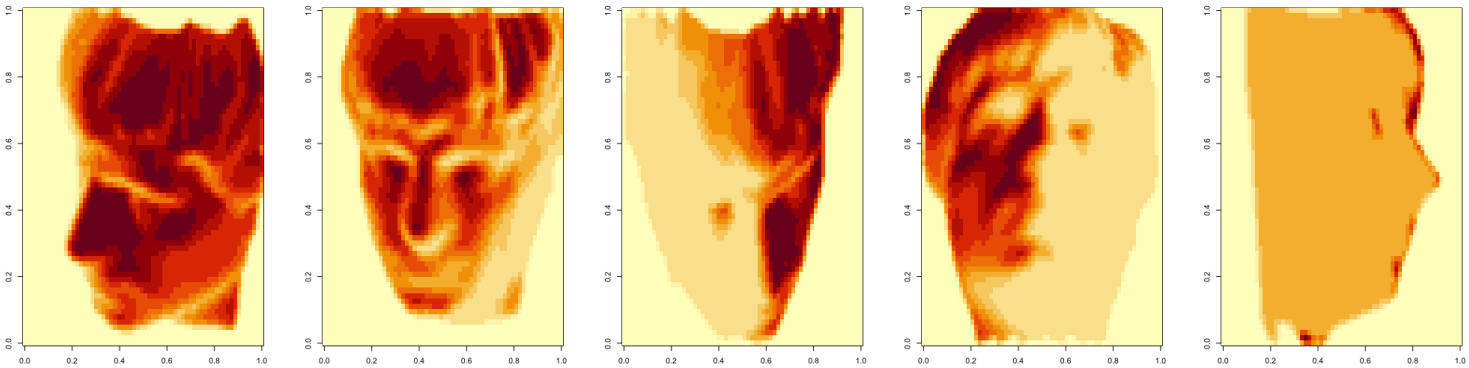
Homework 3

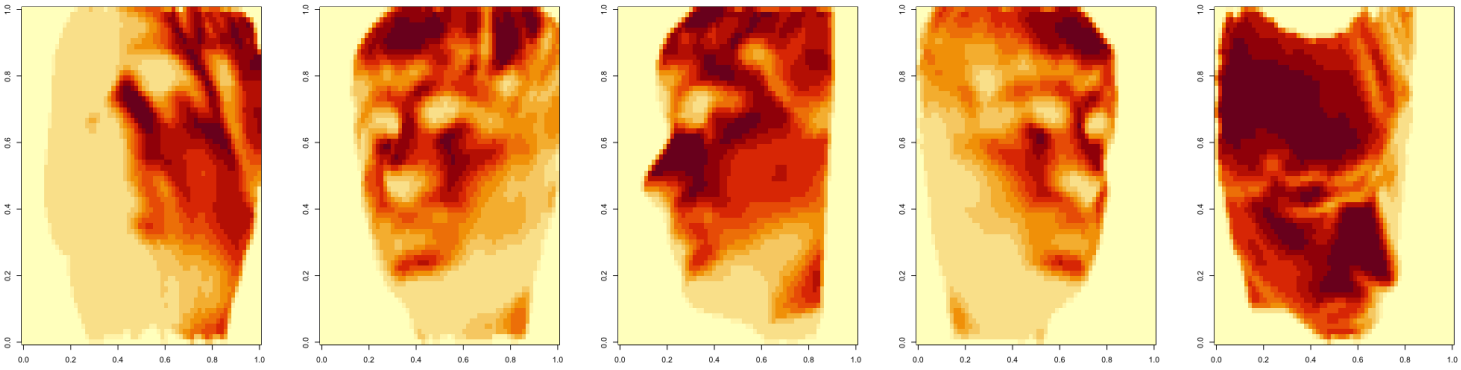
Yiwei He (yh9vhg), Da Lin (dl2de), Ziyue Jin (zj5qj)

Question 1 (Using PCA)

