Page No.: 5
Date: 33156 Babita Yadav

ASSIGNMENT-5

AIM: Perform different operations using R/python

PROBLEM STATEMENT: Perform following operations using R/python on the Amazon book store review & Facebook metrics dataset

a. create data subsets
b. Merge data
c. Sort data
d. Transposing data
e. Melting data to long format
f. Casting data to wide format.

THEORY:

1. SUBSETTING DATA:

R has a powerful indexing library for accessing object elements. These objects can be used to select & exclude variables & observations

2. SELECTING VARIABLES:

```
# select variables
myvars ← c("V1", "V2", "V3")
newdata ← mydata [myvars]

# another method
myvars ← paste(V, 1:3, sep: '')
newdata ← mydata [myvar]
```

3. EXCLUDING VARIABLES:

```
myvars ← names(mydata)
newdata ← mydata [!myvar]

# exclude 3rd & 5th variable
newdata ← mydata [c(-3,-5)]
```

Page No. :
Date :

Delete V3 & V5

mydata $V3 ← mydata $V5 ← NULL

3. SELECT USING SUBSET FUNCTION:

The subset function is the easiest way to select variables & observations

In following sample, we select all rows that have a value of age ≥ 20 or age < 10

newdata ← select(mydata, age >= 20 | age < 10)
select = c(Id, weight)

4. MERGING DATA:

- Adding columns:

To merge 2 dataframes horizontally, use merge formula func^n

total ← merge(dataframe A, dataframe B, by = "Id")

For more than 1 key variable,

total ← merge(dataframe A, dataframe B, by = c("Id", "country"))

- Adding Rows:

To join data vertically, use Rbind function

total ← Rbind(dataframe A, dataframe B)

- CONCLUSION:

In this assignment, we learnt basics of R operations on Facebook metrics dataset.