Project

2022-11-10

## Reading in Data

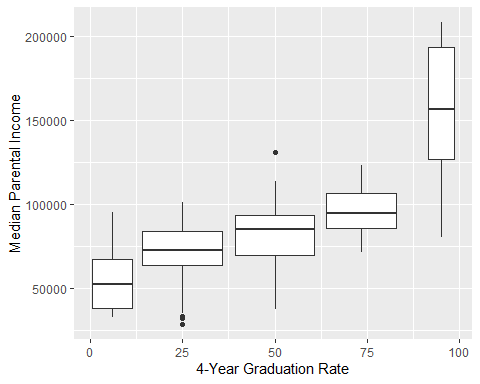
race<- read.csv("HEsegDataviz\_CollegeData\_4-year\_v5.csv")  
mrc<- read.csv("mrc\_table1.csv")  
grad<- read.csv("cc\_institution\_details.csv") %>%  
subset(select = c("chronname","grad\_100\_percentile"))  
state <- read.csv("state\_abbrev\_crosswalk.csv")%>%  
 rename(state = Code)  
grad <- grad %>%  
rename(name = chronname)  
race <- race %>%  
rename(name = inst\_name)  
race\_mrc <- mrc %>%  
 inner\_join(race, by = "name")  
college <- race\_mrc %>%  
 inner\_join(grad, by = "name")%>%  
 filter(year == 2013)%>%   
 filter(state == "AL"| state == "FL"| state == "GA"| state == "AR"| state == "KY"| state == "LA"| state == "MS"| state == "NC"| state == "SC"| state == "TN")  
college$grad\_100\_percentile <- as.numeric(college$grad\_100\_percentile)

## Warning: NAs introduced by coercion

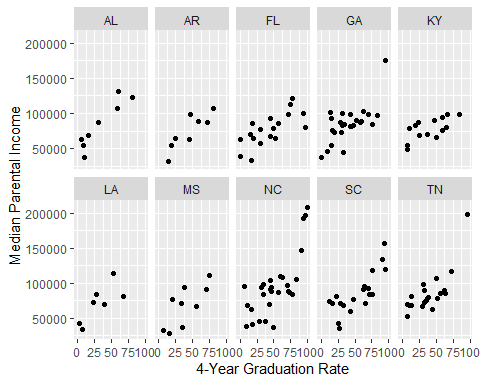
college$grad\_100\_percentile[is.na(college$grad\_100\_percentile)]<-mean(college$grad\_100\_percentile,na.rm=TRUE)  
college <- college%>%  
 left\_join(state, by = "state")  
college <- college %>%  
 rename(NAME = State)

## Initial Graphs

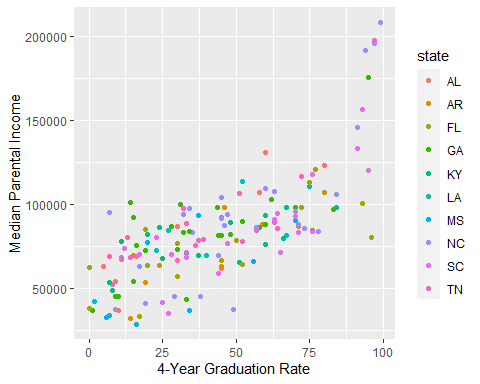
#Box Plot   
college %>%  
 ggplot(mapping = aes(x = grad\_100\_percentile, y = par\_median)) +  
geom\_boxplot(mapping = aes(group = cut\_width(grad\_100\_percentile, 25))) +  
 labs(x = "4-Year Graduation Rate", y = "Median Parental Income")



#Facet Wrap  
college %>%   
 ggplot(mapping = aes(x = grad\_100\_percentile , y = par\_median )) + geom\_point() +  
 facet\_wrap(~state, nrow = 2) +  
 labs(x = "4-Year Graduation Rate", y = "Median Parental Income")



#Scatter Plot  
college %>%   
 ggplot(mapping = aes(x = grad\_100\_percentile , y = par\_median , color = state)) + geom\_point() +  
 labs(x = "4-Year Graduation Rate", y = "Median Parental Income")

 ## Initial Maps

library(tidycensus)  
library(tmap)  
library(sf)

