

Lab3

Jiajun Zhang

April 12, 2019

```
library(RSQLite)
library(DBI)
dbcon = dbConnect(SQLite(), dbname = "stat240Week3lab.sqlite")
dbListTables(dbcon)

## [1] "CA"      "POP2006" "POP2011" "tickets" "zip"

names(dbReadTable(dbcon, "CA"))

## [1] "ID"          "Country"      "Geographic_name" "Region"
## [5] "Province"    "Prov_acr"     "Latitude"        "Longitude"
## [9] "Region_Index"

head(dbReadTable(dbcon, "zip"))

##   zip_code latitude longitude   city state   county
## 1      501 40.92233 -72.63708 Holtsville NY   Suffolk
## 2      544 40.92233 -72.63708 Holtsville NY   Suffolk
## 3      601 18.16527 -66.72258 Adjuntas PR   Adjuntas
## 4      602 18.39310 -67.18095 Aguada   PR   Aguada
## 5      603 18.45591 -67.14578 Aguadilla PR   Aguadilla
## 6      604 18.49352 -67.13588 Aguadilla PR   Aguadilla
```

Question1

```
#a)
names(dbReadTable(dbcon, "zip"))

## [1] "zip_code" "latitude" "longitude" "city"      "state"      "county"

#b)
tail(dbReadTable(dbcon, "CA"))

##           ID Country Geographic_name           Region
## 1635 1635      CA           S9H           Swift Current
## 1636 1636      CA           S9V           Lloydminster
## 1637 1637      CA           S9X           Meadow Lake
## 1638 1638      CA      Y0A Southeastern Yukon (Watson Lake)
## 1639 1639      CA      Y0B      Central Yukon (Dawson City)
## 1640 1640      CA      Y1A           Whitehorse
##           Province Prov_acr Latitude Longitude Region_Index
## 1635 Saskatchewan      SK 50.2834 -107.8013           4
## 1636 Saskatchewan      SK 53.2835 -110.0016           4
## 1637 Saskatchewan      SK 54.1335 -108.4347           4
## 1638      Yukon      YT 61.5793 -131.1481           NA
## 1639      Yukon      YT 64.6450 -137.5360           NA
## 1640      Yukon      YT 60.7161 -135.0537           4

#c)
head(dbReadTable(dbcon, "POP2011"), 3)
```

##	Geographic_name	
## 1	Canada	
## 2	AOA	
## 3	AOB	
##	Incompletely_enumerated_Indian_reserves_and_Indian_settlements__2011	
## 1		1
## 2		NA
## 3		NA
##	Population__2011	Total_private_dwellings__2011
## 1	33476688	14569633
## 2	46297	23950
## 3	20985	12585
##	Private_dwellings_occupied_by_usual_residents__2011	
## 1		13320614
## 2		18701
## 3		8854

Question2

```
zip_df=dbReadTable(dbcon, "zip")  
# unique(zip_df$state)  
head(dbGetQuery(dbcon, "SELECT city, state, zip_code FROM zip WHERE state=='IL'"), 10)
```

```
##           city state zip_code  
## 1         Alden   IL    60001  
## 2        Antioch   IL    60002  
## 3 Arlington Heights IL    60004  
## 4 Arlington Heights IL    60005  
## 5 Arlington Heights IL    60006  
## 6 Elk Grove Village IL    60007  
## 7   Rolling Meadows IL    60008  
## 8 Elk Grove Village IL    60009  
## 9      Barrington   IL    60010  
## 10     Barrington   IL    60011
```

Question3

#a)

```
head(dbGetQuery(dbcon, "SELECT * FROM zip INNER JOIN tickets ON zip.zip_code=tickets.zip_code"),10)
```

