19MID0017

Mothishwaran C.

ELA DA02

```
library(readxl)
library(ggplot2)
```

1. Create a histogram of percadultpoverty - Use a few different bin widths and show the results

```
df_west<-read_excel("west.xlsx")
```

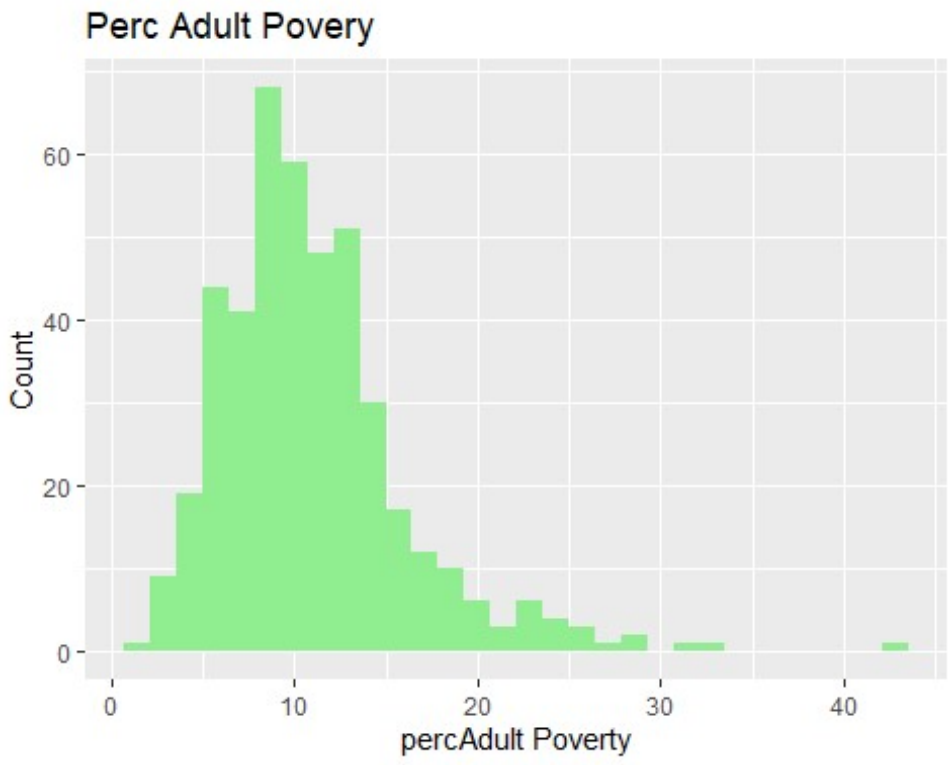
```
## New names:
## * `` -> ...1
```

```
head(df_west)
```

