SharePoint R integration and analysis

In this tutorial, we are going to discuss following important steps.

Step 1: How login into the SharePoint database in r?

Step 2:- How to extract the data from SharePoint?

Step 3:- How to clean the data in r?

Step 4:- Analyze the data in r

Step 5:- Make a report and mail it to the respective person

We are not going to concentrate much on Step 4 and Step 5 because it's subjective and varies based on the requirements.

How to login into the SharePoint database in r?

First, we need to save the user id and password in the mentioned format.

```
Usepassword<-"Sharepoint userid:sharepoint password"
```

Once you store the user id and password, the next step is to set up the **URL**.

First go to the list you want to extract the information and take the **URL** from the browser, for example, if you want to extract the information from the "HR" list, the URL should be something like this.

http://your.sharepoint.websitename//HR//_vti_bin//owssvr.dll?XMLDATA=1&

The next step is to extract the list code, view code, and row limit from the share point database and join to the above URL.

The URL finally looks like mentioned format.

url<-"above URL & list code & view code & Row Limit"

List code, View code and Row Limit looks something like this

LIST={alphanumericcodes}&VIEW={%alphanumericcodes}&RowLimit=something"

Let see how to find the list code? first, you need to click on the list setting and select "**audience target setting**", from the browser URL now you can extract the list codes.

Differences between Association & Correlation

In the same way, you can extract view codes also from the list setting. Go to list settings and under **view** click "**all items"**, now from the URL you can extract view codes.

Finally, the row limit format looks something like this

RowLimit=&RootFolder=%2fmodulename%2fLists%2flistname

For extracting module name and list name, you can just click on the **list** link (the information you want to extract from) and from the browser, can extract the details.

Your URL ready now.

```
url<-"above URL & list code & view code & Row Limit"
```

Getting Data

```
library(xml)
library(xlsx)
data<-getURL(url, userpwd= Usepassword)
xmldata<-xmlParse(data,useInternalNode=TRUE, options=HUGE)
datalist<-xmlToList(xmlroot(xmldata)[["data"]])
mydata<-ldply(datalist,rbind)
```

Now entire columns information's saved in mydata

The column names should look like ows$Title or ows$category etc… now you can do the proper column renaming according to you and select relevant columns.

The cleaned data is ready now and according to your requirements can execute the analysis and make a report.

Conclusion:

Based on the SharePoint centralized database anyone can automate the complete process flow based on R. Really, this will save huge manpower and money.