##	zip_code	latitude	longitude	city	state	county	X
## 1	60647	41.81193	-87.68732	Chicago	IL	Cook	27000000
## 2	60653	41.81193	-87.68732	Chicago	IL	Cook	27000001
## 3	60617	41.81193	-87.68732	Chicago	IL	Cook	27000002
## 4	60641	41.81193	-87.68732	Chicago	IL	Cook	27000003
## 5	60629	41.81193	-87.68732	Chicago	IL	Cook	27000004
## 6	60645	42.00808	-87.72146	Chicago	IL	Cook	27000005
## 7	60639	41.81193	-87.68732	Chicago	IL	Cook	27000006
## 8	46835	41.15670	-85.05722	Fort Wayne	IN	Allen	27000007
## 9	60614	41.81193	-87.68732	Chicago	IL	Cook	27000008
## 10	94609	37.83478	-122.26328	Oakland	CA	Alameda	27000009

##	ticket_number	issue_date	violation_location
## 1	68912194	2017-10-02 10:07:00	5534 S JUSTINE
## 2	68910503	2017-10-02 10:07:00	6515 S YALE
## 3	9192064933	2017-10-02 10:08:00	8122 S LUELLA
## 4	9192430641	2017-10-02 10:08:00	2703 N KEDZIE AVE
## 5	9192412530	2017-10-02 10:08:00	850 N DAMEN
## 6	9192381646	2017-10-02 10:08:00	2745 W ARTHUR
## 7	9192329380	2017-10-02 10:08:00	1605 N MAYFIELD
## 8	9192378785	2017-10-02 10:08:00	740 N WASHTENAW
## 9	9192331677	2017-10-02 10:08:00	538 W MENOMONEE ST
## 10	9192296920	2017-10-02 10:08:00	1150 E 48TH ST

##	license_plate_number
## 1	0e416da71ddb6a9d789ce31f1baace4aa72fe25e0c2743d7f71cb291e1a630bc
## 2	d21b8521e9fb2f938c767265e822dad585daa0609b4e5dc5c09559fd8e3e49e4
## 3	66c348d81a8296e4073ff95dfd19330d3df77a5130e2f248348504850f867a30
## 4	2ed73825fa6a712269bebf3e727f8fc34703d76427d33be6816000a693db7f39
## 5	3f0c45c8683e5dfcd57cd55828ed3530eefa9ded64612ac4017ecb8830217cae
## 6	49d4d613af3f3b3b99bce83c9736f156fb28b8d552fcff0a9ae7d9bef555fbfb
## 7	c655bf9eeb3ca553be8dc699d6822fe259110325d3ff1309156c790384714e07
## 8	746be0558d23649de4355b56c533d24738204a1b4ea7167fe89b3de9ae62fab6
## 9	0f9f3f113297c0eb827c1931df8f9e3f013249d5d3ed6ac81e4a24610d11279c
## 10	2e58ef83a57857ccc7d79310bd87ef2b87632b3d84a1e966bb36161add492232

##	license_plate_state	license_plate_type	zipcode	violation_code
## 1	IL	PAS	606474822.0	0976160B
## 2	IL	PAS	606531302.0	0964040B
## 3	IL	PAS	606171152.0	0976160B
## 4	IL	PAS	60641.0	0964090E
## 5	IL	PAS	606293730.0	0964040B
## 6	IL	PAS	60645.0	0964040B
## 7	IL	PAS	606394013.0	0964040B
## 8	IN	PAS	468359693.0	0964040B
## 9	IL	PAS	606145811.0	0976160B
## 10	IL	PAS	946093006.0	0964040B

