SECOND LECTURE SUMMARY

MAIN CHARACTERISTICS OF QUANTITATIVE DATA

THREE DEFINITIONS OF 'CENTER':
MEAN - THE USUAL AVERAGE ($\bar{x}$ OR :)
MEDIAN - THE 'MIDDLE' OBSERVATION
(MODE - THE MOST FREQUENT OBSERVATION)

THREE DEFINITIONS OF 'SPREAD':

STANDARD DEVIATION - SQUARE ROOT OF
$SS_x \div (n-1)$ FOR SAMPLE (s)
..... $\div N$ FOR POPULATION (F)

WHERE $SS_x = \text{E } x^2 - (\text{E } x)^2 \div n$ *(or* $\div N$*)*

IQR - DISTANCE OF THE TWO QUARTILES

RANGE - DISTANCE FROM THE SMALLEST TO THE LARGEST OBSERVATION

RELATED DEFINITIONS:

COEFFICIENT OF VARIATION (RATIO SCALE NEEDED): $s \div \overline{x}$

VARIANCE: $SS_x \div (n-1)$ (or $\div N$)

BOX (Q_1, Q_2 AND Q_3) AND WHISKER (SMALLEST TO LARGEST) PLOT - MINITAB ALSO IDENTIFIES **OUTLIERS** BY *

GROUPED-DATA VERSION:

MEAN: $\Sigma x f \div n$ ($n = \Sigma f$)

STANDARD DEVIATION: $\sqrt{SS_x \div (n-1)}$

EXCEPT NOW: $SS_x = \Sigma f x^2 - (\Sigma f x)^2 \div n$

NOTE THAT THE $\overline{x} \pm s$ INTERVAL SHOULD CONTAIN THE 'BULK' (50% - 90%) OF THE DATA, WHILE ALMOST ALL (TYPICALLY 95%) WILL BE INSIDE $\overline{x} \pm 2s$

BE ABLE TO:

COMPUTE THE MEAN, MEDIAN, (MODE), STANDARD DEVIATION, IQR, (RANGE) OF A LIST OF NUMBERS (OBSERVATIONS).
SAME FOR GROUPED DATA

PROBABILITY

RANDOM EXPERIMENT - INDIVIDUAL OUTCOMES ARE ELEMENTS OF THE CORRESPONDING SAMPLE SPACE

SUBSETS OF SAMPLE SPACE ARE CALLED EVENTS (*A, B, ...*) - WE CAN TAKE THEIR COMPLEMENTS (*NOT A*), INTERSECTIONS (*A AND B*) AND UNIONS (*A OR B,* USING THE NONEXCLUSIVE *OR*).

TWO (OR MORE) EVENTS CAN BE:

- INDEPENDENT - CLEAR FROM THE EXPERIMENT - ONE HAPPENING (OR NOT) DOES NOT EFFECT PROBABILITY OF THE OTHER

- EXCLUSIVE - 'INCOMPATIBLE', IF ONE HAPPENS, THE OTHER CANNOT

PROBABILITY OF AN EVENT - IN THE SIMPLEST CASE, WHEN ALL OUTCOMES ARE EQUALLY LIKELY:

$\Pr(A)$ = (# OF OUTCOMES CONTRIBUTING TO A) DIVIDED BY (TOTAL # OF OUTCOMES)

A FEW FORMULAS

$$\Pr(\text{NOT } A) = 1 - \Pr(A)$$

$$\Pr(A \text{ AND } B) = \Pr(A) \times \Pr(B) \quad \text{INDEPENDENT}$$

$$\Pr(A \text{ AND } B) = \Pr(A) \times \Pr(B \text{ GIVEN } A) \quad \text{IN GENERAL}$$

(THE LAST PROBABILITY IS CALLED CONDITIONAL)

$$\Pr(A \text{ OR } B) = \Pr(A) + \Pr(B) \quad \text{EXCLUSIVE}$$

$$\Pr(A \text{ OR } B) = \Pr(A) + \Pr(B) - \Pr(A \text{ AND } B) \quad \text{IN GENERAL}$$