Part A -

[a] Write a python script to fit a poisson regression using `data.csv`. Let your code report the fitted coefficients.

[b] Either continue with your python script Or create a spreadsheet (or modify the spreadsheet `poissonreg_fit_calc.xlsx` to estimate the poisson regression coefficients using the Newton Raphson method with the initial value $\hat{\boldsymbol{\beta}}^{(0)} = (0.1, 0.1)^T$

Let your code or spreadsheet report the following:

update	$\hat{\beta}_0$	$\hat{\beta}_1$	log-likelihood	gain in log-likelihood
0	0.1	0.1	-92.6102274746018	
1				
2				
3				
4				
MLE			-68.0744858189719	

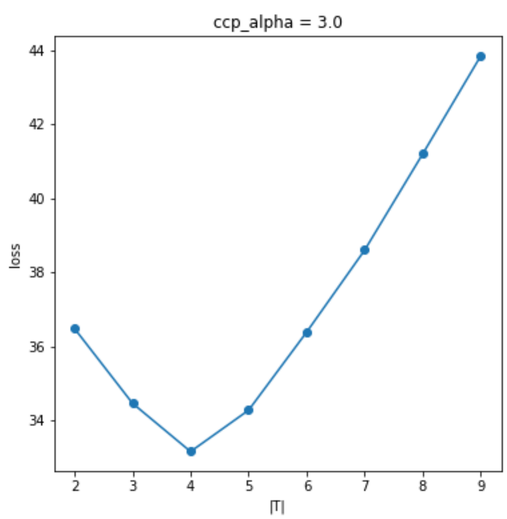
Please submit your work as hw12A.ipnb and hw12A.html and hw12A.xlsx (if applicable) to `Canvas`.

Part B -

Consider the decision tree model, `tree(height` $\sim$ `weight + as.factor(male))`, fitted with the Howell dataset on those aged 18-years or older (please refer to class discussion and HW4).

Write a python script to prune the decision tree using cost-complexity pruning for a range of cost-complexity alpha values (`CCP_ALPHA = np.arange(0.0, 10.0)`)

[a] Validate the optimal tree size for cost complexity pruning when $\alpha = 3.0$. Let your code report the values of the loss function ($MSE + \alpha|T|$) for tree sizes between 2 and 9.



[b] Let your code report, for each α,

- the best tree size: the number of terminal nodes ($|T|$) of the pruned tree which minimizes the loss function

$$MSE + \alpha|T|$$

- the value of the associated loss function for the pruned tree as follows:

ccp_alpha	best tree size	associated loss
0	9.0	16.847466
1	5.0	24.280531
2	4.0	29.162939
3	4.0	33.162939
4	4.0	37.162939
5	3.0	40.465856
6	2.0	42.484992
7	2.0	44.484992
8	2.0	46.484992
9	2.0	48.484992

Please submit your work as hw12B.ipnb and hw12B.html to `Canvas`.