

Welcome back MYP 9 Math!

	Assignment Effort Grade (Circle One)	Comments (What was interesting or challenging?)
Monday Date: <u>4-23</u> Topic: <u>Friday was Quiz 7.1</u>	0 1 2	
Tuesday Date: <u>4-24</u> Topic: <u>17A Comparing Data Sets</u>	0 1 2	
Wednesday Date: _____ Topic: _____	0 1 2	
Thursday Date: _____ Topic: _____	0 1 2	
Friday Date: _____ Topic: _____	0 1 2	

Warm-up: What do you notice? Wonder?

Days of the summer!

Ms. Paulson

Ms. Berg

	12	6	
+	4	11	237
	71	10	6788
	9876653320	9	0345579
	75321	8	22458
	20	7	056
	6	1	

Scale: 0/9/3
means
Ms. Paulson
90 min
and Ms. Berg
93 min



B

COMPARING DATA SETS GRAPHICALLY

Combining the two data sets together makes it easier to visually compare the distribution of each data

- a back-to-back stem-and-leaf plot.

	12	6
4	11	2 3 7
7 1	10	6 7 8 8
9 8 7 6 6 5 3 3 2 0	9	0 3 4 5 5 7 9
7 5 3 2 1	8	2 2 4 5 8
2 0	7	0 5 6
	6	1

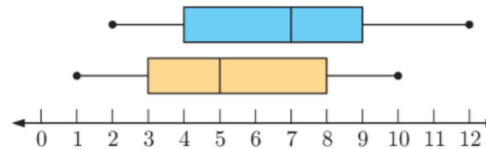
Scale: 0/9/3
means
Ms. Paulson
90 min
and Ms. Berg
93 min

C

PARALLEL BOX-AND-WHISKER PLOTS

In **Chapter 10**, we used a **box-and-whisker plot** to display some of the descriptive statistics of a data set.

If we want to compare two data sets, we can draw a box-and-whisker plot for each data set on the same scale. This is known as a **parallel box-and-whisker plot**.



When comparing data sets which are not of equal size, a parallel box-and-whisker plot is often more useful than the graphs we have studied so far.

a single set of data can be useful, but sometimes to fully understand a situation we need two data sets. We can do this by comparing statistics such as measures of **centre** and **spread** graphically.

Chapter 17

Comparing numerical data

Contents:

- A** Comparing the measures of centre and spread
- B** Comparing data sets graphically
- C** Parallel box-and-whisker plots

Read Scenario:

is a Mexican restaurant. He is about to launch a new advertising campaign to measure its effect. He records the number of customers in his restaurant before he starts advertising, and then for thirty days after he starts advertising. The following data:

Before advertising: 74 82 63 77 89 91 74 71 57 62 83 79 75 62 59
After advertising: 84 93 78 68 94 73 77 84 66 72 61 56 76 88 64
Before advertising: 93 98 108 84 89 76 92 69 93 85 101 78 67
After advertising: 88 70 93 104 108 95 80 103 79 92 98 84 106

Construct a back-to-back stem-and-leaf plot for this data.

Calculate the mean of each data set.

Interpret the effect of the advertising.

74 82 ~~63~~ 77 89 91 74 71 ~~57~~ ~~62~~ 83 79 75 ~~62~~
 84 93 78 ~~68~~ 94 73 77 84 ~~66~~ 72 ~~61~~ ~~56~~ 76 88
 93 98 108 84 89 ~~76~~ 92 ~~69~~ 93 85 101 ~~78~~
 88 ~~70~~ 93 104 108 95 80 103 ~~79~~ 92 98 84

a) Create back-to-back stem and leaf
 Number of customers in a restaurant

Before	Stem	After Ads
9, 7, 6	5	
8, 6, 4, 3, 2, 2, 1	6	7, 9
	7	0, 6, 8, 9
	8	
	9	
	10	

