ESM262-Assignment2-Vela

Alejandro Vela

5/14/2018

Part 1: Import and Tidy

Part 1-1: Load Libraries

```
library(plyr)
library(tidyverse)
```

```
## ── Attaching packages ─────────────────────────────────────────────────────
──────────────────── tidyverse 1.2.1 ──
```

```
## ✔ ggplot2 2.2.1     ✔ purrr   0.2.4
## ✔ tibble  1.4.2     ✔ dplyr   0.7.4
## ✔ tidyr   0.8.0     ✔ stringr 1.3.0
## ✔ readr   1.1.1     ✔ forcats 0.3.0
```

```
## ── Conflicts ────────────────────────────────────────────────────────────
──────────────── tidyverse_conflicts() ──
## ✖ dplyr::arrange()   masks plyr::arrange()
## ✖ purrr::compact()   masks plyr::compact()
## ✖ dplyr::count()     masks plyr::count()
## ✖ dplyr::failwith()  masks plyr::failwith()
## ✖ dplyr::filter()    masks stats::filter()
## ✖ dplyr::id()        masks plyr::id()
## ✖ dplyr::lag()       masks stats::lag()
## ✖ dplyr::mutate()    masks plyr::mutate()
## ✖ dplyr::rename()    masks plyr::rename()
## ✖ dplyr::summarise() masks plyr::summarise()
## ✖ dplyr::summarize() masks plyr::summarize()
```

```
library(readr)
library(tibble)
library(dplyr)
library(tidyr)
library(DBI)
library(sqldf)
```

```
## Loading required package: gsubfn
```

```
## Warning: package 'gsubfn' was built under R version 3.4.4
```

```
## Loading required package: proto
```

```
## Loading required package: RSQLite
```

Part 1-2: Import Data

```
# Import the data we want to store in the SQL database
data = read.csv("Gazetteer.csv", sep="|")
```

Part 1-3: Connect to our Database

```
# Connect to our database
db<- dbConnect(SQLite(), dbname="gaz.db")
```

Part 1-4: Write the Data in our Database

```
# Write the data to our database
dbWriteTable(db, "data", data)
```

Part 2-1: Display Highest Count of Feature Name

```
# Query our database to display us the highest count of feature Name
dbGetQuery(db, 'SELECT Name FROM data GROUP BY Name ORDER BY COUNT(*) DESC LIMIT 1')
```

```
##                Name
## 1 Church of Christ
```

Part 2-2: Display Lowest Count of Feature Class

```
# Query our database to display us the lowest count of feature Class
dbGetQuery(db, 'SELECT Class FROM data GROUP BY Class ORDER BY COUNT(*) ASC LIMIT 1')
```

```
##     Class
## 1 Isthmus
```

Part 2-3: Display Center of Each County

```
# Display center of each county
dbGetQuery(db, 'SELECT County, AVG(Latitude_DEC) AS
           "Approximate Center Latitude", AVG(Longitude_DEC) AS
           "Approximate Center Longitude" FROM data GROUP BY County')
```

```
##             County Approximate Center Latitude
## 1             <NA>                     38.28899
## 2          Alameda                     37.72641
## 3           Alpine                     38.46465
## 4           Amador                     38.24350
## 5            Butte                     39.20366
## 6        Calaveras                     38.02912
## 7           Colusa                     39.04933
## 8     Contra Costa                     37.92377
## 9        Del Norte                     41.50418
## 10       El Dorado                     38.30023
## 11          Fresno                     36.63115
## 12           Glenn                     39.55104
## 13        Humboldt                     40.48476
## 14        Imperial                     32.88554
## 15            Inyo                     36.61158
## 16            Kern                     35.36715
## 17           Kings                     36.16381
## 18            Lake                     38.04076
## 19           Lassen                     40.60123
## 20     Los Angeles                     33.95313
## 21           Madera                     36.37190
## 22            Marin                     37.83550
## 23         Mariposa                     37.32609
## 24       Mendocino                     39.23421
## 25          Merced                     37.22837
## 26           Modoc                     41.33246
## 27            Mono                     37.89973
## 28         Monterey                     35.81198
## 29            Napa                     38.08476
## 30          Nevada                     39.20459
## 31          Orange                     33.58068
## 32          Placer                     38.87233
## 33          Plumas                     39.69230
## 34       Riverside                     33.62548
## 35      Sacramento                     38.37442
## 36      San Benito                     36.13041
## 37  San Bernardino                     34.27944
## 38       San Diego                     32.62458
## 39   San Francisco                     37.76651
## 40     San Joaquin                     37.94056
## 41 San Luis Obispo                     35.31239
## 42       San Mateo                     37.44854
## 43   Santa Barbara                     32.02344
```

```
## 44     Santa Clara                    37.29720
## 45      Santa Cruz                    35.56790
## 46          Shasta                    40.62516
## 47          Sierra                    39.52475
## 48        Siskiyou                    41.42510
## 49          Solano                    38.12828
## 50          Sonoma                    37.57025
## 51      Stanislaus                    37.59442
## 52          Sutter                    38.07167
## 53          Tehama                    40.01780
## 54         Trinity                    40.58469
## 55          Tulare                    36.04082
## 56        Tuolumne                    37.79588
## 57         Ventura                    32.99511
## 58            Yolo                    38.65037
## 59            Yuba                    39.16233
##    Approximate Center Longitude
## 1                     -121.5334
## 2                     -122.1108
## 3                     -119.4155
## 4                     -120.0896
## 5                     -120.1489
## 6                     -120.0978
## 7                     -121.8448
## 8                     -122.0018
## 9                     -123.2159
## 10                    -119.0416
## 11                    -118.8700
## 12                    -122.2146
## 13                    -123.3152
## 14                    -115.3010
## 15                    -117.5558
## 16                    -118.7343
## 17                    -119.7920
## 18                    -119.5217
## 19                    -120.4653
## 20                    -117.7593
## 21                    -116.7716
## 22                    -122.1396
## 23                    -118.9436
## 24                    -122.9063
## 25                    -120.6950
## 26                    -120.0731
## 27                    -118.8492
## 28                    -119.6145
## 29                    -121.1485
## 30                    -120.4911
## 31                    -117.3310
## 32                    -120.1922
## 33                    -120.0633
```

```
## 34                    -116.2180
## 35                    -120.8456
## 36                    -119.5199
## 37                    -116.4048
## 38                    -115.9804
## 39                    -122.4404
## 40                    -121.2930
## 41                    -120.4028
## 42                    -122.1413
## 43                    -111.0647
## 44                    -121.8908
## 45                    -117.2150
## 46                    -121.7631
## 47                    -120.5424
## 48                    -122.1756
## 49                    -121.7599
## 50                    -119.9412
## 51                    -120.9692
## 52                    -118.5862
## 53                    -121.9367
## 54                    -122.8708
## 55                    -118.1389
## 56                    -119.3361
## 57                    -114.3863
## 58                    -121.8380
## 59                    -120.9307
```

Part 2-4: Display all Features if they are ManMande or Natural

```
# Import feature class from previous exercises and join it to our db and store the va
lues of Man/Nature in variable feature
feature<-read.table("feature.txt", header = TRUE)
dbWriteTable(db, "feature", feature)
features<-dbGetQuery(db, "SELECT County, data.Class,
                     feature.Feature FROM data LEFT JOIN feature ON data.Class = feat
ure.Class")[,3]

# Display count of natural and man made features respectively
table(unlist(features))
```

```
##
##    Man Nature
##  82371  39790
```

```
# Display total number of natural and man made features
length(features)
```

```
## [1] 122161
```