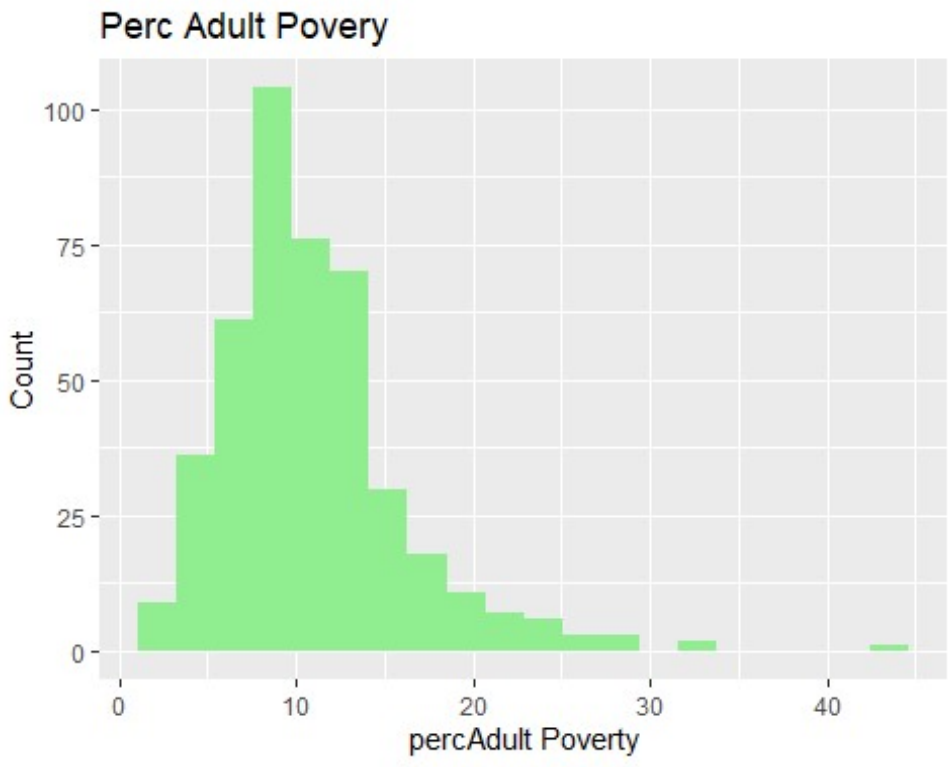
```
## # A tibble: 6 x 29
##    ...1   PID county    state  area poptotal popdensity popwhite popblack
##   <dbl> <dbl> <chr>     <chr> <dbl>    <dbl>      <dbl>    <dbl>    <dbl>
## 1     1   561 ADAMS     IL    0.052    66090      1271.    63917     1702
## 2     2   562 ALEXANDER IL    0.014    10626       759      7054     3496
## 3     3   563 BOND      IL    0.022    14991       681.    14477      429
## 4     4   564 BOONE     IL    0.017    30806      1812.    29344      127
## 5     5   565 BROWN     IL    0.018     5836       324.     5264      547
## 6     6   566 BUREAU    IL    0.05     35688       714.    35157       50
## # ... with 20 more variables: popamerindian <dbl>, popasian <dbl>,
## #   popother <dbl>, percwhite <dbl>, percblack <dbl>, percamerindan <dbl>,
## #   percasian <dbl>, percother <dbl>, popadults <dbl>, perchsd <dbl>,
## #   percollege <dbl>, percprof <dbl>, poppovertyknown <dbl>,
## #   percpovertyknown <dbl>, percbelowpoverty <dbl>, percchildbelowpovert <
dbl>,
## #   percadultpoverty <dbl>, percelderlypoverty <dbl>, inmetro <dbl>,
## #   category <chr>
```

```
ggplot(data=df_west,aes(percadultpoverty))+
  geom_histogram(fill="light green")+
  labs(x="percAdult Poverty", y="Count")+
  ggtitle("Perc Adult Povery ")
```

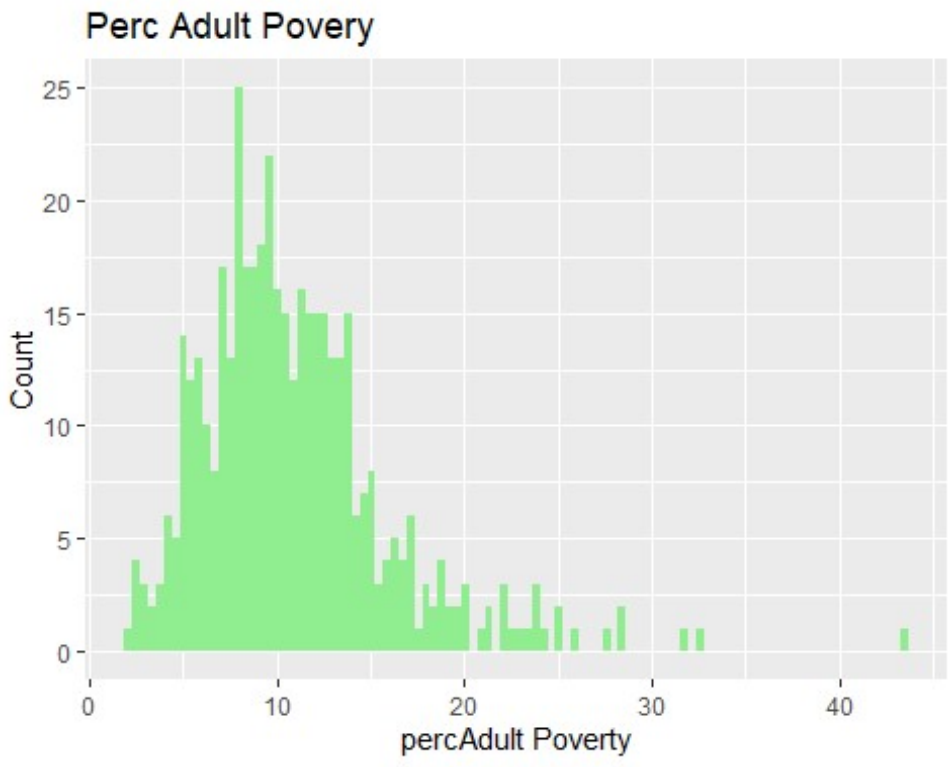
```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```




