```r
# installing the important packages
install.packages("dplyr")
install.packages("ggplot2")

# importing the important packages
library("dplyr")
library("ggplot2")

# loading the datset
df = read.csv("C:\\Users\\JADHAV\\Desktop\\R_project\\Shark Tank India Dataset.csv")

# viewing the top 5 records in the dataset
head(df, 5)

# Viewing the bottom 5 records in the dataset
tail(df, 5)

# datatype of the data
str(df)

# Summary of the data
df%>%summary

# renaming the deal column
df%>%rename("dealdone" = "deal")%>%head(5)

#Creating new column for asked valuation - dealed valuation as difference_valuation
df%>%mutate(difference_valuation = ask_valuation - deal_valuation)%>%head(5)

# Deal done out of the all deals visualization
# pie chart
x = c(df%>%filter(deal == 1)%>%nrow,df%>%filter(deal == 0)%>%nrow)
pp = round(100*x/sum(x), 1)
labels = c(paste("done",pp[1], "%"), paste("not done", pp[2], "%"))
pie(x = x, labels = labels, radius = 1, main = "How many Deals are done", col = c("green", "red"), clockwise = TRUE)

# bar chart
barplot(x, xlab = "Deal", ylab = "count", names.arg = c("done", "not done"), col = c("green", "red"), main = "How may deals are do

# Deals done by each shark in the show
label = c(names(df%>%select(20:25)))
h =c()
for (i in label){
cn = df[i]%>%group_by_all()%>%count
h = append(h, cn[2,2])}
barplot(unlist(h), xlab = "Sharks", ylab = "deals count", names.arg = label, main = "Deals done by each shark")

# amount spend by each shark
a=df%>%select(c("anupam_deal", "amount_per_shark"))%>%filter(anupam_deal == 1)%>%summarize(sum(amount_per_shark))
a2 = df%>%select(c("aman_deal", "amount_per_shark"))%>%filter(aman_deal == 1)%>%summarize(sum(amount_per_shark))
n = df%>%select(c("namita_deal", "amount_per_shark"))%>%filter(namita_deal == 1)%>%summarize(sum(amount_per_shark))
v = df%>%select(c("vineeta_deal", "amount_per_shark"))%>%filter(vineeta_deal == 1)%>%summarize(sum(amount_per_shark))
p = df%>%select(c("peyush_deal", "amount_per_shark"))%>%filter(peyush_deal == 1)%>%summarize(sum(amount_per_shark))
```

```
g = df%>%select(c("ghazal_deal", "amount_per_shark"))%>%filter(ghazal_deal == 1)%>%summarize(sum(amount_per_shark))

amm = c(a[1,1], a2[1,1], n[1,1], v[1,1], p[1,1], g[1,1])
barplot(unlist(amm), xlab = "Sharks", ylab = "amount spend", names.arg = label, main = "Amount Spend by each shark")


# distribution of the data in pitcher_ask_amount

v = df$"pitcher_ask_amount"
hist(v, xlab = i, col = "yellow", border = "blue")

# distribution of the data in "ask_valuation"
v = df$"ask_valuation"
hist(v, xlab = "ask_valuation", col = "yellow", border = "blue")

# Lets see the relation between the deal_equity, deal_valuation with the help of scatter plot
ggplot(df, aes(x= deal_equity, y=deal_valuation)) + geom_point()
```