InClassWork.rmd

Peyton Hall

02/08/2024

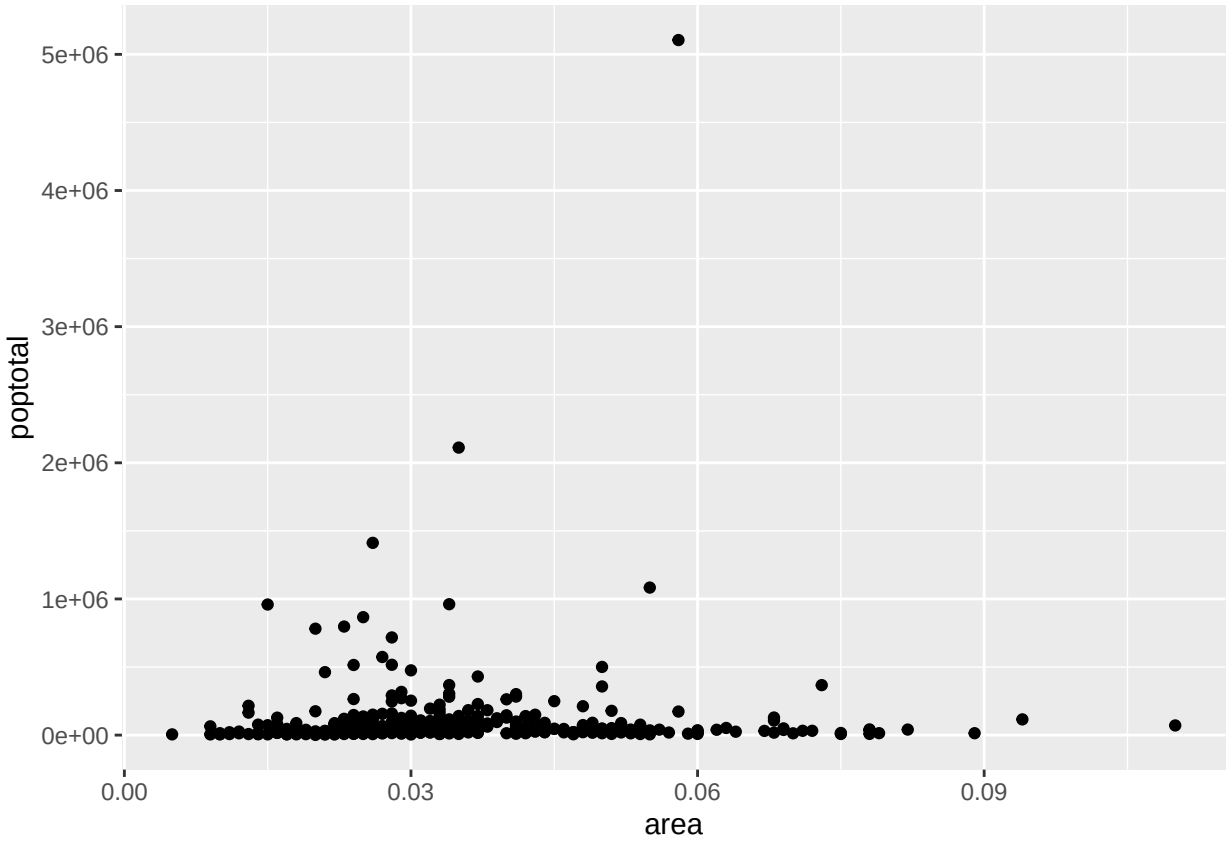
```
data <- USArrests
```

```
# install.packages("ggplot2")
library(ggplot2)
# View(midwest)
str(midwest) # check the structure
```

