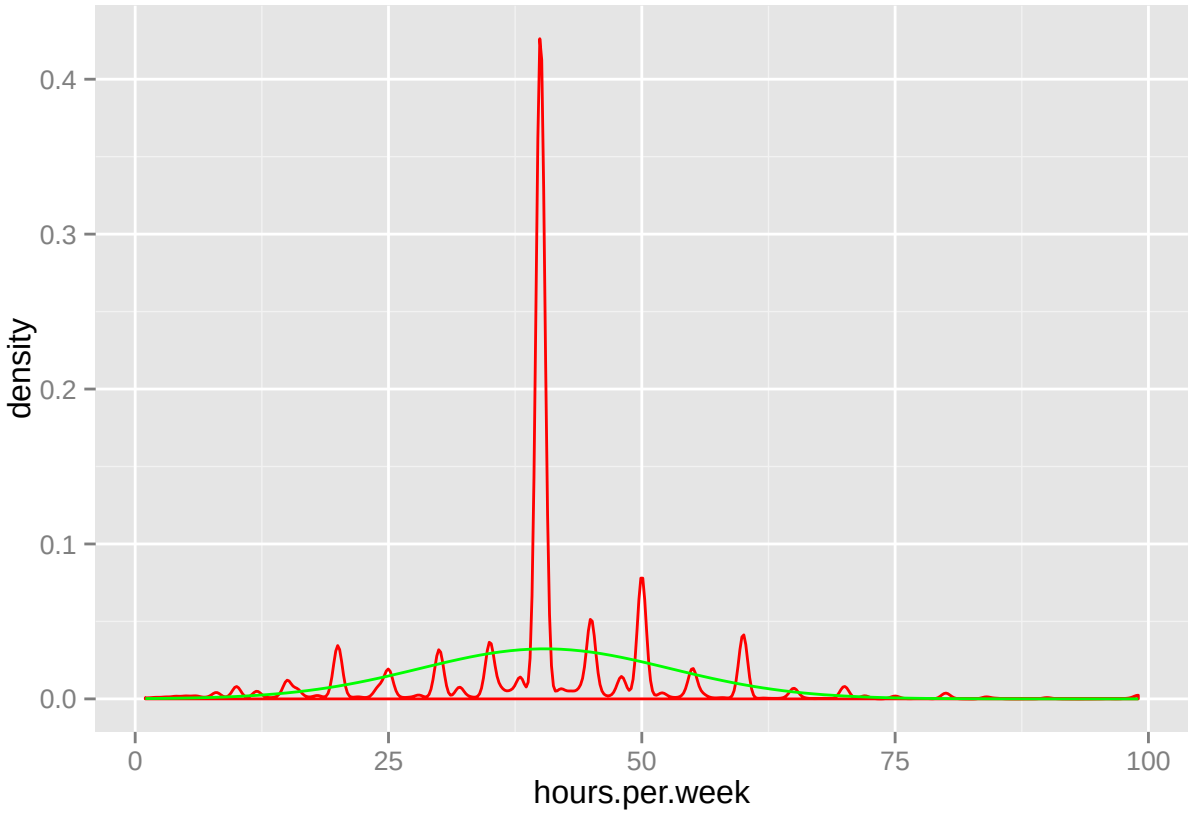
use same file for this assignment which you used for practice assignment for Univariate Statistics

Does your data follow normal distribution?

