R Base Plotting

Instructions: All plots generated for this homework assignment must be a part of the R Base plotting functions. This means no GGPLOTS, plot.ly or any other package plot functions, only R Base

Name: Azar Najafli

Step 1: Import the Global Orders 2016 data set using the read.csv function (5 points)

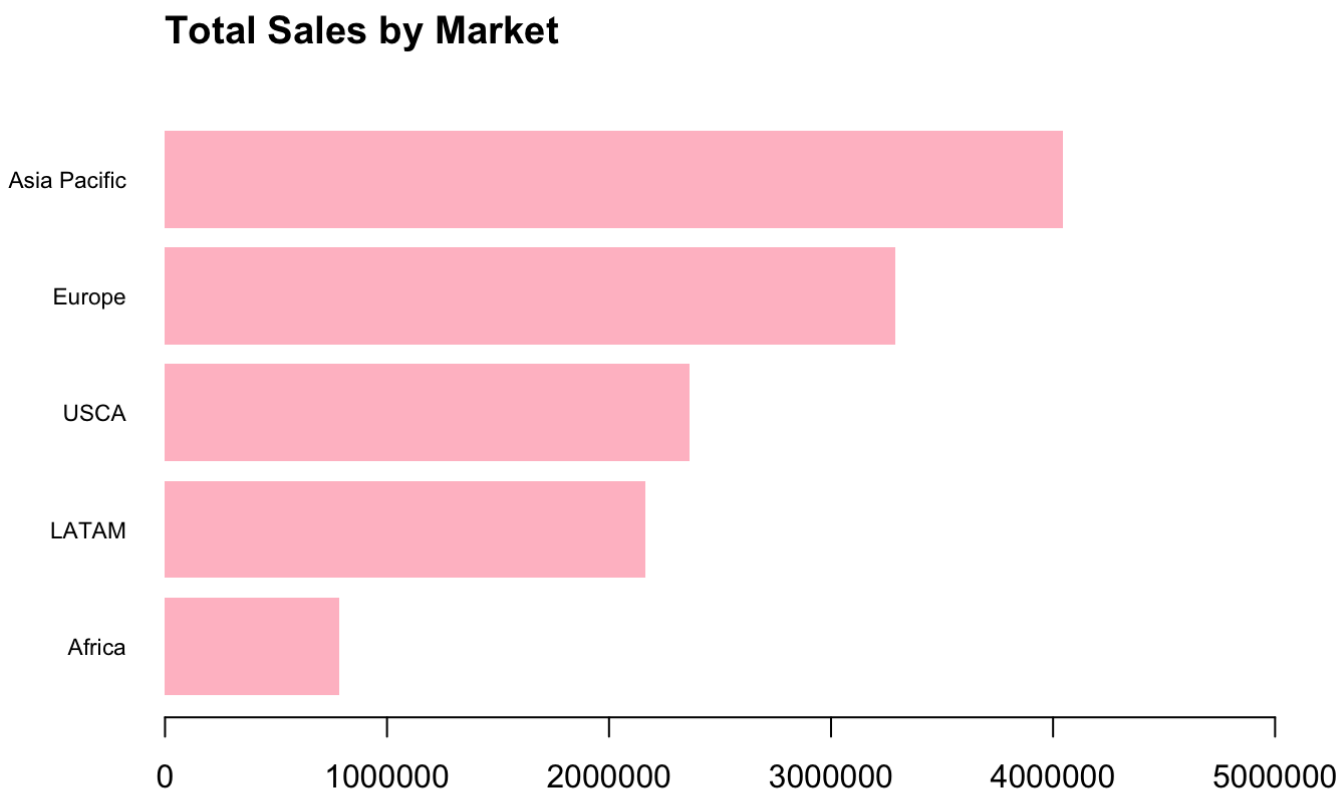
```
setwd("/Users/azarnajafli/Desktop/R")
getwd()
```

```
## [1] "/Users/azarnajafli/Desktop/R"
```

```
data<-read.csv(file = "/Users/azarnajafli/Desktop/R/superstore.csv", header = T)
```

