CODE:

library(ggplot2)

library(dplyr)

data1 <- read.table(file.choose(), header=T, sep=",")

dataframe2 <- as.data.frame(data1)

dataodd1 <- read.table(file.choose(), header=T, sep=",")

dfodd1 <- as.data.frame(dataodd1)

plot1 <- qplot(dataframe2$Home\_Score,

geom="histogram",

binwidth=0.5,

main="Home Goals",

xlab="Home Goals",

ylab="Number of Games",

fill=I("red"),

col=I("black"),

alpha=I(0.2))

plot2 <- qplot(dataframe2$Away\_Score,

geom="histogram",

binwidth=0.5,

main="Away Goals",

xlab="Away Goals",

ylab="Number of Games",

fill=I("blue"),

col=I("black"),

alpha=I(0.2))

df2 <- dataframe2 %>%

mutate(ScoreDiff = Home\_Score-Away\_Score)

plot3 <- qplot(df2$ScoreDiff,

geom="histogram",

binwidth=0.5,

main="Home-Away Goals",

xlab="Home-Away Goals",

ylab="Number of Games",

fill=I("red"),

col=I("black"),

alpha=I(0.2))

dfodd2 <- dfodd1 %>%

mutate(Phomewin = 1/HomeOdd, Pawaywin = 1/AwayOdd, Ptie = 1/TieOdd)

dfodd3 <- dfodd2 %>%

mutate(Phomewinnormal = (Phomewin/(Phomewin+Pawaywin+Ptie)), Pawaywinnormal= (Pawaywin/(Phomewin+Pawaywin+Ptie)),

Ptienormal = (Ptie/(Phomewin+Pawaywin+Ptie)))

dfoddbetsson <- dfodd3 %>% filter(Bookmaker == "Betsson") %>% mutate(homeaway= Phomewin-Pawaywin, homeawaynorm= Phomewinnormal-Pawaywinnormal)

dfoddpinnacle <- dfodd3 %>% filter(Bookmaker == "Pinnacle") %>% mutate(homeaway= Phomewin-Pawaywin, homeawaynorm= Phomewinnormal-Pawaywinnormal)

dfoddbet365 <- dfodd3 %>% filter(Bookmaker == "bet365") %>% mutate(homeaway= Phomewin-Pawaywin, homeawaynorm= Phomewinnormal-Pawaywinnormal)

dfoddbwin <- dfodd3 %>% filter(Bookmaker == "bwin") %>% mutate(homeaway= Phomewin-Pawaywin, homeawaynorm= Phomewinnormal-Pawaywinnormal)

plot4 <- ggplot(dfoddbetsson, aes(x=homeaway) )+

geom\_line(aes(y= Ptie,colour="Bookmaker value"))+

geom\_line(aes(y=Ptienormal, colour="Normalized Value"))+

ylab(label= "Probability of a draw")+

xlab(label= "Probability of home team win - probability of away team win")+

ggtitle("Betsson")+

scale\_color\_manual(values=c("blue","red"))

plot5 <- ggplot(dfoddpinnacle, aes(x=homeaway) )+

geom\_line(aes(y= Ptie, colour="Bookmaker value"))+

geom\_line(aes(y=Ptienormal, colour="Normalized Value"))+

ylab(label= "Probability of a draw")+

xlab(label= "Probability of home team win - probability of away team win")+

ggtitle("Pinnacle")+

scale\_color\_manual(values=c("blue","red"))

plot6 <- ggplot(dfoddbet365, aes(x=homeaway) )+

geom\_line(aes(y= Ptie, colour="Bookmaker value"))+

geom\_line(aes(y=Ptienormal, colour="Normalized Value"))+

ylab(label= "Probability of a draw")+

xlab(label= "Probability of home team win - probability of away team win")+

ggtitle("Bet365")+

scale\_color\_manual(values=c("blue","red"))

