

# NCollin\_607\_Hw7

Noah Collin

10/10/2021

## HTML

(installation code commented out)

```
#install.packages("XML")
library(XML)
tables <- readHTMLTable("BooksILike.html")
#str(tables)
htmldf1 <- as.data.frame( (tables[1]))
#class(htmldf1)
names(htmldf1) <- c("Title","Author","Copywritten")
print(htmldf1)
```

```
##                                     Title
## 1 What If?: Serious Scientific Answers to Absurd Hypothetical Questions
## 2                                     Thing Explainer
##      Author Copywritten
## 1 Randall Munroe      2014
## 2 Randall Munroe      2015
```

## XML

```
xml1 <- xmlParse("BooksILike.xml")
rootNode1 <- xmlRoot(xml1)
rootNode1[1]
```

```
## $row
## <row>
##   <Book>Catch-22</Book>
##   <Author>Joseph Heller</Author>
##   <Copywritten>1961</Copywritten>
## </row>
##
## attr(,"class")
## [1] "XMLInternalNodeList" "XMLNodeList"
```

```
data <- xmlSApply(rootNode1,function(x) xmlSApply(x, xmlValue))
xmldf <- data.frame(t(data),row.names = NULL)
xmldf
```

```
##           Book           Author Copywritten
## 1 Catch-22 Joseph Heller           1961
```

## JSON

```
library(jsonlite)
```

```
##
## Attaching package: 'jsonlite'
```

```
## The following object is masked from 'package:purrr':
##
##      flatten
```

```
jsonDF <- fromJSON("BooksILike.json") %>% as.data.frame()
class(jsonDF)
```

```
## [1] "data.frame"
```

```
jsonDF
```

```
##
##           Book           Author
## 1 The Power of Now: A Guide to Spiritual Enlightenment Eckhart Tolle
##      Copywritten
## 1           1997
```

The three dataframes are essentially identical. The code to generate them was not.