Step 2 - Create a horizontal barplot of Total Sales by Market, ordered descending on total sales (i.e. the market with highest total sales should be on the top) (25 points)

```
data1 <- aggregate.data.frame(data$Sales,by=list(data$Market),sum)
names(data1)<- c("Market","Total_Sales")
A<- data1[order(data1$Total_Sales,decreasing=FALSE),]
barplot(A$Total_Sales,names.arg=A$Market,
       col = c("pink","pink","pink","pink","pink"),
       horiz = TRUE, border=NA,
       xlim=c(0,5000000),
       las =1, cex.names = 0.71)
        options(scipen = 999)
title("Total Sales by Market",adj = 0)
```



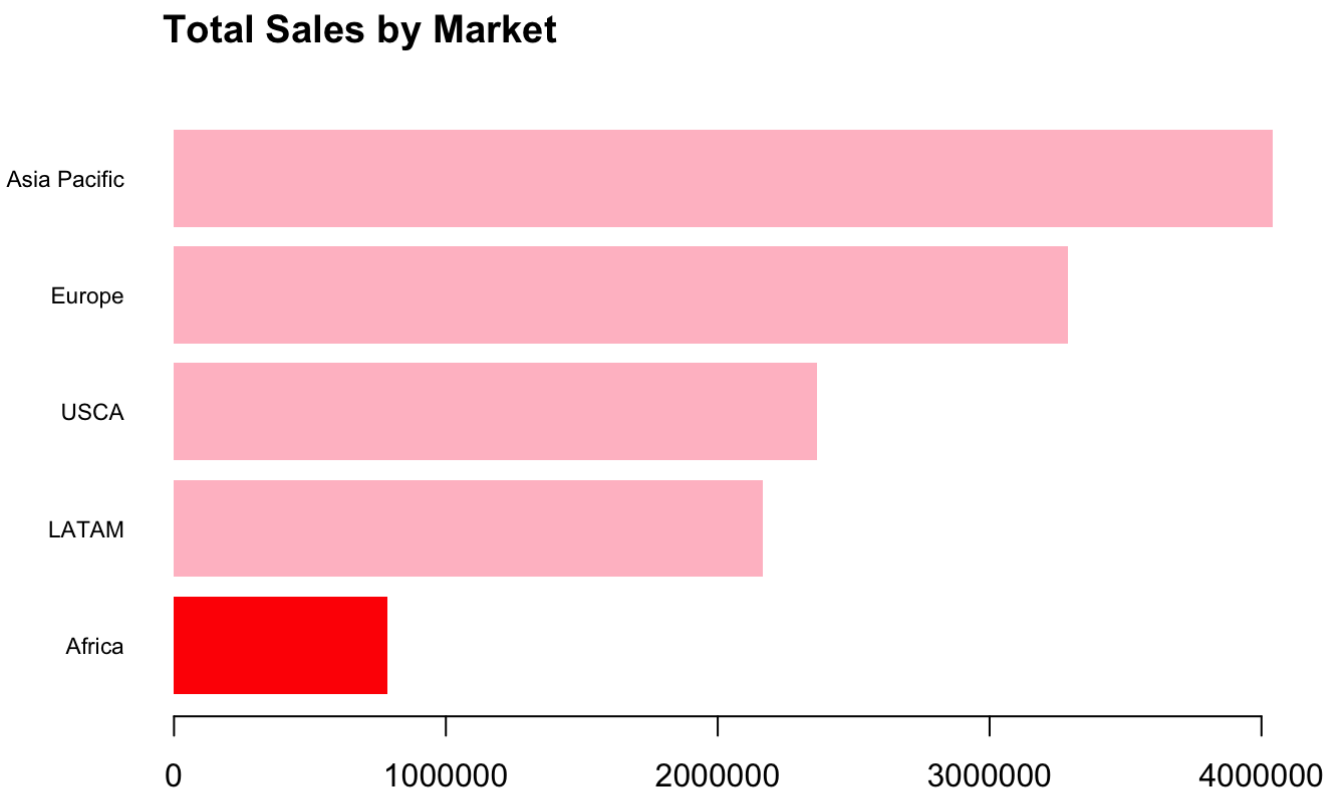
Answer the following question: Does any market appear to be very different from the others for total sales?

