

8). Mean - if no outlier

Median = if there's an outlier

10-2 Spread of Data

Objectives:

- I can find the standard deviation and ~~variance~~ of a data set
- I can find the 5 number summary of a data set and make a box and whisker plot
- I can determine the percent of data within a given interval

Spread: When we use the mean to measure center, we use standard deviation to show how the data is spread out from the mean.

Mean: \bar{X}

Standard Deviation: σ

~~Variance: σ^2~~

Find the standard deviation: Weights in grams of 30 loon chicks

79.5 87.5 88.5 89.2 91.6 84.5 82.1 82.3 85.1 89.8
84.0 84.8 88.2 88.2 82.9 89.8 89.2 94.1 88.0 91.1
91.8 87.0 87.7 88.0 85.4 94.4 91.3 86.3 85.7 86.0

Test scores from a class: 70, 70, 75, 75, 90, 70, 80, 85, 65, 95, 70, 85, 90, 70, 20

Mean: $(\bar{x}) = 74$

Standard Deviation: $(\sigma_x) = 17.0$

~~Variance:~~

Determine the whole number of standard deviations for all data values. The mean price for an iPhone is \$687 with a standard deviation of \$121.

~~\$400, \$601, \$527, \$803, \$410, \$720, \$800~~

566, 929, 324, 445, 1050, 808

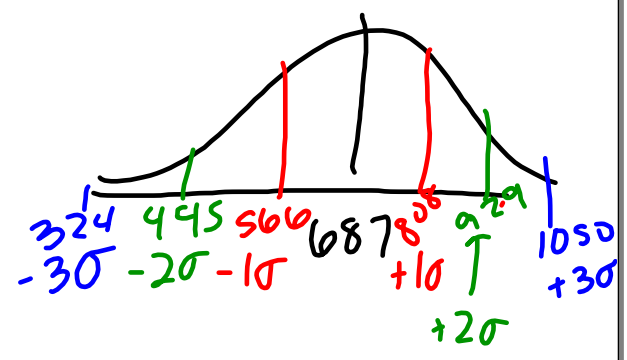
- \bar{x}

$$\begin{array}{r} 929 \\ - 687 \\ \hline 242 \end{array}$$

121 + 121

$$\begin{array}{r} - 566 \\ - 687 \\ \hline - 121 \end{array}$$

-1σ +2σ



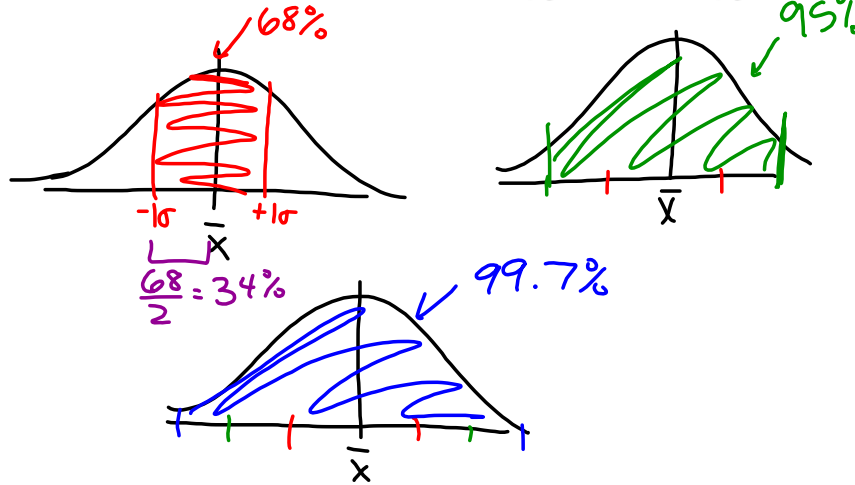
68-95-99.7 Rule

If the data for a population are normally distributed with mean \bar{X} and standard deviation σ then,

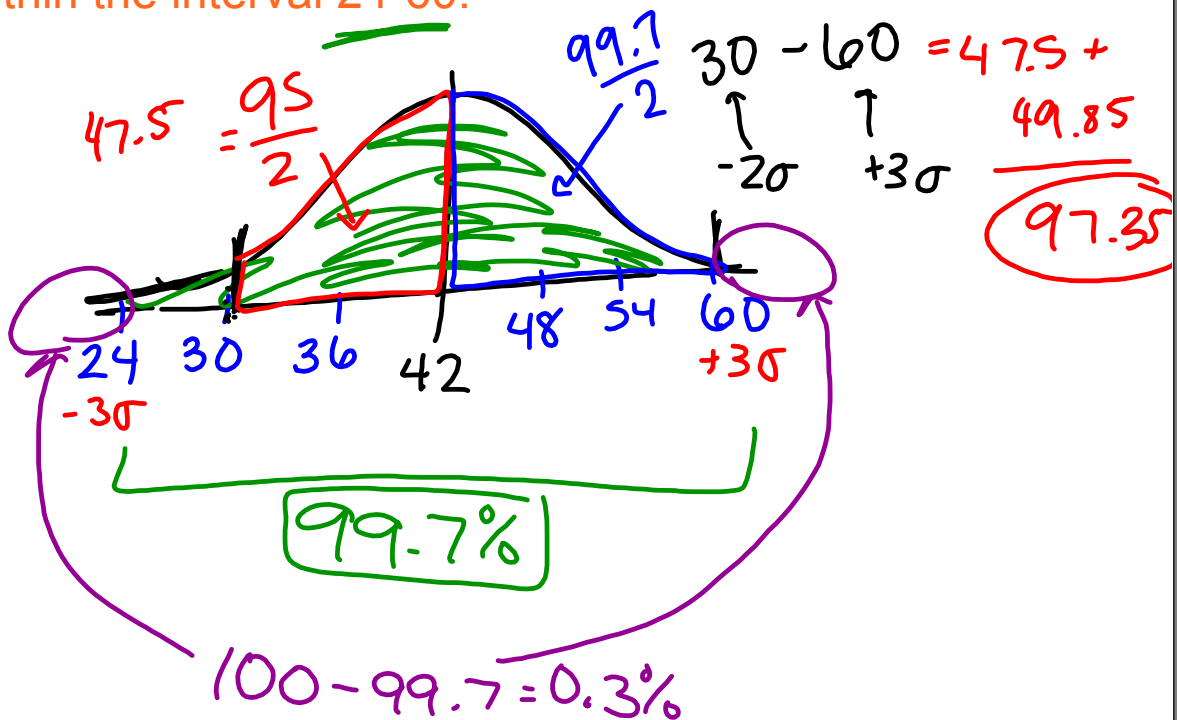
68% of the data lie between $\bar{X} - 1\sigma$ and $\bar{X} + 1\sigma$

