Imran Khan

STAT 515

Professor Carr

April 2nd, 2015

Redesign Project (Rising College Tuitions)

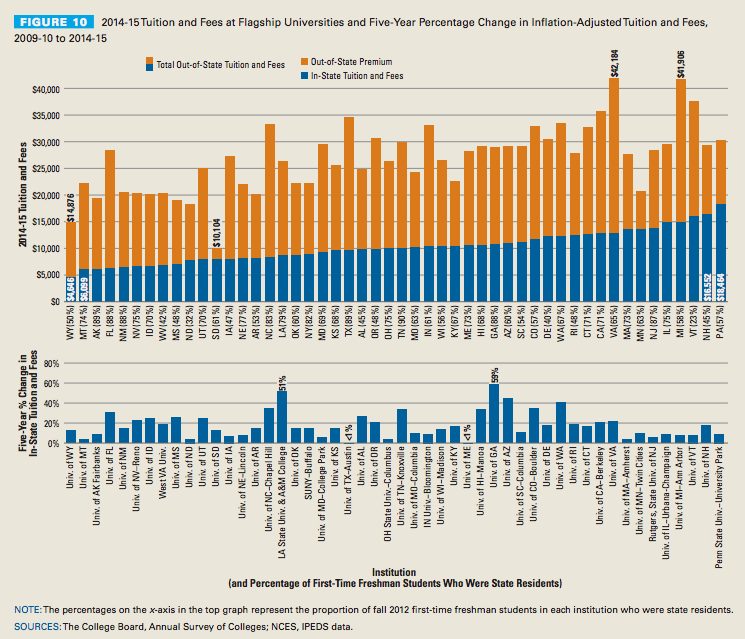
Obama’s recent proposal for free community college for two years will drastically cut the amount that college students will spend during their four years, but it will not affect the cost per year for a four year in state institution. Over the last ten years, college tuition has risen to almost one and a half times the amount from before, far surpassing the inflation rate or the change in household income. According to US News and World Report, one of the most famous sources for information on college rankings and tuition, students seeking bachelor’s degrees have seen the highest increase in fees.

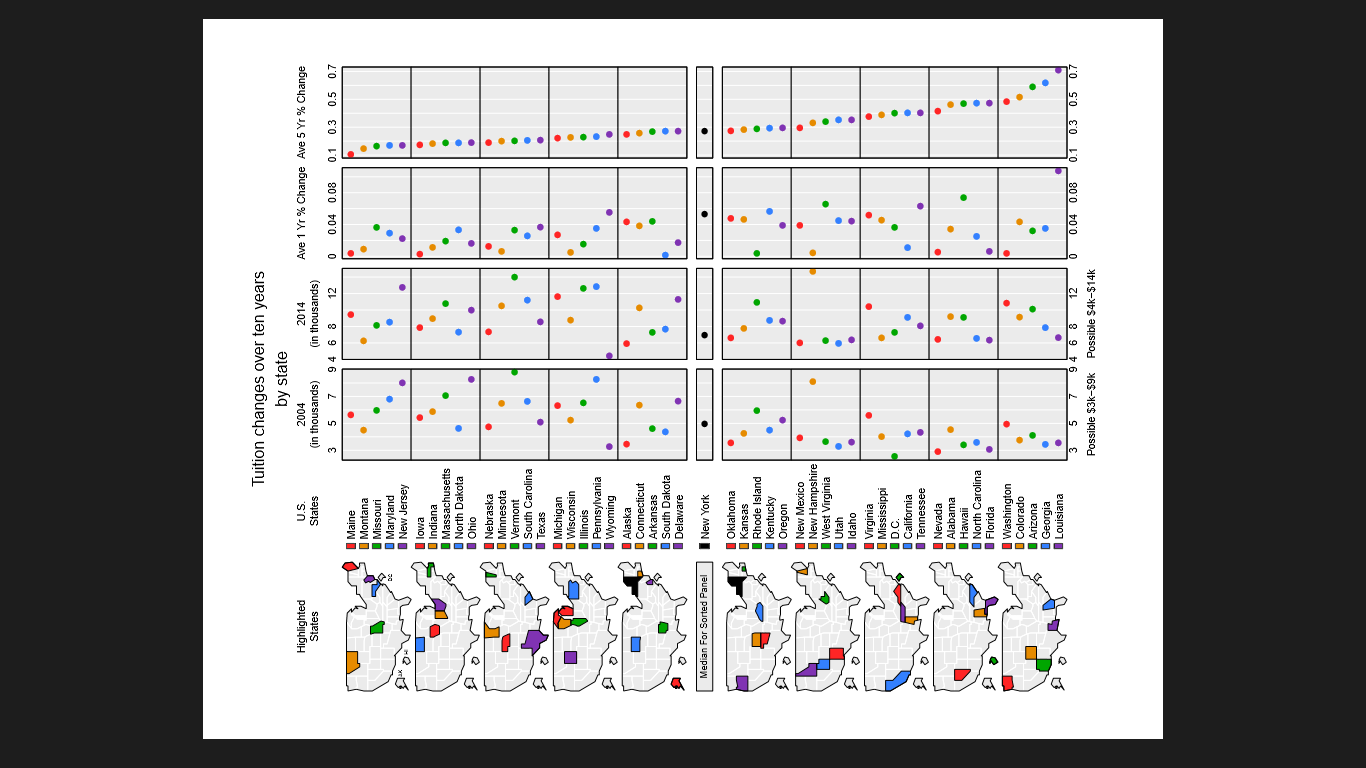
While the data shown in this bar graph is accurate, it is messy and could be reworked into a cleaner graph. One of the major problems with this visualization is that it is including way too much information into a small space. They show both the percent change in the largest school per state and the average tuition per state in each of the states for the 2014-15 school year. The data shown in the graph could also be redone to show data that is perhaps more pertinent to the reader. Finally, the bar graph’s order is a bit hard to understand at first. It is ordered by in-state tuition, but because the graph is continued to show out of state tuition on the same diagram, the order becomes more difficult to recognize.

