BUSINESS ANALYTICS

DR. BRENDA MULLALLY

WHAT ARE THE EXTREME VALUES? SHAPE/DISPERSION OF VARIABLE

- STANDARD DEVIATION IS A STATISTICAL TECHNIQUE USED TO INFORM ABOUT THE VARIATION IN THE VALUES OF A VARIABLE.
- THE STANDARD DEVIATION IS A NUMBER SHOWING BY HOW MUCH A GROUP DIFFERS FROM THE MEAN VALUE OF THE GROUP.
- EXAM RESULTS, HEIGHT, AGE,

EMPIRICAL RULES FOR INTERPRETING STANDARD DEVIATION

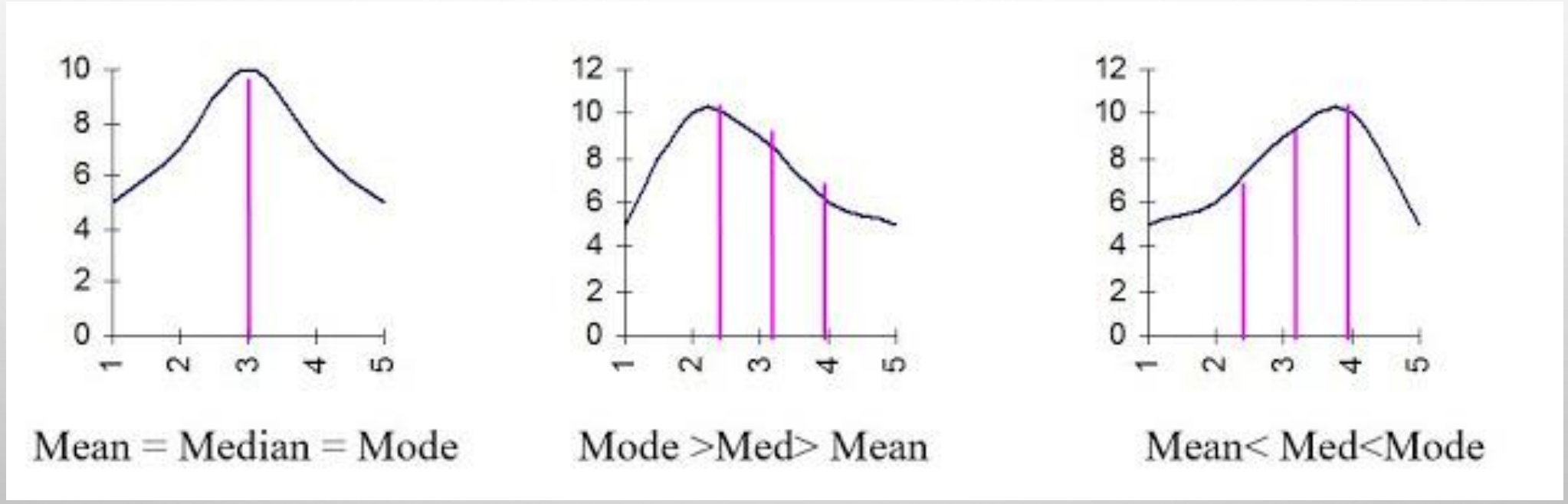
- THE INTERPRETATION OF THE STANDARD DEVIATION CAN BE STATED AS THREE **EMPIRICAL RULES**.
  - IF THE VALUES OF A VARIABLE ARE APPROXIMATELY *NORMALLY* DISTRIBUTED (SYMMETRIC AND BELL-SHAPED), THEN THE FOLLOWING RULES HOLD:
    - APPROXIMATELY 68% OF THE OBSERVATIONS ARE WITHIN ONE STANDARD DEVIATION OF THE MEAN.
    - APPROXIMATELY 95% OF THE OBSERVATIONS ARE WITHIN TWO STANDARD DEVIATIONS OF THE MEAN.
    - APPROXIMATELY 99.7% OF THE OBSERVATIONS ARE WITHIN THREE STANDARD DEVIATIONS OF THE MEAN.

EMPIRICAL RULES FOR BASEBALL SALARIES
(SLIDE 2 OF 3)

- THE EMPIRICAL RULES SHOULD BE APPLIED WITH CAUTION, ESPECIALLY WHEN THE DATA ARE CLEARLY SKEWED, AS ILLUSTRATED BY THE CALCULATIONS FOR BASEBALL SALARIES BELOW.