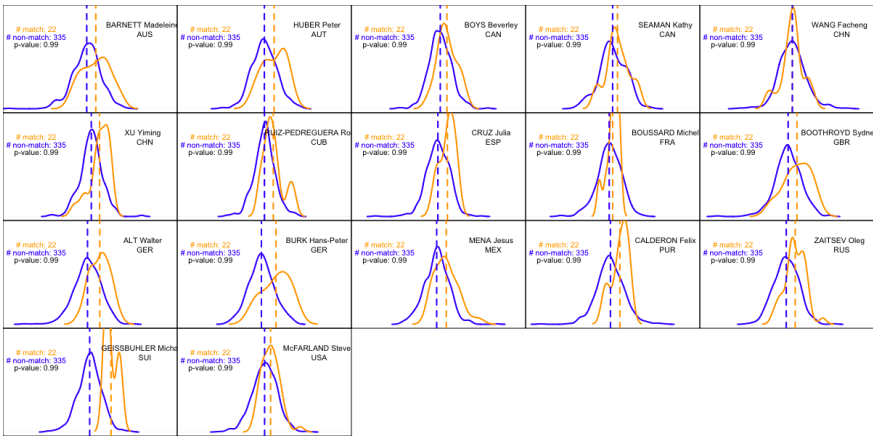
The image corresponded with Coord1 faces from left to right and down to up.



The image corresponded with Coord2 faces from left to right and up to down.

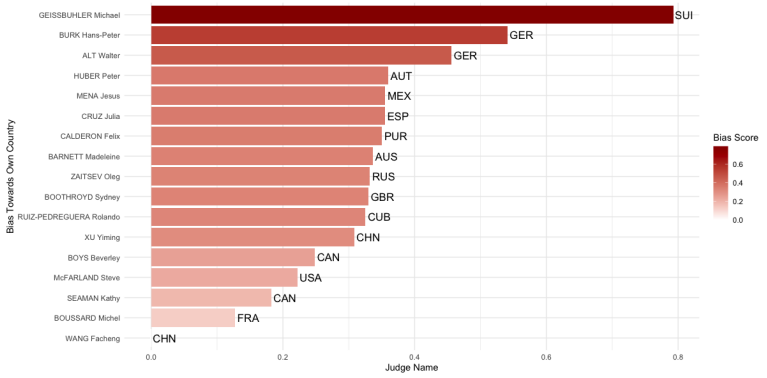


Question 2

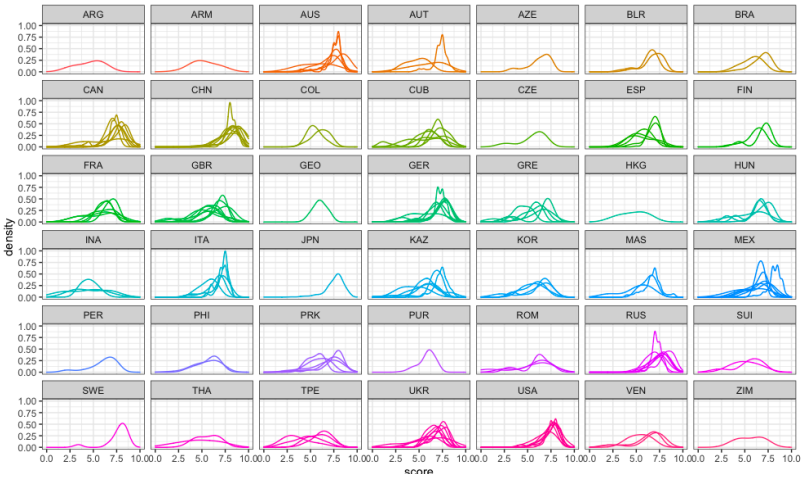


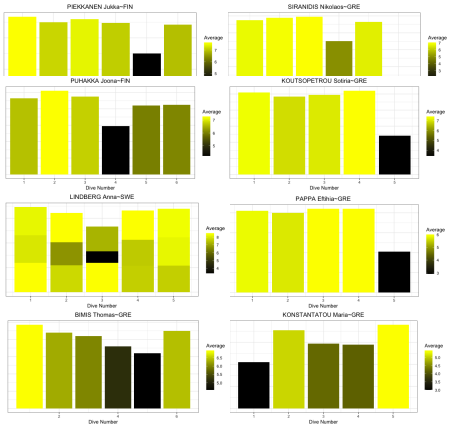
To take an overview of the judges' tendency to have bias towards divers from their own countries, the graph below shows almost all judges have some level of bias towards their own country except one Chinese judge named Wang Facheng.

