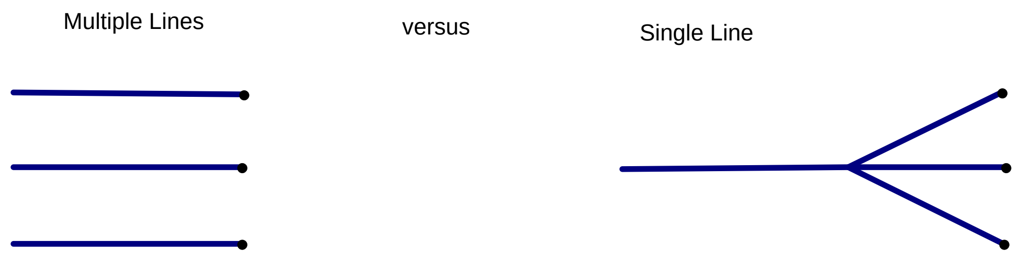
## 3.3 MEASURES OF VARIATION (DISPERSION)

*****Complete the table given below and do a dot plot for both distributions (data is shown below)**

The waiting times of customers (in minutes) is listed below for both banks.

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Bank J<br>single waiting line | 6.5 | 6.6 | 6.7 | 6.8 | 7.1 | 7.3 | 7.4 | 7.7 | 7.7 | 7.7 |
| Bank P<br>multiple waiting lines | 4.2 | 5.4 | 5.8 | 6.2 | 6.7 | 7.7 | 7.7 | 8.5 | 9.3 | 10.0 |

| | median | mean | mode | midrange |
|---|---|---|---|---|
| Bank J | | | | |
| Bank P | | | | |




does not affect the mean but reduces the variation

## RANGE

Highest value – Lowest value

**only affected by 2 numbers

*****find the range of the two distributions shown above**

## STANDARD DEVIATION

Is a measure of variation of values about the mean.

*** In your opinion, considering the two distributions given on the preceding page, which data set has the smallest standard deviation?

## VARIANCE

Is the square of the standard deviation.

| | **standard deviation** | **variance** | |
|---|---|---|---|
| SAMPLE | $s = \sqrt{\frac{\sum(x-\overline{x})^2}{n-1}}$ | $s^2$ | |
| POPULATION | $\sigma = \sqrt{\frac{\sum(x-\mu)^2}{N}}$ | $\sigma^2$ | (sigma)$^2$ |

## FINDING THE STANDARD DEVIATION

*****Find the standard deviation for the waiting times at the Jefferson Valley Bank.**

$\boldsymbol{Recall\ that\ \overline{x} = 7.15}$

| **x** | $\mathbf{x - \overline{x}}$ | $\mathbf{(x - \overline{x})^2}$ |
|---|---|---|
| **6.5** | | |
| **6.6** | | |
| **6.7** | | |
| **6.8** | | |
| **7.1** | | |
| **7.3** | | |
| **7.4** | | |
| **7.7** | | |
| **7.7** | | |
| **7.7** | | |

**Standard Deviation is s =**

*****Follow the same procedure to verify that for the Bank of Providence, s = 1.82 min. Display the information in this table:**

**USEFUL ALTERNATIVE** when the mean is messy $s = \sqrt{\frac{n(\sum x^2)-(\sum x)^2}{n(n-1)}}$

***** Sample of ages (in years) of students cars obtained at the author's college:**
**2 4 4 5 7 8 8 9 10 12 13 16 Find the standard deviation**

| $x$ | $x^2$ |
|---|---|
| 2 | 4 |
| | |
| | |
| | |
| | |
| | |
| | |
| | |
| | |
| | |
| | |
| | |

$\sum x =$ $\sum x^2 =$

***** Sample of ages (in years) of faculty/staff cars obtained at the author's college,**
**7 10 4 13 23 2 7 6 6 3 9 4 <u>Use the calculator</u> to find the st. dev.**

| $L_1$ ($x$) | $L_2 = (L_1)^2$ |
|---|---|
| | |
| | |
| | |
| | |
| | |
| | |
| | |
| | |
| | |
| | |
| | |
| | |

**GROUPED DATA**

$$s = \sqrt{\frac{\sum f\cdot(x-\overline{x})^2}{n-1}} = \sqrt{\frac{n\sum(f\cdot x^2)-(\sum(f\cdot x))^2}{n(n-1)}}$$

***** Find the standard deviation of the car ages for these students:**

| Age | Class Midpoint ($x$) | Students ($f$) | $xf$ | $x^2$ | $x^2 f$ |
|---|---|---|---|---|---|
| 0 – 2 | | 23 | | | |
| 3 – 5 | | 33 | | | |
| 6 – 8 | | 63 | | | |
| 9 – 11 | | 68 | | | |
| 12 – 14 | | 19 | | | |
| 15 – 17 | | 10 | | | |
| 18 – 20 | | 1 | | | |
| 21 – 23 | | 0 | | | |

***** find the standard deviation for the following, <u>with the calculator</u> (see table below)**

| | $L_1$ | $L_2$ | $L_3=L_1*L_2$ | $L_4=(L_1)^2$ | $L_5=L_4*L_2$ |
|---|---|---|---|---|---|
| Temperature | Class Midpoint ($x$) | Frequency ($f$) | $xf$ | $x^2$ | $x^2 f$ |
| 96.5 – 96.8 | | 1 | | | |
| 96.9 – 97.2 | | 8 | | | |
| 97.3 – 97.6 | | 14 | | | |
| 97.7 – 98.0 | | 22 | | | |
| 98.1 – 98.4 | | 19 | | | |
| 98.5 – 98.8 | | 32 | | | |
| 98.9 – 99.2 | | 6 | | | |
| 99.3 – 99.6 | | 4 | | | |
| | | n=sum($L_2$)= | sum($L_3$)= | | sum($L_5$)= |