```
## tibble [437 x 28] (S3: tbl_df/tbl/data.frame)
##  $ PID                 : int [1:437] 561 562 563 564 565 566 567 568 569 570 ...
##  $ county              : chr [1:437] "ADAMS" "ALEXANDER" "BOND" "BOONE" ...
##  $ state               : chr [1:437] "IL" "IL" "IL" "IL" ...
##  $ area                : num [1:437] 0.052 0.014 0.022 0.017 0.018 0.05 0.017 0.027 0.024 0.058 ...
##  $ poptotal            : int [1:437] 66090 10626 14991 30806 5836 35688 5322 16805 13437 173025 ...
##  $ popdensity          : num [1:437] 1271 759 681 1812 324 ...
##  $ popwhite            : int [1:437] 63917 7054 14477 29344 5264 35157 5298 16519 13384 146506 ...
##  $ popblack            : int [1:437] 1702 3496 429 127 547 50 1 111 16 16559 ...
##  $ popamerindian       : int [1:437] 98 19 35 46 14 65 8 30 8 331 ...
##  $ popasian            : int [1:437] 249 48 16 150 5 195 15 61 23 8033 ...
##  $ popother            : int [1:437] 124 9 34 1139 6 221 0 84 6 1596 ...
##  $ percwhite           : num [1:437] 96.7 66.4 96.6 95.3 90.2 ...
##  $ percblack           : num [1:437] 2.575 32.9 2.862 0.412 9.373 ...
##  $ percamerindan       : num [1:437] 0.148 0.179 0.233 0.149 0.24 ...
##  $ percasian           : num [1:437] 0.3768 0.4517 0.1067 0.4869 0.0857 ...
##  $ percother           : num [1:437] 0.1876 0.0847 0.2268 3.6973 0.1028 ...
##  $ popadults           : int [1:437] 43298 6724 9669 19272 3979 23444 3583 11323 8825 95971 ...
##  $ perchsd             : num [1:437] 75.1 59.7 69.3 75.5 68.9 ...
##  $ percollege          : num [1:437] 19.6 11.2 17 17.3 14.5 ...
##  $ percprof            : num [1:437] 4.36 2.87 4.49 4.2 3.37 ...
##  $ poppovertyknown     : int [1:437] 63628 10529 14235 30337 4815 35107 5241 16455 13081 154934 ...
##  $ percpovertyknown    : num [1:437] 96.3 99.1 95 98.5 82.5 ...