Answer below: Yes, Africa

Can you visually indicate this market as a separate color from the others?

```
barplot(A$Total_Sales,names.arg=A$Market,
       col = c("red","pink","pink","pink","pink"),
       horiz = TRUE, border=NA,
       las =1, cex.names = 0.71
```

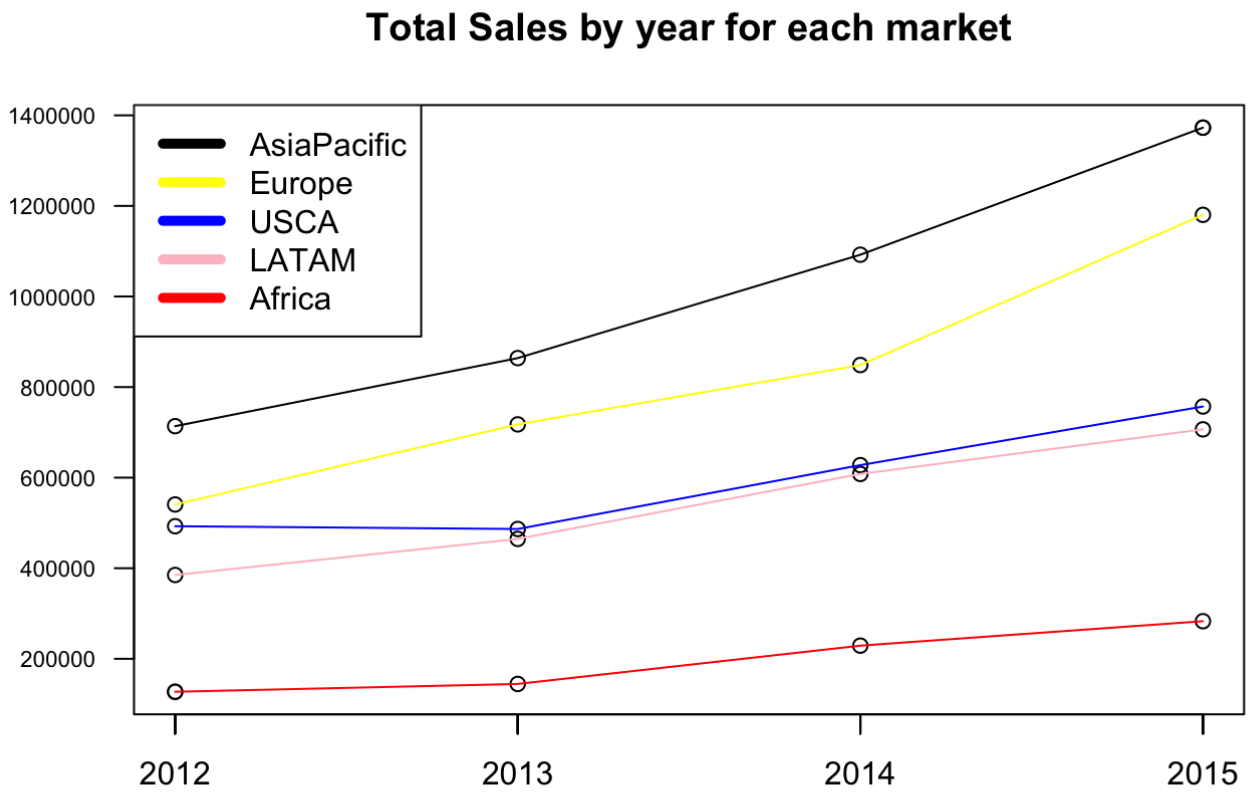
```
      )
title("Total Sales by Market",adj = 0)
```



Step 3 - Create a line chart total sales by year for each market (25 points)

```
data2<-aggregate.data.frame(data$Sales,by=list(data$Market,data$Order.Year),sum)
names(data2) = c("Market","Year","Total_Sales")

data2.AsiaPacific<- data2[data2$Market == "Asia Pacific",]
data2.Europe<- data2[data2$Market == "Europe",]
```

```
data2.USCA<- data2[data2$Market == "USCA",]
data2.LATAM<- data2[data2$Market == "LATAM",]
data2.Africa<- data2[data2$Market == "Africa",]

xrange <- range(data2$Year)
yrange <- range(data2$Total_Sales)

plot(xrange,
     yrange,
     xlab="",ylab="",
     xaxt="n",
     las= 1,
     cex.axis= 0.70,
     main = "Total Sales by year for each market")

axis(1, labels = as.character(data2$Year), at = as.numeric(data2$Year))

points(data2.AsiaPacific$Year, data2.AsiaPacific$Total_Sales)
points(data2.Europe$Year, data2.Europe$Total_Sales)
points(data2.USCA$Year, data2.USCA$Total_Sales)
points(data2.LATAM$Year, data2.LATAM$Total_Sales)
points(data2.Africa$Year, data2.Africa$Total_Sales)


lines(data2.AsiaPacific$Year, data2.AsiaPacific$Total_Sales,col = "black")
lines(data2.Europe$Year, data2.Europe$Total_Sales,col = "yellow")
lines(data2.USCA$Year, data2.USCA$Total_Sales,col = "blue")
lines(data2.LATAM$Year, data2.LATAM$Total_Sales,col = "pink")
lines(data2.Africa$Year, data2.Africa$Total_Sales,col = "red")
legend("topleft", legend=c("AsiaPacific","Europe","USCA","LATAM","Africa"), lwd=c(5,5), col=c("black","yellow","b
lue","pink","red"))
```



