Map

Valerie Chavez

2022-09-23

Wish to explore differences in student experiences in HS across states

We have core survey items that measure different aspects of student/school experience, rigorously designed and tests for validity that measure:

Engagement: engage_fact Academic Challenge: rigor_fact Culture: climate_fact Belonging & Peer Collaboration: peers_fact Relationships: adult_fact College & Career Readiness: prep_fact

Each survey item is compiled into a key factor or overall score.

#libraries

```
library(usmap)
library(ggplot2)
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(tigris)
```

```
## To enable caching of data, set `options(tigris_use_cache = TRUE)`
## in your R script or .Rprofile.
```

```
library(tidyr)
library(plotly)
```

```
##
## Attaching package: 'plotly'

## The following object is masked from 'package:ggplot2':
##
##     last_plot
```

```
## The following object is masked from 'package:stats':
##
##     filter

## The following object is masked from 'package:graphics':
##
##     layout
```

```
library(htmlwidgets)
library(leaflet)
library(ggthemes)
library(stringr)
```

Prepping data

```
HSsegcount<-read.csv("/Users/valerier/Dropbox (CEP)/YouthTruth/Report Production/Reports_22JU/clients/c
HSsegmean<- read.csv("/Users/valerier/Dropbox (CEP)/YouthTruth/Report Production/Reports_22JU/clients/c

#removing variables other than state, and first column
HSsegcount<-HSsegcount[71:106,-1]
HSsegmean<-HSsegmean[71:106, -1]
```

```
#renaming first column as state
names(HSsegcount)[names(HSsegcount) == 'val'] <- 'State'
names(HSsegmean)[names(HSsegmean) == 'val'] <- 'State'
```

```
#Taking only state and factor variables

HSsegcount<-subset(HSsegcount, select= c("State", "engage_fact", "rigor_fact","climate_fact","peers_fac
HSsegmean<-subset(HSsegmean, select= c("State", "engage_fact", "rigor_fact","climate_fact","peers_fact"
```

```
region<- c("arkansas", "california", "colorado", "connecticut", "district of columbia", "florida", "geo


HSsegmean<-cbind(region, HSsegmean)
HSsegcount<-cbind(region, HSsegcount)
```

```
USmap<-map_data("state")
YTmap<-left_join(USmap,HSsegmean, by="region")
#YTmapn<-left_join(HSsegcount,YTmap, by="region")
#YTmap<-merge(x=USmap, YTmap, by ="region")
```

As it is likert data the values tend toward an average of three, peer belonging is the most interesting and will be the one I present. I'd like to develop a map that could click between factors, but perhaps that would be a good shiny app.

```
#"engage_fact", range: 3.208853 4.110526, diff: 0.901673
#"rigor_fact", range: 3.454870 4.217949, diff: 0.763079
#"climate_fact", range: 2.809873 4.190789, diff: 1.380916
#"peers_fact", range: 2.909091 3.757303, diff: 0.848212
#"adult_fact", range: 3.039216 4.311966, diff: 1.27275
```

```
#"prep_fact", range: 2.70276 3.75000, diff: 1.04724

#Interested in looking at Peer belonging

YTUSA<-ggplot(data=YTmap, mapping= aes(x = long, y = lat, group = group))+
  geom_polygon(color="white", aes(fill=peers_fact)) +
  theme_tufte()+
  scale_fill_viridis_c(option = "plasma")+
  scale_colour_gradientn(colours = "plasma")+
    theme(axis.ticks.x = element_blank(),
        axis.text.x = element_blank(),
        axis.ticks.y = element_blank(),
        axis.text.y = element_blank(),
        panel.background = element_blank(),
        axis.title.x=element_blank(),
        axis.title.y=element_blank(),
        panel.grid=element_blank())+
  ggtitle("Belonging & Peer Collaboration by State")+
  labs(subtitle = str_wrap("Average key rating for High School students by state. Values ranged from 1-5
       caption = "Source:YouthTruth Student Survey, Note: Key ratings are a composite score of survey i
       fill= "Belonging Scale")


YTUSA
```