##  $ percbelowpoverty    : num [1:437] 13.15 32.24 12.07 7.21 13.52 ...
##  $ percchildbelowpovert: num [1:437] 18 45.8 14 11.2 13 ...
##  $ percadultpoverty    : num [1:437] 11.01 27.39 10.85 5.54 11.14 ...
##  $ percelderlypoverty  : num [1:437] 12.44 25.23 12.7 6.22 19.2 ...
##  $ inmetro             : int [1:437] 0 0 0 1 0 0 0 0 0 1 ...
##  $ category            : chr [1:437] "AAR" "LHR" "AAR" "ALU" ...
```

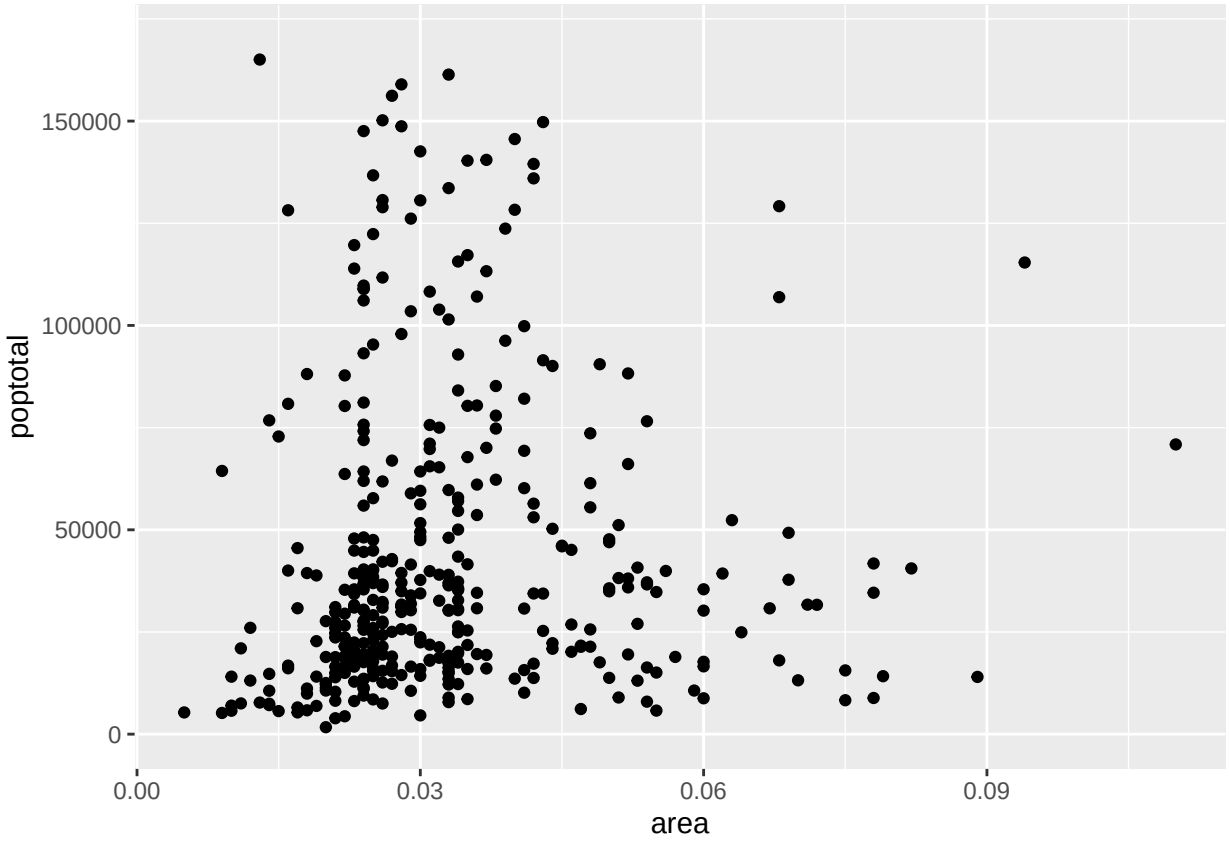
Scatter Plot

```
ggplot(data = midwest, aes(x=area, y=poptotal))+geom_point()
```



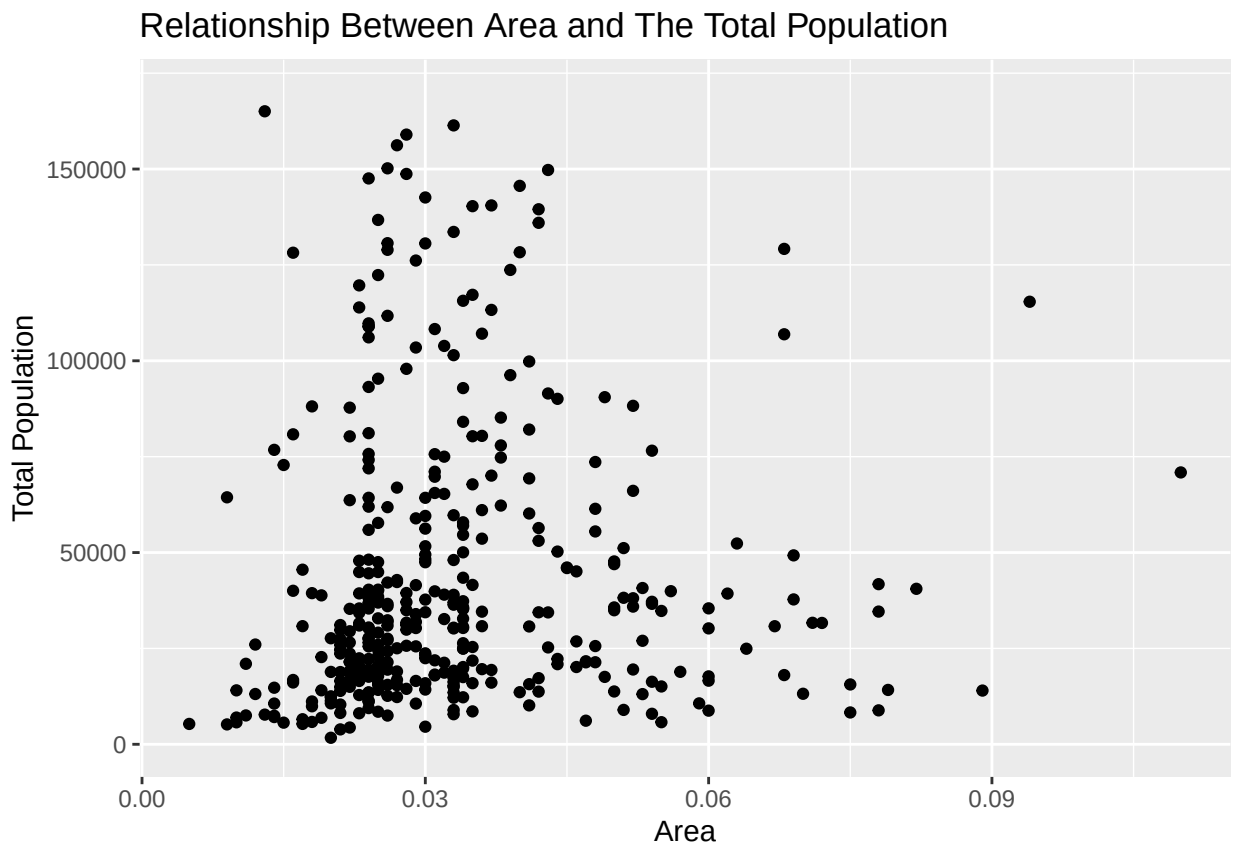

```
ggplot(data = midwest, aes(x=area, y=poptotal))+geom_point()+ylim(c(0,170000))
```

```
## Warning: Removed 43 rows containing missing values (`geom_point()`).
```



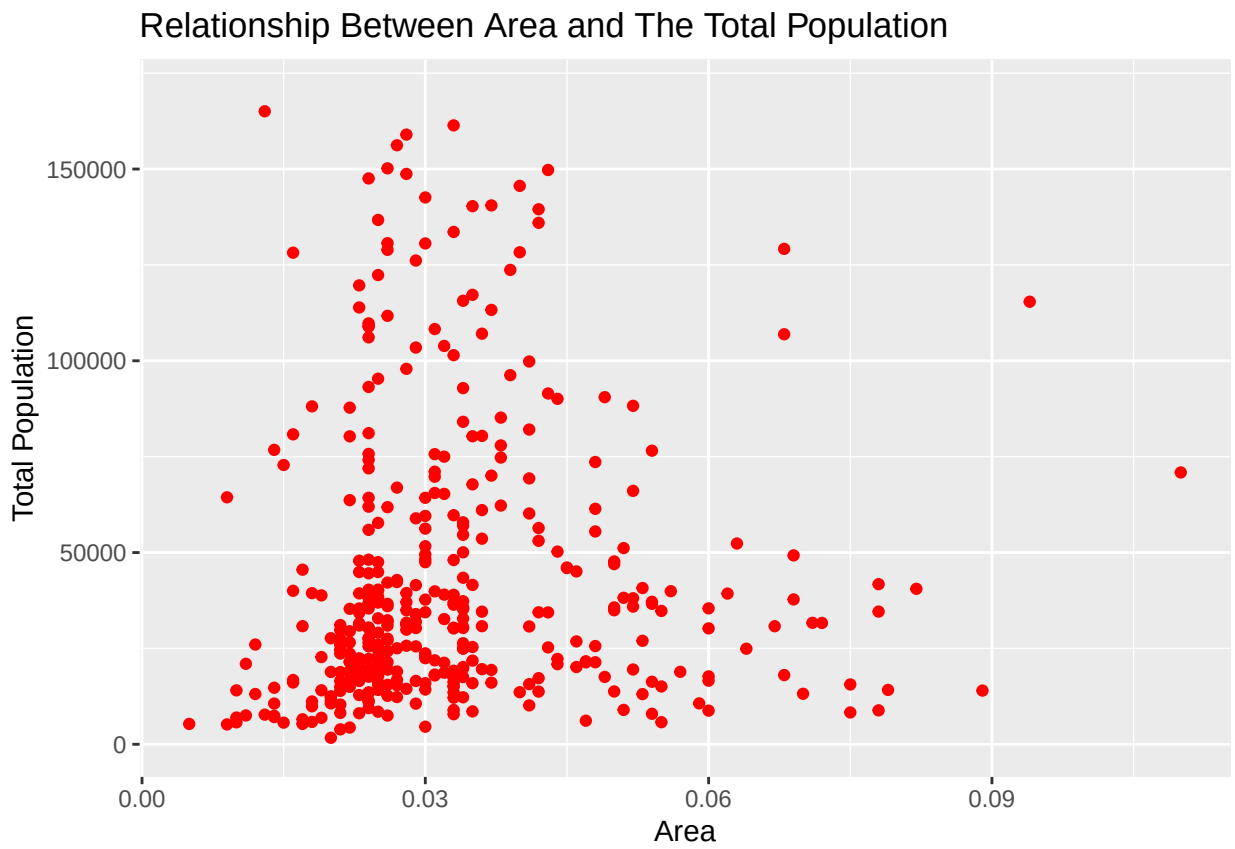

```
ggplot(data=midwest,aes(x=area,y=poptotal))+geom_point()+ylim(c(0,170000))+labs(x="Area", y="Total Popul
```

