

Term work 4

```
#required packages
install.packages("caret")
install.packages("Pdyr")
# Auach the dataset to the environment
data(iris)
help(iris)
iris_dataset<-iris
View(iris_dataset)
head(iris_dataset,7)
colnames(iris_dataset)<-
  c("Sepal.Length","Sepal.Width","Petal.Length","Petal.Width","Species")
head(iris_dataset,5)
library(caret)
# We use the dataset to create a parPPon (80% training 20% tesPng)
index <- createDataPartition(iris_dataset$Species, p=0.80, list=FALSE)
testset <- iris_dataset[-index,]
trainset <- iris_dataset[index,]
dim(trainset)
str(trainset)
summary(trainset)
# Levels of the predicPon column
levels(trainset$Species)
hist(trainset$Sepal.Width)
par(mfrow=c(1,4))
for(i in 1:4) {
  boxplot(trainset[,i], main=names(trainset)[i])
}
install.packages("ggplot2")
library(ggplot2)
# Scauer plot
g <- ggplot(data=trainset, aes(x = Petal.Length, y = Petal.Width))
print(g)
g <-g +
  geom_point(aes(color=Species, shape=Species)) +
  xlab("Petal Length") +
  ylab("Petal Width") +
  ggtitle("Petal Length-Width")+
  geom_smooth(method="lm")
print(g)
## Box Plot
box <- ggplot(data=trainset, aes(x=Species, y=Sepal.Length)) +
  geom_boxplot(aes(fill=Species)) +
```

```

ylab("Sepal Length") +
ggPtle("Iris Boxplot") +
stat_summary(fun.y=mean, geom="point", shape=5, size=4)
print(box)
library(ggthemes)
## Histogram
histogram <- ggplot(data=iris, aes(x=Sepal.Width)) +
  geom_histogram(binwidth=0.2, color="black", aes(fill=Species)) +
  xlab("Sepal Width") +
  ylab("Frequency") +
  ggtitle("Histogram of Sepal Width")
print(histogram)
## FacePng: Producing mulPple charts in one plot
library(ggthemes)
facet <- ggplot(data=trainset, aes(Sepal.Length, y=Sepal.Width, color=Species))+
  geom_point(aes(shape=Species), size=1.5) +
  geom_smooth(method="lm") +
  xlab("Sepal Length") +
  ylab("Sepal Width") +
  ggtitle("FacePng") +

  facet_grid(. ~ Species) # Along rows
print(facet)

```