Question 2

From the task description we get: $P(a \mid r) = 0.3$
$P(o \mid r) = 0.4$
$P(l \mid r) = 0.3$
$P(a \mid g) = 0.3$
$P(o \mid g) = 0.3$
$P(l \mid g) = 0.4$
$P(a \mid b) = 0.5$
$P(o \mid b) = 0.5$
$P(l \mid b) = 0$

We can now determine the following probabilities:
$P(a \cap r) = P(a \mid r) \cdot P(r) = 0.3 \cdot 0.2 = 0.06$
$P(o \cap r) = P(o \mid r) \cdot P(r) = 0.4 \cdot 0.2 = 0.08$
$P(l \cap r) = P(l \mid r) \cdot P(r) = 0.3 \cdot 0.2 = 0.06$
$P(a \cap g) = P(a \mid g) \cdot P(g) = 0.3 \cdot 0.6 = 0.18$
$P(o \cap g) = P(o \mid g) \cdot P(g) = 0.3 \cdot 0.6 = 0.18$
$P(l \cap g) = P(l \mid g) \cdot P(g) = 0.4 \cdot 0.6 = 0.24$
$P(a \cap b) = P(a \mid b) \cdot P(b) = 0.5 \cdot 0.2 = 0.1$
$P(o \cap b) = P(o \mid b) \cdot P(b) = 0.5 \cdot 0.2 = 0.1$
$P(l \cap b) = P(l \mid b) \cdot P(b) = 0 \cdot 0.2 = 0$

(a)

$P(a) = P(a \cap r) + P(a \cap g) + P(a \cap b) = 0.06 + 0.18 + 0.1 = 0.34$
$P(o) = P(o \cap r) + P(o \cap g) + P(o \cap b) = 0.08 + 0.18 + 0.1 = 0.36$
$P(l) = P(l \cap r) + P(l \cap g) + P(l \cap b) = 0.06 + 0.24 + 0 = 0.3$

Thus, the probability of selecting an apple is 0.34.

(b)

$P(r \mid a) = \frac{P(a \cap r)}{P(a)} = \frac{0.06}{0.34} = 0.176$
$P(g \mid a) = \frac{P(a \cap g)}{P(a)} = \frac{0.18}{0.34} = 0.529$
$P(b \mid a) = \frac{P(a \cap b)}{P(a)} = \frac{0.1}{0.34} = 0.294$
$P(r \mid o) = \frac{P(o \cap r)}{P(o)} = \frac{0.08}{0.36} = 0.222$
$P(g \mid o) = \frac{P(o \cap g)}{P(o)} = \frac{0.18}{0.36} = 0.5$
$P(b \mid o) = \frac{P(o \cap b)}{P(o)} = \frac{0.1}{0.36} = 0.278$
$P(r \mid l) = \frac{P(l \cap r)}{P(l)} = \frac{0.06}{0.3} = 0.2$
$P(g \mid l) = \frac{P(l \cap g)}{P(l)} = \frac{0.24}{0.3} = 0.8$
$P(b \mid l) = \frac{P(l \cap b)}{P(l)} = \frac{0}{0.3} = 0$

Thus, if we observe that the selected fruit is in fact an orange, the probability that it came from the green box is 0.5.