Done?
 mean & effect

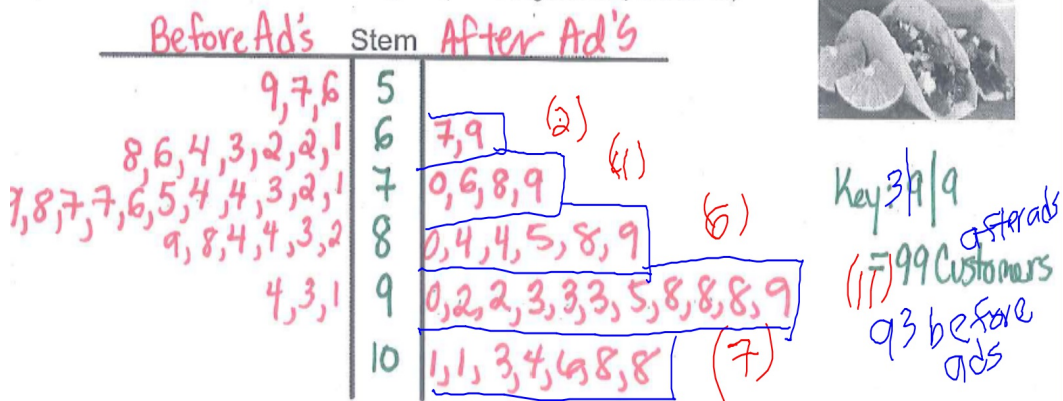
Key
 1 | 7 | 8
 71 customers before ADS.
 78 cust. after



Before advertising:	74 82 63 77 89 91 74 71 57 62 83 79 75 62 59
After advertising:	84 93 78 68 94 73 77 84 66 72 61 56 76 88 64
Before advertising:	93 98 108 84 89 76 92 69 93 85 101 78 67 99 90
After advertising:	88 70 93 104 108 95 80 103 79 92 98 84 106 98 101

a) Create back-to-back stem and leaf Number of customers in a restaurant

a) Construct back-to-back stem and leaf plot. (Don't forget the key and label!!)



Before	74	82	63	77	89	91	74	71	57	62	83	79	75	62	59
advertising:	84	93	78	68	94	73	77	84	66	72	61	56	76	88	64
After	93	98	108	84	89	76	92	69	93	85	101	78	67	99	90
advertising:	88	70	93	104	108	95	80	103	79	92	98	84	106	98	101

b) Find the mean of each data set

c) Discuss the effect of advertising



Before advertising:	74	82	63	77	89	91	74	71	57	62	83	79	75	62	59
After advertising:	84	93	78	68	94	73	77	84	66	72	61	56	76	88	64
Before advertising:	93	98	108	84	89	76	92	69	93	85	101	78	67	99	90
After advertising:	88	70	93	104	108	95	80	103	79	92	98	84	106	98	101

- b) Find the mean of each data set
- c) Discuss the effect of advertising

b) Find the mean of each data set:

Mean before Advertising: 74.4 customers Mean after Advertising: 90.7 customers

c) Discuss the effects of advertising

the number of customers increases after the new advertising campaigns. Personally, I am influenced by advertisements.



Before advertising		After advertising
9 7 6	5	
8 6 4 3 2 2 1	6	7 9
8 7 7 6 5 4 4 3 2 1	7	0 6 8 9
9 8 4 4 3 2	8	0 4 4 5 8 9
4 3 1	9	0 2 2 3 3 3 5 8 8 8 9
	10	1 1 3 4 6 8 8
		6 7 = 67 customers



Before advertising: $\text{mean} = \frac{74 + 82 + 63 + \dots + 88 + 64}{30} = 74.4$ customers

After advertising: $\text{mean} = \frac{93 + 98 + 108 + \dots + 98 + 101}{30} = 90.7$ customers

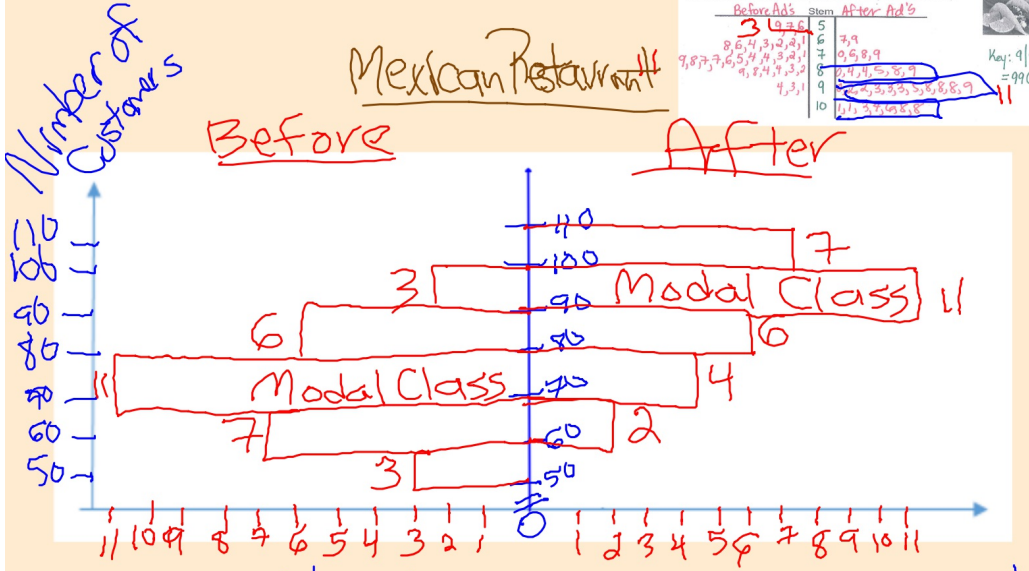
The number of customers has *increased* by approximately 16 customers per night as a result of the advertising.

