FOUNDATION

1.1 Data Cleaning and Outliers (Foundation): 1-hour Pack

Total Marks

EDEXCEL 1ST0 · Calculator allowed · about 60 minutes

Name: _____ Date: ____ / ____ / _____

Answer ALL questions. Show all your working.

1 A sample of 10 measurements is 38, 37, 39, 38, 41, 38, 41, 96, 33, 38. The value 96 is an outlier.

(a) Work out the mean and the median of all 10 values (mean to 1 decimal place if needed).

(b) Work out the mean and the median without the outlier.

(c) Which average is changed less by removing the outlier?

(Total for Question 1 is 4 marks)

---

2 Here are the shoe sizes recorded for 10 people. Valid values are whole numbers from 3 to 12 inclusive.

12, 5, 3, 11, 9, 10, 7, 999, 999, 9

(a) Write down the invalid values.

(b) Work out the mean of all 10 values, to 1 decimal place if needed.

(c) Work out the mean after the invalid values are removed, to 1 decimal place if needed.

(Total for Question 2 is 5 marks)

© Revision Library | revisionlibrary.co.uk

3 A sample of 10 measurements is 23, 58, 26, 26, 21, 23, 25, 20, 23, 23. The value 58 is an outlier.

(a) Work out the mean and the median of all 10 values (mean to 1 decimal place if needed).

(b) Work out the mean and the median without the outlier.

(c) Which average is changed less by removing the outlier?

(Total for Question 3 is 4 marks)

---

4 A sample of 8 measurements is 24, 22, 24, 19, 16, 15, 64, 19. The value 64 is an outlier.

(a) Work out the mean and the median of all 8 values (mean to 1 decimal place if needed).

(b) Work out the mean and the median without the outlier.

(c) Which average is changed less by removing the outlier?

(Total for Question 4 is 4 marks)

---

5 Here are the number of hours of sleep per night recorded for 10 people. Valid values are whole numbers from 3 to 12 inclusive.

3, 8, -1, 1, 11, 12, 5, 11, 4, 4

(a) Write down the invalid values.

(b) Work out the mean of all 10 values, to 1 decimal place if needed.

(c) Work out the mean after the invalid values are removed, to 1 decimal place if needed.

(Total for Question 5 is 5 marks)

© Revision Library | revisionlibrary.co.uk

6 A sample of 6 measurements is 59, 55, 61, 13, 57, 62. The value 13 is an outlier.

(a) Work out the mean and the median of all 6 values (mean to 1 decimal place if needed).

(b) Work out the mean and the median without the outlier.

(c) Which average is changed less by removing the outlier?

(Total for Question 6 is 4 marks)

---

7 A sample of 6 measurements is 62, 59, 118, 58, 61, 65. The value 118 is an outlier.

(a) Work out the mean and the median of all 6 values (mean to 1 decimal place if needed).

(b) Work out the mean and the median without the outlier.

(c) Which average is changed less by removing the outlier?

(Total for Question 7 is 4 marks)

---

8 Here are the ages of club members in years recorded for 9 people. Valid values are whole numbers from 11 to 17 inclusive.

170, 13, 11, 999, 12, 16, 17, 17, 16

(a) Write down the invalid values.

(b) Work out the mean of all 9 values, to 1 decimal place if needed.

(c) Work out the mean after the invalid values are removed, to 1 decimal place if needed.

(Total for Question 8 is 5 marks)

© Revision Library | revisionlibrary.co.uk

9 A sample of 8 measurements is 24, 22, 16, 18, 22, 69, 16, 18. The value 69 is an outlier.

(a) Work out the mean and the median of all 8 values (mean to 1 decimal place if needed).

(b) Work out the mean and the median without the outlier.

(c) Which average is changed less by removing the outlier?

(Total for Question 9 is 4 marks)

---

10 A sample of 10 measurements is 42, 43, 46, 48, 91, 48, 45, 44, 43, 45. The value 91 is an outlier.

(a) Work out the mean and the median of all 10 values (mean to 1 decimal place if needed).

(b) Work out the mean and the median without the outlier.

(c) Which average is changed less by removing the outlier?

(Total for Question 10 is 4 marks)

---

11 Here are the ages of club members in years recorded for 8 people. Valid values are whole numbers from 11 to 17 inclusive.

15, 13, 12, 3, 17, 14, 17, 17

(a) Write down the invalid values.

(b) Work out the mean of all 8 values, to 1 decimal place if needed.

