

Name _____

Period _____

BASEBALL SALARIES: DO YOU GET WHAT YOU PAY FOR?

The American League East has had three excellent teams in the years 2008 – 2010. Here are the combined standings for those three years:

<u>Team</u>	<u>Wins</u>	<u>Losses</u>	<u>Percentage</u>
New York Yankees	287	199	.591
Boston Red Sox	279	207	.574
Tampa Bay Rays	277	209	.570

Is there a great difference between these teams in what they pay their players? Or do they all pay their players about the same? Is the saying true “You get what you pay for”? Below are the team salaries (in thousands) for the 2010 season:

<u>New York Yankees</u>					
Rodriguez	\$32,000	Granderson	\$8,250	Colon	\$900
Sabathia	24,286	Marte	4,000	Ayala	650
Texeira	23,125	Martin	4,000	Gardner	530
Burnett	16,500	Feliciano	3,750	Robertson	460
Rivera	14,912	Hughes	2,700	Cervelli	456
Jeter	14,729	E. Chavez	1,500	G. Molina	455
Posada	13,100	F. Garcia	1,500	Nova	433
Cano	10,000	A. Jones	1,500	Curtis	420
Soriano	10,000	Chamberlain	1,400	Nunez	419
Swisher	9,100	Logan	1,200	Corona	414

<u>Boston Red Sox</u>					
Beckett	\$17,000	A. Gonzalez	\$6,300	Tazawa	\$1,500
Lackey	15,950	Jenks	6,000	Reyes	900
Crawford	14,857	Lester	5,750	Albers	875
Drew	14,000	Pedroia	5,750	Saltalamacchia	750
Ortiz	12,500	Scutaro	5,500	Buchholz	555
Youklis	12,250	Wheeler	3,000	Bard	505
Papelbon	12,000	Ellsbury	2,400	McDonald	470
Matsuzaka	10,333	Varitek	2,000	Lowrie	450
Cameron	7,750	Wakefield	2,000	Doubront	417

<u>Tampa Bay Rays</u>					
Damon	\$5,250	Howell	\$1,100	Joyce	\$427
Upton	4,825	D. Johnson	1,000	Brignac	425
Zobrist	4,687	Peralta	925	Russell	421
Shields	4,250	Sonnanstine	912	Hellickson	418
Shoppach	3,000	Niemann	903	Fuld	418
Farnsworth	2,600	J.Cruz	850	C. Ramos	417
Price	2,085	Davis	434	McGee	415
M. Ramirez	2,020	S. Rodriguez	429	E. Johnson	415
Longoria	2,000	J. Jaso	427		

A great way to compare and contrast two or more distributions is by parallel box plots.

How to Create Parallel Boxplots on the TI-83/84 Calculators

1) Clear lists L1, L2, and L3.

Hit ON. Hit STAT. Hit ENTER. Arrow up, so your cursor is on L1. Hit CLEAR. Hit ENTER. Repeat for lists L2 and L3.

2) Input the Yankee salaries into list L1, the Red Sox salaries into list L2, and the Tampa Bay Ray salaries into list L3.

Hit STAT. Hit ENTER. The cursor is at the top of list L1. Input Alex Rodriguez' salary: 32000. This means \$32 million. Hit ENTER. Input 24286. Hit ENTER. Continue in list L1 all the way down to 414, Corona's salary and the 30th list entry. Arrow RIGHT up to the top of list L2. Input the Red Sox salaries starting with 17000, and ending with 417, the 27th list entry. Finally, input the Tampa Bay Ray salaries into list L3.

3) Display the salary distributions as three separate parallel boxplots.

Hit 2nd QUIT. Hit 2nd STAT PLOT. (It's on the Y= key.) Hit ENTER. Turn Plot1 ON. For TYPE, highlight the fourth icon: the boxplot showing outliers. Hit ENTER. For Xlist: Input L1. For Freq: Input 1. For Mark: Choose the square. Then arrow up and perform similar commands so that Plot 2 displays list L2 and Plot 3 displays list L3.

