1. Machine learning fundamentals
    a. The i.i.d. assumption is not reasonable here because the features are not independent of one another. The presence of a fever may affect whether or not an individual has a headache or one's tiredness, attributing to sluggish behavior. These features are also not identically distributed because each child may react to the presence of COVID differently and the likelihood of each feature is different for each child.
    b. The i.i.d. assumption is not reasonable here because asthma and bronchitis are known to affect lung capacity. Therefore, lung capacity among the children is not identically distributed, because the likelihoods of certain lung capacities will be different for Johnny and Debbie compared to the other children. Assuming that the conditions are non-contagious, lung capacity would be independent because one child's lung has no way of affecting another's.
    c. The i.i.d. assumption is not reasonable here because the chances of getting COVID between Amber and Brenda's children are not independent as they are in the same class. Assuming that these two children spend most of their school time in relatively close proximity with each other, Amber's child getting COVID has a greater measure of effect on Brenda's child getting COVID than Chloe's child getting COVID because Chloe's child is in a different class. Chances of getting COVID will be identically distributed because there are no factors mentioned affecting immunity or particular vulnerability among these three children.
2. Entropy, conditional entropy, mutual information, and information gain
    a.

P(x = sock color)	**$- \Sigma_{X=x}(P(X = x) \log_2(P(X = x)))$**	**= 1.6805**
P(x = red sock) = 0.5	$(0.5) * \log_2(0.5)$	-0.5
P(x = blue sock) = 0.25	$(0.25) * \log_2(0.25)$	-0.5
P(x = yellow sock) = 0.2	$(0.2) * \log_2(0.2)$	-0.4644
P(x = black sock) = 0.05	$(0.05) * \log_2(0.05)$	-0.2161

The entropy of the sock is 1.6805

    b.

P(x = sock color)	**P(y = drawer)**
P(x = red) = 0.5	P(y = top) = 0.67
P(x = blue) = 0.25	P(y = top) = 0.67
P(x = yellow) = 0.2	P(y = top) = 0.67

P(x = black) = 0.05	P(y = top) = 0.67
P(x = red) = 0.5	P(y = bottom) = 0.33
P(x = blue) = 0.25	P(y = bottom) = 0.33
P(x = yellow) = 0.2	P(y = bottom) = 0.33
P(x = black) = 0.05	P(y = bottom) = 0.33

P(x, y)	**P(x | y)**
P(red, top) = 0.5	P(red | top) = 1
P(blue, top) = 0.25	P(blue | top) = 0
P(yellow, top) = 0.2	P(yellow | top) = 0
P(black, top) = 0.05	P(black | top) = 0
P(red, bottom) = 0.5	P(red | bottom) = 0
P(blue, bottom) = 0.25	P(blue | bottom) = 0.5
P(yellow, bottom) = 0.2	P(yellow | bottom) = 0.4
P(black, bottom) = 0.05	P(black | bottom) = 0.1

x = sock color and y = drawer	$-\Sigma_{X=x}(P(X = x, Y = y) \log_2(P(X = x \mid Y = y)))$
x = red and y = top	0.0.5 * 0 = 0
x = blue and y = bottom	0.25 * -1 = -0.25
x = yellow and y = bottom	0.2 * -1.322 = -0.2644
x = black and y = bottom	0.05 * -3.322 = -0.1661

The conditional entropy is 0.6805

c. Information gain is 1.6805 - 0.6805 = 1

3. entropy = 3.3141823231610834
   conditional entropy = 3.3029598816135173