plot7 <- ggplot(dfoddbwin, aes(x=homeaway), legend="TRUE" )+

geom\_line(aes(y= Ptie, colour="Bookmaker value"))+

geom\_line(aes(y=Ptienormal, colour="Normalized Value"))+

ylab(label= "Probability of a draw")+

xlab(label= "Probability of home team win - probability of away team win")+

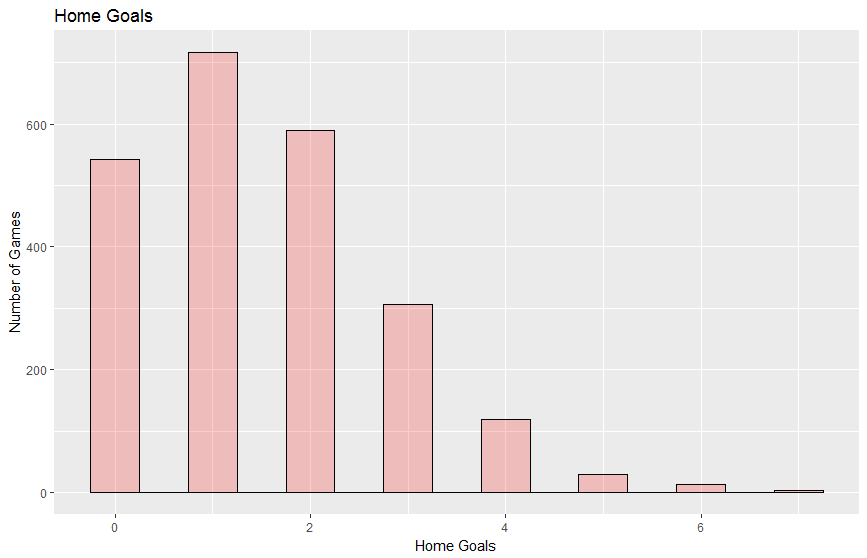
ggtitle("Bwin")+

scale\_color\_manual(values=c("blue","red"))

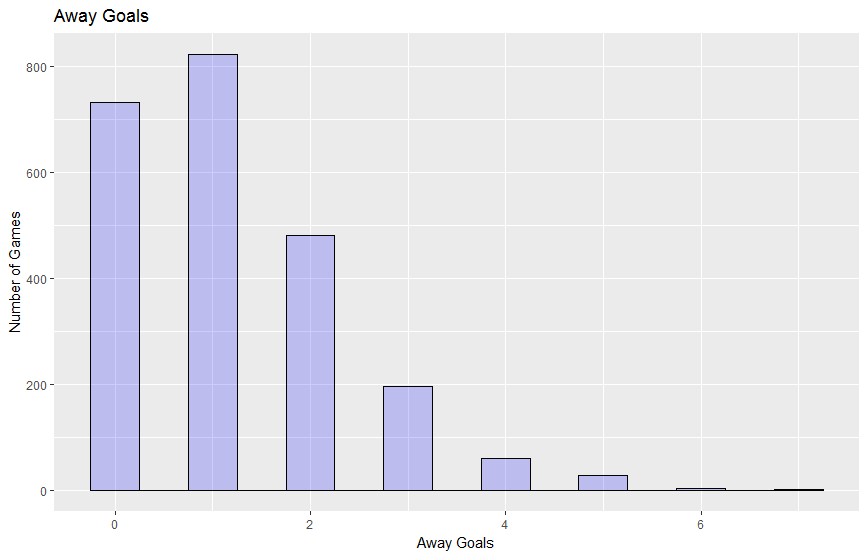
print(plot7)

TASK 1:

