

# IS 607 Assignment 5

*Derek G. Nokes*

*Friday, February 27, 2015*

## Contents

Connect to PostgreSQL . . . . .	2
Query PostgreSQL . . . . .	2
Data Frame Object Properties . . . . .	3
Data Frame Contents . . . . .	3
Disconnect From PostgreSQL . . . . .	5

## Connect to PostgreSQL

We connect to the PostgreSQL database as follows:

```
# load the RPostgreSQL library  
library('RPostgreSQL')
```

```
## Loading required package: DBI
```

```
# connect to PostgreSQL using RPostgreSQL  
driverHandle <- dbDriver("PostgreSQL")  
  
# establish connection to 'flights' database  
connectionHandle <- dbConnect(driverHandle,  
                               dbname="flights",  
                               host="localhost",  
                               port=5432,user="postgres",  
                               password ="tgdnrx78")
```

## Query PostgreSQL

We extract the flights data from any of the three major NYC airports to 'Los Angeles Intl' airport during the period between 2013-01-23 and 2013-03-01 and includes the following information:

- Date of Departure: d\_date
- Hour of departure: hour
- Airport of origin: o\_airport
- Airport of destination: d\_airport
- Carrier: carrier
- Approximate temperature at scheduled time of departure: d\_temp
- Departure delay (in minutes): d\_delay
- Arrival delay (in minutes): a\_delay
- Air time (in minutes): air\_time
- Seating capacity of the airplane: seating

```
# build the query (join the flights,planes,airports,and weather tables)  
query<-"SELECT (f.year::text || '-' || f.month::text || '-' || f.day::text)::date  
as d_date,w.hour,o.name as o_airport,d.name as d_airport,  
f.carrier,w.temp as d_temp,  
f.dep_delay as d_delay,arr_delay as a_delay,air_time,  
p.seats as seating FROM flights f  
INNER JOIN (SELECT * FROM planes) p ON f.tailnum = p.tailnum  
INNER JOIN (SELECT * FROM airports) d ON f.dest =d.faa  
INNER JOIN (SELECT * FROM airports) o ON f.origin = o.faa  
INNER JOIN (SELECT concat_ws('-',weather.year,weather.month,  
weather.day,weather.hour) as w_date_hour,weather.* FROM weather) w  
ON concat_ws('-',f.year,f.month,f.day,f.hour) = w.w_date_hour AND  
f.origin = w.origin  
WHERE (d.name='Los Angeles Intl')  
AND ((f.year::text || '-' || f.month::text || '-' || f.day::text)::date  
BETWEEN '2013-02-23' AND '2013-03-01')"
```

```
# execute the query
data<-dbGetQuery(connectionHandle,query)
```

## Data Frame Object Properties

We can look at the properties of the data frame as follows:

```
# look at the characteristics of the data data.frame object
str(data)
```

```
## 'data.frame':   42 obs. of  10 variables:
## $ d_date   : Date, format: "2013-02-23" "2013-02-23" ...
## $ hour     : int  8 11 13 16 17 6 8 10 13 15 ...
## $ o_airport: chr  "Newark Liberty Intl" "Newark Liberty Intl" "Newark Liberty Intl" "Newark Liberty
## $ d_airport: chr  "Los Angeles Intl" "Los Angeles Intl" "Los Angeles Intl" "Los Angeles Intl" ...
## $ carrier  : chr  "UA" "UA" "UA" "UA" ...
## $ d_temp   : num  37.9 37.9 39.2 39.9 41 ...
## $ d_delay  : int  14 1 5 55 9 4 -1 -2 -4 8 ...
## $ a_delay  : int  6 1 -2 47 -2 -19 -25 -28 -27 -9 ...
## $ air_time : int  336 351 337 335 340 337 330 311 325 325 ...
## $ seating  : int  275 191 178 178 149 149 275 178 178 178 ...
```

```
# look at the first 10 rows of the data frame
head(data,10)
```

```
##      d_date hour      o_airport      d_airport carrier d_temp
## 1  2013-02-23   8 Newark Liberty Intl Los Angeles Intl      UA  37.94
## 2  2013-02-23  11 Newark Liberty Intl Los Angeles Intl      UA  37.94
## 3  2013-02-23  13 Newark Liberty Intl Los Angeles Intl      UA  39.20
## 4  2013-02-23  16 Newark Liberty Intl Los Angeles Intl      UA  39.92
## 5  2013-02-23  17 Newark Liberty Intl Los Angeles Intl      UA  41.00
## 6  2013-02-24   6 Newark Liberty Intl Los Angeles Intl      UA  37.94
## 7  2013-02-24   8 Newark Liberty Intl Los Angeles Intl      UA  37.40
## 8  2013-02-24  10 Newark Liberty Intl Los Angeles Intl      UA  37.94
## 9  2013-02-24  13 Newark Liberty Intl Los Angeles Intl      UA  42.08
## 10 2013-02-24  15 Newark Liberty Intl Los Angeles Intl      UA  42.80
##      d_delay a_delay air_time seating
## 1         14      6      336      275
## 2          1      1      351      191
## 3          5     -2      337      178
## 4         55     47      335      178
## 5          9     -2      340      149
## 6          4    -19      337      149
## 7         -1    -25      330      275
## 8         -2    -28      311      178
## 9         -4    -27      325      178
## 10        8     -9      325      178
```