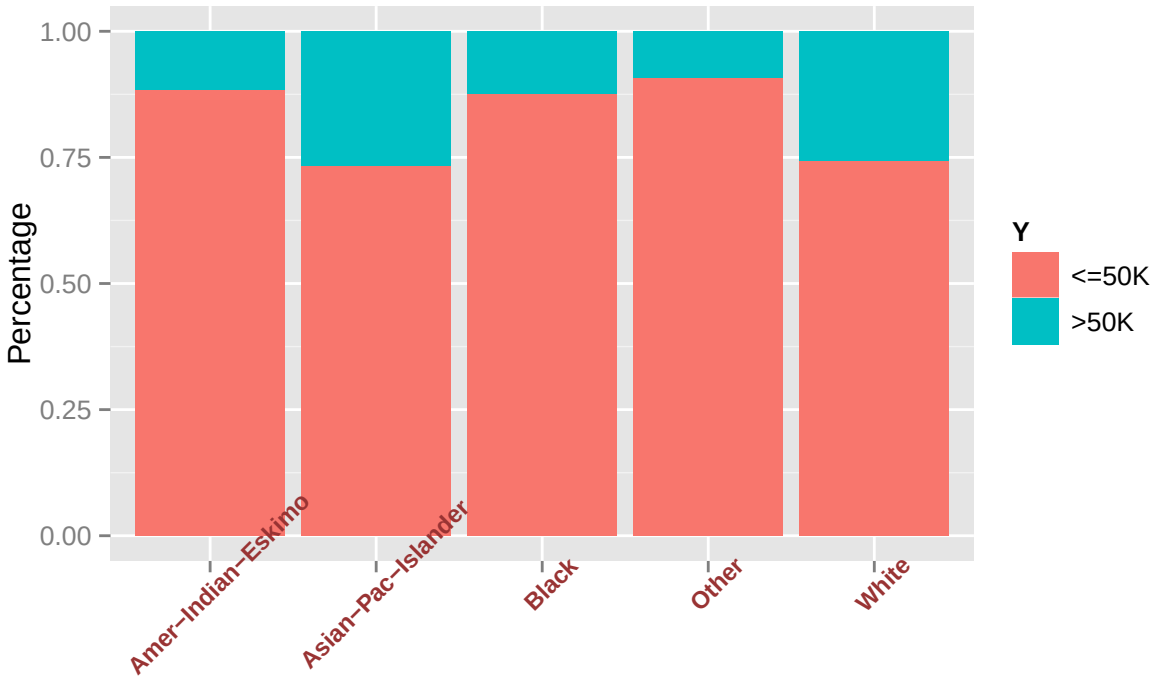
Note: you can revisit this question once you have gone through hypothesis testing module

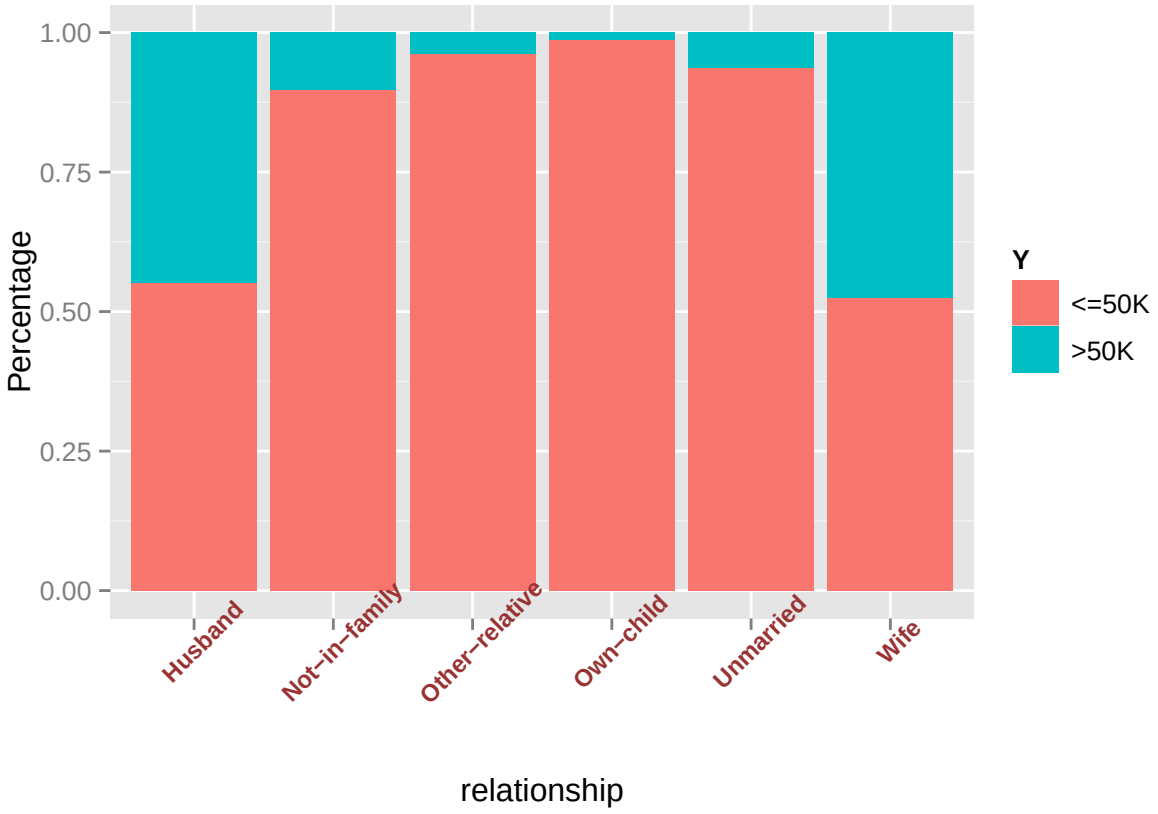
Find out if variable hours.per.week follows normal distribution using density plots. Your plot should like this:



