

Welcome MYP 9 Mathematics!

	Assignment Effort Grade (Circle One)	Comments (What was interesting or challenging?)
Monday Date: <u>4/16</u> Topic: <u>NO SCHOOL</u>	0 1 2	
Tuesday Date: <u>4/17</u> Topic: <u>10FG Five Number Summary & Box Plots</u>	0 1 2	
Wednesday Date: _____ Topic: _____	0 1 2	
Thursday Date: _____ Topic: _____	0 1 2	
Friday Date: _____ Topic: _____	0 1 2	

Class Plan

1. Warm-up

2. Investigation

-How do we interpret the different measures of spread?

-How can we verify we have outliers?

F

MEASURING THE SPREAD OF A DATA SET

3. Practice

C ² O Fair	8:05-9:30
Lunch A Schedule	
1	9:35-10:12
2	10:17-10:54
3	10:59-11:36
LUNCH	11:41-12:11
4	12:16-12:53
5	12:58-1:35
6	1:40-2:17
7	2:22-3:00

C ² O Fair	8:05-9:30
Lunch B Schedule	
1	9:35-10:12
2	10:17-10:54
3	10:59-11:36
4	11:41-12:18
LUNCH	12:23-12:53
5	12:58-1:35
6	1:40-2:17
7	2:22-3:00

C ² O Fair	8:05-9:30
Lunch C Schedule	
1	9:35-10:12
2	10:17-10:54
3	10:59-11:36
4	11:41-12:18
5	12:23-1:00
LUNCH	1:05-1:35
6	1:40-2:17
7	2:22-3:00

Warm-up: Does this data set include an outlier? Why or why not?

Points Scored by Chicago Bulls
Players Who Played over 40 Games
(1997-98 Season)

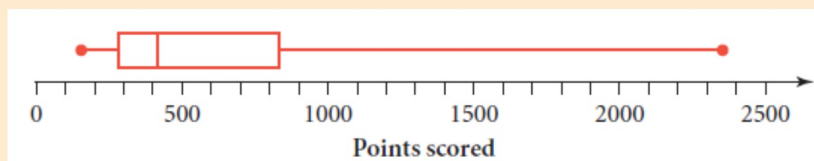
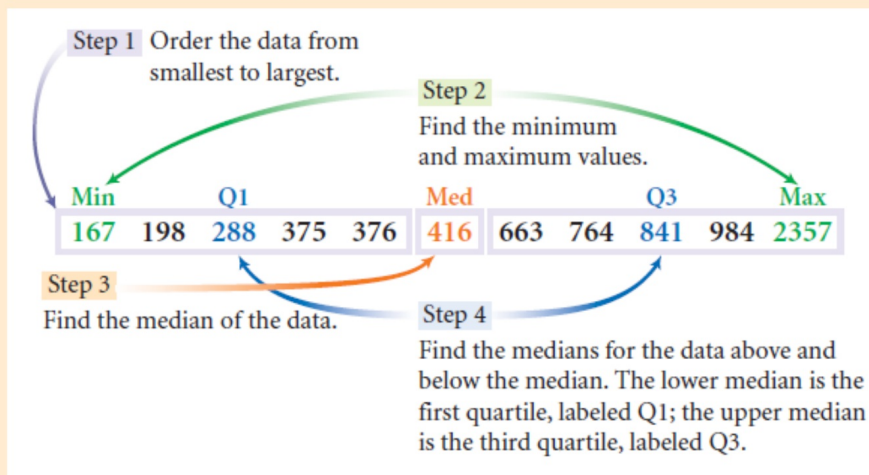
Chicago Bulls	Points
Michael Jordan	2357
Toni Kukoe	984
Scottie Pippen	841
Ron Harper	764
Lue Longley	633
Scott Burrell	416

Chicago Bulls	Points
Steve Kerr	376
Dennis Rodman	375
Randy Brown	288
Jud Bucchler	198

National Basketball Association (www.nba.com)



Chicago Bulls Data Solution



What is an outlier?

Is there a formula to determine if a data point is an outlier?

*****A data point that is distinctly separate from the rest of the data.**

***One definition of outlier is any data point more than 1.5 interquartile ranges (IQRs) below the first quartile or above the third quartile.

$$\text{Outlier} < Q_1 - 1.5IQR$$

$$\text{Outlier} > Q_3 + 1.5IQR$$

Calculating OUTLIERS

“1.5IQR above Q_3 or below Q_1 ”

IQR(Interquartile Range) = $Q_3 - Q_1$

Any point that falls outside the interval
calculated by

$Q_1 - 1.5(IQR)$ and $Q_3 + 1.5(IQR)$

is considered an outlier.

