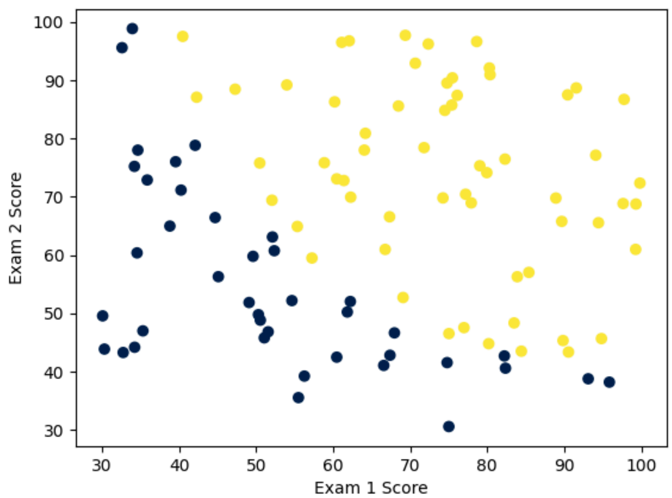
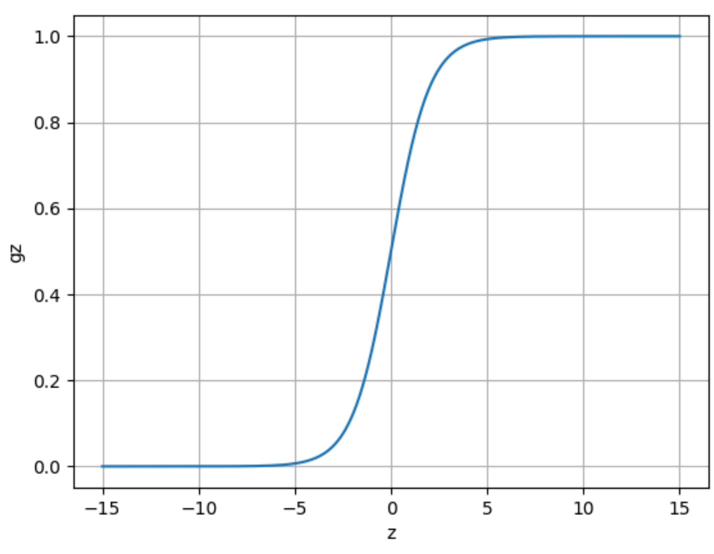
Question 1
Part a)
Text output:
The size of X is: (100, 3)
The size of Y is: (100, 1)

Part b)



Part d)

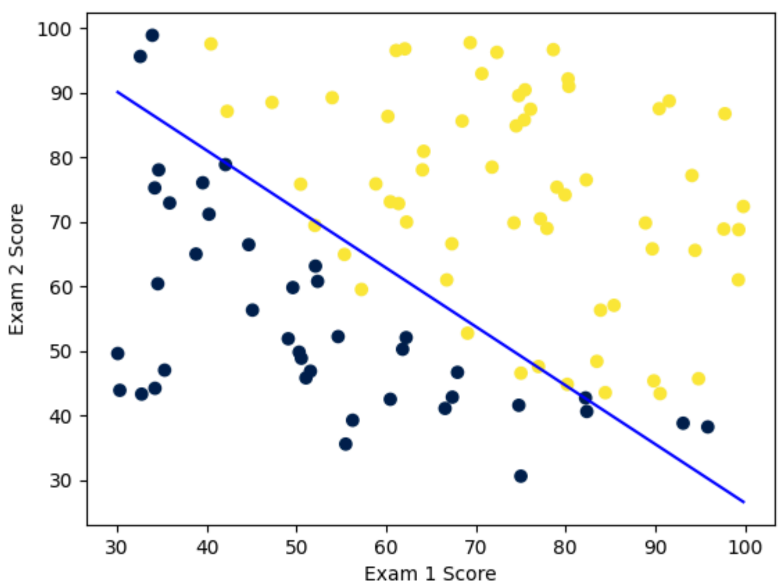


The value the output reaches 0.1 is -ln(9) which is about -2.2

Part e)
Text output:
The cost function using the toy data is: [[1.12692801]]

Part f)
Text output:
The optimal values of theta are: [[-28.53227309   0.220992   0.24304455]]
With the optimal thetas, the cost is: [[0.18785309]]

Part g)



Part i)
Text output:
The admission probability is [[0.62834947]]
The decision should be to admit

Part j)

Part J

$$J(\theta) = \frac{1}{m}\sum_{i=1}^{m}\left[-y^i \log(h_\theta(x^i)) - (1-y^i)(\log(1-h_\theta(x^i))\right]$$

$$\frac{d}{dx}(f(x)\cdot g(x)) = f'(x)\cdot g(x) + f(x)\cdot g'(x)$$

not just +

$$h_\theta(x^i) = \frac{1}{1+e^{-x^i}} =$$

$$\log(h_\theta(x^i)) = \log\left(\frac{1}{1+e^{-x^i}}\right) = -\log\left(1+e^{-x^i}\right)$$

$$\log(1-h_\theta(x^i)) = \log\left(1-\frac{1}{1+e^{-x^i}}\right) = \log\left(\frac{1+e^{-x^i}-1}{1+e^{-x^i}}\right) = \log\left(\frac{e^{-x^i}}{1+e^{-x^i}}\right)$$