Correlation between two categorcial variables

Use stacked barcharts to find out if outcome Y is affected by [behaves differently across] variables race or relationship. Your code should produce following plots:





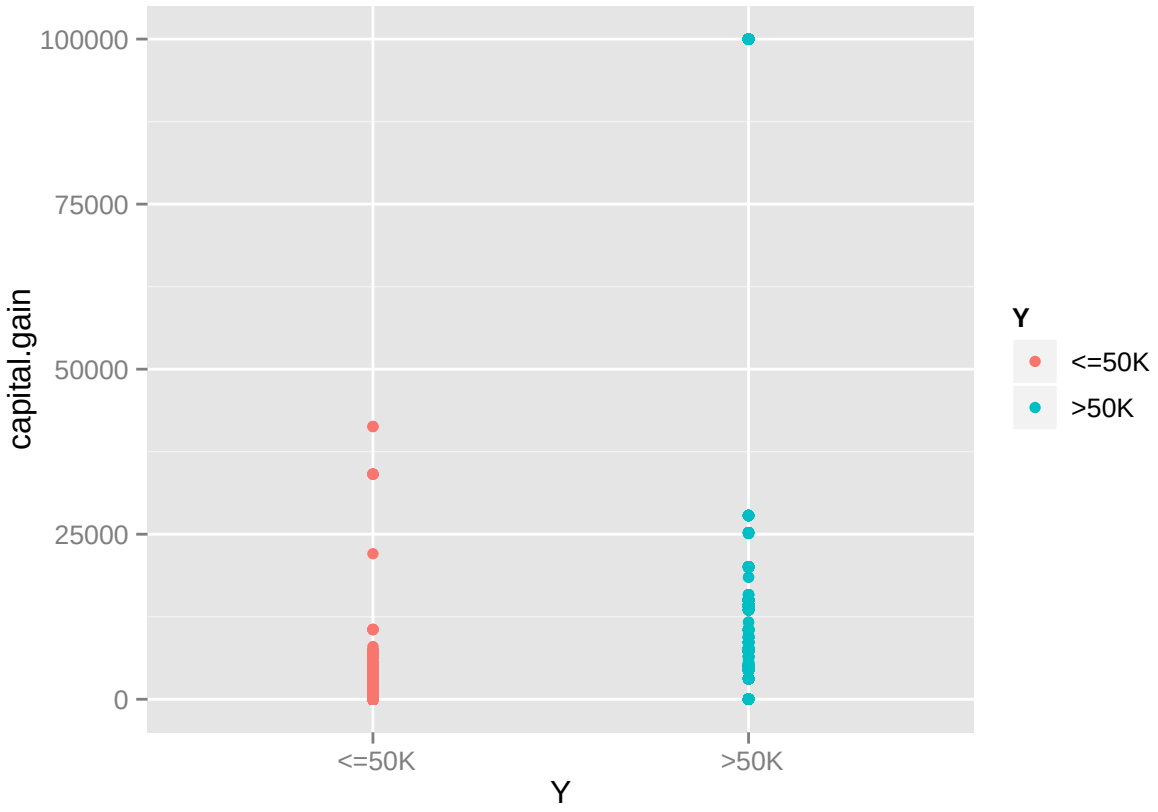
Note : You'll have to explore option position in geom_bar.