```
ggplot(data=df_west,aes(percadultpoverty))+
  geom_histogram(fill="lightgreen",bins=20)+
  labs(x="percAdult Poverty", y="Count")+
  ggtitle("Perc Adult Povery ")
```

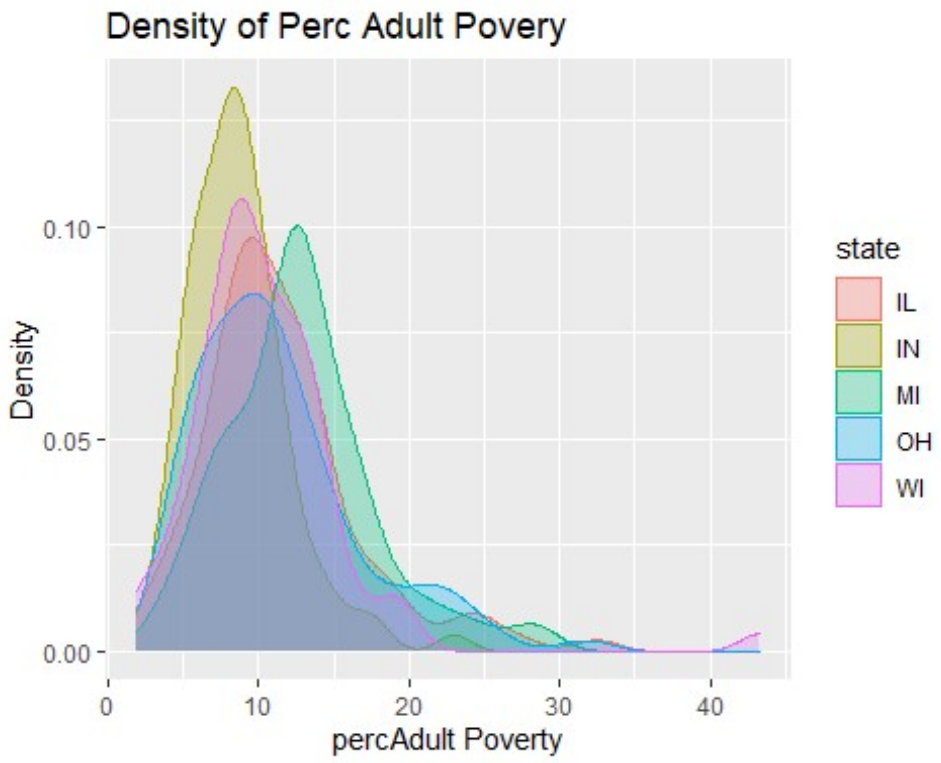