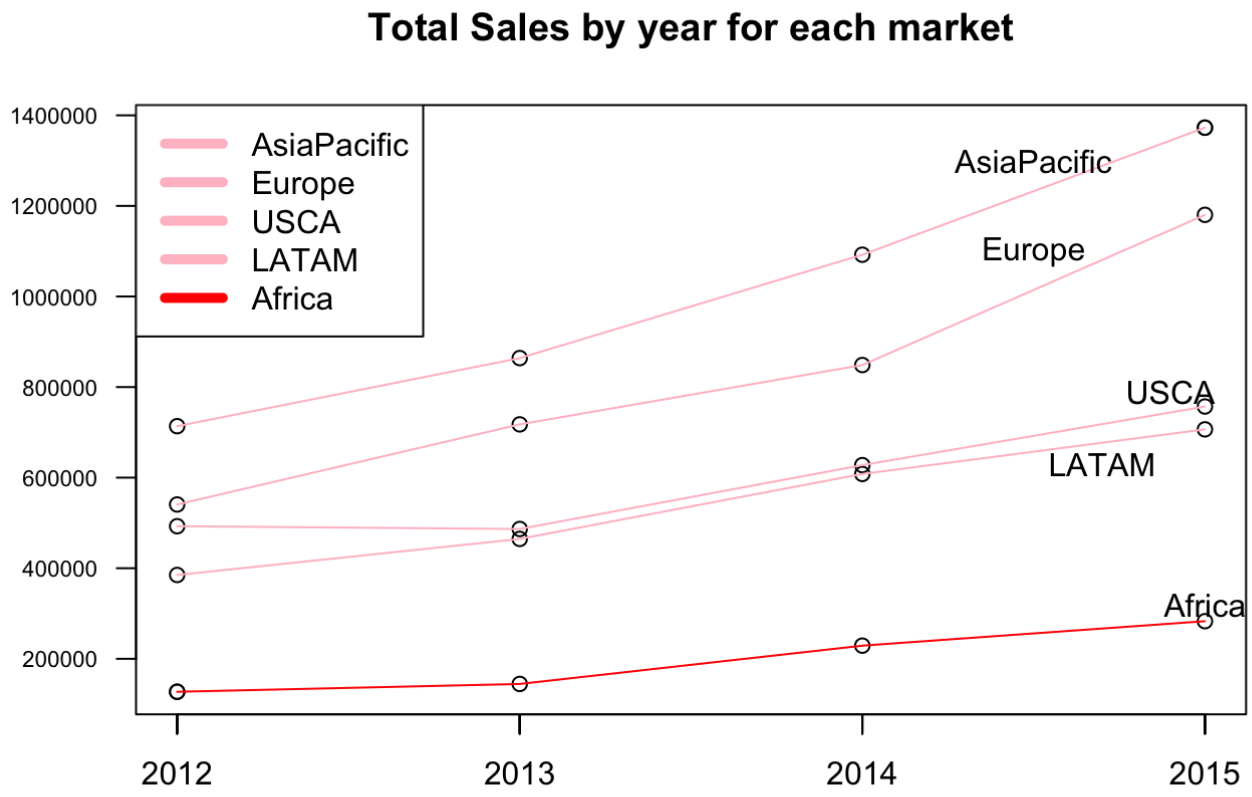
Answer the following question: Does the same market appear to be different in your line graph as well?

