name: Malhar Dharmadhikari
reg no: 20BCE2110

df <- faithful
df_train <- data.frame("eruptions"=df$eruptions[0:260],"waiting"=df$waiting[0:260])
df_test <- data.frame("eruptions"=df$eruptions[261:272], "waiting"=df$waiting[261:272])

chart A
```{r}
reg <- lm(df_train)
q2 = data.frame("waiting"=df_test$waiting, "prediction"= predict(reg, df_test))
plot(q2, pch=20, main="Malhar Dharmadhikari 20BCE2110")
```

chart B
barplot(rbind(df_test$eruptions, q2$prediction), beside = TRUE
,xlab = "index", ylab = "eruption/prediction", col=c("darkblue", "red"), ylim=c(0,7),
main="Malhar Dharmadhikari 20BCE2110")
legend("topright",
legend = c("Test Case", "Prediction"),
fill = c("darkblue", "red"))

chart C
text_arr = c("predictios are higher than actual when passed value is low",
"predictios are lower than actual when passed value is high",
"eruption duration is directly proportional to waiting time")
text_arr

chart D

library("tm")
library("SnowballC")
library("RColorBrewer")
library("wordcloud")
#library(wordcloud2)

text = "predictios are higher than actual when passed value is low predictios are lower than actual when passed value is high eruption duration is directly proportional to waiting time"
#text = read.delim(text_arr)
docs = Corpus(VectorSource(text))
toSpace = content_transformer(function (x, pattern)gsub(pattern, " ", x))
docs1 = tm_map(docs, toSpace, "/")
docs1 = tm_map(docs, toSpace, "@")
docs1 = tm_map(docs, toSpace, "#")

docs1 = tm_map(docs1, content_transformer(tolower))
docs1 = tm_map(docs1, removeNumbers)
docs1 = tm_map(docs1, stripWhitespace)

```
dtm = TermDocumentMatrix(docs)
m = as.matrix(dtm)
v = sort(rowSums(m), decreasing = TRUE)
d = data.frame(word = names(v), freq = v)

wordcloud(words = d$word,
          freq = d$freq,
          min.freq = 1,
          max.words = 2000,
          random.order = FALSE,
          rot.per = 0.35,
          colors = brewer.pal(8, "Dark2"))
```