```
ggplot(data=df_west,aes(percadultpoverty))+
  geom_histogram(fill="lightgreen",bins=100)+
  labs(x="percAdult Poverty", y="Count")+
  ggtitle("Perc Adult Povery ")
```



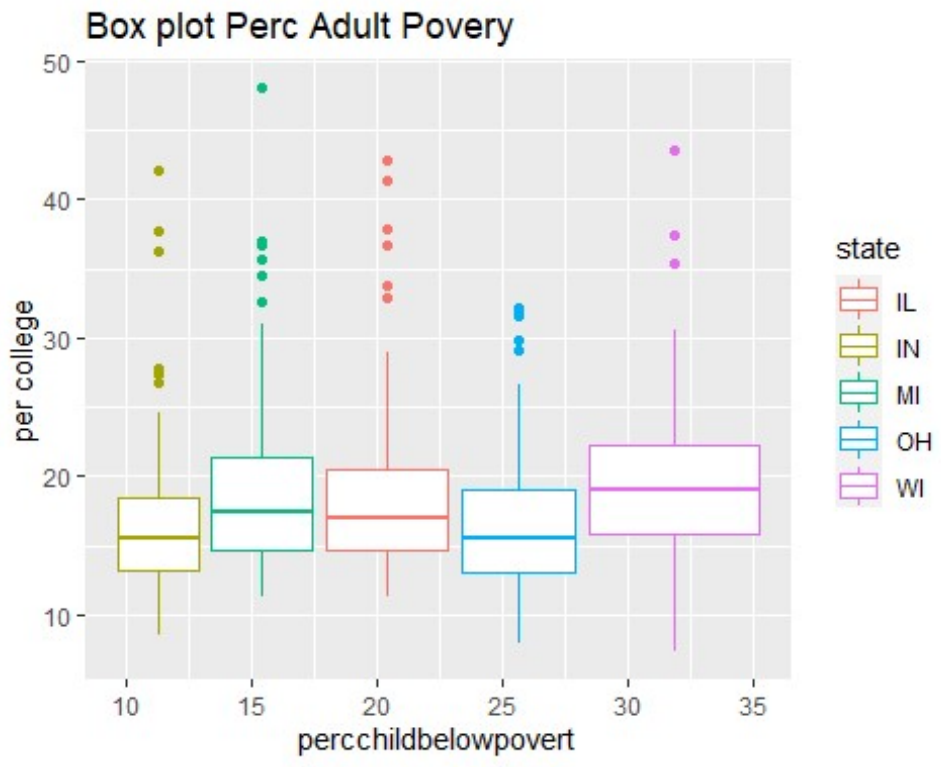
2.Show percadultpoverty using kernel density plot

```
ggplot(data=df_west,aes(percadultpoverty,colour = state,fill=state))+geom_den
sity(alpha = 0.3)+
  labs(x="percAdult Poverty", y="Density")+
  ggtitle("Density of Perc Adult Povery ")
```



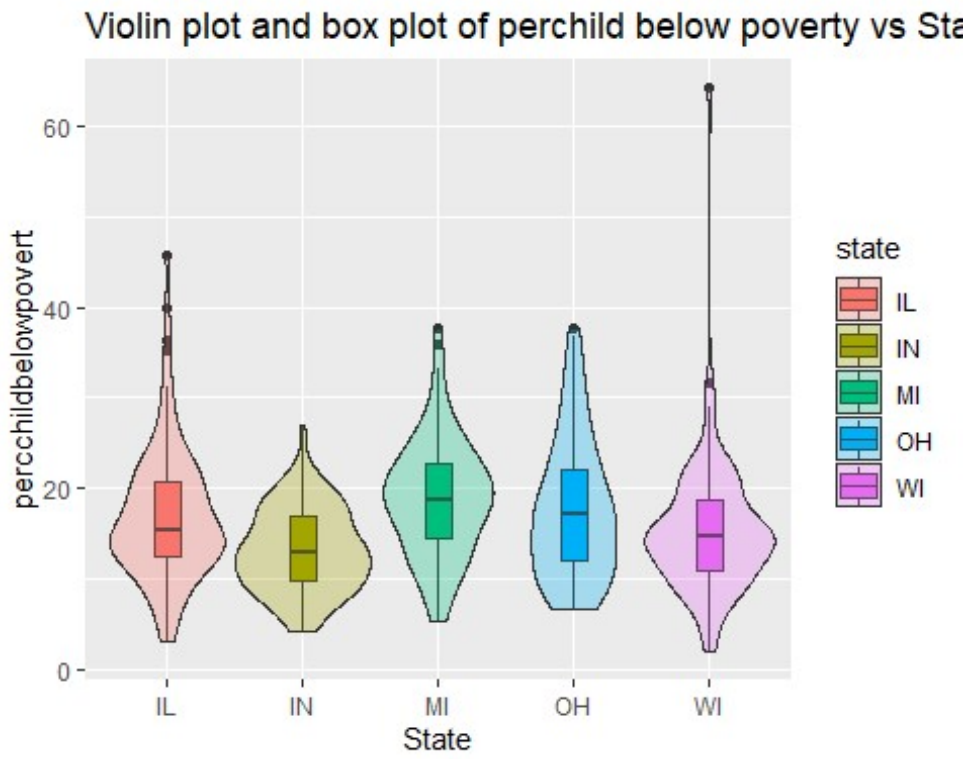
3. Create a box plot of percchildbelowpovert – percentage of population with college degree. Differentiate the states using colors