## Linking to GEOS 3.9.1, GDAL 3.4.3, PROJ 7.2.1; sf\_use\_s2() is TRUE

library(rgdal)

## Loading required package: sp

## Please note that rgdal will be retired by the end of 2023,  
## plan transition to sf/stars/terra functions using GDAL and PROJ  
## at your earliest convenience.  
##   
## rgdal: version: 1.5-32, (SVN revision 1176)  
## Geospatial Data Abstraction Library extensions to R successfully loaded  
## Loaded GDAL runtime: GDAL 3.4.3, released 2022/04/22  
## Path to GDAL shared files: C:/Users/momos/AppData/Local/R/win-library/4.2/rgdal/gdal  
## GDAL binary built with GEOS: TRUE   
## Loaded PROJ runtime: Rel. 7.2.1, January 1st, 2021, [PJ\_VERSION: 721]  
## Path to PROJ shared files: C:/Users/momos/AppData/Local/R/win-library/4.2/rgdal/proj  
## PROJ CDN enabled: FALSE  
## Linking to sp version:1.5-0  
## To mute warnings of possible GDAL/OSR exportToProj4() degradation,  
## use options("rgdal\_show\_exportToProj4\_warnings"="none") before loading sp or rgdal.

library(cartogram)

## Warning: package 'cartogram' was built under R version 4.2.2

library(igraph)

## Warning: package 'igraph' was built under R version 4.2.2

##   
## Attaching package: 'igraph'

## The following objects are masked from 'package:dplyr':  
##   
## as\_data\_frame, groups, union

## The following objects are masked from 'package:purrr':  
##   
## compose, simplify

## The following object is masked from 'package:tidyr':  
##   
## crossing

## The following object is masked from 'package:tibble':  
##   
## as\_data\_frame

## The following objects are masked from 'package:stats':  
##   
## decompose, spectrum

## The following object is masked from 'package:base':  
##   
## union

library(ggraph)

## Warning: package 'ggraph' was built under R version 4.2.2

##   
## Attaching package: 'ggraph'

## The following object is masked from 'package:sp':  
##   
## geometry

us <- get\_acs(geography = "state",   
 variables = "B02001\_001",  
 year=2013,   
 geometry = TRUE,  
 shift\_geo = TRUE)

## Getting data from the 2009-2013 5-year ACS

## Warning: The `shift\_geo` argument is deprecated and will be removed in a future  
## release. We recommend using `tigris::shift\_geometry()` instead.

## Using feature geometry obtained from the albersusa package

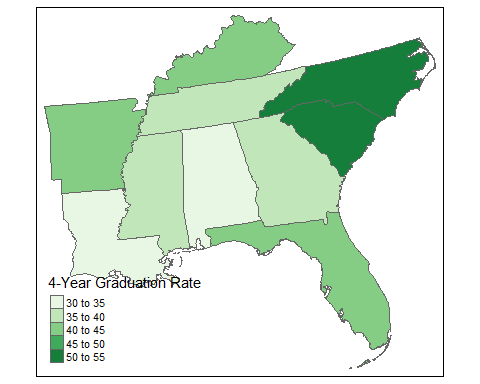
## Please note: Alaska and Hawaii are being shifted and are not to scale.

## old-style crs object detected; please recreate object with a recent sf::st\_crs()

college\_map <- college %>%  
 inner\_join(us, by = "NAME")  
college\_map\_grad <- college\_map %>%  
 group\_by(NAME)%>%  
 summarise(NAME, geometry, `4-Year Graduation Rate` = mean(grad\_100\_percentile))

## `summarise()` has grouped output by 'NAME'. You can override using the  
## `.groups` argument.

st\_sf(college\_map\_grad) %>%  
 tm\_shape() +  
 tm\_polygons("4-Year Graduation Rate", palette = "Greens")



college\_map\_income <- college\_map%>%  
 group\_by(NAME)%>%  
 summarise(NAME, geometry, `Median Parental Income` = mean(par\_median))

## `summarise()` has grouped output by 'NAME'. You can override using the  
## `.groups` argument.

st\_sf(college\_map\_income) %>%  
 tm\_shape() +  
 tm\_polygons("Median Parental Income", palette = "Greens")