4) Prepare to display.

Hit 2nd QUIT. Hit Y=. Clear any functions by hitting CLEAR. Hit 2nd FORMAT. Arrow down and right to AxesOff. Hit ENTER. Hit ZOOM 9.

5) Explore the three box-and-whisker plots.

Hit GRAPH. Hit TRACE. Arrow RIGHT. Arrow LEFT. Arrow DOWN. Arrow UP, etc.

Assignment: 1) Fill in the following summary:

<u>Team</u>	<u>Minimum</u>	<u>Lower Quartile</u>	<u>Median</u>	<u>Upper Quartile</u>	<u>Maximum</u>
Yankees	_____	_____	_____	_____	\$32 million
Red Sox	_____	_____	_____	_____	_____
Rays	_____	_____	_____	_____	_____

2) Describe the shape of these distributions. Compare the lengths of the two boxes in each; compare the lengths of the two whiskers in each. Are these distributions somewhat symmetric, skewed left or skewed right? _____

3) Compare the centers of the three salary data sets by using the median.

4) Why is the median a better measure of center to use than the mean for this comparison?

5) The median salaries from largest to smallest are Red Sox, Yankees, and Rays. Does this imply that the total salaries are in that order? _____ What order of total salaries is suggested by the parallel box plots? _____

6) Which of the three total salaries is much different than the other two? _____

7) Between what Yankee or Red Sox quartiles do we find the highest paid Ray?

8) Define and give the formula for the interquartile range IQR.

9) Calculate the IQR for each data set. Compare the variation of the three data sets using the IQR. Why is the IQR a better measure of variation to use than the standard deviation for this comparison?

10) An upper outlier is defined as an observation x such that $x > Q3 + 1.5 * IQR$. Find the three minimum x values for each team:

Yankees _____

Red Sox _____

Rays _____

11) Look at the data. Which salaries on the three teams are outliers compared to the rest of their team? Show how these salaries exceed the three minima given above.

12) A lower outlier is defined as an observation x such that $x < Q1 - 1.5 * IQR$. Show that none of these teams has lower outliers by finding the maximum x value for each:

Yankees _____

Red Sox _____

Rays _____

13 – 15: We use the median and the IQR as our center and measure of variation, respectively, because they are *resistant* statistics: these are statistics that do not overreact to the presence or absence of outliers. By contrast, the mean and standard deviation are not resistant. Let's observe this contrast using the Rays' salaries.

Their highest paid player, Johnny Damon, has an outlier salary of \$5.25 million. We will observe how the resistant and non-resistant statistics change as we make summary calculations first including and then excluding his outlier salary:

How to Calculate One-Variable Statistics

Hit 2nd QUIT. Hit STAT. Arrow right to CALC. You're on 1-Var Stats. Hit ENTER. Input 2nd L3. (This is where the Ray salaries are listed.) Hit ENTER. Fill in the row "With Johnny Damon."

Hit STAT. Hit ENTER. Arrow to the top so the cursor is on 5250. Hit DEL. Hit STAT. Arrow right to CALC. You're on 1-Var Stats. Hit ENTER. Input 2nd L3. Hit ENTER. Fill in the row "Without Johnny Damon."

	Mean	Median	σ	Q3	Q1	IQR
With Johnny Damon (5250)						
Without Johnny Damon						
Raw Difference						
Percent Decrease						

13) Compare the percent decrease in the mean and median. Which one had the smaller decrease?

14) Compare the percent decrease in the population standard deviation σ and the IQR. Which one had the smaller decrease? _____

15) Because of the small differences in decrease, the resistant measures of center and variation must be the _____ (center) and _____ (variation). Conversely, the non-resistant measures are the _____ (center) and the _____ (variation).

16) Consider the results over the last three years for all three teams. Having displayed and analyzed their different salary distributions, would you conclude that these teams have done equally well in getting what they paid for? Explain. _____