##	violation_description	unit	unit_description
## 1	EXPIRED PLATE OR TEMPORARY REGISTRATION	7	CPD
## 2	STREET CLEANING	7	CPD
## 3	EXPIRED PLATE OR TEMPORARY REGISTRATION	498	DOF
## 4	RESIDENTIAL PERMIT PARKING	498	DOF

```

## 5          STREET CLEANING 498          DOF
## 6          STREET CLEANING 498          DOF
## 7          STREET CLEANING 498          DOF
## 8          STREET CLEANING 498          DOF
## 9 EXPIRED PLATE OR TEMPORARY REGISTRATION 498          DOF
## 10         STREET CLEANING 498          DOF
##      vehicle_make fine_level1_amount fine_level2_amount current_amount_due
## 1      TOYT          60          120          146.4
## 2      CHEV          60          120           0.0
## 3      CHRY          60          120          146.4
## 4      HYUN          75          150          183.0
## 5      LEXU          60          120           0.0
## 6      ACUR          60          120           0.0
## 7      NISS          60          120           0.0
## 8      HOND          60          120           0.0
## 9      KIA          60          120           0.0
## 10     TOYT          60          120           0.0
##      total_payments ticket_queue ticket_queue_date notice_level
## 1          0.0      Notice 2017-10-19 00:00:00      SEIZ
## 2         146.4      Paid 2018-04-26 00:00:00      SEIZ
## 3          0.0      Notice 2017-10-05 00:00:00      SEIZ
## 4          0.0      Notice 2017-10-05 00:00:00      FINL
## 5         60.0      Paid 2017-11-08 00:00:00      VIOL
## 6         60.0      Paid 2017-11-04 00:00:00      VIOL
## 7         60.0      Paid 2017-10-19 00:00:00      VIOL
## 8         60.0      Paid 2017-11-03 00:00:00      VIOL
## 9         60.0      Paid 2017-10-08 00:00:00      VIOL
## 10        60.0      Paid 2017-11-19 00:00:00      DETR
##      hearing_disposition notice_number officer
## 1          5208129220      10582
## 2          5181208890      18851
## 3          5082297980       823
## 4          5207777060       287
## 5          5193607370       129
## 6          5207798220       522
## 7          5199982480       657
## 8          5207774900       114
## 9          5202200770       642
## 10         5189177380       761
##      address zip_code..31
## 1      5500 s justine, chicago, il      60647
## 2      6500 s yale, chicago, il      60653
## 3      8100 s luella, chicago, il      60617
## 4      2700 n kedzie ave, chicago, il      60641
## 5      800 n damen, chicago, il      60629
## 6      2700 w arthur, chicago, il      60645
## 7      1600 n mayfield, chicago, il      60639
## 8      700 n washtenaw, chicago, il      46835
## 9      500 w menomonee st, chicago, il      60614
## 10     1100 e 48th st, chicago, il      94609

```

```

#b)
names(dbReadTable(dbcon, "POP2011"))

```

```
## [1] "Geographic_name"
```

```
## [2] "Incompletely_enumerated_Indian_reserves_and_Indian_settlements__2011"
## [3] "Population__2011"
## [4] "Total_private_dwellings__2011"
## [5] "Private_dwellings_occupied_by_usual_residents__2011"
```

```
names(dbReadTable(dbcon, "POP2006"))
```

```
## [1] "Geographic_name"
## [2] "Incompletely_enumerated_Indian_reserves_and_Indian_settlements__2006"
## [3] "Population__2006"
## [4] "Total_private_dwellings__2006"
## [5] "Private_dwellings_occupied_by_usual_residents__2006"
```

```
head(dbGetQuery(dbcon, "SELECT * FROM POP2011 INNER JOIN POP2006
ON POP2011.Geographic_name = POP2006.Geographic_name"),5)
```

```
## Geographic_name
## 1 Canada
## 2 AOA
## 3 AOB
## 4 AOC
## 5 AOE
## Incompletely_enumerated_Indian_reserves_and_Indian_settlements__2011
## 1 1
## 2 NA
## 3 NA
## 4 NA
## 5 NA
## Population__2011 Total_private_dwellings__2011
## 1 33476688 14569633
## 2 46297 23950
## 3 20985 12585
## 4 12834 8272
## 5 23384 12733
## Private_dwellings_occupied_by_usual_residents__2011 Geographic_name..6
## 1 13320614 Canada
## 2 18701 AOA
## 3 8854 AOB
## 4 5482 AOC
## 5 9659 AOE
## Incompletely_enumerated_Indian_reserves_and_Indian_settlements__2006
## 1 1
## 2 NA
## 3 NA
## 4 NA
## 5 NA
## Population__2006 Total_private_dwellings__2006
## 1 31612897 13576855
## 2 46648 23052
## 3 22099 12284
## 4 13767 7788
## 5 24400 12735
## Private_dwellings_occupied_by_usual_residents__2006
## 1 12435520
## 2 17994
```

