**Practical no 1**

**Aim:Practical of Data collection, Data curation and management for Unstructured data (NoSQL)**

**Download apache couch database**

> install.packages("sofa")

> #devtools::install github("ropensci/sofa")

> library("sofa")

**Create connection object**.

> x<-Cushion$new()

**Check whether object is created or not.**

> x$ping()

$couchdb

[1] "Welcome"

$version

[1] "2.3.1"

$git\_sha

[1] "c298091a4"

$uuid

[1] "6e8f7a7454b2413234383d6d4b2477ca"

$features

$features[[1]]

[1] "pluggable-storage-engines"

$features[[2]]

[1] "scheduler"

$vendor

$vendor$name

[1] "The Apache Software Foundation"

**#create database DS**

> db\_create(x,dbname = 'ds')

$ok

[1] TRUE

> db\_list(x)

[1] "ds"

**Create json doc**

> doc1<-'{"rollno":"01","name":"ABC","GRADE":"A"}'

> doc\_create(x,doc1,dbname = "ds",docid = "a\_1")

$ok

[1] TRUE

$id

[1] "a\_1"

$rev

[1] "1-be7c98bddf8ea7c46f4f401ff387593d"

> doc2<-'{"rollno":"02","name":"PQR","GRADE":"A"}'

> doc\_create(x,doc2,dbname = "ds",docid = "a\_2")

$ok

[1] TRUE

$id

[1] "a\_2"

$rev

[1] "1-1ddcb45704c37893389b050ddbdc440a"

> doc3<-'{"rollno":"03","name":"XYZ","GRADE":"B","REMARK":"PASS"}'

> doc\_create(x,doc3,dbname = "ds",docid = "a\_3")

$ok

[1] TRUE

$id

[1] "a\_3"

$rev

[1] "1-683f5507333722ba7596bf4ad21635ad"

**Changes feed**

> db\_changes(x,"ds")

$results

$results[[1]]

$results[[1]]$seq

[1] "1-g1AAAAF1eJzLYWBg4MhgTmEQTM4vTc5ISXLIyU9OzMnILy7JAUoxJTIkyf\_\_\_z8rkQGPoiQFIJlkT1idA0hdPGF1CSB19QTV5bEASYYGIAVUOp8YtQsgavcTo\_YARO39rAzmRMZcoAC7Raqhgam5ITZ9BE17ADEN5KMsANoweP8"

$results[[1]]$id

[1] "a\_1"

$results[[1]]$changes

$results[[1]]$changes[[1]]

$results[[1]]$changes[[1]]$rev

[1] "1-be7c98bddf8ea7c46f4f401ff387593d"

$results[[2]]

$results[[2]]$seq

[1] "2-g1AAAAGXeJzLYWBg4MhgTmEQTM4vTc5ISXLIyU9OzMnILy7JAUoxJTIkyf\_\_\_z8rkQGPoiQFIJlkT1idA0hdPEhdBnMiYy6Qx55sYGSUnJqITQ8-kxJAJtUTtDGPBUgyNAApoNL5xKhdAFG7nxi1ByBq7yN8Y5FqaGBqbkiibyCmPYCYBvJRFgDah4NV"

$results[[2]]$id

[1] "a\_2"

$results[[2]]$changes

$results[[2]]$changes[[1]]

$results[[2]]$changes[[1]]$rev

[1] "1-1ddcb45704c37893389b050ddbdc440a"

$results[[3]]

$results[[3]]$seq

[1] "3-g1AAAAI7eJyV0EEOwiAQBVC0JurSE-gRoKYtruxNdIBpsMF2oV3rTfQmehO9SaXigibGyGZIhszLzDeEkImOFJnJupFaidzUEoyuD0djv4ZAxLxt21JHQPa2McaM8YIzRaZNpbDYVah-jIuFrWL9EQZvgUvJYk6\_zfyS8k7a9CRJ41gihErbTjr1ruKQpWyZ\_HlVNbKVnO1jkYuXDfKUcQxSrk65eQogFrAKUu5OeXgpI6NJxgKzcdrTaV4-lFKgAvydyhfU77A3"

$results[[3]]$id

[1] "a\_3"

$results[[3]]$changes

$results[[3]]$changes[[1]]

$results[[3]]$changes[[1]]$rev

[1] "1-683f5507333722ba7596bf4ad21635ad"

$last\_seq

[1] "3-g1AAAAI7eJyV0EEOwiAQBVC0JurSE-gRoKYtruxNdIBpsMF2oV3rTfQmehO9SaXigibGyGZIhszLzDeEkImOFJnJupFaidzUEoyuD0djv4ZAxLxt21JHQPa2McaM8YIzRaZNpbDYVah-jIuFrWL9EQZvgUvJYk6\_zfyS8k7a9CRJ41gihErbTjr1ruKQpWyZ\_HlVNbKVnO1jkYuXDfKUcQxSrk65eQogFrAKUu5OeXgpI6NJxgKzcdrTaV4-lFKgAvydyhfU77A3"

