1. Considering the following two-dimensional measurements:

	y_1	y_2
x_1	-2	2
x_2	-1	3
x_3	0	1
x_4	-2	1

(a) What are the maximum likelihood parameters of a multivariate Gaussian distribution for this data set?

(b) What is the Gaussian's shape? Draw its contour plot.

(a)

(b)

2. Consider the following data-set, paired with a query vector $x_{new} = \begin{bmatrix}1 & 1 & 1 & 1 & 1\end{bmatrix}^T$:

	y_1	y_2	y_3	y_4	y_5	z
x_1	1	1	0	1	0	1
x_2	1	1	1	0	0	0
x_3	0	1	1	1	0	0
x_4	0	0	0	1	1	0
x_5	1	0	1	1	1	1
x_6	0	0	1	0	0	1
x_7	0	0	0	0	1	1

(a) Using Bayes' rule, without making any assumptions, compute the posterior probabilities for the query vector. How is it classified?

(b) What is the problem of working without assumptions?

(c) Compute the class for the same query vector under the naive Bayes assumption?

(d) Consider the presence of missings. Under the same naive Bayes assumption, how would you classify the query vector $x_{new} = \begin{bmatrix}1 & ? & 1 & ? & 1\end{bmatrix}^T$?

(a)

(b)

(c)

(d)

3. Considering the following data set, paired with the query vector $x_{new} = \begin{bmatrix}100 & 225\end{bmatrix}^T$:

	y_1	y_2	z
x_1	170	160	0
x_2	80	220	1
x_3	90	200	1
x_4	60	160	0
x_5	50	150	0
x_6	70	190	1

(a) Compute the most probable class for the query vector assuming that the likelihoods are 2-dimensional Gaussians.

(b) Compute the most probable class for the query vector, under the Naive Bayes assumption, using 1-dimensional Gaussians to model the likelihoods.

(a)

(b)

4. Assuming training examples with m features and a binary class:

(a) How many parameters are needed to estimate, considering boolean features and:

i. No assumptions regarding the data's distribution.

ii. Naive Bayes assumption.

(b) How many parameters are needed to estimate, considering numeric-valued features and:

i. Multivariate Gaussian assumption.

ii. Naive Bayes assumption.

(a) i.

ii.

(b) i.

ii.