```
## Warning: Removed 43 rows containing missing values (‘geom_point()‘).
```



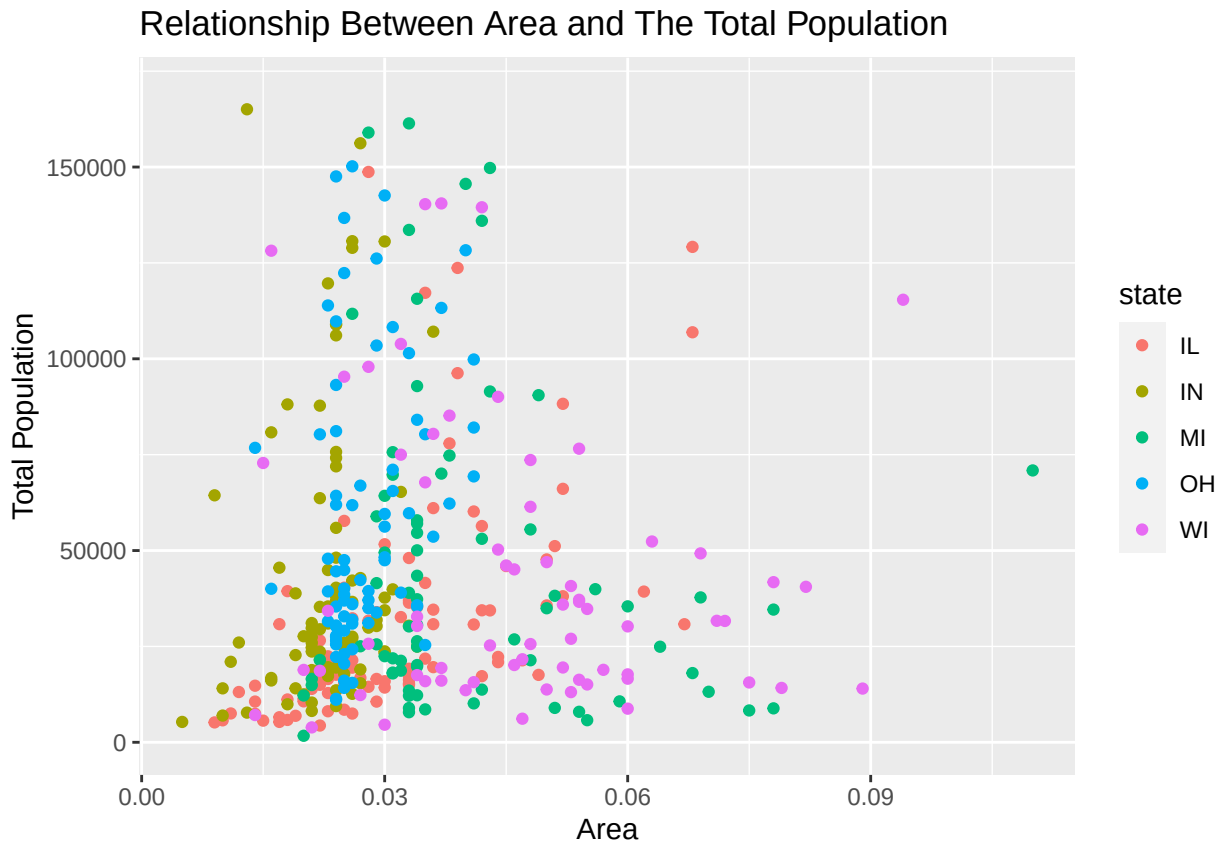

```
ggplot(data=midwest,aes(x=area,y=poptotal))+geom_point(color="red")+ylim(c(0,170000))+labs(x="Area", y=
```