Question 3

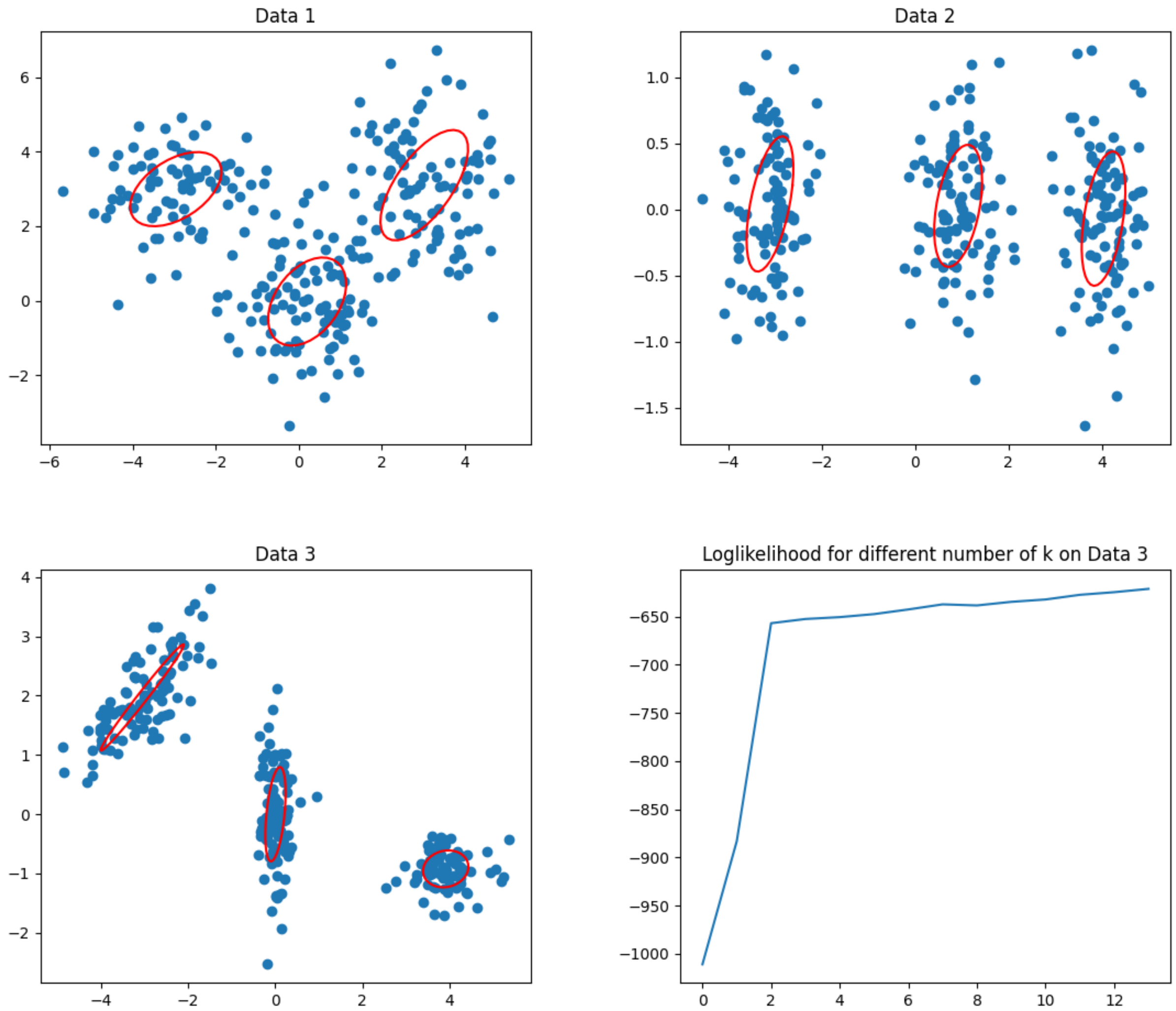
(a)

Question 4

(a)

Question 6

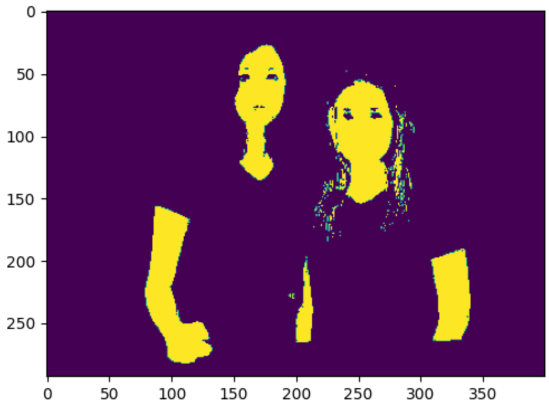
(g)



We can see, that in all three data sets there are 3 clusters (more so in data2 and data3 than in data1). Therefore, $K = 3$ clusters is a very suitable number of clusters for this application.

$K = 2$ yields very under fitting cluster parameters, even for data1. For $K > 3$ we get clusters that are "overlapping" each other, so we see overfitting.

(h)



Increasing K here improves accuracy on the training data set, thus reducing generalization. For $K = 9$, a model is produced that separates the colors in the test image better than the default of $K = 3$. Increasing K further might lead to poor accuracy on other test data. Since the colors are there dimensional, and we are looking at two classifications (skin and no-skin), intuitively speaking, $3 * 2$ clusters should be a good starting point for finding a suitable number of clusters.

Increasing the threshold *theta* increases the required confidence, that a given pixel color represents skin. This leads to a more conservative classification (so rather non-skin than skin), which is more robust to noise, however, can also be less generalized. We increased *theta* to 3 and got the above result.