

1. The data below show recent salaries of Major League Baseball shortstops.

Jose Reyes	22000000	Starlin Castro	6857143	Brendan Ryan	2000000	Nick Franklin	510200	Jose Ramirez	507500
Troy Tulowitzki	20000000	Yunel Escobar	5000000	Ruben Tejada	1880000	Brad Miller	510100	Tyler Ladendorf	507500
Elvis Andrus	15250000	Mike Aviles	3500000	Jose Iglesias	1650000	Gregorio Petit	510000	Marcus Semien	507500
Jhonny Peralta	15000000	Cliff Pennington	3275000	Clint Barmes	1300000	Nick Ahmed	508500	Christian Colon	507500
J.J. Hardy	11500000	Brandon Crawford	3175000	Alexi Amarista	1150000	Tim Beckham	508100	Eduardo Escobar	507500
Jimmy Rollins	11000000	Andrelton Simmons	3142857	Marwin Gonzalez	1062500	Didi Gregorius	507500	Freddy Galvis	507500
Ian Desmond	11000000	Alcides Escobar	3000000	Jean Segura	534000	Andrew Romine	507500	Wilmer Flores	507500
Alexei Ramirez	10000000	Jung Ho Kang	2500000	Pete Kozma	518000	Chris Taylor	507500	Danny Santana	507500
Erick Aybar	8750000	Everth Cabrera	2400000	Xander Bogaerts	517000	Chris Owings	507500	Adeiny Hechavarria	507500
Jed Lowrie	8000000	Zack Cozart	2350000	Jordy Mercer	515500	Herman Perez	507500	Jonathan Villar	507500

- Compute the mean, median, and mode for the data
- Choose one measure of central tendency as the “best” measure of central tendency and explain why it is best.
- What salary would represent the 20th percentile?
- The standard deviation of the data is \$5,350,277. What portion of the shortstops earn within two standard deviations of the mean?
- What is the z-score for Tim Backham’s salary?
- What is the z-score for a player earning the minimum salary?
- Organize the data into classes and draw a bar graph with ten bars.

2. A class of students were asked, *For how many years did you play organized sports?* The data had a mean of 9.04 and standard deviation of 6.04 years. In the Fall 2015 term, the mean and standard deviation were 6.93 and 5.93 respectively. Make two conclusions based upon the change in each of these statistics (one conclusion about the dataset for mean, one for standard deviation; do not try to explain the changes).
3. Which of the following statements is more accurate? Explain your reasoning.
- a. *If \$100K is the 80th percentile for salaries in a dataset, then \$50K is the 40th percentile.*
 - b. *The percentile ranks for SAT math scores of 650 and 530 were 90 and 60, respectively. Therefore a score of 590 is the 75th percentile.*
4. Using the data from 15 automobile accidents, the correlation coefficient between the combined speeds of the cars (x) in an accident and the amount of damage done (y) is computed to be a strong positive value. The regression equation for the two variables is $y = 801.518 + 162.845x$.
- a. Predict the amount of damage done in an accident in which the combined speeds of the car involved was 100 mph.
 - b. A car traveling at 28 mph is involved in an accident with another car which did \$9,780 in damages to the cars. Estimate how fast the other car was driving.