	H	I	J	K	L	M	N	O
1	**Do empirical rules apply?**							
2		Lower endpoint	Upper endpoint	# below lower	# above upper	% below lower	% above upper	% between
3	Rule 1	-$1,229,688	$7,839,797	0	108	0%	13.20%	86.80%
4	Rule 2	-$5,764,430	$12,374,539	0	54	0%	6.60%	93.40%
5	Rule 3	-$10,299,172	$16,909,281	0	19	0%	2.32%	97.68%

IS THE CHART SYMMETRICAL? MEASURES OF SHAPE



IS THE CHART SYMMETRICAL? MEASURES OF SHAPE

- **SKEWNESS** OCCURS WHEN THERE IS A LACK OF SYMMETRY.
  - A VARIABLE CAN BE **SKEWED TO THE RIGHT** (OR **POSITIVELY SKEWED**) BECAUSE OF SOME REALLY *LARGE* VALUES (E.G., REALLY LARGE BASEBALL SALARIES).
  - OR IT CAN BE **SKEWED TO THE LEFT** (OR **NEGATIVELY SKEWED**) BECAUSE OF SOME REALLY *SMALL* VALUES (E.G., TEMPERATURE LOWS IN ANTARCTICA).
- IN EXCEL, A MEASURE OF SKEWNESS CAN BE CALCULATED WITH THE *SKEW* FUNCTION.

SKEWNESS

Skewness

The coefficient of Skewness is a measure for the degree of symmetry in the variable distribution.

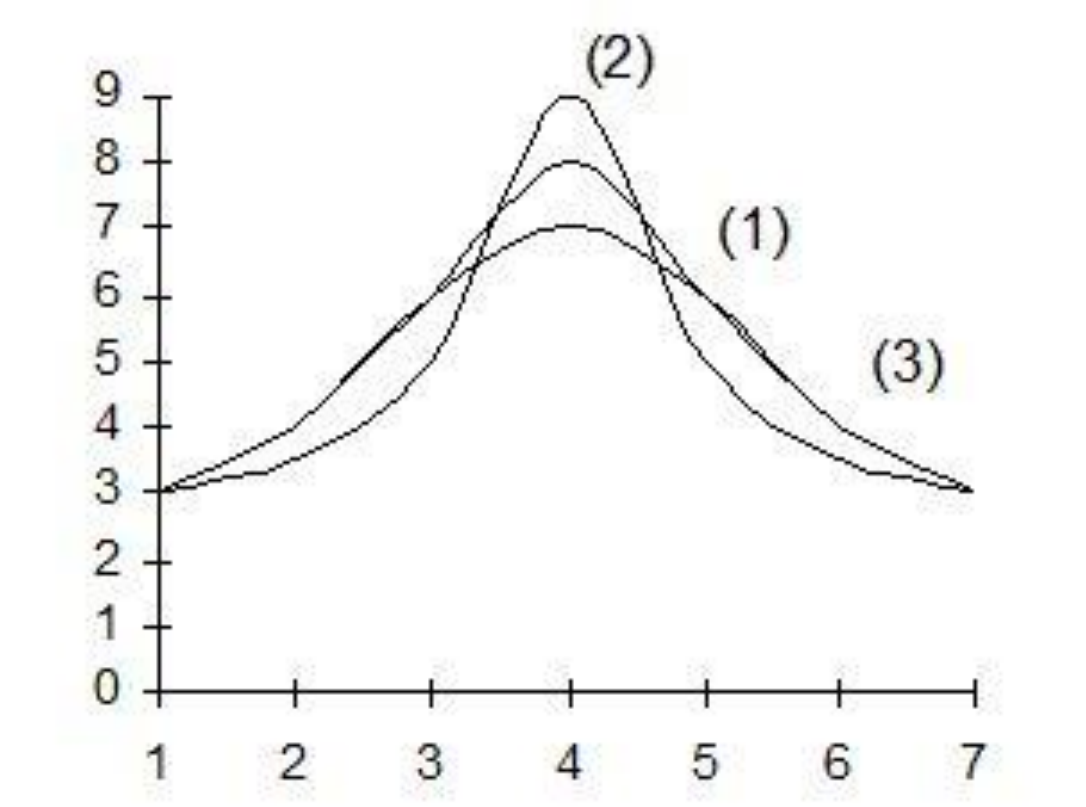
Negatively skewed distribution
or Skewed to the left
Skewness <0

Normal distribution
Symmetrical
Skewness = 0

Positively skewed distribution
or Skewed to the right
Skewness > 0

IS THE CHART SYMMETRICAL? MEASURES OF SHAPE

- **KURTOSIS** HAS TO DO WITH THE "FATNESS" OF THE TAILS OF THE DISTRIBUTION RELATIVE TO THE TAILS OF A NORMAL DISTRIBUTION. IT CAN ALSO BE SAID THE DATA LOOKS FLATTER COMPARED TO NORMAL DISTRIBUTION.
- A DISTRIBUTION WITH HIGH KURTOSIS (POSITIVE NUMBER) HAS MANY MORE EXTREME VALUES, ONE WITH LOW KURTOSIS (NEGATIVE NUMBER) HAS FEW EXTREME VALUES.
- IN EXCEL, KURTOSIS CAN BE CALCULATED WITH THE *KURT* FUNCTION.



KURTOSIS

- CURVE 1 IS NORMAL
- CURVE 2 IS A LEADING CURVE
- CURVE 3 IS A FLAT CURVE

- ALL THREE ARE SYMMETRICAL AROUND THE MEAN.