library(ggplot2)

setwd("~/Documents/GitHub/dsc520")

load("data/Batting.Rdata")

load("data/People.Rdata")

triples <- c(Batting$X3B)

trip\_per <- c(Batting$X3B/Batting$H)

trip\_per <- na.omit(trip\_per)

triples

trip\_per

ggplot(Batting, aes(x=triples))+geom\_histogram(binwidth = 1)

ggplot(Batting, aes(x=trip\_per))+geom\_histogram()

height <- c(People$height)

weight <- c(People$weight)

height

weight

ggplot(People, aes(x=height))+geom\_histogram(binwidth = 1)

ggplot(People, aes(x=height))+geom\_density()

ggplot(People, aes(x=weight))+geom\_histogram(binwidth = 5)

ggplot(People, aes(x=weight))+geom\_density()

ggplot(Batting, aes(x=X3B))+stat\_ecdf(aes(colour=X3B))

ggplot(Batting, aes(x=X3B)) + stat\_function(fun = dlnorm)

summary(triples)

summary(trip\_per)

summary(height)

summary(weight)

combine <- c(Batting$playerID == People$playerID )

combine

comb <- merge(Batting, People, by="playerID")

comb\_clean <- subset(comb, select = c("playerID","H","X3B","weight", "height"))

head(comb\_clean)

trip\_hit <- comb\_clean$X3B/comb\_clean$H

ggplot(comb\_clean, aes(x=weight, y= X3B))+geom\_point()

ggplot(comb\_clean, aes(x=height, y= X3B))+geom\_point()

ggplot(comb\_clean, aes(x=height, y= X3B/H))+geom\_point()

ggplot(comb\_clean, aes(x=weight, y= X3B/H))+geom\_point()

cor.test(comb\_clean$height,comb\_clean$X3B, method=c("pearson"))

cov(comb\_clean$height,comb\_clean$X3B, method=c("pearson"))

cor.test(comb\_clean$weight,comb\_clean$X3B, method=c("pearson"))

cov(comb\_clean$weight,comb\_clean$X3B, method=c("pearson"))

cor.test(comb\_clean$weight,comb\_clean$X3B/comb\_clean$H, method=c("pearson"))

cov(comb\_clean$weight,comb\_clean$X3B/comb\_clean$H, method=c("pearson"))

cor.test(comb\_clean$height,comb\_clean$X3B/comb\_clean$H, method=c("pearson"))

cov(comb\_clean$height,comb\_clean$X3B/comb\_clean$H, method=c("pearson"))

t.test(comb\_clean$height,comb\_clean$X3B)

t.test(comb\_clean$weight,comb\_clean$X3B)

t.test(comb\_clean$height,comb\_clean$X3B/comb\_clean$H)

t.test(comb\_clean$weight,comb\_clean$X3B/comb\_clean$H)

fit <- lm(comb\_clean$X3B ~ comb\_clean$weight + comb\_clean$height + comb\_clean$X3B/comb\_clean$H, data=comb\_clean)

summary(fit)