$10^Y = e^{-x^i}$

$$= \log(e^{-x^i}) - \log(1+e^{-x^i})$$

$$J(\theta) = \frac{1}{m}\sum\left[-y^i\left(-\log(1+e^{-x^i})\right) - (1-y^i)\left(\log(e^{-x^i}) - \log(1+e^{-x^i})\right)\right]$$

let $u = \log(1+e^{-x^i})$ $w = \log(e^{-x^i})$

$$J(\theta) = \frac{1}{m}\sum\left[-y^i(-u) - (1-y^i)(w-u)\right]$$

$$= \frac{1}{m}\sum\left[\cancel{y^i u} - w + u + y^i w - \cancel{y^i u}\right]$$

$$= \frac{1}{m}\sum\left[\log(1+e^{-x^i}) + y^i\log(e^{-x^i}) - \log(e^{-x^i})\right]$$

$$J(\theta) = \frac{1}{m}\sum\left[y^i\log(e^{-x^i}) + \log(1+e^{-x^i}) - \log(e^{-x^i})\right]$$

$$= \frac{1}{m}\sum\left[y^i\log(e^{-x^i}) + \log\left(\frac{1+e^{-x^i}}{e^{-x^i}}\right)\right]$$

$$= \frac{1}{m}\sum\left[y^i\log(e^{-x\theta^i}) + \log(e^{x\theta^i}+1)\right]$$

$$\frac{\partial J}{\partial \theta} = \frac{1}{m}\sum\left[y^i\frac{1}{e^{-x\theta^i}\ln 10} + \frac{1}{(e^{x\theta^i}+1)\ln 10}\right]$$

$$\frac{1}{m}\sum\left[y^i\theta x^i + \log(e^{x\theta^i}+1)\right]$$

$$\frac{\partial J}{\partial \theta} = \frac{1}{m}\sum\left[y^i x^i + \frac{1}{e^{x\theta^i}+1}\left(x e^{\theta x^i}\right)\right]$$

$$= \frac{1}{m}\sum\left[y^i x^i + \frac{x^i e^{\theta x^i}}{e^{x\theta^i}+1}\right]$$

with $h_\theta(x^i) = \frac{1}{1+e^{-x\theta^i}}\cdot\frac{e^{\theta x^i}}{e^{\theta x^i}} = \frac{e^{\theta x^i}}{e^{x\theta^i}+1}$

$$\frac{\partial J}{\partial \theta} = \frac{1}{m}\sum\left[y^i + h_\theta(x^i)\right]x^i$$

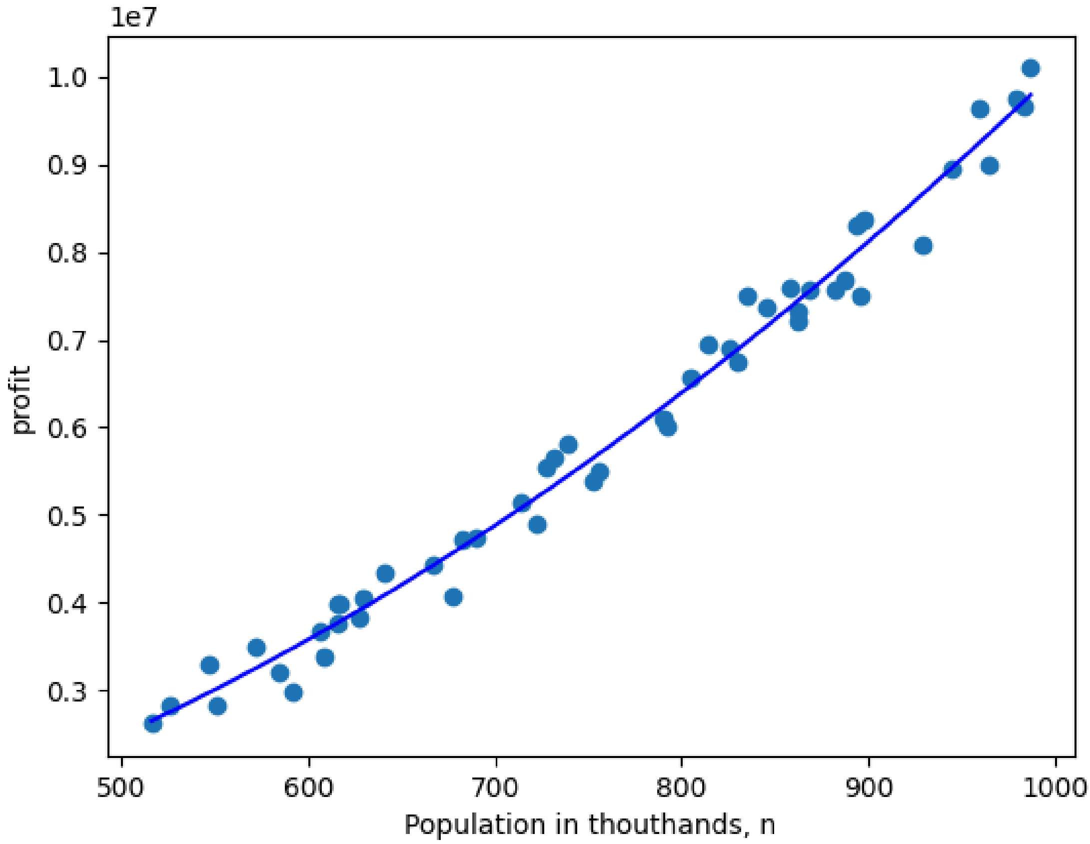
Question 2

Part a)

Text output:

The optimal values of theta are: [[2.19256106e+05 -7.75884747e+02 1.06170499e+01]]

Part b)



1e7
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
profit
500
600
700
800
900
1000
Population in thouthands, n