Question 4

Load the XML with the DTD into R using validation.

```
library("XML")

xml_data <- xmlParse("CustomersAndOrders.xml",validate = T)
```

Question 5

Execute an XPath expression that returns the names of all customers that do not live in "MA"

```
cust_name <- xpathApply(xml_data,"/Root/Customers/Customer[FullAddress/Region
                        != 'MA']/CompanyName",xmlValue)

customer_names <- unlist(cust_name)

print(customer_names)
```

```
## [1] "Great Lakes Food Market"    "Hungry Coyote Import Store"
## [3] "Lazy K Kountry Store"       "Let's Stop N Shop"
```

Question 6

Using the result returned in (5), count how many customers there are.

```
print(length(customer_names))
```

```
## [1] 4
```

Question 7

Using a combination of R and XPath, calculate the total amount paid for freight for all orders placed by customer with ID "GREAL".

```
freight_values <- as.numeric(xpathApply(xml_data, "/Root/Orders
                                        /Order[CustomerID = 'GREAL']
                                        /ShipInfo/Freight" , xmlValue))


freight_values <- unlist(freight_values)

print(freight_values)
```

```
##  [1]   3.35   4.42 116.53  18.53  57.15  76.13 719.78  33.68  25.19  18.84
## [11]  14.01
```

```
total_freight_amount <- sum(freight_values)

print(total_freight_amount)
```

```
## [1] 1087.61
```

Question 8

Using a combination of R and XPath, calculate the average amount paid for freight for all orders shipped to the USA.

```
a <- as.numeric( xpathApply(xml_data,"/Root/Orders
                            /Order[ShipInfo/ShipCountry='USA']
                            /ShipInfo/Freight",xmlValue) )

a <- unlist(a)

print(mean(a))
```

```
## [1] 68.91818
```