Answer below: Yes *Can you visually indicate this market as a separate color from the other markets in your graph?*

Can you tell your markets apart from one another in your graph? If not consider what you need to add to your graph so that you can tell them apart
Yes, by labeling lines.

```
plot(xrange,
     yrange,
     xlab="",ylab="",
     xaxt="n",
```

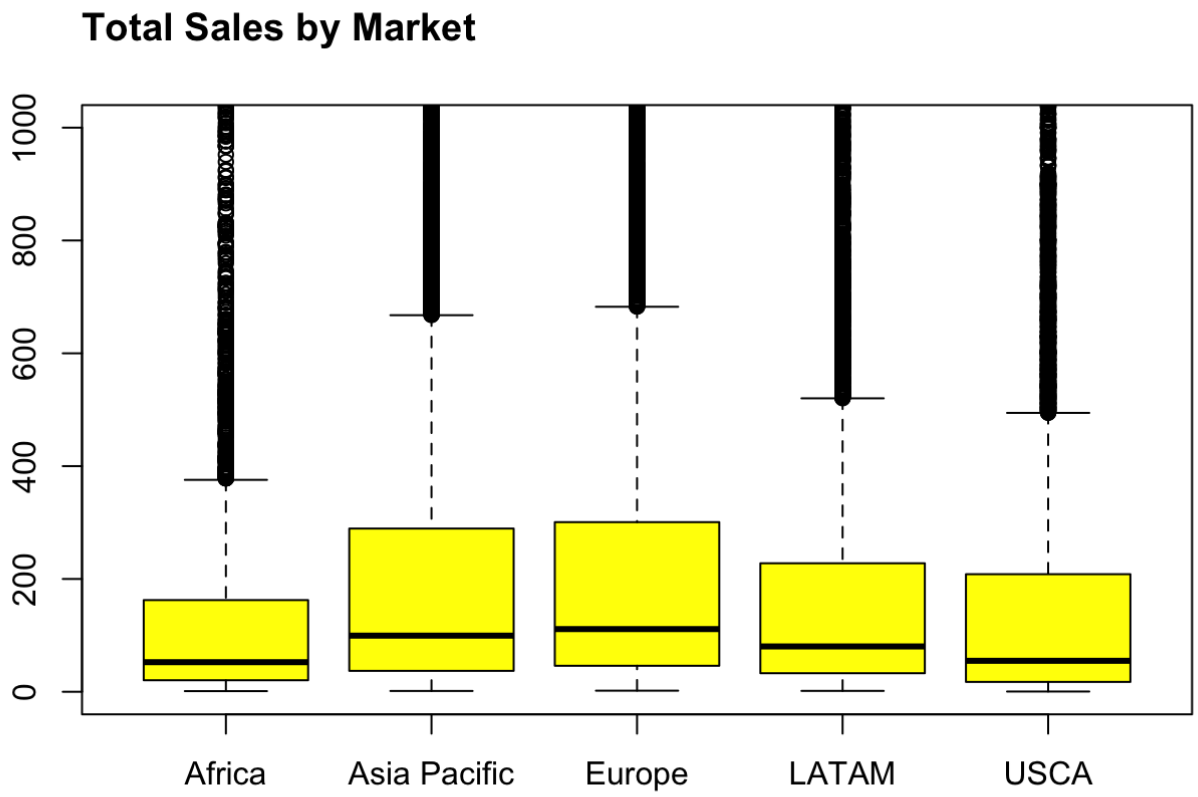
```
     las= 1,
     cex.axis= 0.70,
     main = "Total Sales by year for each market")

axis(1, labels = as.character(data2$Year), at = as.numeric(data2$Year))

points(data2.AsiaPacific$Year, data2.AsiaPacific$Total_Sales)
points(data2.Europe$Year, data2.Europe$Total_Sales)
points(data2.USCA$Year, data2.USCA$Total_Sales)
points(data2.LATAM$Year, data2.LATAM$Total_Sales)
points(data2.Africa$Year, data2.Africa$Total_Sales)


lines(data2.AsiaPacific$Year, data2.AsiaPacific$Total_Sales,col = "pink")
lines(data2.Europe$Year, data2.Europe$Total_Sales,col = "pink")
lines(data2.USCA$Year, data2.USCA$Total_Sales,col = "pink")
lines(data2.LATAM$Year, data2.LATAM$Total_Sales,col = "pink")
lines(data2.Africa$Year, data2.Africa$Total_Sales,col = "red")
legend("topleft", legend=c("AsiaPacific","Europe","USCA","LATAM","Africa"), lwd=c(5,5), col=c("pink","pink","pin
k","pink","red"))
text(2015, 320000, "Africa")
text(2014.7, 630000, "LATAM")
text(2014.9, 790000, "USCA")
text(2014.5, 1100000, "Europe")
text(2014.5, 1300000, "AsiaPacific")
```