```
## Warning: Removed 43 rows containing missing values (‘geom_point()‘).
```



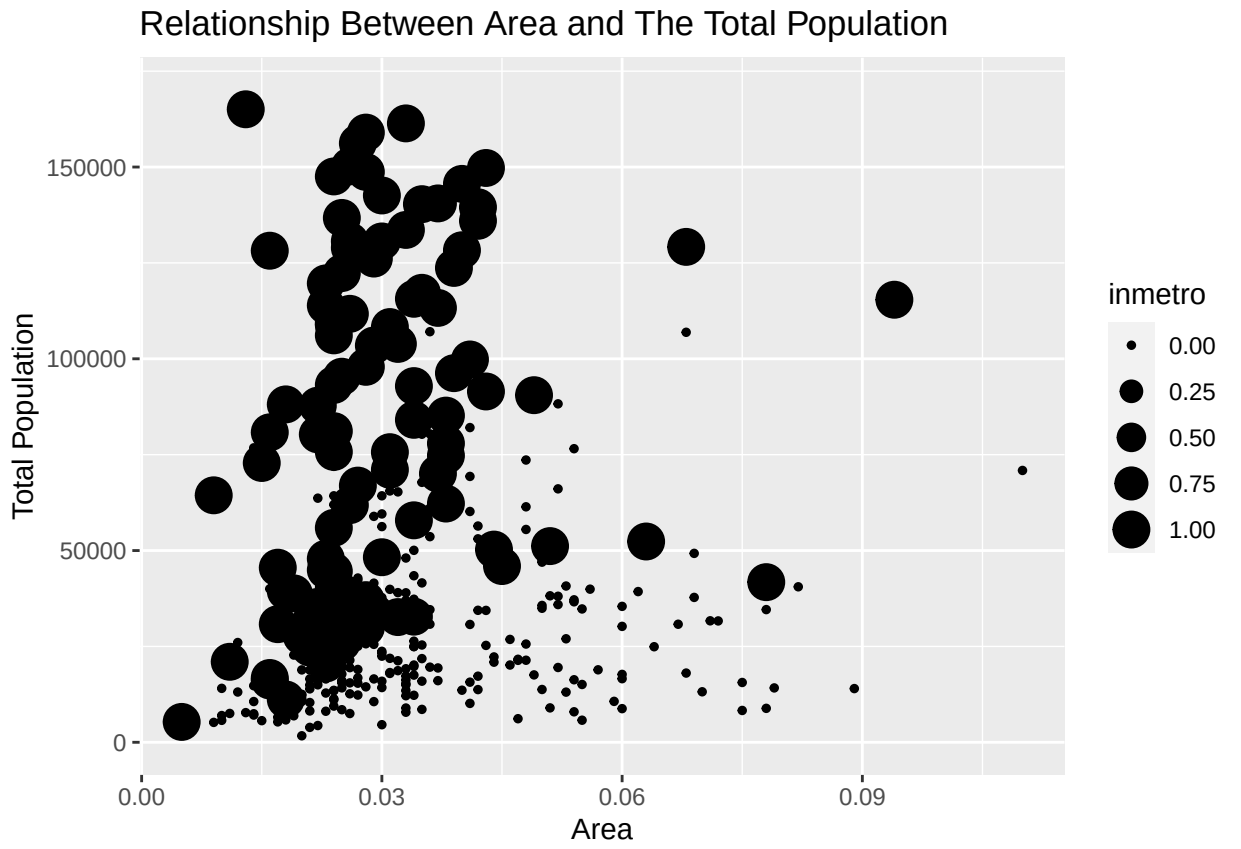

```
ggplot(data=midwest,aes(x=area,y=poptotal))+geom_point(aes(color=state))+ylim(c(0,170000))+labs(x="Area
```

