1.Question 1

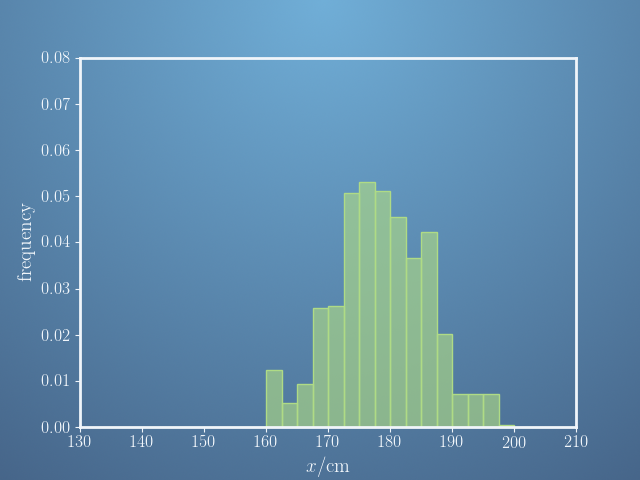
In this exercise, we shall see how it is often convenient to use vectors in machine learning. These could be in the form of data itself, or model parameters, and so on.

The purpose of this exercise is to set the scene for Linear Algebra and the rest of the maths we will cover in the specialization. If this is confusing right now - stick with us! We'll build up your skills throughout the rest of the course. For this reason we've set a low pass mark for this quiz, but even if you don't pass in one go, reading the feedback from a wrong answer can often give more insight than guessing a correct answer!

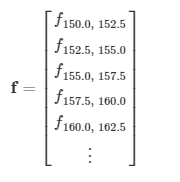
\* \* \*∗∗∗

The problem we shall focus on in this exercise is the distribution of heights in a population.

If we do a survey of the heights of people in a population, we may get a distribution like this:

  
This histogram indicates how likely it is for anyone in the survey to be in a particular height range. (6 ft is around 183 cm)

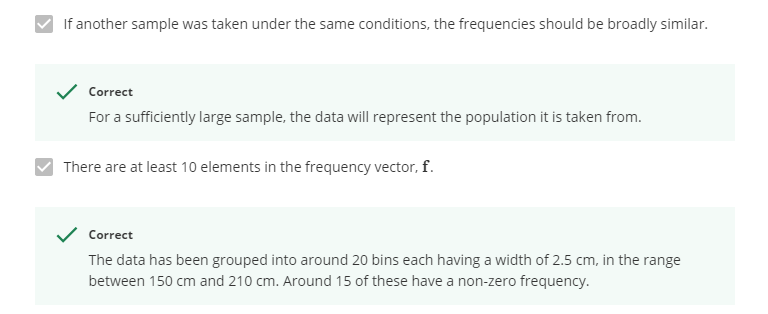
This histogram can also be represented by a vector, i.e. a list of numbers. In this case, we record the frequency of people with heights in little groups at 2.5 cm intervals, i.e. between 150 cm and 152.5 cm, between 152.5 cm and 155 cm, and so on. We can define this as the vector \mathbf{f}**f** with components,



These vector components are then the sizes of each bar in the histogram.

Of the following statements, select all that you think are true.

