Error bar Plot, Error bars are visual representations of the variability of data and used on graphs to suggest the error in a reported measurement.

They give a general idea of how precise a measurement is, or conversely, how far from the reported value the true value might be.

Error bars regularly constitute one standard deviation uncertainty, one standard error, or a 95% confidence interval.

Error bar mainly communicates how the data spread around the mean, for example, a small sd bar indicates lower spread, and a higher sd indicates higher spread. In other words, a smaller sd indicates more reliability and a higher sd indicates less reliability.

The standard deviation error bars on a chart can be used to get an idea for significant differences exists or not.

When standard deviation error bars overlap quite a bit, it provides a hint that the difference is not statistically significant.

When standard deviation error bars overlap even less, it provides the hint that the difference is probably not statistically significant.

When standard deviation error bars do not overlap, it provides the hint that the difference may be significant, but cannot be sure.

Before making the decision based on an error bar chart, one need to perform a statistical test to draw a conclusion.

Standard deviation is the measure of the variability, for testing the significant difference sample size also needs to account.

Recommended, need to perform an appropriate statistical test to draw a conclusion about significant differences.

Here we are going to discuss how to create error bar plots with help of ggplot.

Following function will help us to summarize the dataset.

data\_summary <- function(data, varname, groupnames){

require(plyr)

summary\_func <- function(x, col){

c(mean = mean(x[[col]], na.rm=TRUE),

sd = sd(x[[col]], na.rm=TRUE))

}

data\_sum<-ddply(data, groupnames, .fun=summary\_func,

varname)

data\_sum <- rename(data\_sum, c("mean" = varname))

return(data\_sum)

}

df2 <- data\_summary(ToothGrowth, varname="len",

groupnames=c("supp", "dose"))

df2$dose=as.factor(df2$dose)

head(df2)

supp dose len sd

1 OJ 0.5 13.23 4.459709

2 OJ 1 22.70 3.910953

3 OJ 2 26.06 2.655058

4 VC 0.5 7.98 2.746634

5 VC 1 16.77 2.515309

6 VC 2 26.14 4.797731

**Error Bar Plot**

library(ggplot2)

p<- ggplot(df2, aes(x=dose, y=len, fill=supp)) +

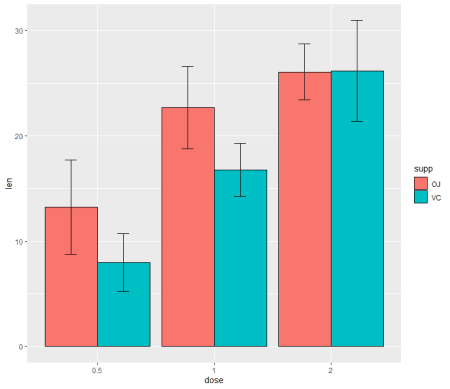
geom\_bar(stat="identity", color="black",

position=position\_dodge()) +

geom\_errorbar(aes(ymin=len-sd, ymax=len+sd), width=.2,

position=position\_dodge(.9))

print(p)



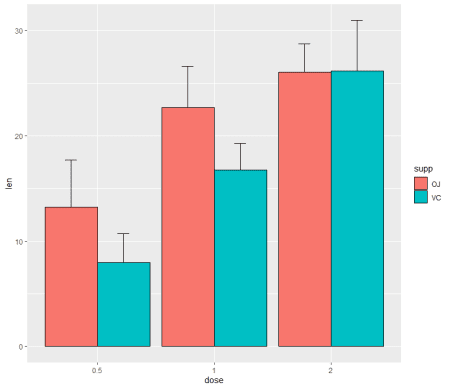
You can create plot based on only upper error bars

ggplot(df2, aes(x=dose, y=len, fill=supp)) +

geom\_bar(stat="identity", color="black", position=position\_dodge()) +

geom\_errorbar(aes(ymin=len, ymax=len+sd), width=.2,

position=position\_dodge(.9))



**Line Error Bar Plot**

p<- ggplot(df2, aes(x=dose, y=len, group=supp, color=supp)) +

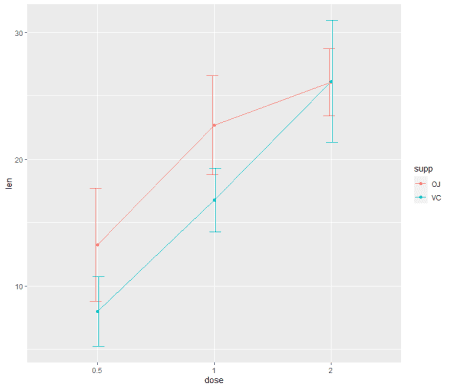
geom\_line() +

geom\_point()+

geom\_errorbar(aes(ymin=len-sd, ymax=len+sd), width=.2,

position=position\_dodge(0.05))

print(p)



**Error Bar Plot with CI**

ggplot(df2, aes(x = dose, y = len, fill = supp)) +

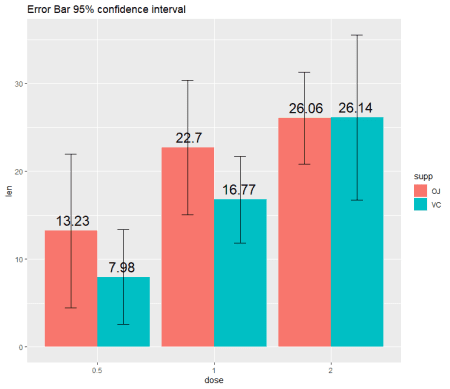
geom\_bar(stat = "identity", position = "dodge") +

ggtitle("Error Bar 95% confidence interval") + ylab("len") +

geom\_errorbar(aes(ymin = len - 1.96 \* sd, ymax = len + 1.96 \* sd), width = 0.2, position = position\_dodge(0.9)) +

geom\_text(aes(label = round(len, 2)), size = 6,

position = position\_dodge(0.85), vjust = -0.5)



Multiple R functions available to calculate sd, se and mean, so you can make utilize the same.

Based on aggregate function also can summarize the data set easily.

mydata<-ToothGrowth

stderr <- function(x) sqrt(var(x)/length(x))

datatoplotse<-aggregate(mydata$len, list(mydata$supp, mydata$dose), stderr)

datatoplotsd<-aggregate(mydata$len, list(mydata$supp, mydata$dose), sd,na.rm=TRUE)

datatoplot<-aggregate(mydata$len, list(mydata$supp, mydata$dose), mean,na.rm=TRUE)