## COMPUTING THE STANDARD DEVIATION ON THE TI-83

**Non-grouped data**

17 38 14 18 34

Step 1: STAT 1: Edit △ CLEAR ENTER, .....clears out $L_1$.
Step 2: Enter the data in $L_1$.
Step 3: 2nd QUIT to go to the home screen
2nd STAT [LIST] →→ MATH 7:stdDev($L_1$) ENTER

...............OR................

Step 3: STAT → CALC 1:1-Var Stats ENTER $L_1$ ENTER
will display descriptive statistics for this list.

**Grouped Frequency Table Values**

| Weight (kg) | $x$ (class midpoints) | Frequency |
|---|---|---|
| 0.0 - 4.9 | | 43 |
| 5.0 - 9.9 | | 57 |
| 10.0 - 14.9 | | 35 |
| 15.0 - 19.9 | | 45 |
| 20.0 - 24.9 | | 20 |

Step 1: STAT 1: Edit △ CLEAR ENTER, .....clears out $L_1$.
Step 2: Enter the class midpoints in $L_1$.
Step 3: Clear $L_2$.
Step 4: Enter class frequencies in $L_2$.
Step 5: 2nd QUIT to go to the home screen
2nd STAT [LIST] →→ MATH 7:stdDev($L_1$,$L_2$) ENTER

...............OR................

Step 5: STAT → CALC 1:1-Var Stats ENTER $L_1$, $L_2$ ENTER
will display descriptive statistics for this grouped data.

# UNDERSTANDING STANDARD DEVIATION

– **It measures the variation among scores.**

– Scores **close together** yield a **small** standard deviation.

– Scores spread **farther apart** yield a **larger** standard deviation.

– Standard deviation has the same units of measurement (such as dollars or grams or minutes) as the original data values.

– Units of variation are different from the original data. If the original data is in dollars, variation would be expressed as square dollars.

– **Range rule of thumb:** range ~ 4 standard deviations ($r = 4s$)
minimum ~ mean – 2 standard deviations
maximum ~ mean + 2 standard deviations

– Most (95%) of the data is within the interval:
(mean – 2 standard deviations, mean + 2 standard deviations)

– A value is considered *unusual* if it is further than 2 standard deviations from the mean.

*****If the range of a certain distribution is 17, approximate the standard deviation**

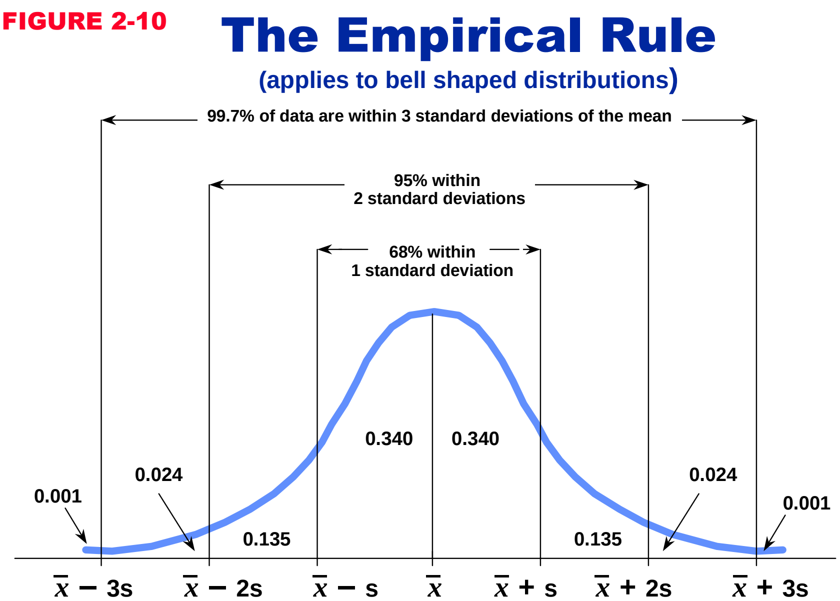
*****do #32, page 113**

**EMPIRICAL RULE** (or 68-95-99.7 rule)

If Distribution is approximately bell shaped, then

- about 68% of the scores fall within 1 standard deviation of the mean
- about 95% of the scores fall within 2 standard deviations of the mean
- about 99.7% of the scores fall within 3 standard deviations of the mean

FIGURE 2-10




*****do #33, pg. 113**

**CHEBYSHEV'S THEOREM**

Proportion of data lying within $k$ standard deviations of the mean is at least $1 - \frac{1}{k^2}$, where $k > 1$

- $k = 2$ ➡ $1 - \frac{1}{4} = \frac{3}{4}$

  (at least 3/4 or 75% of all scores fall within 2 standard deviations of the mean)

- $k = 3$ ➡ $1 - \frac{1}{9} = \frac{8}{9}$

  (at least 8/9 or 89% of all scores fall within 3 standard deviations of the mean)

– It applies to any set of data, but its results are very approximate

– For typical data sets, it is unusual for a score to differ from the mean by more than 2 or 3 standard deviations