Definitions:

1. Be able to classify variables as either

    (a) Population vs. Sample
    (b) Quantitative vs. Qualitative
    (c) Discrete vs. Continuous

2. Be able to take raw data and make a frequency distribution table and a histogram

    Ex : The ages of 20 children are given as followed:

    $$\begin{matrix} 7 & 4 & 5 & 6 & 7 & 5 & 7 & 4 & 6 & 3 \\ 9 & 7 & 7 & 5 & 4 & 7 & 4 & 3 & 3 & 8 \end{matrix}$$

    (a) Create a frequency distribution table

Age	Frequency
3	3
4	4
5	3
6	2
7	6
8	1
9	1

    (b) Create a histogram for the data

3. Finding mean, median and mode for ungrouped data

    (a) Mean and Variance for Population data :

    $$\mu = \frac{\Sigma x}{N} \qquad \sigma^2 = \frac{\Sigma x^2 - \frac{1}{N}(\Sigma x)^2}{N}$$

    (b) Mean and Variance for Sample data :

    $$\overline{x} = \frac{\Sigma x}{n} \qquad s^2 = \frac{\Sigma x^2 - \frac{1}{n}(\Sigma x)^2}{n-1}$$

4. Finding mean and standard deviation for grouped data.

Commuting Time (in minutes)	f	m	mf	$m^2 f$
0 to less than 10	3	5	15	75
10 to less than 20	5	15	75	1125
20 to less than 30	8	25	200	5000
30 to less than 40	11	35	385	13475
40 to less than 50	7	45	315	14175
	$N = 34$		$\Sigma mf = 990$	$\Sigma m^2 f = 33850$

And so

$$\mu = \frac{\Sigma mf}{N} = \frac{990}{34} \approx 29.12,$$

and

$$
\begin{aligned}
\sigma^2 &= \frac{\Sigma m^2 f - \frac{1}{N}\left(\Sigma mf\right)^2}{N} \\
&= \frac{33850 - \frac{1}{34}\left(990\right)^2}{34} \\
&\approx \frac{33850 - 28826.47}{34} \\
&\approx 147.75.
\end{aligned}
$$

Therefore, the standard deviation is

$$\sigma = \sqrt{147.75} \approx 12.15$$

5. Stem and Leaf displays. Given raw data, make a stem and leaf display. Also, given a stem and leaf display, find the raw data.

   Example : Create a stem and leaf display for the following data:

17	43	53	26	34	35	27	42	26	33
19	47	27	25	41	37	14	32	33	18

1	4	7	8	9	
2	5	6	6	7	7
3	2	3	3	4	5	7
4	1	2	3	7	
5	3				

   Example : The following is a stem and leaf display. Write the data set that is represented by the display

0-2	3	3	4	5	*	2	3	*
3-5	*	2	4	6	*	2		
6-8	1	4	*	*	2			

3	3	4	5	12	13	42
44	46	52	61	64	72	

6. Chebyshev's Theorem and Empirical Rule

(a) Chebyshev's Theorem : For any $k \geq 1$, at least

$$1 - \frac{1}{k^2}$$

of the values are within k standard deviations of the mean.

(b) Empirical Rule : For a data set with a *bell-shaped distribution*

- Approximately 68 percent of the data lie within 1 standard deviation of the mean.
- Approximately 95 percent of the data lie within 2 standard deviations of the mean.
- Approximately 99.7 percent of the data lie within 3 standard deviations of the mean.

7. Probability : Experiments, Outcomes, Sample Spaces, Conditional Probability, Intersections and Unions of Events

Example :

	A	$\overline{A}$	Total
B	3	4	7
$\overline{B}$	6	8	14
Total	9	12	21

(a) Find $P(A), P(B), P(\overline{B}), P(A \cap \overline{B}), P(A \mid B), P(\overline{A} \mid B), P(A \cup B)$.

    i. $P(A) = \frac{9}{21} = \frac{3}{7}$

    ii. $P(B) = \frac{7}{21} = \frac{1}{3}$

    iii. $P(\overline{B}) = \frac{14}{21} = \frac{2}{3}$

    iv. $P(A \cap \overline{B}) = \frac{3}{7}$

    v. $P(A \mid B) = \frac{3}{21} = \frac{1}{7}$

    vi. $P(\overline{A} \mid B) = \frac{8}{21}$

    vii. $P(A \cup B) = \frac{13}{21}$

(b) Are the events A and B independent?

Since $P(A) = P(A \mid B)$, the events A and B are independent.

(c) Are the events A and B mutually exclusive?

Since $A \cap B$ is not empty, events A and B are *not* mutually exclusive.

Example : Toss a coin three times. Let A be the event that there was at least one head, and let B be the event there was exactly one tail.

(a) Find the sample space for this experiment.

$$S = \{HHH, HHT, HTH, HTT, THH, THT, TTH, TTT\}$$

(b) Find $P(A)$.

The event $A = \{HHH, HHT, HTH, HTT, THH, THT, TTH\}$ and so

$$P(A) = \frac{\text{number of outcomes in } A}{\text{number of outcomes in } S} = \frac{7}{8}$$

(c) Find $P(B)$.

The event $B = \{HHT, HTH, THH\}$ and so

$$P(B) = \frac{\text{number of outcomes in } B}{\text{number of outcomes in } S} = \frac{3}{8}$$

(d) Find $P(A \cap B)$.

The event $A \cap B = \{HHT, HTH, THH\}$ and so $P(A \cap B) = 3/8$.

(e) Find $P(A \cup B)$.

The event $A \cup B = \{HHH, HHT, HTH, HTT, THH, THT, TTH\}$ and so $P(A \cup B) = 7/8$.

Example : It is known that 75 percent of the population has mad cow disease. An experiment consists of selecting two members of the population and noting whether or not they have mad cow.

(a) Find the probability exactly one selected member has mad cow.

Exactly one will have mad cow in the two outcomes NM and MN (where M stands for the event that the person selected has mad cow, and N stands for the event that the person does not). We find these probabilities by making a tree diagram to both be 3/16. And so the probability that exactly one has mad cow is 6/16.

(b) Find the probability both selected members have mad cow.

By making a tree diagram, we see that the probability that both have mad cow disease is

$$P(\text{first has mad cow})P(\text{second has mad cow}) = \frac{3}{4} \cdot \frac{3}{4} = \frac{9}{16}$$

8. Probability Distribution Functions of a Discrete Random Variable

Example. In the previous experiment, let x be the discrete random variable that counts the number of people that we selected who have mad cow disease. Find the probability density function for this random variable x. Find the mean and standard deviation for this probability density.

x	$P(x)$	$xP(x)$	$x^2 P(x)$
0	1/16	0	0
1	6/16	6/16	6/16
2	9/16	18/16	36/16
		$\Sigma xP(x) = 24/16$	$\Sigma x^2 P(x) = 42/16$

And so
$$\mu = \Sigma x P(x) = \frac{24}{16} = \frac{3}{2}$$
and
$$\begin{aligned}
\sigma^2 &= \Sigma x^2 P(x) - \mu^2 \\
&= \frac{42}{16} - \left(\frac{3}{2}\right)^2 \\
&= \frac{21}{8} - \frac{9}{4} \\
&= \frac{3}{8}.
\end{aligned}$$

Therefore, $\sigma = \sqrt{3/8} \approx .61$.