#You have to include the following packages, in order to connect to POSTGRESQL Database

#library("RPostgreSQL")

#library("sqldf")

#If the packages "RPostgreSQL" and "sqldf" are not available, please install using the following commands (uncomment the following lines)

#install.packages("RPostgreSQL")

#install.packages("sqldf")

#Including the required packages

library("RPostgreSQL")

library("sqldf")

#Creating db driver object

drv <- dbDriver("PostgreSQL")

#Creating a connection object...You have to modify the host, port, username, passwd as per your postgres database configuration. Go to ToolS -> Server Configuration in PGADMINIII, to find the details of port number etc.

con <- dbConnect(drv, dbname="flights",host="localhost",port=5432,user="postgres",password="XXXXXXX")

#Executing the query. The query's results are collected to a cursor object or results set object

rs <- dbSendQuery(con, "

SELECT \* FROM

(

SELECT

A.ORIGIN AIRPORT\_OF\_ORIGIN,

A.CARRIER CARRIER,

C.TEMP TEMPERATURE,

A.DEP\_DELAY DEPARTURE\_DELAY,

A.ARR\_DELAY ARRIVAL\_DELAY,

A.AIR\_TIME AIR\_TIME,

B.SEATS SEATING\_CAPACITY,

TO\_DATE(TO\_CHAR(A.YEAR,'9999')||'-'||

LTRIM(TO\_CHAR(A.MONTH,'99'))

||'-'||

LTRIM(TO\_CHAR(A.DAY,'99')),'YYYY-MM-DD') AS DEP\_DATE

FROM FLIGHTS A, PLANES B, WEATHER C

WHERE A.DEST = 'LAX'

AND A.TAILNUM = B.TAILNUM

AND A.YEAR = C.YEAR

AND A.MONTH = C.MONTH

AND A.DAY = C.DAY

AND A.HOUR = C.HOUR

) AS TAB

WHERE DEP\_DATE BETWEEN '2/23/2013' AND '3/1/2013'

"

)

#Collecting the results to a data.frame "NY\_to\_LA\_Flights".

NY\_to\_LA\_Flights<- fetch(rs,n=-1)

#Disconnecting the database connection

dbDisconnect(con)

#To display all the data.frame contents, you may just list the data.frame name or use the "print"command

NY\_to\_LA\_Flights

#or use the print command as shown below

print(NY\_to\_LA\_Flights)

#The summary command will give the details of various columns/variables available in the data frame.

#For numeric variables, it shows the details like Mean, 25th percentile, 75th percentile, Min, Max, Median.

summary(NY\_to\_LA\_Flights)

#The head command will display the top rows (only some rows) of the data.frame

head(NY\_to\_LA\_Flights)

#The tail command will display the bottom rows (only some rows) of the data.frame

tail(NY\_to\_LA\_Flights)

#To display the number of rows of data frame, use "nrow" command

nrow(NY\_to\_LA\_Flights)

#The names command will display the names of the variables in the data frame.

names(NY\_to\_LA\_Flights)

#If you want to eliminate a column from the data frame, use NULL, as shown below.

#For example, in the data frame we created "NY\_to\_LA\_Flights", we included a column called

#dep\_date, but this column was not asked in the question/assignment. So we can delete this column as shown below:

NY\_to\_LA\_Flights$dep\_date <- NULL

#Let us display the names again, to verify if the dep\_date variable was deleted.

names(NY\_to\_LA\_Flights)

#You can observe that the dep\_date is no more present