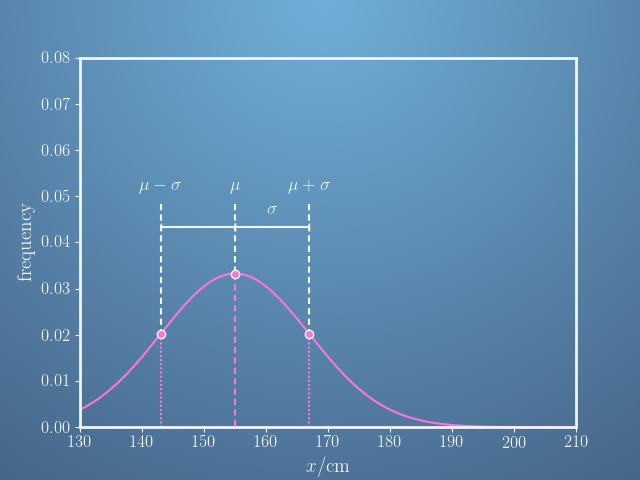
**Ans.**

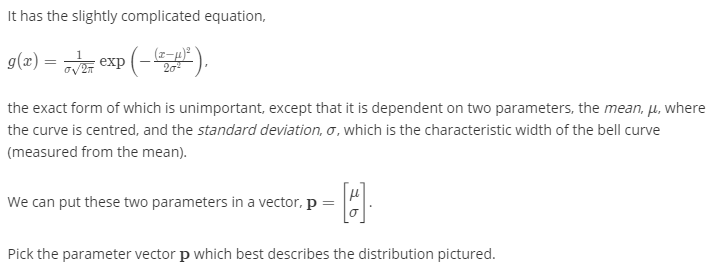
****

2.Question 2

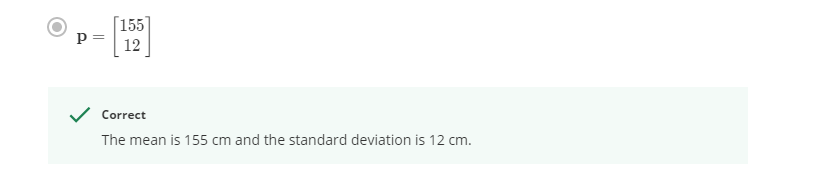
One of the tasks of machine learning is to fit a model to data in order to represent the underlying distribution.

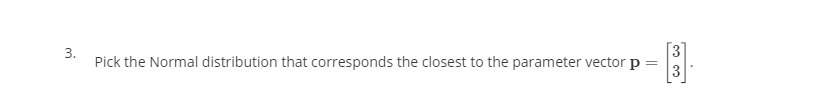
For the heights of a population, a model we may use to predict frequencies is the Normal (or Gaussian) distribution. This is a model for a bell-shaped curve, which looks like this,



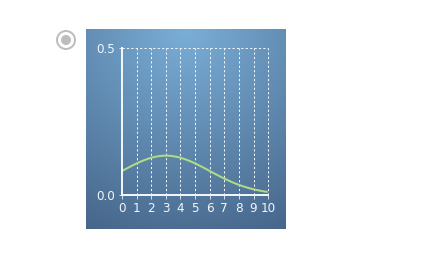
****

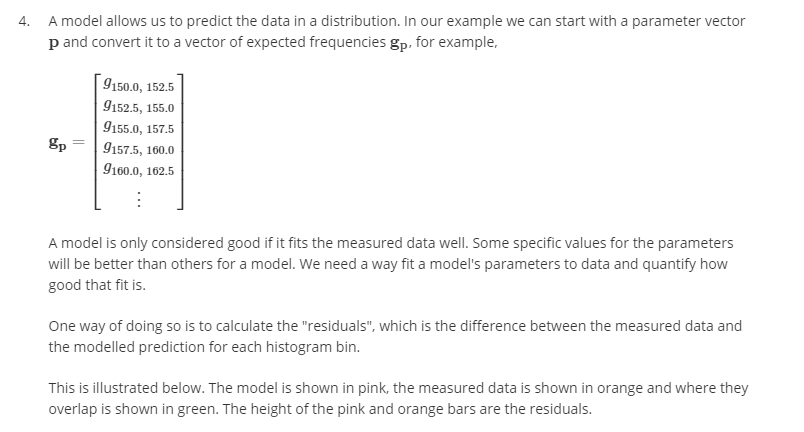
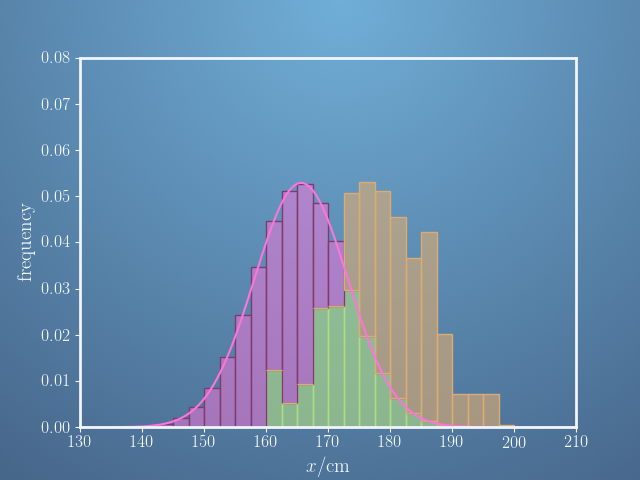
Ans.