By listing all mean bias scores, we can see that Geissbuhler Michael, the judge from Switzerland, has the highest average bias score, which means this judge tends to give a higher score to his own country diver.



We are interested in divers' performance. We can see that Chinese divers made no severe mistake based on distribution plots. However, divers from Finland, Sweden, and Greece seem to have several mistakes during their dives as there are outliers in the distribution curve. We used a bar chart to find out when these divers made mistakes and we found that most mistakes are made mostly at dive number 3, 4 and 5. This shows that diver might get tired and therefore made mistakes





Appendix

```
# Question 1

library(ggplot2)

load("/Users/yhe/Desktop/3280 Homework 3/Full_Faces_Data.rda")

X <- scale(face,center=TRUE,scale=TRUE)


##### MDS

d <- dist(X)

fit <- isoMDS(d, k=2)

Coord1 <- fit$points[,1]

Coord2 <- fit$points[,2]


##### PCA

zero.column <- which(is.na(colMeans(X)))

dt <- X[,-zero.column]

PCA <- prcomp(dt)

Coord1 <- PCA$x[,1]

Coord2 <- PCA$x[,2]


## retrieve index

dqt1 <- outer(Coord1, quantile(Coord1), function(Coord1,y) sqrt( (Coord1-y)^2) )

quantile1 <- apply(dqt1, 2, which.min)
```

```
dqt2 <- outer(Coord2, quantile(Coord2), function(Coord2,y) sqrt( (Coord2-y)^2) )
quantile2 <- apply(dqt2, 2, which.min)

face11 <- face[quantile1[1],]
face12 <- face[quantile1[2],]
face13 <- face[quantile1[3],]
face14 <- face[quantile1[4],]
face15 <- face[quantile1[5],]

face21 <- face[quantile2[1],]
face22 <- face[quantile2[2],]
face23 <- face[quantile2[3],]
face24 <- face[quantile2[4],]
face25 <- face[quantile2[5],]

## image matrix
mat11 <- matrix(face11,nrow=64,ncol=64)
mat12 <- matrix(face12,nrow=64,ncol=64)
mat13 <- matrix(face13,nrow=64,ncol=64)
mat14 <- matrix(face14,nrow=64,ncol=64)
mat15 <- matrix(face15,nrow=64,ncol=64)

mat21 <- matrix(face21,nrow=64,ncol=64)
```

```
mat22 <- matrix(face22,nrow=64,ncol=64)

mat23 <- matrix(face23,nrow=64,ncol=64)

mat24 <- matrix(face24,nrow=64,ncol=64)

mat25 <- matrix(face25,nrow=64,ncol=64)


########## Coord 1 image

par(mfrow=c(1,5))


mat11=apply(t(mat11),2,rev)

mat11=apply(t(mat11),2,rev)

mat11=apply(t(mat11),2,rev)

mat12=apply(t(mat12),2,rev)

mat12=apply(t(mat12),2,rev)

mat12=apply(t(mat12),2,rev)

mat13=apply(t(mat13),2,rev)

mat13=apply(t(mat13),2,rev)

mat13=apply(t(mat13),2,rev)

mat14=apply(t(mat14),2,rev)

mat14=apply(t(mat14),2,rev)

mat14=apply(t(mat14),2,rev)

mat15=apply(t(mat15),2,rev)

mat15=apply(t(mat15),2,rev)

mat15=apply(t(mat15),2,rev)
```

```
image(mat11)

image(mat12)

image(mat13)

image(mat14)

image(mat15)


########## Coord 2 image

par(mfrow=c(1,5))


mat21=apply(t(mat21),2,rev)

mat21=apply(t(mat21),2,rev)

mat21=apply(t(mat21),2,rev)

mat22=apply(t(mat22),2,rev)

mat22=apply(t(mat22),2,rev)

mat22=apply(t(mat22),2,rev)