In my redesign, I chose to focus only on public in-state tuitions by state because it is the category that the majority of current college students (including myself) fall in to. In my redesign, I also chose to omit the second part of the first graph- because it seemed like a misrepresentation to categorize an entire state by just a single school. Instead, the data I found was for the average student paying in state tuition in each state. This way, I am able to take in to account all of the public schools in a state and I can also take into account the size (population) of each school.

For the redesign, I used what we learned in week 5 of class, the micromapST package. I found the data I needed on the College Board website. It took a little while to clean up the data so that I could read it into R, and from there I cleaned it again using the micropmapSTprep function given from the week 5 assignment so that it could be used by the micromaps package.

From the redesigned project, a (slight) spatial pattern becomes visible in the data. It seems that the more south a state is, the more its tuition has seen an increase over the last ten years. You can see that Maine and Montana (both along the US’s northern coast) are the first two on the list, while Arizona, Louisiana, and Georgia, are a part of the last group. Although the trend is only slight, it would not have been visible at all in the graph given by the US News report. I also noticed a few minor discrepancies between my graph and the redesigned graph. For example, in the original graph, Georgia had the highest percent change over time and Louisiana was second, while in my graph, it was Louisiana first and Georgia second. I believe these differences may have been due to where the data was collected from and how the calculations were made.

**Original Graph** (taken from <http://www.usnews.com/cmsmedia/c1/a5/e917865840f588eff5cf7420de44/141023-eduf10-graphic.35.46%20PM.png>):

**Final Project:**

**Code:**

setwd("~/SP 15/STAT 515/Redesign")

library(micromapST)

micromapSTprep <- function(stateDF,stateId=NULL, ref=stateNamesFips){

if(is.null(stateId)) nam <- row.names(stateDF) else

nam <- stateDF[,stateId]

nam <- ifelse(nam=="District of Columbia","D.C.",nam)

check <- match(nam,row.names(ref))

bad <- is.na(check)

good <- !bad

nbad <- sum(bad)

if(nbad>0){

warning(paste(nbad,"Unmatch Names Removed",nam[bad]))

stateDF <- stateDF[!bad,]

nam <- nam[!bad]

check <- check[!bad]

good <- good[!bad]

}

ngood <- sum(good)

if(ngood < 51)warning(paste("Only",ngood,"State Ids"))

row.names(stateDF) <- ref[check,2]

return(stateDF)

}

data <- read.csv(file = 'cleaned data.csv', header = T, as.is = TRUE)

feesbystate <- micromapSTprep(data, "State")

feesbystate[3:13]<- (feesbystate[3:13])/(1000)

colNumbers <- 1:ncol(data)

names(colNumbers)=colnames(data)

colNumbers

panelDesc <- data.frame(

type=c('map','id','dot','dot','dot','dot'),

lab1=c('' ,'','2004','2014','Ave 1 Yr % Change','Ave 5 Yr % Change'),

lab2=c('' ,'','(in thousands)','(in thousands)','',''),

lab3=c('','','Possible $3k-$9k','Possible $4k-$14k',

' ',''),

col1 = c(NA,NA,3,12,"oneyear","fiveyear"))

t(panelDesc)

pdf("trial1.pdf",

width=7.5,height=10)

micromapST(feesbystate, panelDesc,

sortVar=15,ascend=TRUE,

title=c("Tuition changes over ten years",

"by state"))

dev.off()