3
4
5

8799
5598
9575

Question4

```
head(dbGetQuery(dbcon, "SELECT * FROM zip INNER JOIN tickets ON
                        zip.zip_code=tickets.zip_code WHERE tickets.fine_level1_amount>100"),6)
```

```
##      zip_code latitude longitude      city state county      X ticket_number
## 1      60614 41.81193 -87.68732 Chicago    IL    Cook 27000019      9192331678
## 2      60617 41.81193 -87.68732 Chicago    IL    Cook 27000022      9192064934
## 3      60623 41.81193 -87.68732 Chicago    IL    Cook 27000030      9191906374
## 4      60639 41.81193 -87.68732 Chicago    IL    Cook 27000033      9192448302
## 5      60647 41.81193 -87.68732 Chicago    IL    Cook 27000034      9192411382
## 6      61603 40.77985 -89.55406 Peoria     IL Peoria 27000063      9192351940
##
##      issue_date  violation_location
## 1 2017-10-02 10:09:00  538 W MENOMONEE ST
## 2 2017-10-02 10:09:00      8122 S LUELLA
## 3 2017-10-02 10:10:00  1812 S MILLARD AVE
## 4 2017-10-02 10:10:00      2757 N NEWLAND
## 5 2017-10-02 10:10:00  2201 N CENTRAL PARK
## 6 2017-10-02 10:11:00      40 W LAKE ST
##
##      license_plate_number
## 1 0f9f3f113297c0eb827c1931df8f9e3f013249d5d3ed6ac81e4a24610d11279c
## 2 66c348d81a8296e4073ff95dfd19330d3df77a5130e2f248348504850f867a30
## 3 23567beaea72dd68f638b07571d3711aec7bcf87f9ed7dd26aed13fb94f04443
## 4 94e9971fb6fb7130ba7b81658178c83fcd21fe698a3dc92feee9b1d3b7a48f25
## 5 78c44d3d6617eabcb87135e69b0edf11132e8fd8778f247fa7a547b5ef7304412
## 6 89c056ac4b4de5368d900eccdc2de77bc73c90ad5e75268163f4cf8df58dd51e
##      license_plate_state license_plate_type      zipcode violation_code
## 1                      IL                      PAS 606145811.0      0964125B
## 2                      IL                      PAS 606171152.0      0964125B
## 3                      IL                      PAS 606233753.0      0964125B
## 4                      IL                      PAS      60639.0      0964125B
## 5                      IL                      PAS 606472315.0      0964125B
## 6                      IL                      TMP      61603.0      0964190C
##
##      violation_description unit unit_description
## 1 NO CITY STICKER VEHICLE UNDER/EQUAL TO 16,000 LBS. 498      DOF
## 2 NO CITY STICKER VEHICLE UNDER/EQUAL TO 16,000 LBS. 498      DOF
## 3 NO CITY STICKER VEHICLE UNDER/EQUAL TO 16,000 LBS. 498      DOF
## 4 NO CITY STICKER VEHICLE UNDER/EQUAL TO 16,000 LBS. 498      DOF
## 5 NO CITY STICKER VEHICLE UNDER/EQUAL TO 16,000 LBS. 498      DOF
## 6 NON PYMT/NON-COM VEH PARKED IN COM LOADING ZONE 498      DOF
##      vehicle_make fine_level1_amount fine_level2_amount current_amount_due
## 1      KIA      200      400      0.00
## 2      CHRY      200      400      488.00
## 3      DODG      200      400      373.66
## 4      MAZD      200      400      488.00
## 5      CHEV      200      400      0.00
## 6      CHRY      140      280      140.00
##      total_payments ticket_queue      ticket_queue_date notice_level
## 1      200.00      Paid 2017-10-08 00:00:00      VIOL
## 2      0.00      Notice 2017-10-05 00:00:00      SEIZ
## 3      114.34      Notice 2017-10-05 00:00:00      SEIZ
## 4      0.00      Notice 2017-10-05 00:00:00      FINL
## 5      200.00      Paid 2017-11-13 00:00:00      DETR
## 6      0.00      Define 2017-10-03 00:00:00
```



```
##   hearing_disposition notice_number officer
## 1                      5202200770      642
## 2                      5082297980      823
## 3                      5014482510      219
## 4                      5207777290      638
## 5                      5178722170      737
## 6                      0             297
##                                address zip_code..31
## 1 500 w menomonee st, chicago, il      60614
## 2   8100 s luella, chicago, il      60617
## 3 1800 s millard ave, chicago, il      60623
## 4   2700 n newland, chicago, il      60639
## 5 2200 n central park, chicago, il      60647
## 6    1 w lake st, chicago, il      61603
```

```
dim(dbGetQuery(dbcon, "SELECT * FROM zip INNER JOIN tickets ON
                        zip.zip_code=tickets.zip_code WHERE tickets.fine_level1_amount>100"))
```

```
## [1] 64916    31
```