(c) Work out the mean after the invalid values are removed, to 1 decimal place if needed.

(Total for Question 11 is 5 marks)

© Revision Library | revisionlibrary.co.uk

12 A sample of 8 measurements is 41, 43, 38, 38, 42, 110, 46, 43. The value 110 is an outlier.

(a) Work out the mean and the median of all 8 values (mean to 1 decimal place if needed).

(b) Work out the mean and the median without the outlier.

(c) Which average is changed less by removing the outlier?

(Total for Question 12 is 4 marks)

---

13 Here are the number of hours of sleep per night recorded for 9 people. Valid values are whole numbers from 3 to 12 inclusive.

-1, 10, 12, 10, 3, 6, 3, -1, 12

(a) Write down the invalid values.

(b) Work out the mean of all 9 values, to 1 decimal place if needed.

(c) Work out the mean after the invalid values are removed, to 1 decimal place if needed.

(Total for Question 13 is 5 marks)

© Revision Library | revisionlibrary.co.uk

Mark scheme · 1.1 Data Cleaning and Outliers (Foundation): 1-hour Pack

Question 1

mean 43.9, median 38	**B1**
removes 96 and recalculates	**M1**
mean 38.1, median 38	**A1**
the median	**B1**

Answer: (a) mean 43.9, median 38; (b) mean 38.1, median 38; (c) median

Question 2

invalid values 999, 999	**B1**
sum of all values divided by 10	**M1**
206.4 cao	**A1**
sum of valid values divided by 8	**M1**
8.3 cao	**A1**

Answer: (a) 999, 999; (b) 206.4; (c) 8.3

Question 3

mean 26.8, median 23	**B1**
removes 58 and recalculates	**M1**
mean 23.3, median 23	**A1**
the median	**B1**

Answer: (a) mean 26.8, median 23; (b) mean 23.3, median 23; (c) median

Question 4

mean 25.4, median 20.5	**B1**
removes 64 and recalculates	**M1**
mean 19.9, median 19	**A1**
the median	**B1**

Answer: (a) mean 25.4, median 20.5; (b) mean 19.9, median 19; (c) median

Question 5

invalid values -1, 1	**B1**
sum of all values divided by 10	**M1**
5.8 cao	**A1**
sum of valid values divided by 8	**M1**
7.3 cao	**A1**

Answer: (a) -1, 1; (b) 5.8; (c) 7.3

Question 6

mean 51.2, median 58	**B1**
removes 13 and recalculates	**M1**
mean 58.8, median 59	**A1**
the median	**B1**

Answer: (a) mean 51.2, median 58; (b) mean 58.8, median 59; (c) median

Question 7

mean 70.5, median 61.5	**B1**
removes 118 and recalculates	**M1**
mean 61, median 61	**A1**
the median	**B1**

Answer: (a) mean 70.5, median 61.5; (b) mean 61, median 61; (c) median

Question 8

invalid values 170, 999	**B1**
sum of all values divided by 9	**M1**
141.2 cao	**A1**
sum of valid values divided by 7	**M1**
14.6 cao	**A1**

Answer: (a) 170, 999; (b) 141.2; (c) 14.6

Question 9

mean 25.6, median 20	**B1**
removes 69 and recalculates	**M1**
mean 19.4, median 18	**A1**
the median	**B1**

Answer: (a) mean 25.6, median 20; (b) mean 19.4, median 18; (c) median

Question 10

mean 49.5, median 45	**B1**
removes 91 and recalculates	**M1**
mean 44.9, median 45	**A1**
the median	**B1**

Answer: (a) mean 49.5, median 45; (b) mean 44.9, median 45; (c) median

Question 11

invalid values 3	**B1**
sum of all values divided by 8	**M1**
13.5 cao	**A1**
sum of valid values divided by 7	**M1**
15 cao	**A1**

Answer: (a) 3; (b) 13.5; (c) 15

Question 12

mean 50.1, median 42.5	**B1**
removes 110 and recalculates	**M1**
mean 41.6, median 42	**A1**
the median	**B1**

Answer: (a) mean 50.1, median 42.5; (b) mean 41.6, median 42; (c) median

Question 13

invalid values -1, -1	**B1**
sum of all values divided by 9	**M1**
6 cao	**A1**
sum of valid values divided by 7	**M1**
8 cao	**A1**

Answer: (a) -1, -1; (b) 6; (c) 8

© Revision Library | revisionlibrary.co.uk