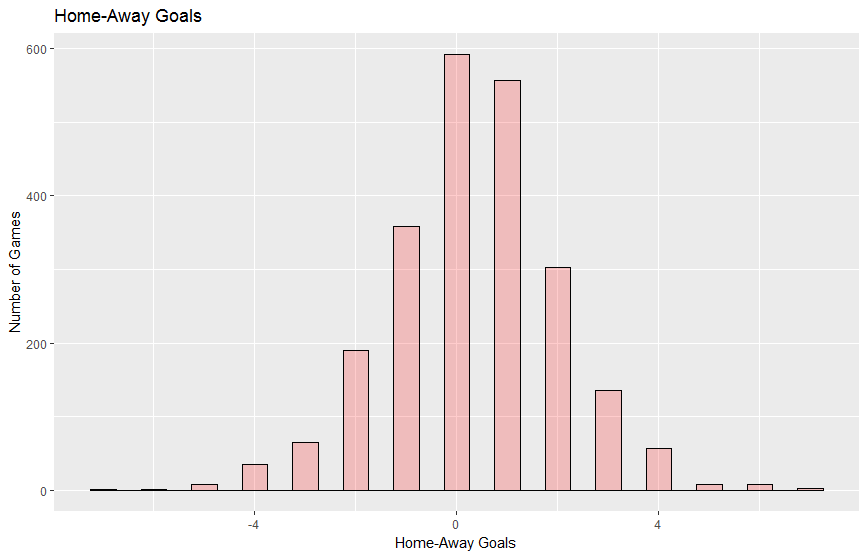
A)



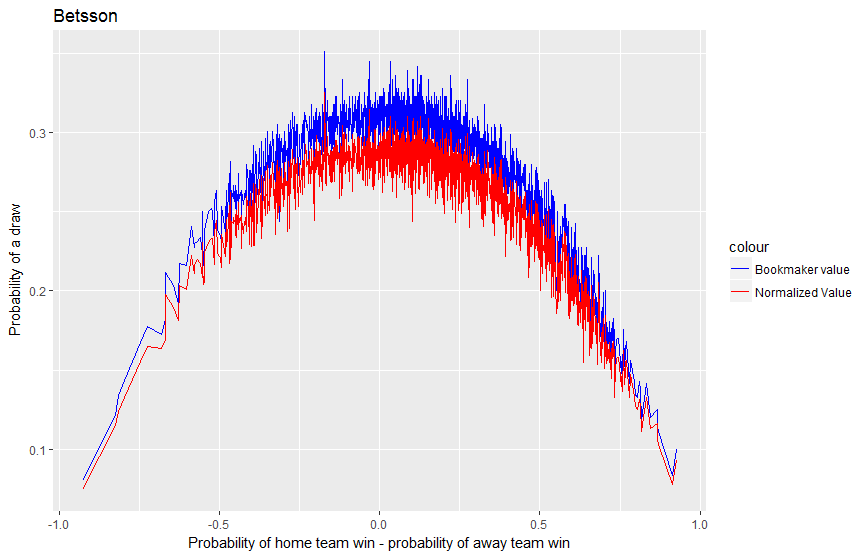
B)

