Attention: three persons worked on this script, therefore file names may differ.

setwd("C:\\Dokumente und Einstellungen\\csae2104\\Desktop\\PhD\\PhD-Kurse\\Adv. Stat. (raw data)\\Vortrag\\")

data <- read.csv(file="C:\\Dokumente und Einstellungen\\csae2104\\Desktop\\PhD\\PhD-Kurse\\Adv. Stat. (raw data)\\Vortrag\\data\_paper.csv",sep=";",dec=",",header=TRUE)

attach(data)

group1 <- subset(data,CLU4\_Pcord=="1")

group2 <- subset(data,CLU4\_Pcord=="17")

group3 <- subset(data,CLU4\_Pcord=="27")

group4 <- subset(data,CLU4\_Pcord=="49")

par(mfrow=c(2,2))

attach(group1)

hist(altitude,breaks="Sturges",xlim=c(500,3000),ylim=c(0,100),col="lightgreen",main="Altitude in group 1",xlab="Altitude (m)")

attach(group2)

hist(altitude,breaks="Sturges",xlim=c(1500,3000),ylim=c(0,40),col="lightgreen",main="Altitude in group 2",xlab="Altitude (m)")

attach(group3)

hist(altitude,breaks="Sturges",xlim=c(1400,2200),ylim=c(0,15),col="lightgreen",main="Altitude in group 3",xlab="Altitude (m)")

attach(group4)

hist(altitude,breaks="Sturges",xlim=c(1400,2400),ylim=c(0,15),col="lightgreen",main="Altitude in group 4",xlab="Altitude (m)")

library(calibrate)

attach(data)

model\_alt <- lm(No\_of\_sp~altitude)

qqnorm(model\_alt$residuals)

qqline(model\_alt$residuals,col="red")

attach(group1)

mean(altitude)

median(altitude)

sd(altitude)

var(altitude)

attach(group2)

mean(altitude)

median(altitude)

sd(altitude)

var(altitude)

attach(group3)

mean(altitude)

median(altitude)

sd(altitude)

var(altitude)

attach(group4)

mean(altitude)

median(altitude)

sd(altitude)

var(altitude)

par(mfrow=c(2,2))

attach(group1)

hist(angle,breaks="Sturges",xlim=c(0,50),ylim=c(0,80),col="lightblue",main="Slope in group 1",xlab="Slope (°)")

attach(group2)

hist(angle,breaks="Sturges",xlim=c(0,40),ylim=c(0,40),col="lightblue",main="Slope in group 2",xlab="Slope (°)")

attach(group3)

hist(angle,breaks="Sturges",xlim=c(0,40),ylim=c(0,30),col="lightblue",main="Slope in group 3",xlab="Slope (°)")

attach(group4)

hist(angle,breaks="Sturges",xlim=c(0,50),ylim=c(0,15),col="lightblue",main="Slope in group 4",xlab="Slope (°)")

library(calibrate)

attach(data)

model\_slo <- lm(No\_of\_sp~angle)

qqnorm(model\_slo$residuals)

qqline(model\_slo$residuals,col="red")

attach(group1)

mean(angle)

median(angle)

sd(angle)

var(angle)

attach(group2)

mean(angle)

median(angle)

sd(angle)

var(angle)

attach(group3)

mean(angle)

median(angle)

sd(angle)

var(angle)

attach(group4)

mean(angle)

median(angle)

sd(angle)

var(angle)

par(mfrow=c(2,2))

attach(group1)

hist(pH\_2,breaks="Sturges",xlim=c(3.5,6.5),ylim=c(0,200),col="yellow",main="pH in group 1",xlab="pH")

attach(group2)

hist(pH\_2,breaks="Sturges",xlim=c(3.5,5.5),ylim=c(0,100),col="yellow",main="pH in group 2",xlab="pH")

attach(group3)

hist(pH\_2,breaks="Sturges",xlim=c(4,7.5),ylim=c(0,80),col="yellow",main="pH in group 3",xlab="pH")

attach(group4)

hist(pH\_2,breaks="Sturges",xlim=c(4.2,4.6),ylim=c(0,50),col="yellow",main="pH in group 4",xlab="pH")

library(calibrate)

attach(data)

model\_pH <- lm(No\_of\_sp~pH\_2)

qqnorm(model\_pH$residuals)

qqline(model\_pH$residuals,col="red")

plot(pH\_2,No\_of\_sp,xlab="pH",ylab="Plant species",main="Species abundance plotted against pH",type="p")

attach(group1)

mean(pH\_2)

median(pH\_2)

sd(pH\_2)

var(pH\_2)

attach(group2)

