# Baseball and Normal Distribution

**Strand:**  Data Analysis

**Topic:**  Normal Distribution

**Primary SOL:** AFDA.7 The student will

a) identify and describe properties of a normal distribution;

b) interpret and compare z-scores for normally distributed data; and

c) apply properties of normal distributions to determine probabilities associated with areas under the standard normal curve.

**Related SOL:** AFDA.6

## Materials

* Baseball Payroll activity sheet (attached)
* Graphing utility or spreadsheet software

## Vocabulary

*curve, standard deviation, mean, normal distribution, percentile, probability, z-score*

## Student/Teacher Actions

*Time: 90 minutes*

1. Begin by displaying examples of normal distribution curves (such as heights of adults, IQ scores, SAT scores).
2. Distribute the Baseball Payroll activity sheet. Group students into pairs or fours while completing the activity. Students will need to use a graphing utility or spreadsheet software to compute z-scores, mean, and standard deviation.
3. After students complete the z-score chart, discuss percentages between the standard deviation ranges before having students complete questions 6 through 10 on the activity sheet.

## Assessment

### Questions

* + What does a z-score with a large positive value, large negative value, or zero describe for a data point?
	+ What do we know about a data set that has a small standard deviation compared with a data set with a large standard deviation?

### Journal/writing prompts

* + Describe how a data set with a large mean but small standard deviation would compare with a data set with a small mean but large standard deviation. Describe with the context of data sets of income.
	+ Describe how a data set of incomes of people in the same career would compare with a data set of incomes of people in different careers as it relates to mean and standard deviation.

### Other Assessments

* + Give students multiple graphs of standard distributions and have them determine which sets have larger standard deviations, larger means, etc.
	+ Create an exit slip on which students need to calculate the mean, standard deviation, and z-scores for a data set including 10 or fewer data points.

## Extensions and Connections

* Have students research general admission ticket prices for the different Major League Baseball (MLB) teams and compare those to the teams’ payroll using z-scores.
* Have students compare the mean, standard deviation, and z-scores between other professional sports leagues. Have students compare why some sports may have higher total payrolls and how we could equally compare the leagues.

## Strategies for Differentiation

* Split students into the baseball leagues (National and American) and have them find the mean, standard deviation, and z-scores for one league. Have students compare leagues as a whole class.
* Use vocabulary cards for related vocabulary listed above.
* Have students create histograms of the team payrolls to better visualize the differences in salary.

**Note: The following pages are intended for classroom use for students as a visual aid to learning.**

**Baseball Payroll**

Using the 2017 total team payroll and records for each team in Major League Baseball (MLB) on the table on the next page:

1. Use a spreadsheet or graphing utility to find the mean and standard deviation for the payroll data, as well as the z-score for each team.
2. Are there any teams whose payroll is more than two standard deviations away from the mean?
3. Some people say that it is not fair for the Yankees to pay so much and the Rays to pay so little in salaries. Do you think it is fair? If you do think it is fair, explain why. If not, how could you make it fairer? What could Major League Baseball do to try to maintain equity in competition?
4. Find the mean and standard deviation for the number of wins data and z-score for each team. How do the z-scores from that data set compare to the payroll z-scores? Does a team that pays more in salary lead to more wins in a season? Explain your reasoning.
5. Look at the data and determine which team seemed to get the best return on their payroll investment. Justify your answer.

**Baseball 2017 Payroll and Wins Data Table**

|  |
| --- |
| Data from [Spotrac.com](http://www.spotrac.com/mlb/payroll/) and [Baseball-Reference.com](https://www.baseball-reference.com/)  |
| **Team** | **Payroll** | **z-score** | **Wins** | **z-score** |
| New York Yankees | $224,458,752  |   | 84 |   |
| Boston Red Sox | $222,552,008  |   | 93 |   |
| New York Mets | $154,829,658  |   | 87 |   |
| Los Angeles Angels | $176,337,209  |   | 74 |   |
| Chicago White Sox | $99,889,688  |   | 78 |   |
| Los Angeles Dodgers | $265,149,292  |   | 91 |   |
| Seattle Mariners | $171,255,830  |   | 86 |   |
| Chicago Cubs | $182,400,336  |   | 103 |   |
| Detroit Tigers | $198,716,188  |   | 86 |   |
| Baltimore Orioles | $182,523,427  |   | 89 |   |
| St. Louis Cardinals | $149,454,185  |   | 86 |   |
| San Francisco Giants | $191,065,209  |   | 87 |   |
| Philadelphia Phillies | $116,874,208  |   | 71 |   |
| Houston Astros | $149,964,163  |   | 84 |   |
| Atlanta Braves | $115,455,675  |   | 68 |   |
| Toronto Blue Jays | $199,430,487  |   | 89 |   |
| Oakland Athletics | $85,977,680  |   | 69 |   |
| Minnesota Twins | $138,629,177  |   | 59 |   |
| Milwaukee Brewers | $83,488,679  |   | 73 |   |
| Cincinnati Reds | $115,323,803  |   | 68 |   |
| Texas Rangers | $185,899,040  |   | 95 |   |
| Kansas City Royals | $158,275,155  |   | 81 |   |
| Cleveland Indians | $139,165,884  |   | 94 |   |
| San Diego Padres | $91,963,878  |   | 68 |   |
| Colorado Rockies | $146,651,941  |   | 75 |   |
| Arizona Diamondbacks | $119,898,775  |   | 69 |   |
| Pittsburgh Pirates | $109,840,330  |   | 78 |   |
| Washington Nationals | $189,292,654  |   | 95 |   |
| Miami Marlins | $117,557,599  |   | 79 |   |
| Tampa Bay Rays | $92,491,605  |   | 68 |   |
|  | Mean | Mean |
| Standard Deviation  | Standard Deviation |

6. What happens to the z-score of other teams if the New York Yankees are removed from the data set for payroll? Additionally, what happens when the Boston Red Sox are removed?

7. Using the z-scores obtained in this activity, what do you notice about the z-scores for payroll and wins?

8. What percentage of teams falls below the mean total team payroll? What percentage falls above?

9. Calculate the team payroll that would equate to a z-score of one and minus one. What percentage of teams would fall between these values?

10. What is the probability that a team chosen randomly would have a total team payroll greater than one standard deviation above the mean?