Assignment_2

Juliette Verstaen

5/2/2019

Contents

A. Import and Tidy

SB county parcels data

```
###load packages
library(tidyverse)
```

```
## -- Attaching packages ---------------------------------------------------------------------- tidyverse
```

```
## v ggplot2 3.1.0     v purrr   0.2.5
## v tibble  2.0.1     v dplyr   0.7.8
## v tidyr   0.8.2     v stringr 1.3.1
## v readr   1.3.1     v forcats 0.3.0
```

```
## -- Conflicts ------------------------------------------------------------------------------- tidyverse_confl
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
```

```
###set working directory to Assignment_1 folder
getwd()
```

```
## [1] "/GitHub/ESM262/Assignment_2"
```

```
###load in raw data
parcels_raw <-
  read_csv("data/Santa_Barbara_County_parcels_2011.csv",
          col_types = cols(.default = col_character())) %>%
  as_tibble()

###select only columns of interest
parcels <-
  transmute(parcels_raw,
    APN       = APN,
    Situs1    = Situs1,
    Situs2    = Situs2,
    Acreage   = parse_double(Acreage, na = "0"),
    UseCode   = UseCode,
    NonTaxCode= NonTaxCode,
```

```
    AgPres    = AgPres,
    LandValue = parse_double(LandValue, na = "0"),
    Net_Impr  = parse_double(Net_Impr, na = "0"),
    Net_AV    = parse_double(Net_AV, na = "0"),
    M_Address1 = M_Address1,
    M_Address2 = M_Address2)

###Convert all blanks in tibble to NAs
parcels[is.na(parcels)] <- NA
```

Use code data

```
###load in use code data
use_code <- read_delim("data/UseCodes.csv",
                       delim="|",
                       quote= "",
                       col_types = cols(.default = col_character()))

use_code[is.na(use_code)] <- NA
```

Creating a connection to SQLite and a parcels database

```
library(DBI)
  ###establish connection
db <- dbConnect(RSQLite::SQLite(),
                dbname = "parcels.db")

  ###add in parcels data table
dbWriteTable(conn = db, name = "parcels", value = parcels, overwrite = TRUE)

  ###add in use code data table
dbWriteTable(conn = db, name = "use_code", value = use_code, overwrite = TRUE)

  ### arguments within the dbWriteTable()
  ### the first argument is the connection handle;

  ### the second is the name for the table inside the database;

  ### the third is the dataframe with all your beautifully processed data.

dbListTables(db) ### double checking that tables are there. YAY it works
```

```
## [1] "parcels"  "use_code"
```
```
#dbDisconnect(db)
```

B. Analyze the data using SQL

1.What are the 10 most-frequently-occuring land uses (in descending order)?

```
SELECT parcels.UseCode, use_code.CodeDesc
FROM   parcels JOIN use_code
ON     parcels.UseCode = use_code.UseCode
Group BY use_code.CodeDesc;
```

```sql
Select count(CodeDesc);


SELECT parcels.UseCode, use_code.CodeDesc, COUNT(use_code.CodeDesc) from table group by use_code.CodeDes
FROM   parcels JOIN use_code
ON     parcels.UseCode = use_code.UseCode;
GROUP BY use_code.CodeDesc

select name, count(*) from table group by name
```

Table 1: Displaying records 1 - 10

UseCode	CodeDesc
2413	NA
5631	Ag pre, cont. restrict., Race
5632	Ag pre, cont. restrict., Show
5633	Ag pre. cont. restrict., Boarding
5634	Ag pre. cont. restrict., Riding Stables
5635	Ag pre., cont. restrict., Other
4447	Ag preserve, Notice non-renewal, cash rent $200-250ac
4445	Ag preserve, Notice non-renewal, cash rent $300-400ac
4444	Ag preserve, Notice non-renewal, cash rent over $400ac
4448	Ag preserve, Notice non-renewal, cash rent under$200ac

2. How many acres are in agricultural preserves?

```sql
SELECT sum(Acreage)
FROM   parcels JOIN use_code
ON     parcels.UseCode = use_code.UseCode
WHERE AgPres IS NOT NULL;
```

Table 2: 1 records

sum(Acreage)
549563.3511

3. What is the mean net assessed value per acre of the entire county?

```sql
SELECT round(sum(Net_AV)/sum(Acreage), 2)
FROM   parcels JOIN use_code
ON     parcels.UseCode = use_code.UseCode
WHERE Net_AV >=0;
```

Table 3: 1 records

round(sum(Net_AV)/sum(Acreage), 2)
65159.04

4. What is the total net assessed value of all non-taxable parcels?

Assume non-taxable parcels are indicated by non-NA values for NonTaxCode.

```sql
SELECT sum(Net_AV)
FROM   parcels JOIN use_code
ON     parcels.UseCode = use_code.UseCode
WHERE  NonTaxCode IS NOT NULL AND Net_AV >=0 ;
```

Table 4: 1 records

sum(Net__AV)
1095897708

5. What are the 10 largest property holders, by acreage?

```sql
SELECT parcels.M_Address1, parcels.M_Address2, sum(Acreage)
FROM   parcels JOIN use_code
ON     parcels.UseCode = use_code.UseCode
WHERE M_Address1 IS NOT NULL OR M_Address2 IS NOT NULL
GROUP BY parcels.M_Address1, parcels.M_Address2
Order By sum(Acreage) DESC
LIMIT 10;
```

Table 5: Displaying records 1 - 10

M__Address1	M__Address2	sum(Acreage)
785 MARKET ST	SAN FRANCISCO CA 94103	37860.00
166 PARADISE RD	SANTA BARBARA CA 93105	30109.71
PO BOX 573	LAKE ARROWHEAD CA 92352	24262.89
870 MARKET ST 1100	SAN FRANCISCO CA 94102	23833.67
2491 BULL CANYON RD	SANTA MARIA CA 93454	18594.80
201 MISSION ST 4TH FLR	SAN FRANCISCO CA 94105 1831	16640.00
PO BOX 340	SAN LUCAS CA 93954	15373.61
PO BOX 9	SAN JUAN CAPISTRANO CA 92693	10170.02
2300 EAST VALLEY RD	SANTA BARBARA CA 93108	9800.95
PO BOX 338	SANTA YNEZ CA 93460	9189.45

6. What are the 10 largest property holders, by net assessed value?

```sql
SELECT parcels.M_Address1, parcels.M_Address2, sum(Net_AV)
FROM   parcels JOIN use_code
ON     parcels.UseCode = use_code.UseCode
WHERE M_Address1 IS NOT NULL OR M_Address2 IS NOT NULL
GROUP BY parcels.M_Address1, parcels.M_Address2
Order By sum(Net_AV) DESC
LIMIT 10;
```

Table 6: Displaying records 1 - 10

M_Address1	M_Address2	sum(Net_AV)
737 GARDEN ST	SANTA BARBARA CA 93101	242575885
735 ANACAPA ST	SANTA BARBARA CA 93101	190784490
21 E VICTORIA ST 200	SANTA BARBARA CA 93101	179642731
PO BOX 20130	SANTA BARBARA CA 93120 0130	177735038
280 CHESTNUT	WESTMONT IL 60559	157323864
PO BOX 660248	DALLAS TX 75266 0248	151925292
1260 CHANNEL DR	SANTA BARBARA CA 93105	120001340
645 FIFTH AVE 8	NEW YORK NY 10022	118000000
500 STEVENS AVE 100	SOLANA BEACH CA 92075	107926369
2235 FARADAY AVE O	CARLSBAD CA 92008	104533906

Close the sql database connection dbDisconnect(db)