1ST WEDNESDAY @2:30 * FACING HOUSING CHALLENGES OR KNOW PEOPLE IN DIFFICULT HOUSING HOMELESS; CONTACT PRENTISS! WILL DIRECT TO RESOURCES.

2-4-19

=> $\bar{x}$ (OR ANY VAR) W A BAR OVER IT IS THE MEAN = $\bar{x}$

- IQR = $Q_3 - Q_1$ SPREAD OUT MIDDLE 50% OF DATA
  -> "Q_1 TO Q_3" GIVES US A RANGE OF TYPICAL VALUES

2-6-19

DETERMINING OUTLIERS

* PRIMARILY FOR SKEWED DATA WHEN WE USE MEDIAN & IQR

- UPPER FENCE (BOUNDARY): $Q_3 + 1.5(IQR)$
  => REFERENCE "WESTERN STATES SPENDING" DATA

WEST $11,240.50 + 1.5(11,240.50 - 8737) = 14,995.75$ ①

EAST $14,551 + 1.5(14,551 - 9430) = 22,232.5$ ②

- LOWER FENCE: $Q_1 - 1.5(IQR)$

WEST: $8737 - 1.5(11240.50 - 8737) = 4981.75$ ①

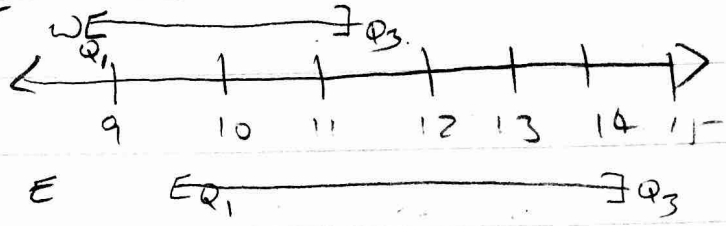
EAST $9430 - (1.5)(14,551 - 9430) = 1748.50$ ②

① IF WE HAVE DATA IN WESTERN PLOT:
ANYTHING ABOVE 14995.75 IS AN "OUTLIER"
ANYTHING BELOW 4981.75 IS AN "OUTLIER"

② " " EASTERN PLOT
ANYTHING ABOVE 22,232.5 => "OUTLIER"
ANYTHING BELOW 1748.50 => OUTLIER

2-6-19

COMPARE "TYPICAL RANGES" USING BOX PLOT



⇒ LITTLE OVERLAP BETWEEN THE TWO

THE STANDARD DEVIATION (WHAT ABOUT SPREAD?)

- MORE POWERFUL MEASURE THAN IQR
  - ↳ TAKES INTO ACCOUNT HOW FAR VALUE IS FROM MEAN ($\bar{x}$)

SLIDE 19

★ SYMMETRIC DISTRIBUTION ⇒ USE MEAN & STANDARD DEVIATION (STD DEV, σ BABY SIGMA)

- STD DEV. MEASURES SPREAD ABOUT THE MEAN

x	DEVIATIONS FROM MEAN $x-\bar{x}$	$(x-\bar{x})^2$
3	$3-5=-2$	4
4	-1	1
5	0	0
6	1	1
7	2	4
$\bar{x}=5$	$\Sigma=0$ ALWAYS!	10

$\bar{x} = \frac{3+4+5+6+7}{5} = \frac{25}{5} = \boxed{5}$

MEAN

$\bar{x} = \frac{\sum x}{n}$ ← SIGMA "SUM"; → DATA POINT

n = SAMPLE SIZE

★ $\frac{(x-\bar{x})^2}{n-1} = s^2$

$s^2 = 10/4$

2-6-19

- STANDARD DEVIATION CALCULATION

$s = \sqrt{10/4} \simeq \boxed{1.58}$ → AVG DISTANCE FROM THE MEAN

- ✱ STANDARD DEVIATION FORMULA: SLIDE 21

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

- TYPICAL RANGE FOR SYMMETRICAL DATA

$\bar{x} \pm s$ (MEAN ± STD DEV)

=> FROM EXAMPLE.

RANGE: $5 - 1.58 = 3.42$

$5 + 1.58 = 6.58$

- ✱ TYPICAL RANGE USUALLY HOLDS 68% OF THE DATA WHEN DISTRIBUTION IS SYMMETRICAL