d) Construct a Back-to-back Histogram

2) Construct back-to-back stem and leaf plot. (Don't forget the key and label!)

Before Ads	Stem	After Ads
3	5	3, 9
8, 6, 7, 3, 1, 2, 3, 1	6	9, 6, 8, 9
9, 8, 7, 7, 6, 5, 7, 4, 3, 2	7	11, 11, 5, 10, 9, 9
4, 2, 4, 4, 3, 2	8	11, 11, 10, 8, 7
4, 3, 1	9	
	10	

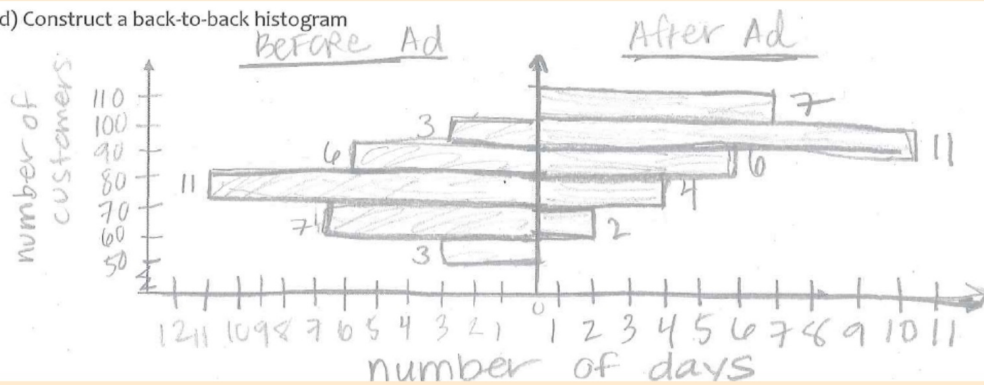
Key: 9|9 = 99 Customers



Number of days (Frequency)