mean(pH\_2)

median(pH\_2)

sd(pH\_2)

var(pH\_2)

attach(group3)

mean(pH\_2)

median(pH\_2)

sd(pH\_2)

var(pH\_2)

attach(group4)

mean(pH\_2)

median(pH\_2)

sd(pH\_2)

var(pH\_2)

par(mfrow=c(2,2))

attach(group1)

hist(land\_use\_intensity,breaks="Sturges",xlim=c(0,3),ylim=c(0,150),col="brown",main="Land use intensity in group 1",xlab="Land use intensity")

attach(group2)

hist(land\_use\_intensity,breaks="Sturges",xlim=c(0,3),ylim=c(0,80),col="brown",main="Land use intensity in group 2",xlab="Land use intensity")

attach(group3)

hist(land\_use\_intensity,breaks="Sturges",xlim=c(0,3),ylim=c(0,60),col="brown",main="Land use intensity in group 3",xlab="Land use intensity")

attach(group4)

hist(land\_use\_intensity,breaks="Sturges",xlim=c(0,3),ylim=c(0,30),col="brown",main="Land use intensity in group 4",xlab="Land use intensity")

library(calibrate)

attach(data)

model\_int <- lm(No\_of\_sp~land\_use\_intensity)

qqnorm(model\_int$residuals)

qqline(model\_int$residuals,col="red")

attach(group1)

mean(land\_use\_intensity)

median(land\_use\_intensity)

sd(land\_use\_intensity)

var(land\_use\_intensity)

attach(group2)

mean(land\_use\_intensity)

median(land\_use\_intensity)

sd(land\_use\_intensity)

var(land\_use\_intensity)

attach(group3)

mean(land\_use\_intensity)

median(land\_use\_intensity)

sd(land\_use\_intensity)

var(land\_use\_intensity)

attach(group4)

mean(land\_use\_intensity)

median(land\_use\_intensity)

sd(land\_use\_intensity)

var(land\_use\_intensity)

par(mfrow=c(2,2))

attach(group1)

hist(land\_use\_4types,breaks=c(0,1,2,3,4),xlim=c(0,4),ylim=c(0,200),col="red",main="4 Land use types in group 1",xlab="Land use",labels=TRUE)

attach(group2)

hist(land\_use\_4types,breaks=c(0,1,2,3,4),xlim=c(0,4),ylim=c(0,50),col="red",main="4 Land use types in group 2",xlab="Land use",labels=TRUE)

attach(group3)

hist(land\_use\_4types,breaks=c(0,1,2,3,4),xlim=c(0,4),ylim=c(0,50),col="red",main="4 Land use types in group 3",xlab="Land use",labels=TRUE)

attach(group4)

hist(land\_use\_4types,breaks=c(0,1,2,3,4),xlim=c(0,4),ylim=c(0,25),col="red",main="4 Land use types in group 4",xlab="Land use",labels=TRUE)

attach(group1)

model\_1 <- lm(No\_of\_sp~altitude)

summary(model\_1)

attach(group2)

model\_2 <- lm(No\_of\_sp~altitude)

summary(model\_2)

attach(group3)

model\_3 <- lm(No\_of\_sp~altitude)

summary(model\_3)

attach(group4)

model\_4 <- lm(No\_of\_sp~altitude)

summary(model\_4)

library(calibrate)

par(mfrow=c(2,2))

attach(group1)

qqnorm(model\_int$residuals)

qqline(model\_int$residuals,col="red")

attach(group2)

qqnorm(model\_int$residuals)

qqline(model\_int$residuals,col="red")

attach(group3)

qqnorm(model\_int$residuals)

qqline(model\_int$residuals,col="red")

attach(group4)

qqnorm(model\_int$residuals)

qqline(model\_int$residuals,col="red")

par(mfrow=c(2,2))

library(calibrate)

attach(group1)

model\_alt <- lm(No\_of\_sp~altitude)

plot(model\_alt$fitted.values,model\_alt$residuals)

attach(group2)

model\_alt <- lm(No\_of\_sp~altitude)

plot(model\_alt$fitted.values,model\_alt$residuals)

attach(group3)

model\_alt <- lm(No\_of\_sp~altitude)

plot(model\_alt$fitted.values,model\_alt$residuals)

attach(group4)

model\_alt <- lm(No\_of\_sp~altitude)

plot(model\_alt$fitted.values,model\_alt$residuals)

library(AER)

attach(group1)

model\_1 <- lm(No\_of\_sp~altitude)

bptest(model\_1)

attach(group2)

model\_2 <- lm(No\_of\_sp~altitude)

bptest(model\_2)

attach(group3)

model\_3 <- lm(No\_of\_sp~altitude)

bptest(model\_3)

attach(group4)

model\_4 <- lm(No\_of\_sp~altitude)

bptest(model\_4)

attach(group1)

coeftest(model\_1,vcov=vcovHC)

attach(group2)

coeftest(model\_2,vcov=vcovHC)

attach(group3)

coeftest(model\_3,vcov=vcovHC)

attach(group4)

coeftest(model\_4,vcov=vcovHC)

