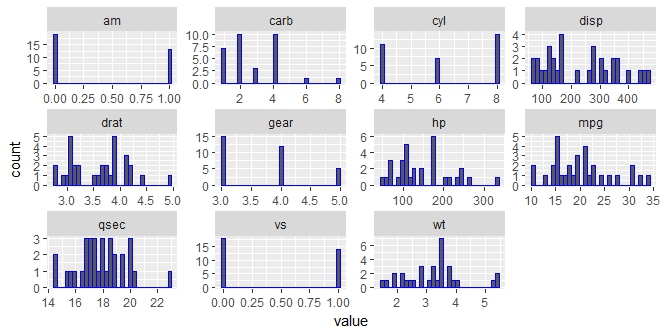
# Problem Statement

1. Histogram for all variables in a dataset mtcars. Write a program to create histograms for all columns.

|  |
| --- |
| > View(mtcars)  > library(purrr)  > library(tidyr)  > library(ggplot2)  > mtcars %>%  + keep(is.numeric) %>%  + gather() %>%  + ggplot(aes(value)) +  + facet\_wrap(~ key,scales = "free") +  + geom\_histogram(col="blue")  `stat\_bin()` using `bins = 30`. Pick better value with `binwidth`. |
|  |
| |  | | --- | | > | |



|  |
| --- |
| > #or we can do individually too  > #Histogram for all variables in a dataset mtcars  > hist(mtcars$mpg ,xlab = "Mpg", ylab = "Frequency",main="Histogram of Mpg",col="red")  > hist(mtcars$cyl ,xlab = "cyl", ylab = "Frequency",main="Histogram of cyl",col="blue")  > hist(mtcars$disp ,xlab = "disp", ylab = "Frequency",main="Histogram of disp",col="yellow")  > hist(mtcars$hp ,xlab = "hp", ylab = "Frequency",main="Histogram of hp",col="darkblue")  > hist(mtcars$drat ,xlab = "drat", ylab = "Frequency",main="Histogram of drat",col="pink")  > hist(mtcars$wt ,xlab = "wt", ylab = "Frequency",main="Histogram of wt",col="purple")  > hist(mtcars$qsec ,xlab = "qsec", ylab = "Frequency",main="Histogram of qsec",col="blue")  > hist(mtcars$vs ,xlab = "vs", ylab = "Frequency",main="Histogram of vs",col="green")  > hist(mtcars$am ,xlab = "am", ylab = "Frequency",main="Histogram of am",col="grey")  > hist(mtcars$gear ,xlab = "gear", ylab = "Frequency",main="Histogram of gear",col="blue")  > hist(mtcars$carb ,xlab = "carb", ylab = "Frequency",main="Histogram of carb",col="red") |
|  |
| |  | | --- | | > | |

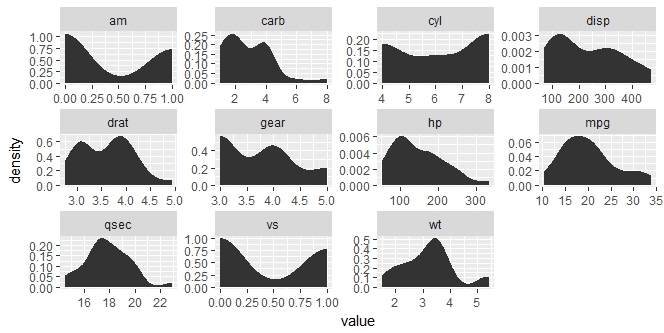
1. Check the probability distribution of all variables in mtcars

|  |
| --- |
| > library(ggplot2)  > mtcars %>%  + keep(is.numeric) %>%  + gather() %>%  + ggplot(aes(value)) +  + facet\_wrap(~ key,scales = "free") +  + stat\_density() |
|  |
| |  | | --- | |  | |

#stat\_bin is suitable only for continuous x data.

#If our x data is discrete, we probably want to use stat\_count.

