EXERCISE #2 - Probability

1. What is probability?

2. What is the probability that a flip of a fair coin will come up heads?

3. If a fair coin comes up tails four times in a row, what is the probability that it will come up heads on the fifth toss?

4. In a pool of 24 workers, one is a corporate spy. The pool is divided into teams, with four workers per team. If you were to randomly pick one of these teams, what is the probability that the team you picked would contain a spy?

PERMUTATIONS

5. Write out all the possible three-digit permutations of the numbers 7,8,9, then show mathematically how many permutations there should be.

6. Next, write out all the possible two-digit permutations of the numbers 7,8,9, *with repetition*, then show mathematically how many permutations there should be.

COMBINATIONS

7. Of the five projects you completed last year, only three will fit on your resume.
    How many different combinations of projects are there?

INTERSECTIONS, UNIONS & COMPLEMENTS
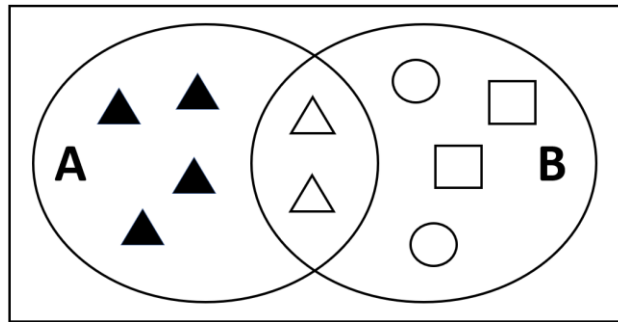
8. In the above Venn Diagram, if **A** represents triangles, and **B** represents hollow figures,
    express the Intersection (A and B), the Union (A or B) and the Complement of A (not A)
    mathematically.

CONDITIONAL PROBABILITY

9. A company conducts a historical study of employees. Out of 250 hires, 180 are still with the company, 120 had been offered promotions, and 75 are both with the company and had been offered promotions.
   Given that an employee was offered a promotion, what is the probability that they are still with the company?

BAYES THEOREM

10. A spam filtering program uses Bayes Theorem to assign a probability that an incoming email that contains the word "special" is spam (an unwanted message) as opposed to ham (a valid message). The following assumptions are made:

   - The initial probability $P(S)$ that an incoming message is spam is 0.50 (ie, it is equally likely to be spam as ham)

   - During a learning phase where known spam and ham messages were tested for the occurrence of the word, it is found that 4% of spam messages contain the word, and 1% of ham messages have it.

   What is the probability that a message containing the word is spam?

$$P(S|W) = \frac{P(W|S) \cdot P(S)}{P(W)} = \frac{P(W|S) \cdot P(S)}{P(W|S) \cdot P(S) + P(W|H) \cdot P(H)}$$

$P(S|W)$ = probability of being spam if containing the word

$P(W|S)$ = probability of containing the word if spam

$P(S)$ = probability of being spam

$P(W)$ = probability of containing the word

$P(W|H)$ = probability of containing the word if ham

$P(H)$ = probability of being ham