Tan.CS5200.CompXPath

Jiaxing Tan

2023-04-02

```
library(pacman)
p_load(XML)
```

Question 4

Load the XML with the DTD into R using validation

```
xmlURL <- "CustomersAndOrders.xml"
xmlObj <- xmlParse(xmlURL, validate=T)
root <- xmlRoot(xmlObj)
```

Question 5

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

```
xPS <- "//Root/Customers/Customer/FullAddress[Region!='MA']/../ContactName"
res <- xpathSApply(xmlObj, xPS, xmlValue)
res
```

```
## [1] "Howard Snyder" "Yoshi Latimer" "John Steel"    "Jaime Yorres"
```

Question 6

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

```
length(res)
```

```
## [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”.

```
xPS <- "sum(//Root/Orders/Order[CustomerID='GREAL']/ShipInfo/Freight)"
res <- xpathSApply(xmlObj, xPS, xmlValue)
res
```

```
## [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.

```
xPS1 <- "sum(//Root/Orders/Order/ShipInfo[ShipCountry='USA']/Freight)"
xPS2 <- "count(//Root/Orders/Order/ShipInfo[ShipCountry='USA']/Freight)"
freight_sum <- xpathSApply(xmlObj, xPS1, xmlValue)
freight_tally <- xpathSApply(xmlObj, xPS2, xmlValue)
freight_sum / freight_tally
```

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