mat23=apply(t(mat23),2,rev)

mat23=apply(t(mat23),2,rev)

mat23=apply(t(mat23),2,rev)

mat24=apply(t(mat24),2,rev)

mat24=apply(t(mat24),2,rev)

mat24=apply(t(mat24),2,rev)

mat25=apply(t(mat25),2,rev)
```

```
mat25=apply(t(mat25),2,rev)
mat25=apply(t(mat25),2,rev)

image(mat21)
image(mat22)
image(mat23)
image(mat24)
image(mat25)

#Question 2
x <- read.csv("Diving2000.csv", stringsAsFactors = FALSE)

jinfo <- x[!duplicated(x$judge), 9:10]
jinfo <- jinfo[order(jinfo$jcountry),]
rownames(jinfo) <- NULL
jinfo$numscores <- 0
jinfo$nummatch <- 0
jinfo$numother <- 0
jinfo$ownbias <- 0
for (i in 1:nrow(jinfo)) {
  thisjudge <- jinfo$judge[i]
  thiscountry <- jinfo$jcountry[i]
  y <- x[x$judge==thisjudge,]
```

```
  jinfo$numscores[i] <- nrow(y)
  jinfo$nummatch[i] <- sum(y$match)
  jinfo$numother[i] <- sum(!y$match)
  jinfo$ownbias[i] <- bias(y, thisjudge, thiscountry)
}
par(mfrow=c(4,5),mar=c(0,0,0,0))
for (i in 1:nrow(jinfo)) {
  thisjudge <- jinfo$judge[i]
  thiscountry <- jinfo$jcountry[i]
  if (jinfo$nummatch[i]>0) {
    y <- x[x$judge==thisjudge,]
    plot(density(y$diff[!y$match]), xlim=c(-3,3), ylim=c(0,1.25),
         xlab="", main="", ylab="", lwd=2, yaxt="n", xaxt="n",col="blue")
    lines(density(y$diff[y$match]), col="orange", lwd=2)
    abline(v=mean(y$diff[y$match]), col="orange", lwd=2,lty=2)
    abline(v=mean(y$diff[!y$match]), lwd=2,lty=2,col="blue")
    text(-2, 1, paste("# match:", result$n1), col="orange")
    text(-2, 0.9, paste("# non-match:", result$n2),col="blue")
    text(-2, 0.8, paste("p-value:", round(result$p.value,digits=3)))
    text(2, 1, paste(thisjudge, "\n", thiscountry))

  }
}
```

```
jinfo$ownbias

jinfo$judge

x1=data.frame(jinfo$ownbias,jinfo$judge,jinfo$jcountry)

x2=x1[order(x1$jinfo.ownbias),]

x2=na.omit(x2)

basicplot=ggplot(x2,aes(x=reorder(jinfo.judge,jinfo.ownbias),y=jinfo.ownbias,fill=jinfo.ownbias))+geom_bar(stat='identity')+coord_flip()

label=labs(x="Bias Towards Own Country", y="Judge Name")

finalplot=basicplot+label+scale_fill_gradient(low="white",high="darkred","Bias Score")+geom_text(aes(label=x2$jinfo.jcountry),hjust=-0.1, size=4.5)+theme_minimal()

finalplot




swe=data.frame(x[x$dcountry=="SWE",])

gre=data.frame(x[x$dcountry=="GRE",])

fin=data.frame(x[x$dcountry=="FIN",])

above9=data.frame(x[x$score>=9,])


basicplot=ggplot(above9,aes(x=reorder(factor(dcountry),score),y=score,fill=jinfo.ownbias))+geom_bar(stat='identity')+coord_flip()


ggplot(data=x, aes(x=score, group=factor(diver), color=factor(dcountry))) +
```

```
  geom_density(adjust=1.5) +
  facet_wrap(~factor(dcountry))
par(mfrow=c(4,4),mar=c(0,0,0,0))
ggplot(fin[fin$diver=="PIEKKANEN Jukka",],aes(x=factor(divenum),y=avg,fill=avg))+
  geom_bar(stat='identity')+
  labs(x="Dive Number")+
  scale_fill_gradient(low="Black",high="yellow","Average")+
  theme(axis.title.y=element_blank(),