```
ggplot(df_west, aes(x=percadultpoverty, y=percollege, color=state)) +
  geom_boxplot()+
  labs(x="percchildbelowpovert", y="per college")+
  ggtitle("Box plot Perc Adult Povery ")
```



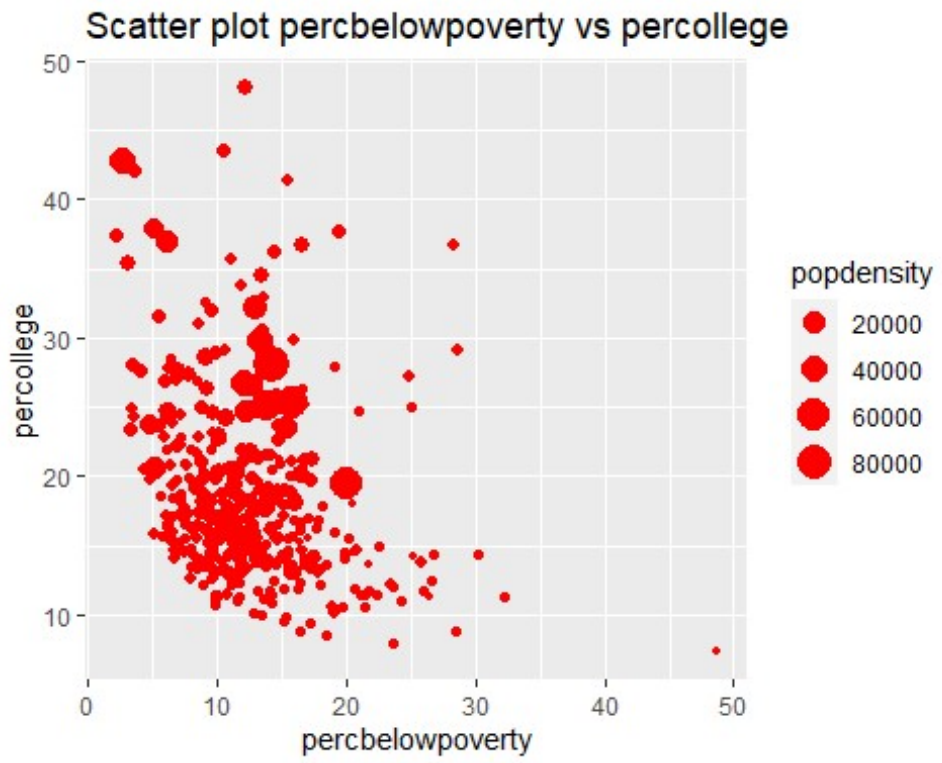
4.Create a plot that combines both violin plot and box plot for percchildbelowpovert

```
ggplot(df_west, aes(x=state,y=percchildbelowpovert,fill=state)) +
  geom_boxplot(width=0.2)+
  geom_violin(width=1,alpha=0.3)+
  labs(x="State", y="percchildbelowpovert")+
  ggtitle("Violin plot and box plot of perchild below poverty vs State ")
```



#5.Create a scatter plot between precollege and percbelowpoverty. Use size channel to show the popdensity

```
ggplot(df_west, aes(x=percbelowpoverty,y=percollege,size=popdensity))+
  geom_point(color="red")+
  ggtitle("Scatter plot percbelowpoverty vs percollege ")
```


Scatter plot percbelowpoverty vs percollege
percollege
percbelowpoverty
popdensity
20000
40000
60000
80000
50
40
30
20
10
0
10
20
30
40
50