d) Construct a Back-to-back Histogram

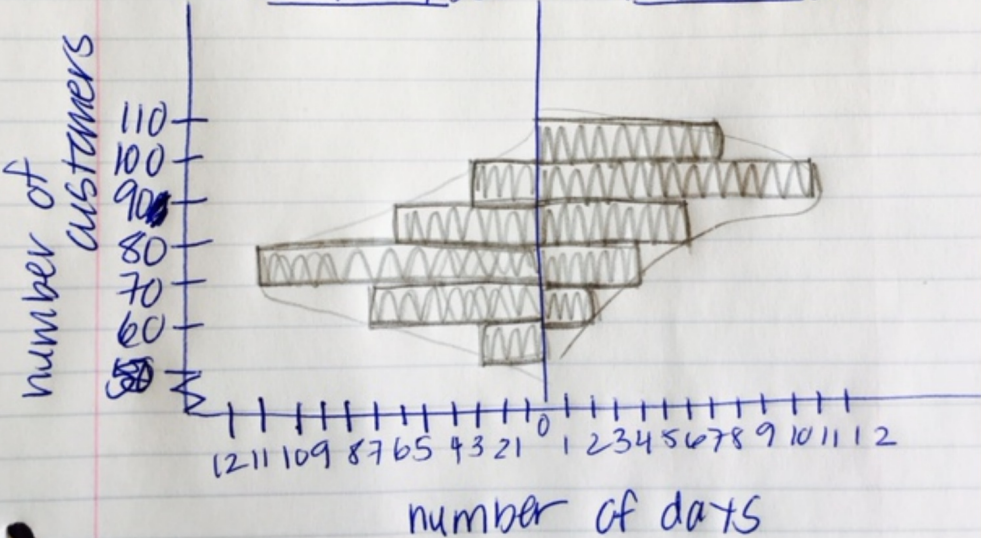
d) Construct a back-to-back histogram



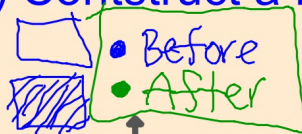
PEOPLE AT RESTAURANT

BEFORE ADS

AFTER ADS



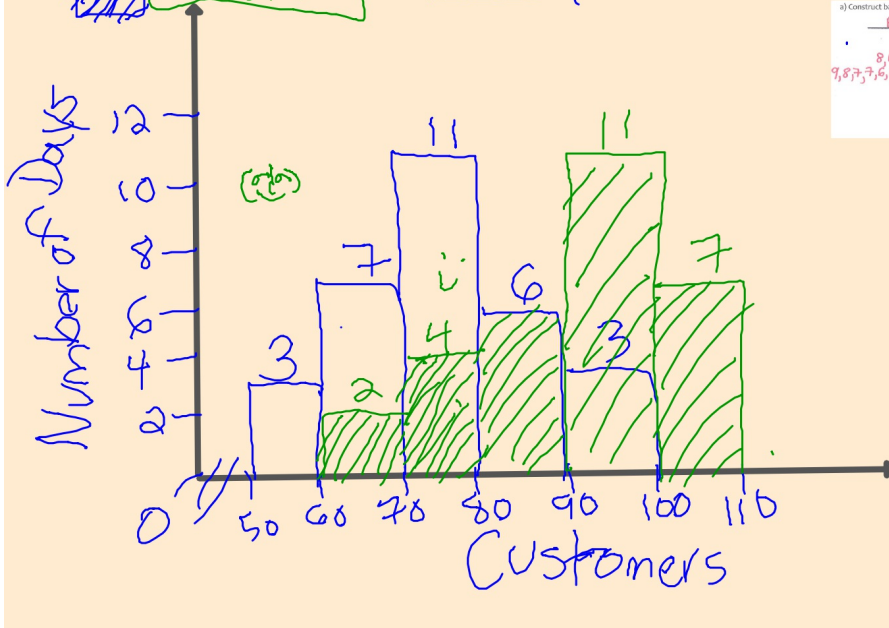
d) Construct a Back-to-back Histogram



Before | After Advertising

a) Construct back-to-back stem and leaf plot. (Don't forget the stem!)

Before Ad's	Stem	After Ad
9, 7, 6	5	7, 9
8, 6, 4, 3, 2, 2, 1	6	9, 6, 8, 9
7, 8, 7, 7, 6, 5, 4, 3, 2, 1	7	2, 4, 4, 5, 8
4, 3, 1	8	8, 2, 2, 3, 3, 2
	9	1, 1, 3, 4, 4, 1



Exercises...

EXERCISE 17B

#7

EXERCISE 17C

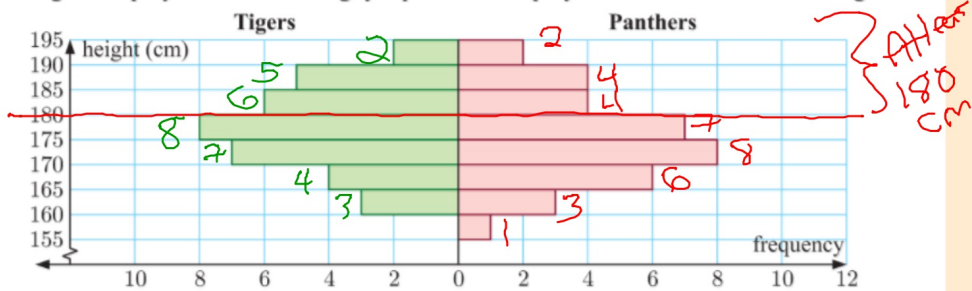
#3,

#4 (Also create a back-to-back histogram!)

Exercises...

EXERCISE 17B

- 7 The heights of players from two rugby squads are displayed on a back-to-back histogram.