Question5

```
pop2006_sql=dbGetQuery(dbcon, "SELECT * FROM POP2006 INNER JOIN CA ON  
POP2006.Geographic_name = CA.Geographic_name WHERE POP2006.Population_2006 < 27368")
```

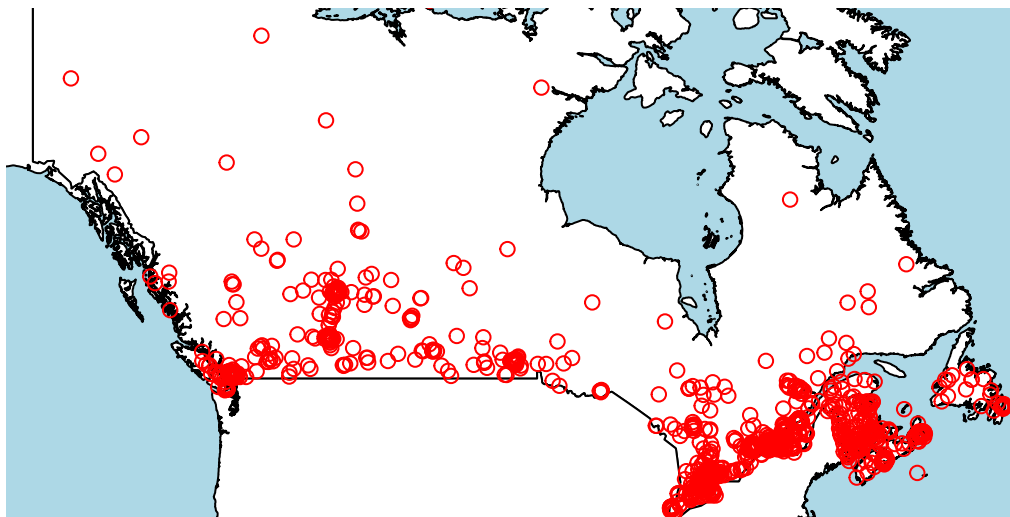
```
library(sp); library(rworldmap); library(rworldxtra)
```

```
## ### Welcome to rworldmap ###
```

```
## For a short introduction type : vignette('rworldmap')
```

```
worldmap = getMap(resolution = "high")  
NrthAm = worldmap[which(worldmap$REGION == "North America"), ]  
plot(NrthAm, col = "white", bg = "lightblue", xlim = c(-140, -55), ylim = c(50, 60),  
      main="Population in Canada in 2006 \n where the Population is Less Than 27368")  
points(x=pop2006_sql$Longitude, y=pop2006_sql$Latitude, col=2)
```

