8$



Data Analysis and Modelling

21 What type of bias does this sampling approach introduce?

Goal: Student opinion on final exams
Sampling approach: Ask, "Do you agree that stressful final exams should be eliminated?"

- | | |
|---|-------------------------|
| A | Measurement bias |
| B | Voluntary response bias |
| C | Undercoverage bias |
| D | Selection bias |

22 Is this sample representative of the population?

Population: All homes in a province
Sample: 1,000 homes built within the past five years

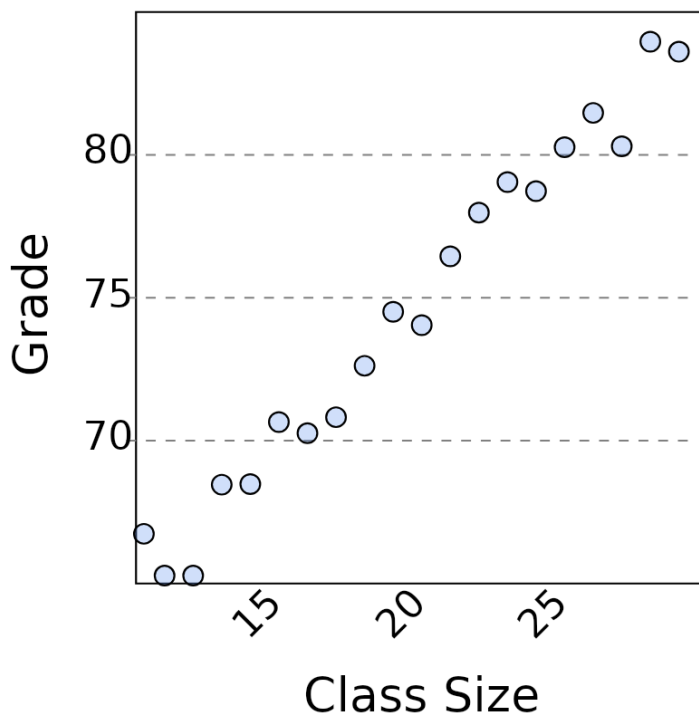
- | | |
|---|--|
| A | Yes, it is representative of the population |
| B | No, it is not representative of the population |

23 Which sample is most representative of this population?

Population: All moviegoers in Canada

- | | |
|---|---|
| A | 1,000 people attending opening weekend screenings |
| B | 1,000 moviegoers randomly selected from theatres nationwide |

24 If class size is 23, what would a reasonable value of grade be?



A

78

B

66