Min	Q1			Med	Q3			Max		
167	198	288	375	376	416	663	764	841	984	2357

IQR =

$$1.5(\underline{\hspace{2cm}}) = \underline{\hspace{2cm}}$$

Q1 - 1.5IQR =

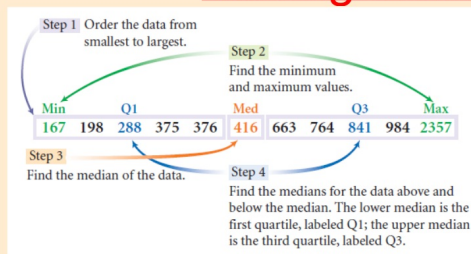
Q3 + 1.5IQR =

Did anyone score a number of points outside of -541.5 to 1670.5?

Michael Jordan	2357
----------------	------

Chicago Bulls Data Solution

Chicago Bulls Data Solution



$$IQR = 841 - 288 = 553$$

$$1.5(553) = 829.5$$

$$Q1 - 1.5IQR = 288 - 829.5 = -541.5$$

$$Q3 + 1.5IQR = 841 + 829.5 = 1670.5$$

**Michael Jordan is outside of the
-541.5 to 1670.5 outlier range.**

Michael Jordan	2357
-----------------------	-------------

Investigation: Minnesota Lynx Data

Minnesota Lynx Players and their minutes played (2016 season)



1. Five Number Summary
2. Outliers
3. Standard Deviation (Measure of Spread)



Interpret & Analyze Measures

Maya Moore	263
Lindsay Whalen	213
Sylvia Fowles	251
Seimone Augustus	215
Rebekkah Brunson	234
Natasha Howard	104
Jia Perkins	141
Renee Montgomery	101
Keisha Hampton	8
Anna Cruz	51
Janel McCarville	19

*****Come back together at end to go over solution & discuss standard deviation**

4. When done:
Help others!

Investigation: Minnesota Lynx Data

1) Calculate using the Lynx player's time.

Minimum 8 Maximum 263 Range 255

2) Analyze and interpret what the range says about the player's time on the basketball court.

Maya Moore	263
Lindsay Whalen	213
Sylvia Fowles	251
Seimone Augustus	215
Rebekkah Brunson	234
Natasha Howard	104
Jia Perkins	141
Renee Montgomery	101
Keisha Hampton	8
Anna Cruz	51
Janel McCarville	19

Investigation: Minnesota Lynx Data SOLUTION

Minnesota Lynx Players and their minutes played (2016 season)

Maya Moore	263
Lindsay Whalen	213
Sylvia Fowles	251
Seimone Augustus	215
Rebekkah Brunson	234
Natasha Howard	104
Jia Perkins	141
Renee Montgomery	101
Keisha Hampton	8
Anna Cruz	51
Janel McCarville	19

1) Calculate using the Lynx player's time.

Minimum 8 Maximum 263 Range 255

2) Analyze and interpret what the range says about the player's time on the basketball court.

There are 255 minutes between the players who have the most and least number of minutes played on the court.

Investigation: Minnesota Lynx Data

3) Identify any possible outliers and explain how they could affect the data.

```

1-Var Stats
x̄=145.4545455
Σx=1600
Σx²=322444
Sx=94.71891431
σx=90.31094128
n=11
    
```

```

1-Var Stats
n=11
minX=8
Q1=51
Med=141
Q3=234
maxX=263
    
```

Maya Moore	263
Lindsay Whalen	213
Sylvia Fowles	251
Seimone Augustus	215
Rebekkah Brunson	234
Natasha Howard	104
Jia Perkins	141
Renee Montgomery	101
Keisha Hampton	8
Anna Cruz	51
Janel McCarville	19

4) Calculate using the Lynx player's time.

Min 8 Q₁ 51 Median 141 Q₃ 234 Max 263 IQR = 183 min

Q₃ - Q₁ = IQR

5) Analyze and interpret what the IQR says about the player's time on the basketball court.

Investigation: Minnesota Lynx Data SOLUTION

3) Identify any possible outliers and explain how they could affect the data.

Maya Moore, Keisha Hampton, and Jane McCarvill could affect the average time playing in the season

4) Calculate using the Lynx player's time.

Min 8 Q_1 51 Median 141 Q_3 234 Max 263 IQR = $234 - 51 = 183$

5) Analyze and interpret what the **IQR** says about the player's time on the basketball court.

50% of the Lynx players have played between 51 minutes and 234 minutes in the season.
183 minutes of time is represented in this 50%.