## Data Frame Contents

The following data table contains the data:

```
# creat the table
```

```
knitr::kable(data, caption = 'NYC flights to Los Angeles Intl Between 2013-02-23 and 2013-03-01')
```

d_date	hour	o_airport	d_airport	carrier	d_temp	d_delay	a_delay	air_time	seating
2013-02-23	8	Newark Liberty Intl	Los Angeles Intl	UA	37.94	14	6	336	275
2013-02-23	11	Newark Liberty Intl	Los Angeles Intl	UA	37.94	1	1	351	191
2013-02-23	13	Newark Liberty Intl	Los Angeles Intl	UA	39.20	5	-2	337	178
2013-02-23	16	Newark Liberty Intl	Los Angeles Intl	UA	39.92	55	47	335	178
2013-02-23	17	Newark Liberty Intl	Los Angeles Intl	UA	41.00	9	-2	340	149
2013-02-24	6	Newark Liberty Intl	Los Angeles Intl	UA	37.94	4	-19	337	149
2013-02-24	8	Newark Liberty Intl	Los Angeles Intl	UA	37.40	-1	-25	330	275
2013-02-24	10	Newark Liberty Intl	Los Angeles Intl	UA	37.94	-2	-28	311	178
2013-02-24	13	Newark Liberty Intl	Los Angeles Intl	UA	42.08	-4	-27	325	178
2013-02-24	15	Newark Liberty Intl	Los Angeles Intl	UA	42.80	8	-9	325	178
2013-02-24	17	Newark Liberty Intl	Los Angeles Intl	UA	48.92	31	14	317	178
2013-02-25	6	Newark Liberty Intl	Los Angeles Intl	UA	33.08	2	-15	335	149
2013-02-25	8	Newark Liberty Intl	Los Angeles Intl	UA	32.00	1	-10	335	275
2013-02-25	10	Newark Liberty Intl	Los Angeles Intl	UA	32.00	2	-2	328	178
2013-02-25	13	Newark Liberty Intl	Los Angeles Intl	UA	39.02	2	-7	334	178
2013-02-25	15	Newark Liberty Intl	Los Angeles Intl	UA	42.08	-4	5	346	178
2013-02-25	17	Newark Liberty Intl	Los Angeles Intl	UA	44.06	23	40	334	178
2013-02-26	6	Newark Liberty Intl	Los Angeles Intl	UA	35.96	4	-22	333	149
2013-02-26	8	Newark Liberty Intl	Los Angeles Intl	UA	33.98	-1	-30	319	149
2013-02-26	13	Newark Liberty Intl	Los Angeles Intl	UA	39.02	-1	-35	309	178
2013-02-26	15	Newark Liberty Intl	Los Angeles Intl	UA	44.06	16	-32	298	178
2013-02-26	17	Newark Liberty Intl	Los Angeles Intl	UA	44.06	19	-7	312	178
2013-02-27	5	Newark Liberty Intl	Los Angeles Intl	UA	37.94	-1	-40	315	149
2013-02-27	8	Newark Liberty Intl	Los Angeles Intl	UA	37.94	-4	-45	305	149
2013-02-27	10	Newark Liberty Intl	Los Angeles Intl	UA	39.92	-1	-22	295	191
2013-02-27	13	Newark Liberty Intl	Los Angeles Intl	UA	42.80	-3	-32	311	178
2013-02-27	15	Newark Liberty Intl	Los Angeles Intl	UA	44.60	-1	-36	314	178
2013-02-27	17	Newark Liberty Intl	Los Angeles Intl	UA	44.96	43	1	299	178
2013-02-28	5	Newark Liberty Intl	Los Angeles Intl	UA	42.98	-3	-51	318	149
2013-02-28	8	Newark Liberty Intl	Los Angeles Intl	UA	42.98	24	-38	292	275
2013-02-28	10	Newark Liberty Intl	Los Angeles Intl	UA	42.08	19	-17	285	178
2013-02-28	13	Newark Liberty Intl	Los Angeles Intl	UA	44.06	-1	-62	285	178
2013-02-28	16	Newark Liberty Intl	Los Angeles Intl	UA	50.00	77	24	292	178
2013-02-28	17	Newark Liberty Intl	Los Angeles Intl	UA	48.92	9	-38	296	178

d_date	hour	o_airport	d_airport	carrier	d_temp	d_delay	a_delay	air_time	seating
2013-02-28	18	Newark Liberty Intl	Los Angeles Intl	UA	48.92	-3	-35	308	178
2013-02-28	20	Newark Liberty Intl	Los Angeles Intl	UA	46.04	-3	-62	303	178
2013-03-01	6	Newark Liberty Intl	Los Angeles Intl	UA	35.96	-3	-68	294	149
2013-03-01	9	Newark Liberty Intl	Los Angeles Intl	UA	35.96	33	-1	299	275
2013-03-01	10	Newark Liberty Intl	Los Angeles Intl	UA	35.06	0	-43	304	149
2013-03-01	13	Newark Liberty Intl	Los Angeles Intl	UA	42.08	-5	-41	298	178
2013-03-01	15	Newark Liberty Intl	Los Angeles Intl	UA	42.08	-2	-32	308	178
2013-03-01	17	Newark Liberty Intl	Los Angeles Intl	UA	44.96	6	-23	304	275

Table 1: NYC flights to Los Angeles Intl Between 2013-02-23 and 2013-03-01

## Disconnect From PostgreSQL

Disconnect from the database:

```
# Close PostgreSQL connection
dbDisconnect(connectionHandle)
```

```
## [1] TRUE
```