Step 4 - Create a box plot of total sales by market (25 points)

Hint: use the ylim parameter to restrict your graph so that you tell the boxes apart

```
boxplot(Sales~Market,
        data=data,
        ylim = range(0:1000),
        col= c("yellow","yellow","yellow","yellow","yellow" ) )
title("Total Sales by Market",adj=0)
```



Does the general pattern you observe match that of the earlier steps 2 and 3?

Answer below: Yes, we can see Africa contains the values mostly below other markets.

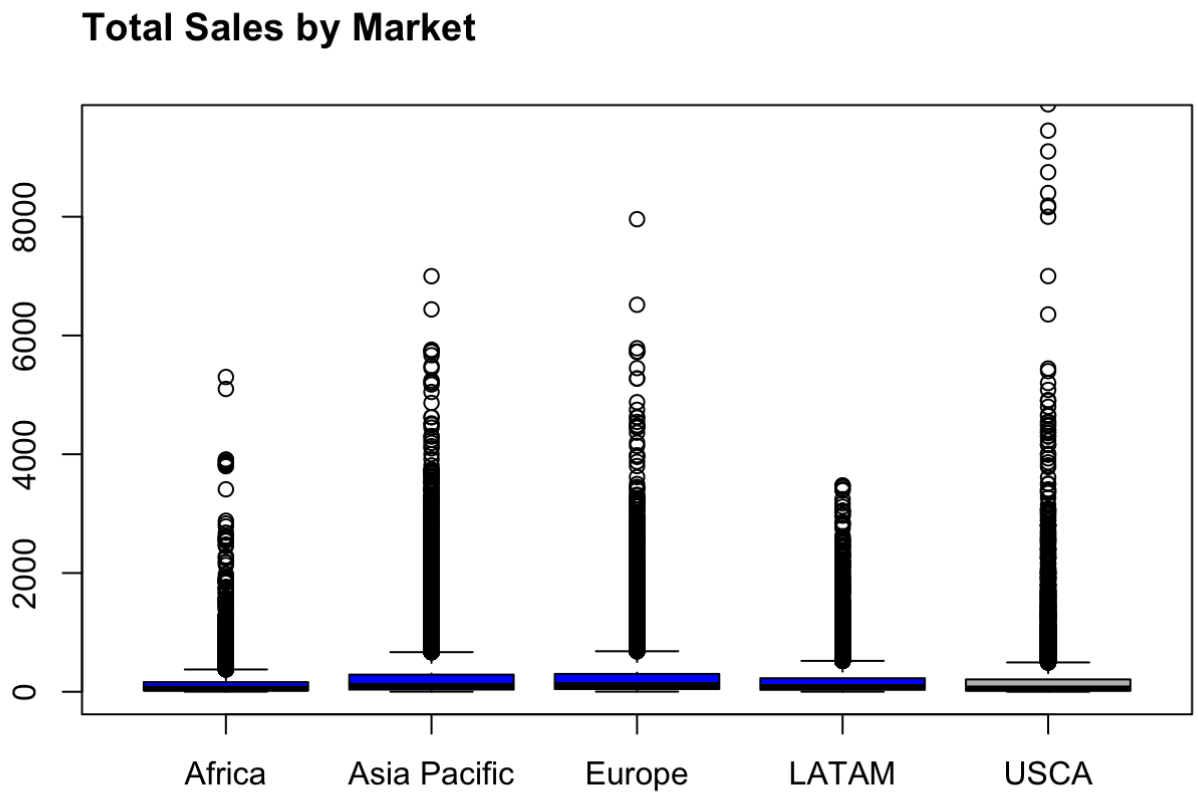
What other insights can you draw from your box plot above?

Answer below:

It is visible that Asia and Europe markets are mainly distributed 100-300k total amount of sales in dollars although highest values are more than 600.