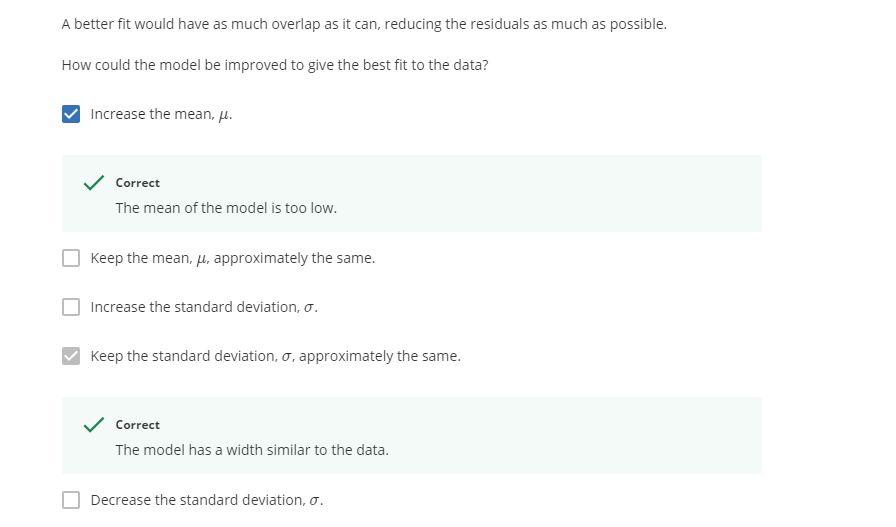


Ans.



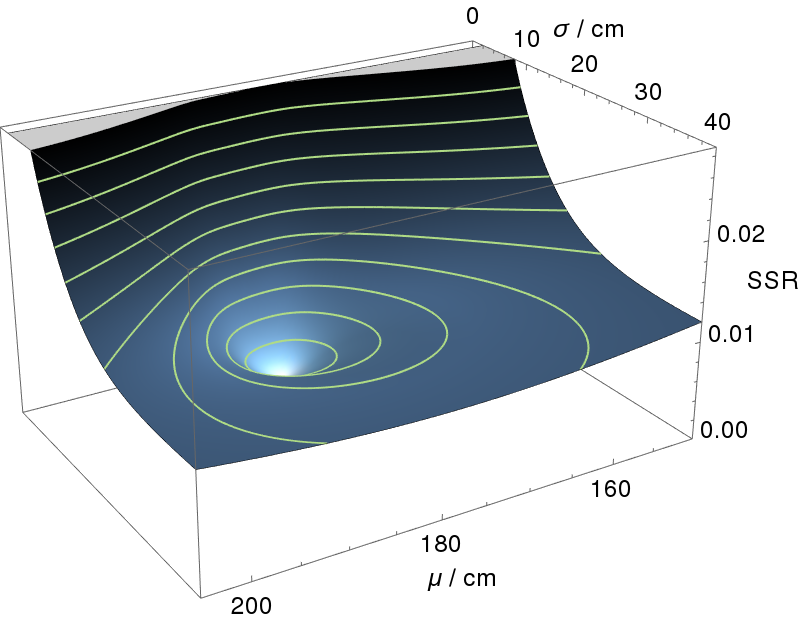
 

Ans.



5. Since each parameter vector  represents a different bell curve, each with its own value for the sum of squared residualsSSR, we can draw the surface of \mathrm{SSR}SSR values over the space spanned by \mathbf{p}**p**, such as \mu*μ*and \sigma*σ* in this example.

Here is an illustration of this surface for our data.



Every point on this surface represents the SSR of a choice of parameters, with some bell curves performing better at representing the data than others.

We can take a ‘top-down’ view of the surface, and view it as a contour map, where each of the contours (in green here) represent a constant value for the \mathrm{SSR}SSR.