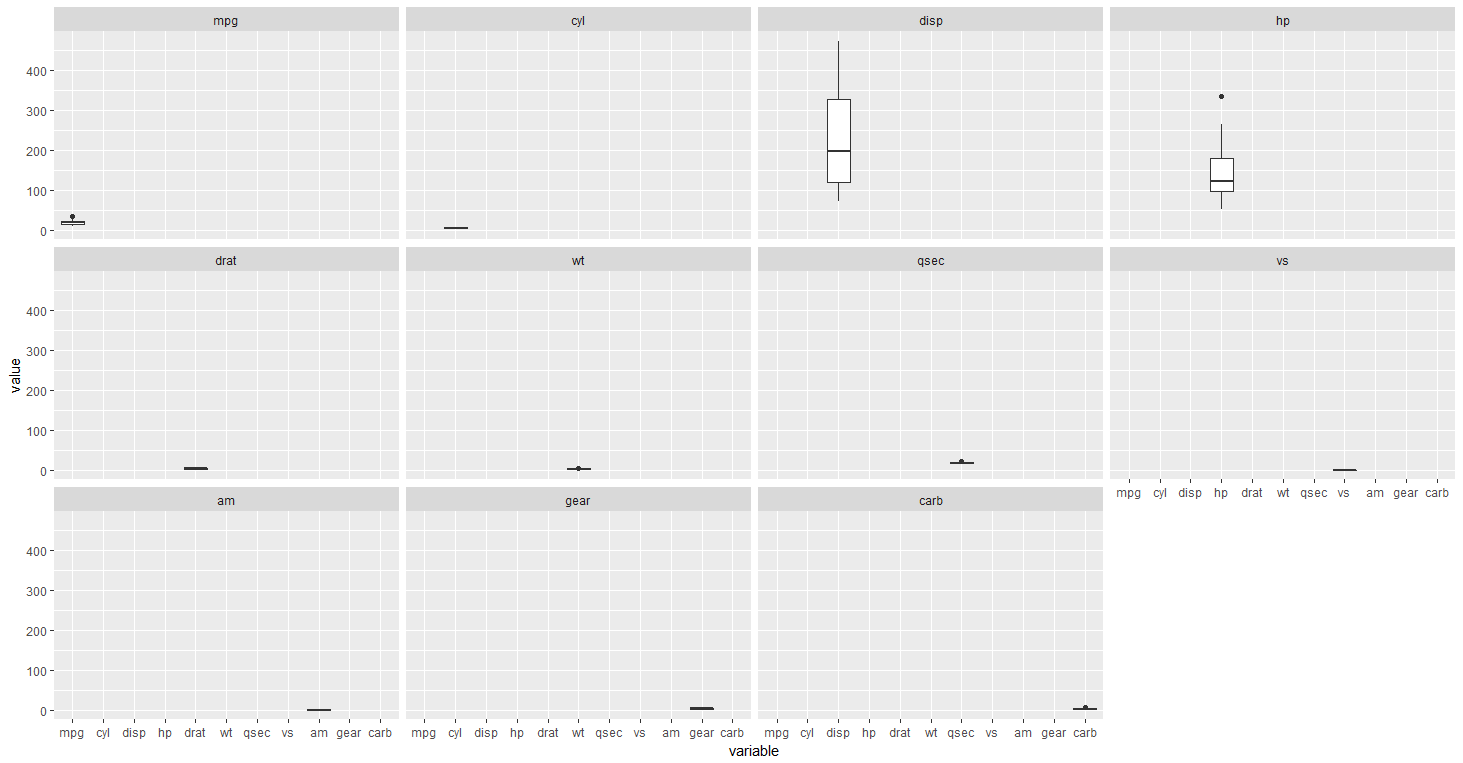
#we can also use geom\_density function



|  |
| --- |
| > #or like this too  > #we just do freq=FALSE and we get the Probability Distribution/Density of our variables  > hist(mtcars$mpg ,freq = F,xlab = "Mpg", ylab = "Probability Distribution/Density",main="Histogram of Mpg",col="red")  > hist(mtcars$cyl ,freq = F,xlab = "cyl", ylab = "Probability Distribution/Density",main="Histogram of cyl",col="blue")  > hist(mtcars$disp ,freq = F,xlab = "disp", ylab = "Probability Distribution/Density",main="Histogram of disp",col="yellow")  > hist(mtcars$hp ,freq = F,xlab = "hp", ylab = "Probability Distribution/Density",main="Histogram of hp",col="darkblue")  > hist(mtcars$drat ,freq = F,xlab = "drat", ylab = "Probability Distribution/Density",main="Histogram of drat",col="pink")  > hist(mtcars$wt ,freq = F,xlab = "wt", ylab = "Probability Distribution/Density",main="Histogram of wt",col="purple")  > hist(mtcars$qsec ,freq = F,xlab = "qsec", ylab = "Probability Distribution/Density",main="Histogram of qsec",col="blue")  > hist(mtcars$vs ,freq = F,xlab = "vs", ylab = "Probability Distribution/Density",main="Histogram of vs",col="green")  > hist(mtcars$am ,freq = F,xlab = "am", ylab = "Probability Distribution/Density",main="Histogram of am",col="grey")  > hist(mtcars$gear ,freq = F,xlab = "gear", ylab = "Probability Distribution/Density",main="Histogram of gear",col="blue")  > hist(mtcars$carb ,freq = F,xlab = "carb", ylab = "Probability Distribution/Density",main="Histogram of carb",col="red") |
|  |
| |  | | --- | | > | |

1. Write a program to create boxplot for all variables.

|  |
| --- |
| > library(ggplot2)  > library(reshape)  Attaching package: ‘reshape’  The following objects are masked from ‘package:tidyr’:  expand, smiths  > m1 <- melt(mtcars)  Using as id variables  > ggplot(m1,aes(x = variable,y = value)) + facet\_wrap(~variable) + geom\_boxplot() |
|  |
| |  | | --- | | > | |



|  |
| --- |
| > #or like this too individually  > boxplot(mtcars$mpg ,xlab = "Box plot", ylab = "Mpg",main="Box plot of Mpg",horizontal = T,col="red")  > boxplot(mtcars$cyl ,xlab = "Box plot", ylab = "cyl",main="Box plot of cyl",horizontal = T,col="blue")  > boxplot(mtcars$disp ,xlab = "Box plot", ylab = "disp",main="Box plot of disp",horizontal = T,col="yellow")  > boxplot(mtcars$hp ,xlab = "Box plot", ylab = "hp",main="Box plot of hp",horizontal = T,col="darkblue")  > boxplot(mtcars$drat ,xlab = "Box plot", ylab = "drat",main="Box plot of drat",horizontal = T,col="pink")  > boxplot(mtcars$wt ,xlab = "Box plot", ylab = "wt",main="Box plot of wt",horizontal = T,col="purple")  > boxplot(mtcars$qsec ,xlab = "Box plot", ylab = "qsec",main="Box plot of qsec",horizontal = T,col="blue")  > boxplot(mtcars$vs ,xlab = "Box plot", ylab = "vs",main="Box plot of vs",horizontal = T,col="green")  > boxplot(mtcars$am ,xlab = "Box plot", ylab = "am",main="Box plot of am",horizontal = T,col="grey")  > boxplot(mtcars$gear ,xlab = "Box plot", ylab = "gear",main="Box plot of gear",horizontal = T,col="blue")  > boxplot(mtcars$carb ,xlab = "Box plot", ylab = "carb",main="Box plot of carb",horizontal = T,col="red") |
|  |
| |  | | --- | | > | |