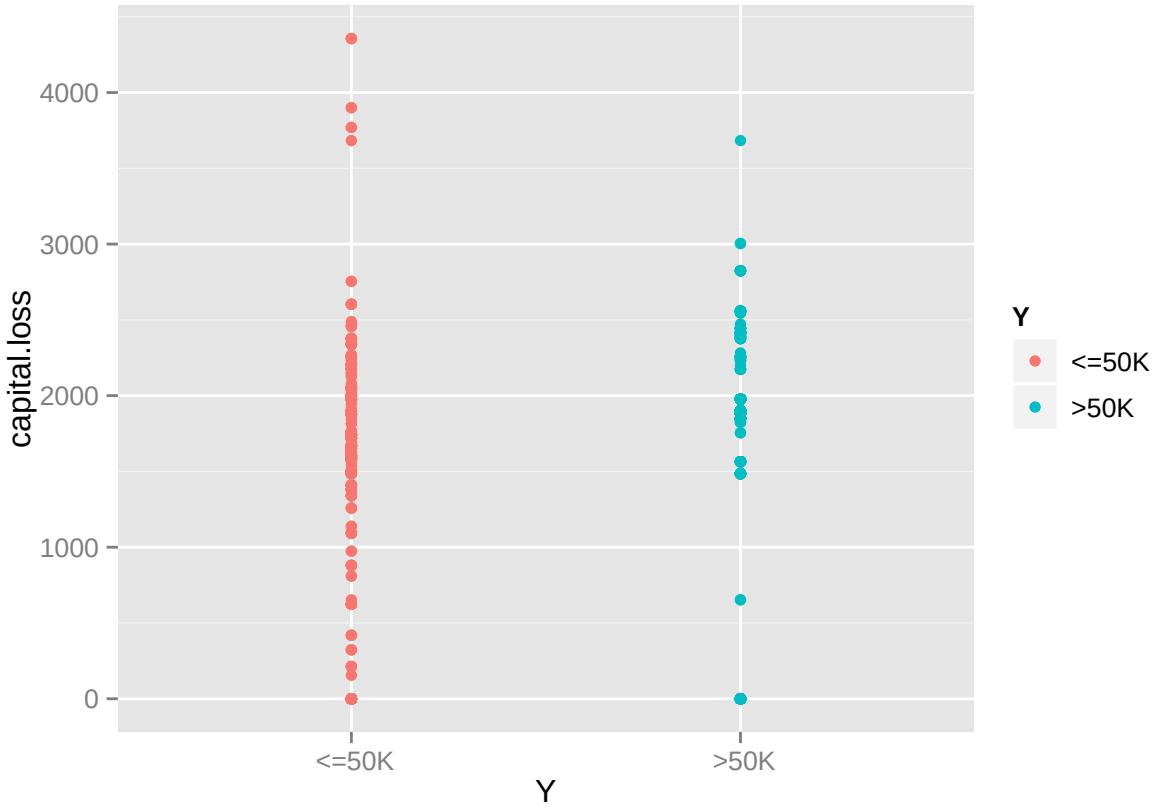
Use smoothing to see natural patterns

Show visually that hours per week vary similarly across age for both the sexes. Your code should produce following plot:

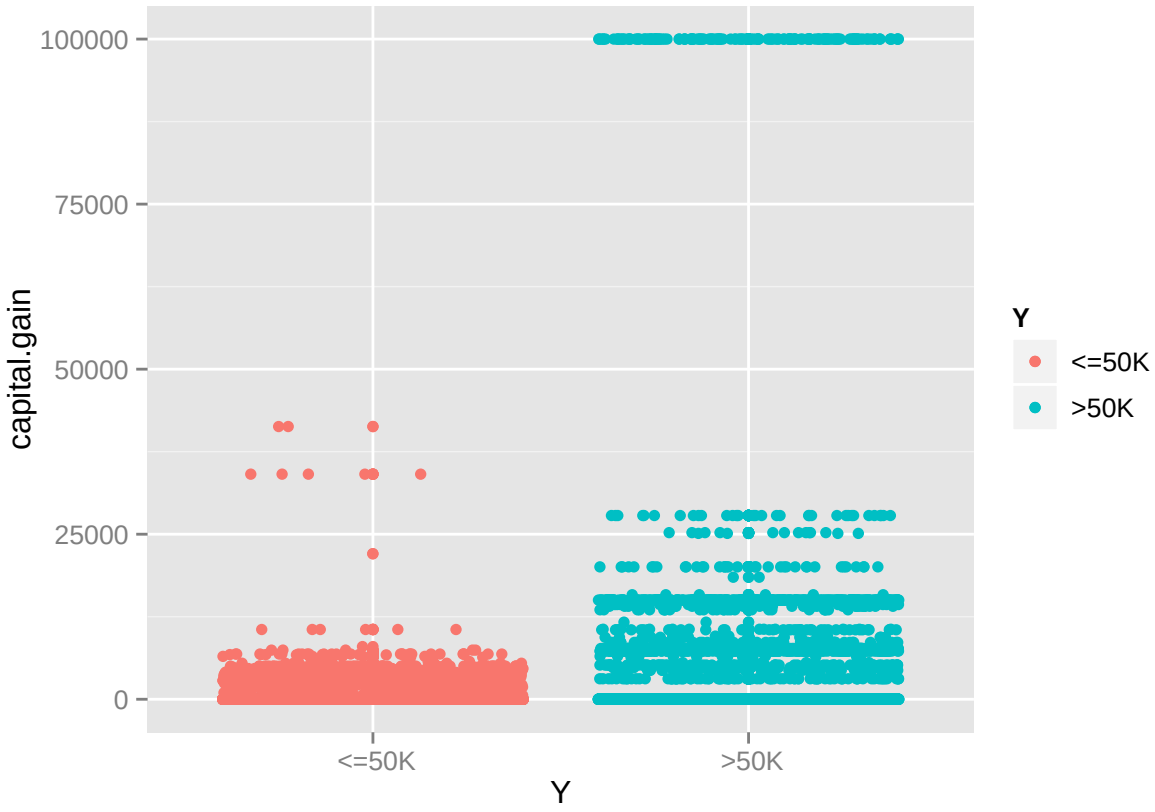
Using point plots with jitter instead of boxplots

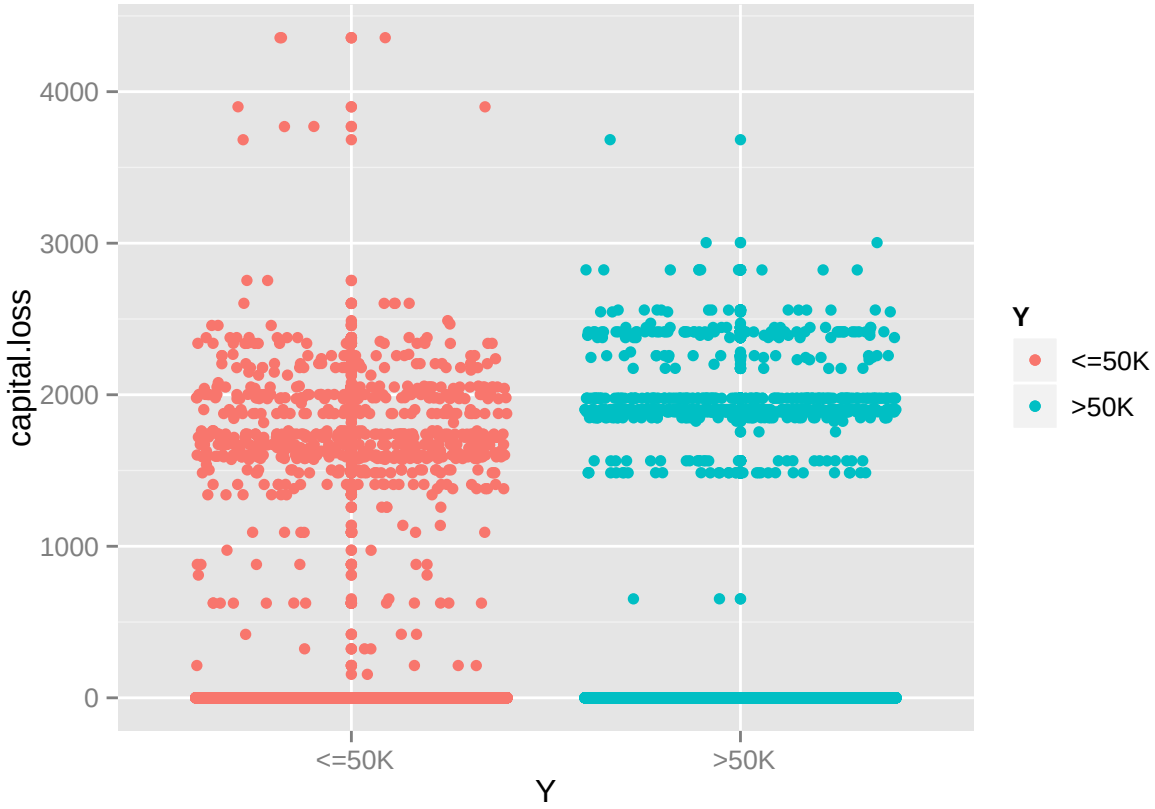
Prepare following plots for capital.gain and capital.loss with outcome Y to examine their behaviour across both outcomes.