**Population in Canada in 2006
where the Population is Less Than 27368**



```
dbGetQuery(dbcon, "SELECT * FROM POP2011 WHERE Geographic_name LIKE 'V5%' ")
```

##	Geographic_name	
## 1	V5A	
## 2	V5B	
## 3	V5C	
## 4	V5E	
## 5	V5G	
## 6	V5H	
## 7	V5J	
## 8	V5K	
## 9	V5L	
## 10	V5M	
## 11	V5N	
## 12	V5P	
## 13	V5R	
## 14	V5S	
## 15	V5T	
## 16	V5V	
## 17	V5W	
## 18	V5X	
## 19	V5Y	
## 20	V5Z	
##	Incompletely_enumerated_Indian_reserves_and_Indian_settlements__2011	
## 1		NA
## 2		NA
## 3		NA
## 4		NA
## 5		NA
## 6		NA
## 7		NA
## 8		NA
## 9		NA
## 10		NA
## 11		NA
## 12		NA
## 13		NA
## 14		NA
## 15		NA
## 16		NA
## 17		NA
## 18		NA
## 19		NA
## 20		NA
##	Population__2011	Total_private_dwellings__2011
## 1	23803	9092
## 2	22312	8770
## 3	28373	12301
## 4	26291	10868
## 5	19205	7212
## 6	35943	16953
## 7	20331	7484
## 8	22545	8760
## 9	18454	10597

```
## 10      22729      8211
## 11      33157     14285
## 12      32082     10912
## 13      45899     17022
## 14      27736      9938
## 15      21905     12641
## 16      23132      9055
## 17      19219      7321
## 18      28509      9240
## 19      13160      6644
## 20      21413     10852
## Private_dwellings_occupied_by_usual_residents__2011
## 1      8620
## 2      8332
## 3     11587
## 4     10280
## 5      6938
## 6     16058
## 7      7129
## 8      8098
## 9      9524
## 10     7706
## 11     13293
## 12     10084
## 13     16063
## 14      9646
## 15     11991
## 16      8461
## 17      6750
## 18      8525
## 19      5828
## 20     10286
```

Question6

```
# names(dbReadTable(dbcon, "tickets"))
parking_df=dbGetQuery(dbcon, " SELECT * FROM tickets WHERE violation_description LIKE '%parking%' ")
dim(parking_df)
```

```
## [1] 117375      25
```

```
table(parking_df$vehicle_make)
```

```
##
## ACUR ALFA AUDI BENE BENT BMTA BMW BUIC CADI CHEC CHEV CHRY
## 1864 8 1479 6 26 1 2380 1986 1386 63 13779 2694
## DAEW DATS DODG DUCA EGIL FERR FIAT FLNR FORD FTWD GELY GENU
## 10 1 4936 2 2 9 141 838 12515 1 1 19
## GEO GMC HARL HINO HOND HUMM HYUN IND INFI INTL ISU JAGU
## 25 2262 24 185 10394 73 4788 5 1394 529 390 192
## JEEP KAWA KENW KIA KTM KYMC LAMO LCE LEXU LINC LING LNCI
## 5450 26 62 2616 3 8 2 2 2195 986 1 3
## LNDR LONT MACK MASE MAZD MERC MERK MERZ MILA MINI MITS MOGU
## 469 3 6 25 2098 884 2 2126 1 274 1208 1
## MTCH NISS OLDS OTHR PACK PETR PIAG PLYM POLS PONT PORS PTRB
```