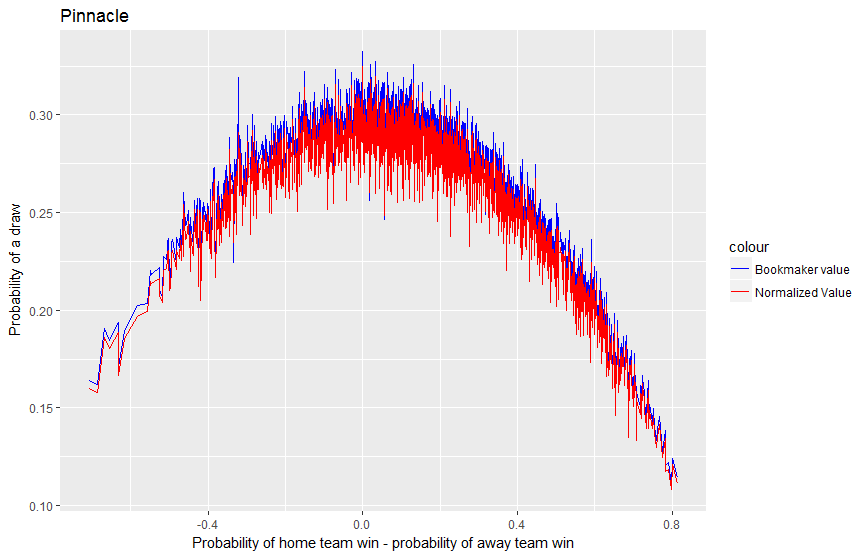


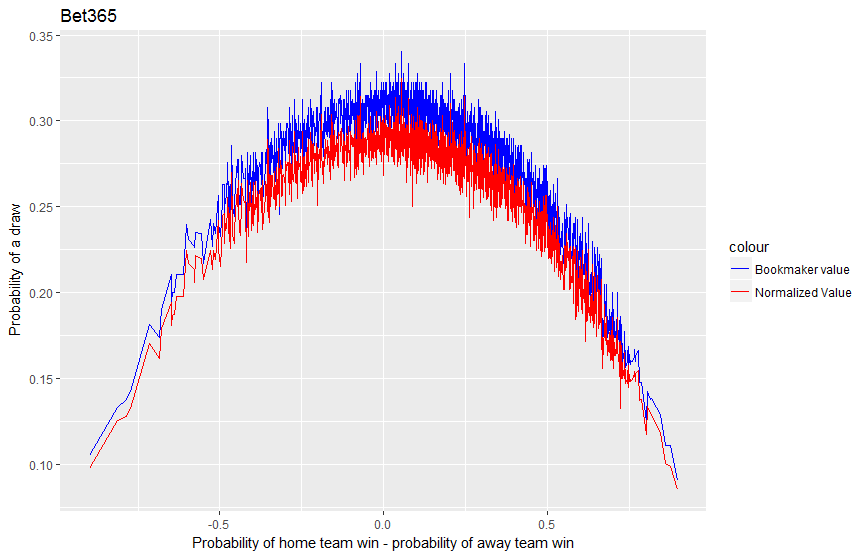
C)

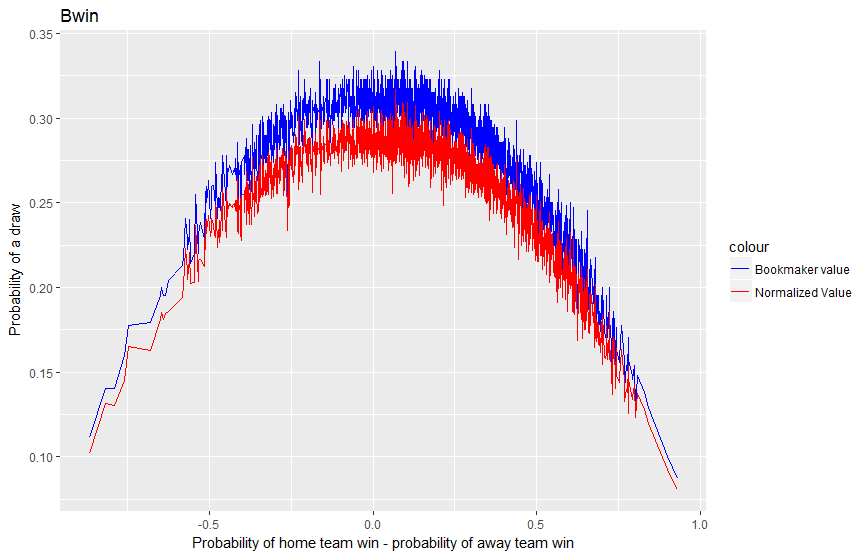


TASK 2:









Comment:

As we see from the graphs, except Pinacle all the bookmakers are adding significant margin to odds. This is consisted with if we are to bet constantly we would loose in the long term.

We could thinnk that Pinacle is a small company. It does lowers its margin to gain competitve advantage against other 3 competitors.