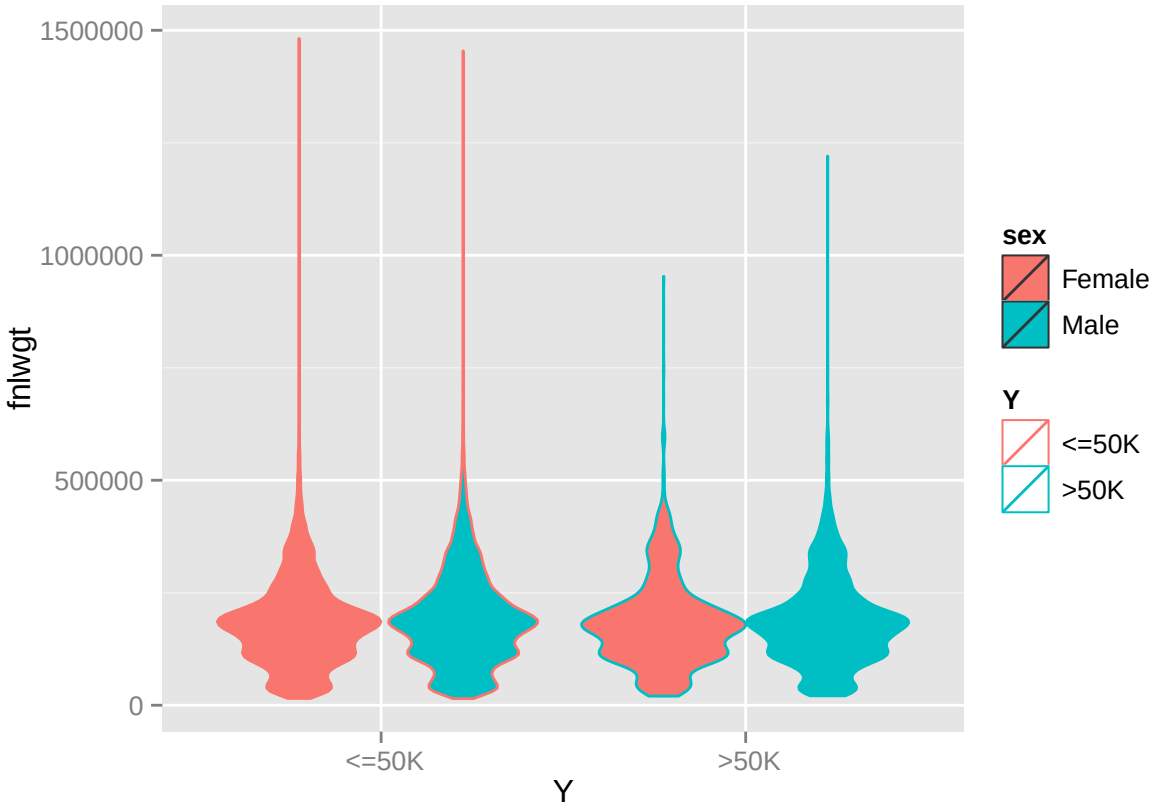
As you can see these plots give a fair idea about ranges of capital.gain and capital.loss but tell nothing about density or count of observations at certain levels which is a crucial set of information. We can remove this drawback by adding some jitter to the point positions. you code should prdocue following plots:


100000
75000
50000
25000
0
capital.gain
<=50K
>50K
Y
Y
<=50K
>50K


4000
3000
2000
1000
0
capital.loss
<=50K
>50K
Y
Y
<=50K
>50K

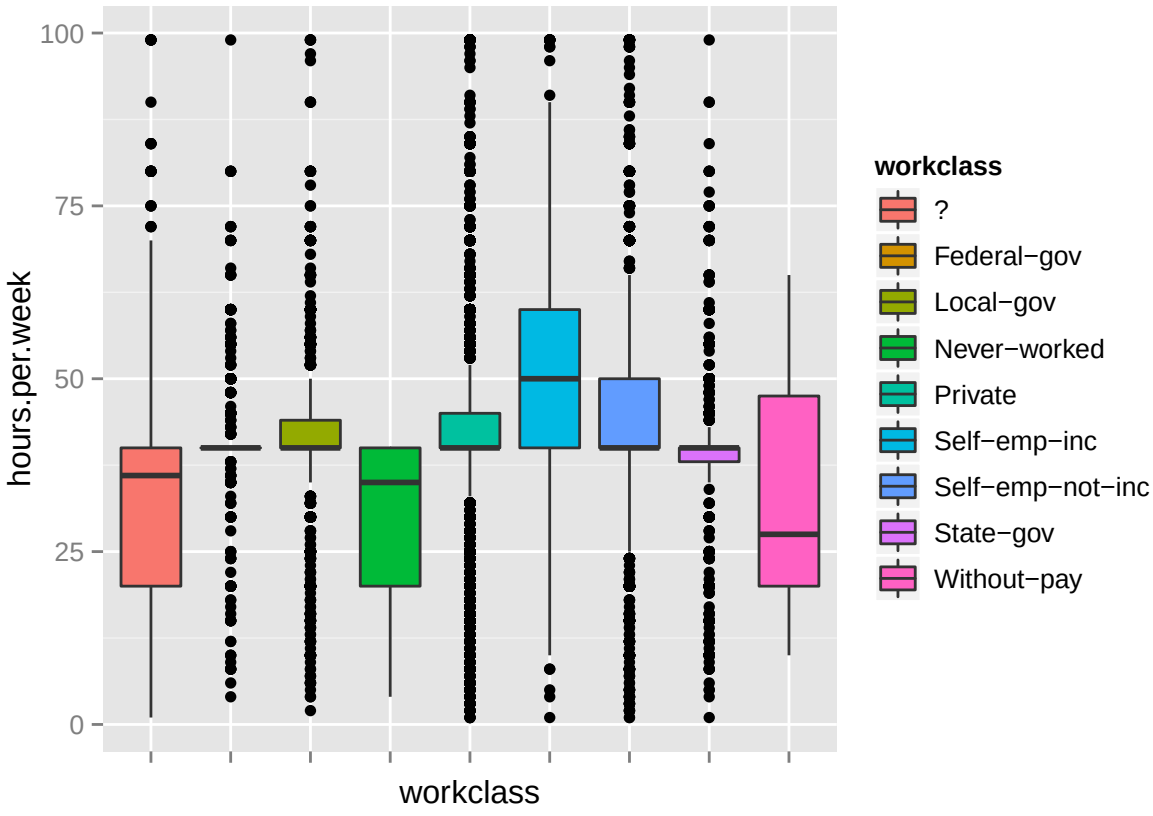
Violin Plots

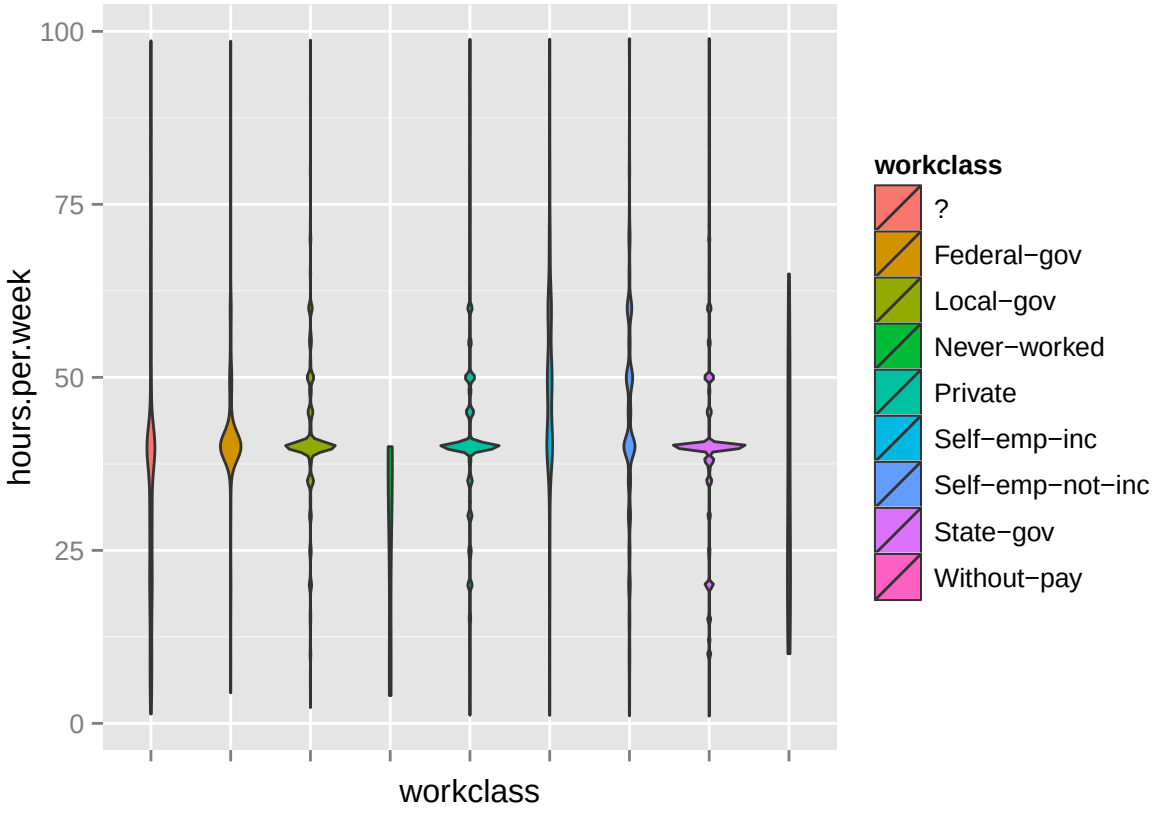
examine behaviour of fnlwgt for classes of Y [and within that for both the sexes]with . your code should produce following plots:



Which visualisation to select

Examine hours.per.week by workclass. which one will be a better choice , a box plot , a violin plot or simple point plot with jitter, think and discuss pros and cons associated with each.


100
75
50
25
0
hours.per.week
workclass
workclass
?
Federal−gov
Local−gov
Never−worked
Private
Self−emp−inc
Self−emp−not−inc
State−gov
Without−pay


100
75
50
25
0
hours.per.week
workclass
workclass
?
Federal−gov
Local−gov
Never−worked
Private
Self−emp−inc
Self−emp−not−inc
State−gov
Without−pay


100
75
50
25
0
hours.per.week
workclass
workclass
?
Federal−gov
Local−gov
Never−worked
Private
Self−emp−inc
Self−emp−not−inc
State−gov
Without−pay