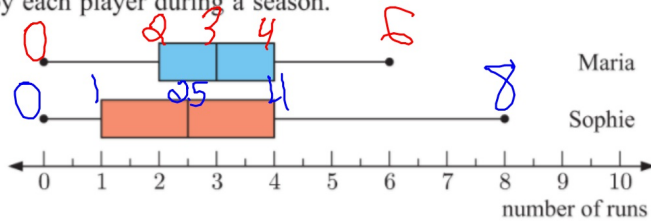


- How many players from each squad are at least 180 cm tall? $T = 13$ $P = 10$
- Find the modal class for each squad.
- Describe the distribution of each data set.
- Estimate the mean of each data set.
- Which team appears to have the greater average height? Discuss the reliability of your answer.

Exercises...

EXERCISE 17C

Maria and Sophie play in the same softball team. They are fierce but friendly rivals when it comes to scoring the most runs. The parallel box-and-whisker plot below shows the numbers of runs scored by each player during a season.



- Find the five-number summary for each player.
- For each player, find the:
 - range
 - interquartile range.
- Compare the performances of the players.



Exercises...

EXERCISE 17C

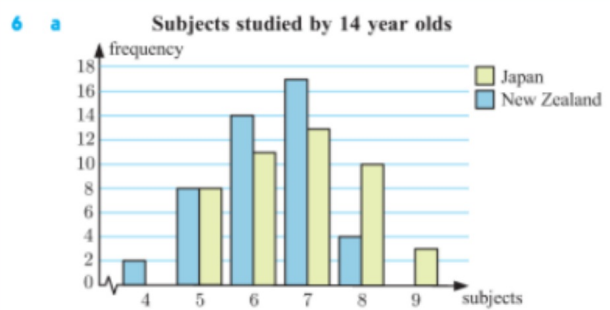
The data below shows the number of test matches played by players from the Zimbabwean Indian cricket teams.

Zimbabwe: 19 4 4 11 11 8 25 4 2 19 22 12 8
India: 81 19 6 6 90 23 6 4 17 3 20 53 9

- a Find the five-number summary for each data set.
- b Draw a parallel box-and-whisker plot for the data sets.
- c Which team's players have generally played more games?
- d Compare the spread of the data sets. **And back-to-back histogram!**



Solutions 17B



- b** New Zealand: mean ≈ 6.29 subjects, median = 6 subjects;
 Japan: mean ≈ 6.76 subjects, median = 7 subjects
- c** New Zealand: 4 subjects; Japan: 4 subjects
- d** "The selected students from *Japan* generally studied more subjects."
- 7 a** Tigers: 13 players; Panthers: 10 players
- b** Tigers: $175 \leq h < 180$ cm;
 Panthers: $170 \leq h < 175$ cm
- c** Tigers: approximately symmetric;
 Panthers: approximately symmetric
- d** Tigers: ≈ 177 cm; Panthers: ≈ 175 cm
- e** The Tigers appear to have a slightly greater average height. This conclusion may be unreliable as we can only estimate the means.

Solutions 17C

EXERCISE 17C

- 1 a i Set B, as the box-and-whisker plot is further to the right.
ii Set A, as the box-and-whisker plot covers a greater spread of values.
- b i Set B, as the box-and-whisker plot is further to the right.
ii Set B, as the box-and-whisker plot covers a greater spread of values.
- 2 a i 50% ii 75% b bus B

3 a

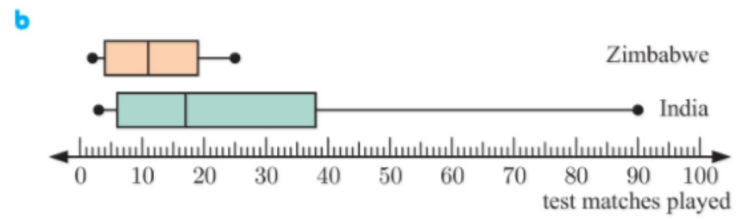
	Maria	Sophie
min.	0	0
Q ₁	2	1
med.	3	2.5
Q ₃	4	4
max.	6	8

- b Maria: range = 6 runs, IQR = 2 runs;
Sophie: range = 8 runs, IQR = 3 runs
- c Maria was more consistent (her scores had less spread) and had a higher median than Sophie. Sophie had a higher maximum but her scores were positively skewed.

Solutions 17C

4 a

	Zimbabwe	India
min.	2	3
Q ₁	4	6
med.	11	17
Q ₃	19	38
max.	25	90



- c India d India has a far greater spread than Zimbabwe.