Spread

Investigation: Minnesota Lynx Data

6) Verify that the possible outliers you identified in #3) *are in fact* outliers.

Outlier $< Q_1 - 1.5IQR$
Outlier $> Q_3 + 1.5IQR$

$$IQR = 183 \quad 1.5(183) = 274.5$$

$$Q_1 = 51 \text{ min}$$

$$Q_1 - 1.5IQR$$

$$51 - 274.5$$

$$= -223.5 \text{ min}$$

Below? No
(lower)

$$Q_3 = 234$$

$$234 + 274.5$$

$$= 508.5 \text{ min}$$

Anyone above??

No outliers

Investigation: Minnesota Lynx Data SOLUTION

6) Verify that the possible outliers you identified in #3) *are in fact* outliers.

$$1.5IQR = 1.5(183) = 274.5$$

$$\text{Outlier} < Q_1 - 1.5IQR$$

$$\text{Outlier} > Q_3 + 1.5IQR$$

$$Q_1 - 1.5IQR$$

$$51 - 274.5$$

$$\underline{-223.5}$$

No players below -223.5

$$Q_3 + 1.5IQR$$

$$234 + 274.5$$

$$\underline{508.5}$$

No players above 508.5 minutes

Conclusion

No outliers of time played

Investigation: Minnesota Lynx Data

Finding the Standard Deviation

What is a standard deviation?

Deviation just means how far from the normal

Standard Deviation

The Standard Deviation is a measure of how spread out numbers are.

Its symbol is σ (the greek letter sigma)

Investigation: Minnesota Lynx Data SOLUTION

Finding the Standard Deviation

What is a standard deviation?

A measure of how far the data points are
on average from the mean. σ : Sigma

Investigation: Minnesota Lynx Data

How can we find the average distance from the mean ("norm")?

Challenge Yourself.

Find the algorithm for standard deviation by hand.

(average distance from the mean...)

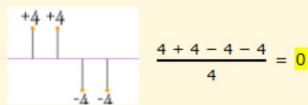
Challenge Yourself.

Algorithm for standard deviation.

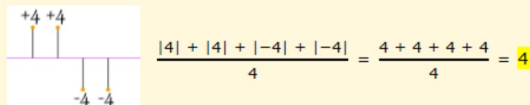
(average distance from the mean...)

*Footnote: Why square the differences?

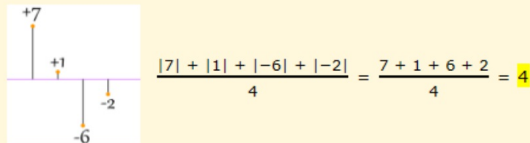
If we just add up the differences from the mean ... the negatives cancel the positives:



So that won't work. How about we use [absolute values](#)?



That looks good (and is the [Mean Deviation](#)), but what about this case:



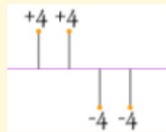
Oh No! It also gives a value of 4, Even though the differences are more spread out.

Challenge Yourself.

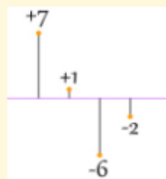
Algorithm for standard deviation.

(average distance from the mean...)

So let us try squaring each difference (and taking the square root at the end):



$$\sqrt{\left(\frac{4^2 + 4^2 + 4^2 + 4^2}{4}\right)} = \sqrt{\left(\frac{64}{4}\right)} = 4$$



$$\sqrt{\left(\frac{7^2 + 1^2 + 6^2 + 2^2}{4}\right)} = \sqrt{\left(\frac{90}{4}\right)} = 4.74\dots$$

That is nice! The Standard Deviation is bigger when the differences are more spread out ... just what we want.

In fact this method is a similar idea to [distance between points](#), just applied in a different way.

And it is easier to use algebra on squares and square roots than absolute values, which makes the standard deviation easy to use in other areas of mathematics.

Standard deviation by hand.

(average distance from the mean...)

Test Scores	Mean \bar{x}	$x_i - \bar{x}$	$(x_i - \bar{x})^2$
72	69	3	9
55	69	-14	196
53		-16	256
51		-18	324
74		5	25
71		2	4
83		14	196
95		26	676
67		-2	4
		Sum	

Subtract

$$n = 9$$

$$\bar{x} = 69$$

$$\frac{\text{sum}}{n} =$$

$$\sigma = \sqrt{\frac{\text{sum}}{n}} =$$

Standard deviation =

Standard deviation by hand. (average distance from the mean...)