#Clusteranalisis#

data<-read.table(file="C:\\Users\\data.txt", sep="\t",header=TRUE)

d<-dist(data)

cluster<-hclust(d, method="ward")

plot(cluster)

clust4<-cutree(cluster,4)

describe(clust4)

#Indicator Species Analisis#

multipatt(data, clust4, func="IndVal.g",duleg=TRUE)

#Bonferroni-Test#

attach(data)

pairwise.t.test(altitude,CLU4\_Pcord,p.adj="bonf")

pairwise.t.test(No\_of\_sp,CLU4\_Pcord,p.adj="bonf")

pairwise.t.test(angle,CLU4\_Pcord,p.adj="bonf")

#USB#

setwd("F:\\Advanced Statistics\\Paper")

data <- read.table(file="F:\\Advanced Statistics\\Paper\\discan.csv",sep = ";",header=TRUE,dec=",")

#USB (Spss data)#

setwd("F:\\Advanced Statistics\\Paper")

library(foreign)

data <- read.spss("modificado2.sav",to.data.frame = TRUE)

#Remove all missing values listwise

data.withoutna<-na.omit(data)

#Group variable

data\_grouping<-data.withoutna[,2]

data\_grouping

#dim(data)

#Possible independent variables

variables<-data.withoutna[,11:345]

variables

#Data set for discriminant analysis

ldadataset<-cbind(data\_grouping,variables)

ldadataset

#Discriminant analysis as SPSS does it (excluded variables by SPSS, denoted by -)

library(MASS)

#data from CSV#

model\_lda <- lda(data\_grouping ~., data=ldadataset, prior=c(255/471,100/471,76/471,40/471))

model\_lda

attributes(model\_lda)

model\_lda$prior

model\_lda$counts

model\_lda$means

model\_lda$scaling

model\_lda$lev

model\_lda$svd

model\_lda$N

model\_lda$call

model\_lda$terms

model\_lda$xlevels

model\_lda$lda

prediction<-predict(model\_lda)$class

#Classification Table

ct <- table(data\_grouping, prediction)

ct

diag(prop.table(ct, 1))

# total percent correct

hit\_ratio<-sum(diag(prop.table(ct)))

hit\_ratio

###SECOND PART###

#Remove everything

rm(list=ls(all=TRUE))

#USB#

setwd("F:\\Advanced Statistics\\Paper")

data <- read.table(file="F:\\Advanced Statistics\\Paper\\discan.csv",sep = ";",header=TRUE,dec=",")

#USB (Spss data)#

setwd("F:\\Advanced Statistics\\Paper")

library(foreign)

data <- read.spss("modificado2.sav",to.data.frame = TRUE)

####Creating a subdata of random samples (a 80% of the total)####

set.seed(110)

data80 <- data[sample(472, 385), ]

data80

#Remove all missing values listwise

data80.withoutna<-na.omit(data80)

#Group variable

data\_grouping80<-data80.withoutna[,2]

data\_grouping80

#dim(data80)

#Possible independent variables

variables80<-data80.withoutna[,11:345]

variables80

#Data set for discriminant analysis

ldadataset80<-cbind(data\_grouping80,variables80)

ldadataset80

#Discriminant analysis as SPSS does it (excluded variables by SPSS, denoted by -)

library(MASS)

model\_lda80<-lda(data\_grouping80 ~. ,data=ldadataset80, prior=c(255/471,100/471,76/471,40/471))

model\_lda80<-lda(data\_grouping80 ~. ,data=ldadataset80)

model\_lda80

attributes(model\_lda80)

model\_lda80$prior

model\_lda80$counts

model\_lda80$means

model\_lda80$scaling

model\_lda80$lev

model\_lda80$svd

model\_lda80$N

model\_lda80$call

model\_lda80$terms

model\_lda80$xlevels

model\_lda80$lda

prediction80<-predict(model\_lda80)$class

#Classification Table

ct80 <- table(data\_grouping80, prediction80)

ct80

diag(prop.table(ct80, 1))

# total percent correct

hit\_ratio80<-sum(diag(prop.table(ct80)))

hit\_ratio80