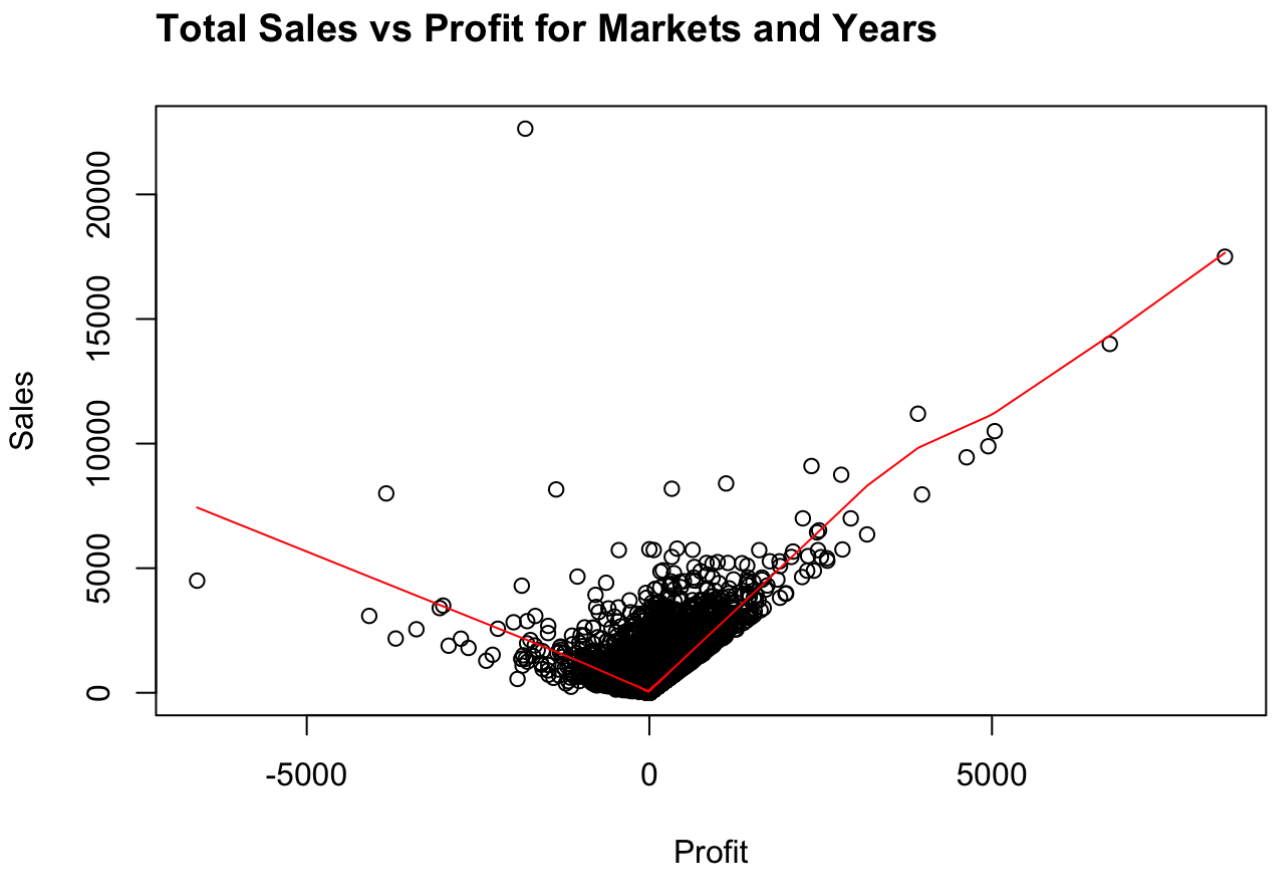
Can you visually indicate this market as a separate color from the others in your boxplot graph?

```
boxplot(Sales~Market,
        data=data,
        ylim = range(0:9500),
        col= c("blue","blue","blue","blue","grey" ) )
title("Total Sales by Market",adj=0)
```



Step 5 - Generate a different kind of graph other than what was produced in steps 2, 3, & 4 (20 points)

```
data3 <- data[c('Profit', 'Sales')]
plot(x = data3$Profit, y = data3$Sales,
   xlab = "Profit",
   ylab = "Sales")
title("Total Sales vs Profit for Markets and Years",adj=0)
 lines(lowess(data3$Profit,data3$Sales), col="red")
```



This means not a bar plot, line graph, or box plot. You need to use R Base plots for this question, not GGPLOT or any other plotting functions, only R Base. We can observe an interesting trend from the scatter plot. In general market does not make a profit as main bubbles gathered around 0 profit. However, while the sales are in high volume, we have more positive profit.