Assignment 6

Sophia Telles & Samantha Driscoll

October 23, 2017

```
library(gudatavizfa17)
library(dplyr)
```

```
##
## Attaching package: 'dplyr'
```

```
## The following objects are masked from 'package:stats':
##
##     filter, lag
```

```
## The following objects are masked from 'package:base':
##
##     intersect, setdiff, setequal, union
```

```
library(tidyr)
library(googleVis)
```

```
## Warning: package 'googleVis' was built under R version 3.4.2
```

```
## Creating a generic function for 'toJSON' from package 'jsonlite' in package 'googleVis'
```

```
##
## Welcome to googleVis version 0.6.2
##
## Please read Google's Terms of Use
## before you start using the package:
## https://developers.google.com/terms/
##
## Note, the plot method of googleVis will by default use
## the standard browser to display its output.
##
## See the googleVis package vignettes for more details,
## or visit http://github.com/mages/googleVis.
##
## To suppress this message use:
## suppressPackageStartupMessages(library(googleVis))
```

1. Please look at relationships between different variables and identify variables that appear to be strongly related to having a higher rank in the survey. (10 points)

```
cityrank <- bcities[c("City","Rank","Population")]
cityrank[c("City")][is.na(cityrank[c("City")])]<- "Washington"
cityrank
```

```
           City Rank Population
```

1 Los Angeles 50 3811518 2 Anchorage 49 297018 3 Omaha 48 416855 4 St. Louis 47 304219 5 Cleveland 46 396441 6 Chesapeake 45 223454 7 Phoenix 44 1418687 8 Scottsdale 43 214770 9 Reno 42 225561 10 Dallas 41 1199739 11 Virginia Beach 40 438243 12 Charlotte 39 745596 13 Tulsa 38 395176 14 Indianapolis 37 831943 15 Colorado Springs 36 421350 16 Tampa 35 340509 17 Lincoln 35 259068 18 San Jose 33 956368 19 Rochester 32 213178 20 Oakland 31 401036 21 San Antonio 30 1365256 22 Baltimore 29 612701 23 Chicago 28 2679998 24 Arlington 27 208143 25 Milwaukee 26 593545 26 Lexington 25 299520 27 Philadelphia 24 1522648

28 Oklahoma City 23 588053 29 Houston 22 2131940 30 Cincinatti 21 292050 31 Columbus 20 796520 32 Honolulu 19 399124 33 Raleigh 18 405462 34 Madison 17 231999 35 Atlanta 16 410606 36 Kansas City 15 458064 37 New Orleans 14 349773 38 Nashvile 13 603394 39 Minneapolis 12 388229 40 Pittsburgh 11 308090 41 St. Paul 10 288263 42 San Diego 9 1319558 43 Austin 8 797215 44 New York 7 8110206 45 Denver 6 597466 46 Portland 5 598205 47 Boston 4 615462 48 Washington 3 607731 49 Seattle 2 624070 50 San Francisco 1 808854

```
Gauge<- gvisGauge(cityrank,options= list(min=50,max=1,greenFrom=16,greenTo=1,yellowFrom=33,yellowTo=15,
plot(Gauge)
```

```
## starting httpd help server ... done
```

```
Bubble <- gvisBubbleChart(cityrank, idvar="City", xvar="Population",yvar="Rank",sizevar = "Population",
plot(Bubble)
```

```
Table <- gvisTable(bcities,formats=list(Population="#,###"),options=list(page='enable',width=600,height=
plot(Table)
```

bars

2. Plot the cities on a map of the US. (5 points)

```
GeoCities <- gvisGeoChart(bcities,locationvar = "City", sizevar = "Population",options=list(region="US"
plot(GeoCities)
```

3. Identify the States in which these cities lie and then develop a choropleth that colors different states based on the number of cities that are present in each State. (5 points)

```
Ccount=bcities%>%group_by(City)%>%summarize(number=length(City))
Ccount
```

A tibble: 50 x 2

```
     City number
   <fctr>  <int>
```

1 Anchorage 1 2 Arlington 1 3 Atlanta 1 4 Austin 1 5 Baltimore 1 6 Boston 1 7 Charlotte 1 8 Chesapeake 1 9 Chicago 1 10 Cincinatti 1 # ... with 40 more rows

```
Ccount$States <-1:50
Ccount
```

A tibble: 50 x 3

```
     City number States
   <fctr>  <int>  <int>
```

1 Anchorage 1 1 2 Arlington 1 2 3 Atlanta 1 3 4 Austin 1 4 5 Baltimore 1 5 6 Boston 1 6 7 Charlotte 1 7 8 Chesapeake 1 8 9 Chicago 1 9 10 Cincinatti 1 10 # ... with 40 more rows

```
Ccount$States <- c(1:50)
Ccount$States=recode(1:50, "Alaska","Washington DC", "Georgia", "Texas","Maryland", "Massachusets", "No
Ccount
```

A tibble: 50 x 3

```
     City number         States
   <fctr>  <int>          <chr>
```

1 Anchorage 1 Alaska 2 Arlington 1 Washington DC 3 Atlanta 1 Georgia 4 Austin 1 Texas 5 Baltimore 1 Maryland 6 Boston 1 Massachusets 7 Charlotte 1 North Carolina 8 Chesapeake 1 Maryland 9 Chicago 1 Illinois 10 Cincinatti 1 Ohio # ... with 40 more rows

```
Citycount=Ccount%>%group_by(States)%>%summarize(number=length(City))
Citycount
```

A tibble: 29 x 2

```
   States number
    <chr>  <int>
```

1 Alaska 1 2 Arizona 2 3 California 5 4 Colorado 2 5 Florida 1 6 Georgia 1 7 Hawaii 1 8 Illinois 1 9 Indiana 1 10 Kentucky 1 # ... with 19 more rows

```
Numcities <- gvisGeoChart(Citycount, "States", "number",
                          options=list(region="US", displayMode="regions", resolution="provinces",
                                       width=600, height=400))
plot(Numcities)
```