SESSION 4: FOUNDATIONAL R PROGRAMMING-II

Assignment 1

Problem Statement

1.

df1 = data.frame(CustId = c(1:6), Product = c(rep("TV", 3), rep("Radio", 3)))

df2 = data.frame(CustId = c(2, 4, 6), State = c(rep("Texas", 2), rep("NYC", 1)))

df1 #left table

df2 #right table

For the above given data frames and tables perform the following operations:

• Return only the rows in which the left table have match.

```
Library(sqldf)
df4 <- sqldf("SELECT CustId , Product , State
FROM df1
LEFT JOIN df2 USING(CustId)")
```

• Returns all rows from both tables, join records from the left which have matching keys in the right table.

```
library(plyr)
join(df1, df2,
    type = "inner")
```

• Return all rows from the left table, and any rows with matching keys from the right table.

```
solSpecs <- list(
  merge=list(testFuncs=list(left =function(df1,df2,key) merge(df1,df2,key,all.x=T)
```

• Return all rows from the right table, and any rows with matching keys from the left table.

```
library(dplyr)
```

right_join(df1, df2)

2. Perform the below operations on above given data frames and tables:

• Return a long format of the datasets without matching key.

• Keep only observations in df1 that match in df2.

List.df2 <- dput(as.character(df2$Header))

Df1 %% select(one_of(list.df2))

• Drop all observations in df1 that match in df2.

```
library("nycflights13")
flights2 <- flights %>% select(year:day, hour, origin, dest, tailnum, carrier)
flights2 %>%
  left_join(airlines)
```