        axis.text.y=element_blank(),
        axis.ticks.y=element_blank(),
        plot.title = element_text(hjust = 0.5))+
  ggtitle("PIEKKANEN Jukka~FIN")

ggplot(fin[fin$diver=="PUHAKKA Joona",],aes(x=factor(divenum),y=avg,fill=avg))+
  geom_bar(stat='identity')+
  labs(x="Dive Number")+
  scale_fill_gradient(low="Black",high="yellow","Average")+
  theme(axis.title.y=element_blank(),
        axis.text.y=element_blank(),
        axis.ticks.y=element_blank(),
        plot.title = element_text(hjust = 0.5))+
  ggtitle("PUHAKKA Joona~FIN")
```

```
ggplot(swe[swe$diver=="LINDBERG Anna",],aes(x=divenum,y=avg,fill=avg))+
  geom_bar(stat='identity')+
  labs(x="Dive Number")+
  scale_fill_gradient(low="Black",high="yellow","Average")+
  theme(axis.title.y=element_blank(),
        axis.text.y=element_blank(),
        axis.ticks.y=element_blank(),
        plot.title = element_text(hjust = 0.5))+
  ggtitle("LINDBERG Anna~SWE")

ggplot(gre[gre$diver=="BIMIS Thomas",],aes(x=divenum,y=avg,fill=avg))+
  geom_bar(stat='identity')+
  labs(x="Dive Number")+
  scale_fill_gradient(low="Black",high="yellow","Average")+
  theme(axis.title.y=element_blank(),
        axis.text.y=element_blank(),
        axis.ticks.y=element_blank(),
        plot.title = element_text(hjust = 0.5))+
  ggtitle("BIMIS Thomas~GRE")

ggplot(gre[gre$diver=="SIRANIDIS Nikolaos",],aes(x=divenum,y=avg,fill=avg))+
  geom_bar(stat='identity')+
  labs(x="Dive Number")+
```

```
  scale_fill_gradient(low="Black",high="yellow","Average")+
  theme(axis.title.y=element_blank(),
        axis.text.y=element_blank(),
        axis.ticks.y=element_blank(),
        plot.title = element_text(hjust = 0.5))+
  ggtitle("SIRANIDIS Nikolaos~GRE")

ggplot(gre[gre$diver=="KOUTSOPETROU Sotiria",],aes(x=divenum,y=avg,fill=avg))+
  geom_bar(stat='identity')+
  labs(x="Dive Number")+
  scale_fill_gradient(low="Black",high="yellow","Average")+
  theme(axis.title.y=element_blank(),
        axis.text.y=element_blank(),
        axis.ticks.y=element_blank(),
        plot.title = element_text(hjust = 0.5))+
  ggtitle("KOUTSOPETROU Sotiria~GRE")

ggplot(gre[gre$diver=="PAPPA Eftihia",],aes(x=divenum,y=avg,fill=avg))+
  geom_bar(stat='identity')+
  labs(x="Dive Number")+
  scale_fill_gradient(low="Black",high="yellow","Average")+
  theme(axis.title.y=element_blank(),
        axis.text.y=element_blank(),
```

```
      axis.ticks.y=element_blank(),
      plot.title = element_text(hjust = 0.5))+
  ggtitle("PAPPA Eftihia~GRE")

ggplot(gre[gre$diver=="KONSTANTATOU Maria",],aes(x=divenum,y=avg,fill=avg))+
  geom_bar(stat='identity')+
  labs(x="Dive Number")+
  scale_fill_gradient(low="Black",high="yellow","Average")+
  theme(axis.title.y=element_blank(),
      axis.text.y=element_blank(),
      axis.ticks.y=element_blank(),
      plot.title = element_text(hjust = 0.5))+
  ggtitle("KONSTANTATOU Maria~GRE")
```