```
## Warning: Removed 43 rows containing missing values (`geom_point()`).
```




```
ggplot(data=midwest,aes(x=area,y=poptotal))+geom_point(aes(size=inmetro))+ylim(c(0,170000))+labs(x="Area
```

```
## Warning: Removed 43 rows containing missing values (`geom_point()`).
```




```
midwest2<-midwest
midwest2$metro<-as.factor(midwest2$inmetro)
str(midwest2)
```

```
## tibble [437 x 29] (S3: tbl_df/tbl/data.frame)
##  $ PID                : int [1:437] 561 562 563 564 565 566 567 568 569 570 ...
##  $ county             : chr [1:437] "ADAMS" "ALEXANDER" "BOND" "BOONE" ...
##  $ state              : chr [1:437] "IL" "IL" "IL" "IL" ...
##  $ area               : num [1:437] 0.052 0.014 0.022 0.017 0.018 0.05 0.017 0.027 0.024 0.058 ...
##  $ poptotal           : int [1:437] 66090 10626 14991 30806 5836 35688 5322 16805 13437 173025 ...
##  $ popdensity         : num [1:437] 1271 759 681 1812 324 ...
##  $ popwhite           : int [1:437] 63917 7054 14477 29344 5264 35157 5298 16519 13384 146506 ...
##  $ popblack           : int [1:437] 1702 3496 429 127 547 50 1 111 16 16559 ...
##  $ popamerindian      : int [1:437] 98 19 35 46 14 65 8 30 8 331 ...
##  $ popasian           : int [1:437] 249 48 16 150 5 195 15 61 23 8033 ...
##  $ popother           : int [1:437] 124 9 34 1139 6 221 0 84 6 1596 ...
##  $ percwhite          : num [1:437] 96.7 66.4 96.6 95.3 90.2 ...
##  $ percblack          : num [1:437] 2.575 32.9 2.862 0.412 9.373 ...
##  $ percamerindan      : num [1:437] 0.148 0.179 0.233 0.149 0.24 ...
##  $ percasian          : num [1:437] 0.3768 0.4517 0.1067 0.4869 0.0857 ...
##  $ percother          : num [1:437] 0.1876 0.0847 0.2268 3.6973 0.1028 ...
##  $ popadults          : int [1:437] 43298 6724 9669 19272 3979 23444 3583 11323 8825 95971 ...
##  $ perchsd            : num [1:437] 75.1 59.7 69.3 75.5 68.9 ...
##  $ percollege         : num [1:437] 19.6 11.2 17 17.3 14.5 ...
##  $ percprof           : num [1:437] 4.36 2.87 4.49 4.2 3.37 ...
##  $ poppovertyknown    : int [1:437] 63628 10529 14235 30337 4815 35107 5241 16455 13081 154934 ...
```

```
##  $ percpovertyknown    : num [1:437] 96.3 99.1 95 98.5 82.5 ...
##  $ percbelowpoverty    : num [1:437] 13.15 32.24 12.07 7.21 13.52 ...
##  $ percchildbelowpovert: num [1:437] 18 45.8 14 11.2 13 ...
##  $ percadultpoverty    : num [1:437] 11.01 27.39 10.85 5.54 11.14 ...
##  $ percelderlypoverty  : num [1:437] 12.44 25.23 12.7 6.22 19.2 ...
##  $ inmetro             : int [1:437] 0 0 0 1 0 0 0 0 0 1 ...
##  $ category            : chr [1:437] "AAR" "LHR" "AAR" "ALU" ...
##  $ metro               : Factor w/ 2 levels "0","1": 1 1 1 2 1 1 1 1 1 2 ...
```

```
ggplot(data=midwest2,aes(x=area,y=poptotal))+geom_point(aes(size=metro))+ylim(c(0,170000))+labs(x="Area
```

```
## Warning: Using size for a discrete variable is not advised.
```

```
## Warning: Removed 43 rows containing missing values (`geom_point()`).
```