$pending

[1] 0

**#search forid > null so all docs will display**

> db\_query(x,dbname = "ds",selector = list('\_id'=list('$gt'=NULL)))$docs

[[1]]

[[1]]$`\_id`

[1] "a\_1"

[[1]]$`\_rev`

[1] "1-be7c98bddf8ea7c46f4f401ff387593d"

[[1]]$rollno

[1] "01"

[[1]]$name

[1] "ABC"

[[1]]$GRADE

[1] "A"

[[2]]

[[2]]$`\_id`

[1] "a\_2"

[[2]]$`\_rev`

[1] "1-1ddcb45704c37893389b050ddbdc440a"

[[2]]$rollno

[1] "02"

[[2]]$name

[1] "PQR"

[[2]]$GRADE

[1] "A"

[[3]]

[[3]]$`\_id`

[1] "a\_3"

[[3]]$`\_rev`

[1] "1-683f5507333722ba7596bf4ad21635ad"

[[3]]$rollno

[1] "03"

[[3]]$name

[1] "XYZ"

[[3]]$GRADE

[1] "B"

[[3]]$REMARK

[1] "PASS"

**#search for students with grade A**

> db\_query(x,dbname = "ds",selector = list(GRADE="A"))$docs

[[1]]

[[1]]$`\_id`

[1] "a\_1"

[[1]]$`\_rev`

[1] "1-be7c98bddf8ea7c46f4f401ff387593d"

[[1]]$rollno

[1] "01"

[[1]]$name

[1] "ABC"

[[1]]$GRADE

[1] "A"

[[2]]

[[2]]$`\_id`

[1] "a\_2"

[[2]]$`\_rev`

[1] "1-1ddcb45704c37893389b050ddbdc440a"

[[2]]$rollno

[1] "02"

[[2]]$name

[1] "PQR"

[[2]]$GRADE

[1] "A"

**#search for students with remark pass**

> db\_query(x,dbname = "ds",selector = list(REMARK="PASS"))$docs

[[1]]

[[1]]$`\_id`

[1] "a\_3"

[[1]]$`\_rev`

[1] "1-683f5507333722ba7596bf4ad21635ad"

[[1]]$rollno

[1] "03"

[[1]]$name

[1] "XYZ"

[[1]]$GRADE

[1] "B"

[[1]]$REMARK

[1] "PASS"

**#return only certain fields where rollno >2**

> db\_query(x,dbname = "ds",selector = list(rollno=list('$gt'='02')),fields = c("name","GRADE"))$docs

[[1]]

[[1]]$name

[1] "XYZ"

[[1]]$GRADE

[1] "B"

**#convert the result of a query into a data frame using jsonlite**

> library("jsonlite")

> res<-db\_query(x,dbname = "ds",selector = list('\_id'=list('$gt'=NULL)),fields = c("name","rollno","GRADE","REMARK"),as="json")

**#display json doc**

> fromJSON(res)$doc

name rollno GRADE REMARK

1 ABC 01 A <NA>

2 PQR 02 A <NA>

3 XYZ 03 B PASS

**#deleting the doc doc\_delete(cusion,dbname,docid)**

|  |
| --- |
| > doc\_delete(x,dbname = "ds",docid = "a\_2")  $ok  [1] TRUE  $id  [1] "a\_2"  $rev  [1] "2-82f1879cc7d73bef5574cc5cdf7c4094"  > doc\_get(x,dbname = "ds",docid = "a\_2")  Error: (404) – deleted  **#Updating the doc**  > doc2<-'{"name":"Sfood","biryani":"TEST","note":"yummy","note2":"yay"}'  > doc\_update(x,dbname = "ds",doc=doc2,docid = "a\_3",rev = "1-683f5507333722ba7596bf4ad21635ad")  $ok  [1] TRUE  $id  [1] "a\_3"  $rev  [1] "2-c8f7eac1d54375cf3371e975a944fe3c"  > doc3<-'{"rollno":"01","name":"UZMA","GRADE":"A"}'  > doc\_update(x,dbname = "ds",doc=doc3,docid = "a\_1",rev="1-be7c98bddf8ea7c46f4f401ff387593d")  $ok  [1] TRUE  $id  [1] "a\_1"  $rev  [1] "2-8e881d6a3e0fbdfd735da8ff70cff6cc"  **#Displaying after updating**  > db\_query(x,dbname = "ds",selector = list('\_id'=list('$gt'=NULL)))$docs  [[1]]  [[1]]$`\_id`  [1] "a\_1"  [[1]]$`\_rev`  [1] "2-8e881d6a3e0fbdfd735da8ff70cff6cc"  [[1]]$rollno  [1] "01"  [[1]]$name  [1] "UZMA"  [[1]]$GRADE  [1] "A"  [[2]]  [[2]]$`\_id`  [1] "a\_3"  [[2]]$`\_rev`  [1] "2-c8f7eac1d54375cf3371e975a944fe3c"  [[2]]$name  [1] "Sfood"  [[2]]$biryani  [1] "TEST"  [[2]]$note  [1] "yummy"  [[2]]$note2  [1] "yay" |
|  |
|  |
| |  | | --- | |  | |