95% of the data lie between $\bar{X} - 2\sigma$ and $\bar{X} + 2\sigma$

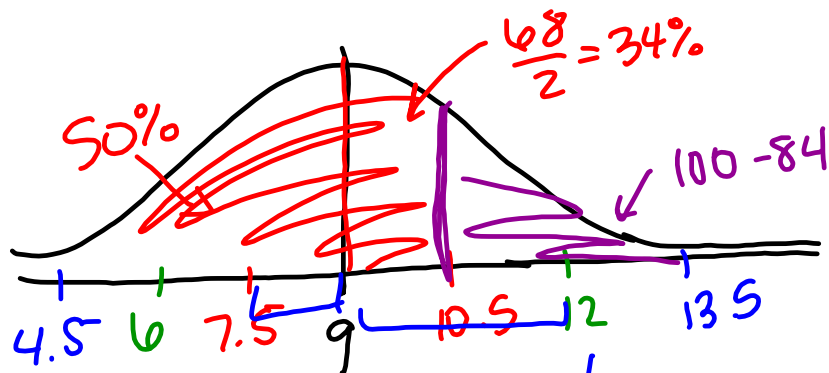
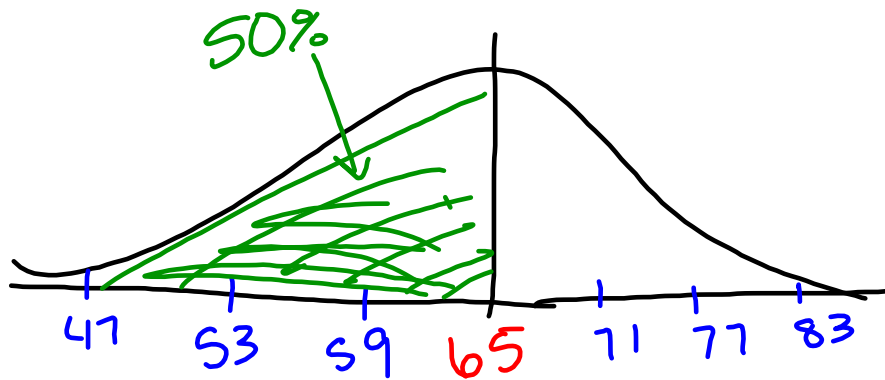
99.7% of the data lie between $\bar{X} - 3\sigma$ and $\bar{X} + 3\sigma$



A set of data has a normal distribution with a mean of 42 and a standard deviation of 6. Find the percent of data within the interval 24-60.



A set of data has a normal distribution with a mean of 65 and a standard deviation of 6. Find the percent of data that is less than 65.



⑤ larger than 10.5

$$50 + 34 = 84\%$$

$$100 - 84 = 16\%$$

⑥



$$100 - 50 - 47.5 =$$

Spread: When we use the median to measure center, we use 5-Number Summary

Range = maximum - minimum

Quartiles split the data into **fourths**

First Quartile (Q_1) = the median of the lower half of the data

Second Quartile = the median

Third Quartile (Q_3) = the median of the upper half of the data

Interquartile Range (IQR) measures the spread between Q_1 and Q_3

$$IQR = Q_3 - Q_1$$

Five number summary = {minimum, Q_1 , median, Q_3 , maximum}

Find the five number summary for the male and female life expectancies in South American nations and compare.

♂ males: {59.0, 60.5, 61.5, 66.7, 67.9, 68.5, 69.0, 70.3, 71.4, 71.9, 72.1, 72.6}

females: {66.2, 66.7, 67.7, 72.8, 74.3, 74.4, 74.6, 76.5, 76.6, 78.8, 79.0, 79.4}

Min: 59

Q_1 : 64.1

Med: 68.75

Q_3 : 71.65

Max: 72.6

A **box plot** (sometimes called box and whisker plot) is a graph that depicts the five number summary of a data set.

To Construct:

1. Draw a rectangular box from Q_1 to Q_3 with a vertical line for the median
2. Draw line segments (whiskers) that extend from the end of the box to the max and mins respectively

Create a box and whisker plot based on the 5 number summaries found for the life expectancy of Males and Females in South American Countries.

