

1. Find the **mean, mode, median** and **range** for the following data sets and compare each measurement

- (1) Inauguration ages of U.S. presidents:

57 61 57 57 58 57 61 54 68 51 49 64 50 48 65 52 56
 46 54 49 51 47 55 55 54 42 51 56 55 51 54 51 60 62
 43 55 56 61 52 69 64 46 54

- (2) Scores when a die is thrown 40 times:

2 4 5 5 1 3 4 6 2 5 2 4 6 1 2 5 4 4 1 1
 3 4 6 5 5 2 3 3 1 6 5 4 2 1 3 3 2 1 6 6

2. Ten teenagers in a school were given a general knowledge test on AIDS and another ten a test on drug-related behaviour. The scores on the two tests were as follows:

Knowledge of AIDS: 8 1 8 6 12 9 6 5 11 13
 Knowledge of drugs: 91 42 98 30 18 73 84 92 45 92

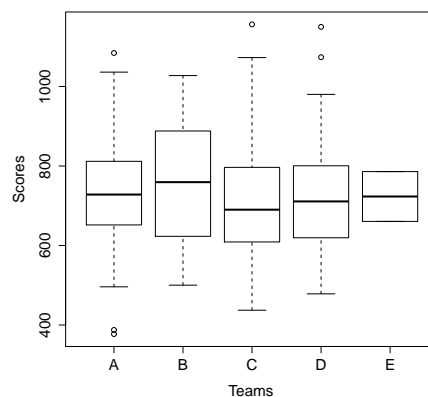
- (a) Calculate the mean score of the group on each test.
 (b) Calculate the variance and standard deviation of the two tests.
 (c) On which test is there more variability?
 (d) Find the median for each test.
 (e) For each test say whether it is positively skewed, negatively skewed, or symmetrical.
3. Prove that

$$r = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{(n-1)s_x s_y} = \frac{\sum_{i=1}^n x_i y_i - n\bar{x}\bar{y}}{\sqrt{(x_i^2 - n\bar{x}^2)(y_i^2 - n\bar{y}^2)}}$$

where \bar{xy} is the sample mean of data points $x_i y_i$ and, s_x and s_y are the sample standard deviations of datasets x and y respectively.

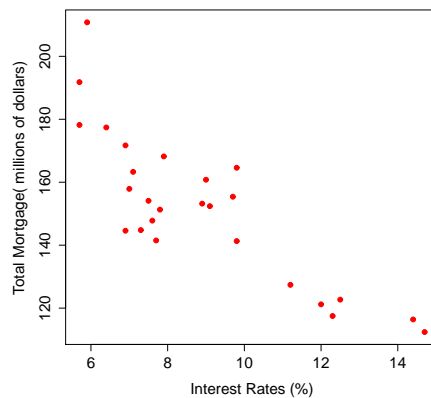
Hint: $\sqrt{ab} = \sqrt{a}\sqrt{b}$, $\sqrt{\frac{a}{b}} = \frac{\sqrt{a}}{\sqrt{b}}$, and $\sqrt{a^2} = a$, for numbers a and $b \neq 0$.

4. Look at the boxplots of scores for five teams in a game



- (a) Overall, which team you think performs better in the game? Why?
 (b) How you would describe the shape of the distribution for team D? and team E?

- (c) what would be an “atypical” score for group A?. Give approximate values.
- (d) What team shows more variability in its scores?
- (e) What team shows the largest IQR?.
5. Since 1980, average mortgage interest rates have fluctuated from low of under 6% to a higher of over 14 %. Is there a relationship between the amount of money people borrow and the interest rate that’s offered?. Here is a scatterplot of Total Mortgages in USA (in million of dollars) versus Interest rates (in percentage) at various times over the last 26 years. The correlation is $r = -0.84$



- (a) Describe the relationship between *Total Mortgages* and *Interest Rates*.
- (b) If we were to measure *Total Mortgages* in thousands of dollars instead of millions of dollars, how would the correlation coefficient change?
- (c) Suppose in another year, interest rates were 11%, and mortgages totaled \$250 million. How would including that year with these data affect the correlation coefficient?
- (d) Do these data provide proof that if mortgage rates are lowered, people will take out more mortgages? Explain.

Hint: Recall how to construct the correlation coefficient.