Test Scores	Mean \bar{x}	$x_i - \bar{x}$ Sum=0	$(x_i - \bar{x})^2$
72	69	3	9
55	69	-14	196
53	69	-16	256
51	69	-18	324
74	69	5	25
71	69	2	4
83	69	14	196
95	69	26	676
67	69	-2	4
		Sum	1690

$$n = 9$$

$$\bar{x} = 69$$

$$\frac{\text{sum}}{n} = \frac{1690}{9}$$

$$\sigma = \sqrt{\frac{\text{sum}}{n}} = \sqrt{187.8}$$

Standard deviation = 13.7

Standard deviation by calculator:

```

x̄=69
Σx=621
Σx²=44539
sx=14.53444185
σx=13.70320319
    
```

Investigation: Minnesota Lynx Data

7) Calculate using the Lynx player's time. Mean = 145.5 min

8) Use the steps below to find the standard deviation of the player's time using a calculator. (TI-83/84)

i. **STAT, 1:EDIT**, then enter the data into a list.

ii. **STAT, CALC, 1:1-var stats, ENTER**

You should see a picture like:

Find $\sigma =$ _____ 90.3

9) Interpret what this measure of spread means.

```
1-Var Stats
x̄=145.4545455
Σx=1600
Σx²=322444
Sx=94.71091431
σx=90.31094128
n=11
```


Investigation: Minnesota Lynx Data SOLUTION

7) Calculate using the Lynx player's time. Mean = 145.45 minutes

8) Use the steps below to find the standard deviation of the player's time using a calculator. (TI-83/84)

i. **STAT, 1:EDIT**, then enter the data into a list.

ii. **STAT, CALC, 1:1-var stats, ENTER**

You should see a picture like:

Find $\sigma =$ 90.3 minutes

```
1-Var Stats
x̄=145.4545455
Σx=1600
Σx²=322444
Sx=94.71891431
σx=90.31094128
↓n=11
```

9) Interpret what this measure of spread means.

On average, each player is about 90 minutes from the average time of 145.45 minutes.

Investigation: Minnesota Lynx Data

10) Suppose an all-star team is compiled of players that are within 1 standard deviation **above** the mean. Which Lynx players would qualify for the all-star team?

$$1\sigma \approx 90 \text{ min} \quad \bar{x} \approx 145 \text{ min}$$

$$> 1\sigma + \bar{x}$$

$$90 + 145 = 235$$

Who is on the court more than 235 min? Moore, Fowels

Investigation: Minnesota Lynx Data

SOLUTION

10) Suppose an all-star team is compiled of players that are within 1 standard deviation **above** the mean. Which Lynx players would qualify for the all-star team?

$1 \text{ sd} = 90.3$ $\text{mean} + 1\sigma = 145.45 + 90.3 \approx 235.75$
Sylvia Fowles and Maya Moore have minutes over 235.75.

Exercises...

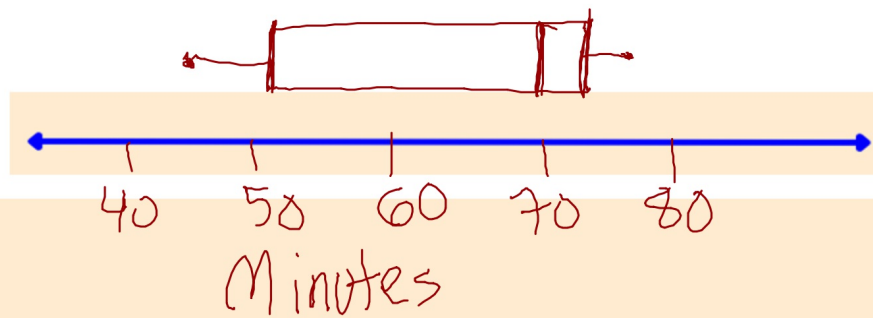
Music Playlist Lengths

1. The lengths in minutes of nine Spotify playlists are 45, 63, 74, 69, 72, 53, 72, 73, and 50.

a) Find the 5 Number Summary

Order the data:

Minimum: 45 Q1: 51.5 Median: 69 Q3: 72.5 Maximum: 74



Exercises...

b) Calculate and interpret the IQR. Use the IQR to verify that there are **no outliers**.

$$\text{Outlier} < Q_1 - 1.5IQR$$

$$\text{Outlier} > Q_3 + 1.5IQR$$

Exercises...

c) The standard deviation of the playlists is 10.6 minutes. ($\sigma \approx 10.6$ minutes).

What does this mean?

d) Suppose playlists that are 1 standard deviation **above** the mean require too much memory on our phones. Which playlist(s) would be too **big** for our phone?