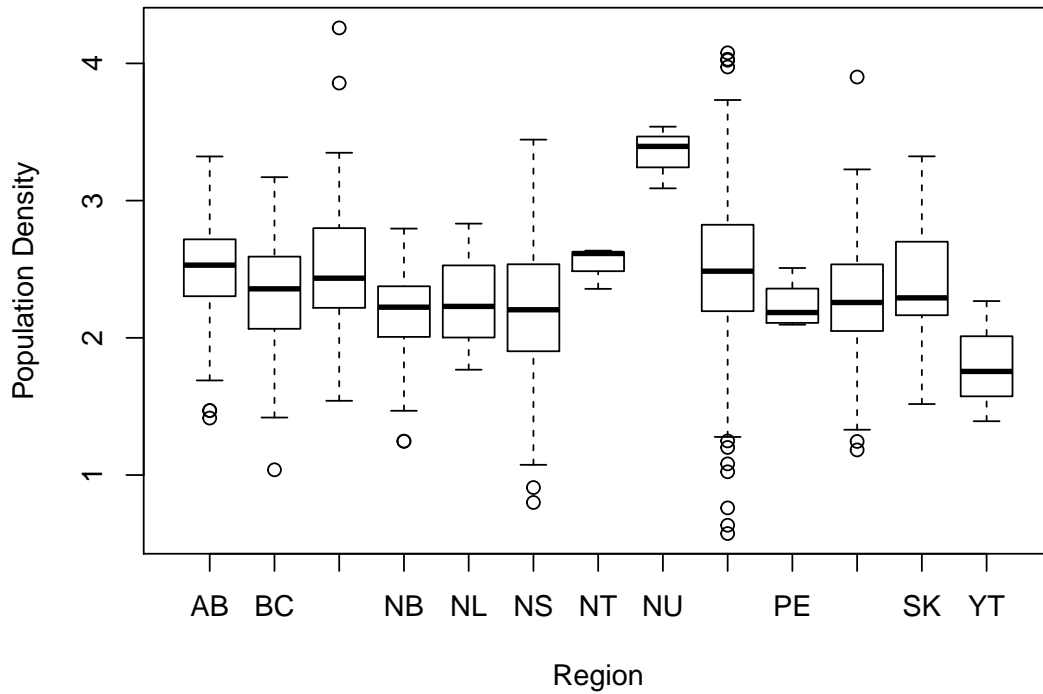
##	2	10044	358	64	1	26	8	54	2	1672	241	13
##	RAMB	ROL	ROV	RROV	SAAB	SANY	SATR	SCHW	SCIO	SMRT	SPEC	SPNT
##	5	5	138	111	134	1	725	1	481	42	3	11
##	STRG	SUBA	SUZI	SYMP	TAIZ	TAOT	TESL	THMP	TOYT	TRAL	TRIU	UNKN
##	13	1595	251	2	1	1	113	5	14253	12	15	22
##	VESP	VOLK	VOLV	VPGP	WORH	WSTR	YAMA					
##	13	3209	770	2	75	1	21					

Question7

```
POP2006=dbGetQuery(dbcon, "SELECT * FROM POP2006 INNER JOIN CA ON
                           POP2006.Geographic_name = CA.Geographic_name WHERE POP2006.Population__2006>=100")
par(mfrow=c(2,1))
pop_density=POP2006$Population__2006/POP2006$Total_private_dwellings__2006
boxplot(pop_density~POP2006$Prov_acr, ylab="Population Density", xlab="Region",
        main="Canadian Population Density within \n Each Geographical Region in 2006")

boxplot(POP2006$Population__2006~POP2006$Prov_acr, ylab="Population", xlab="Region",
        main="Canadian Population Level within \n Each Geographical Region in 2006")
abline(h=5000, lty=2, col=2)
```

**Canadian Population Density within
Each Geographical Region in 2006**



**Canadian Population Level within
Each Geographical Region in 2006**

