Final Project Draft Analysis
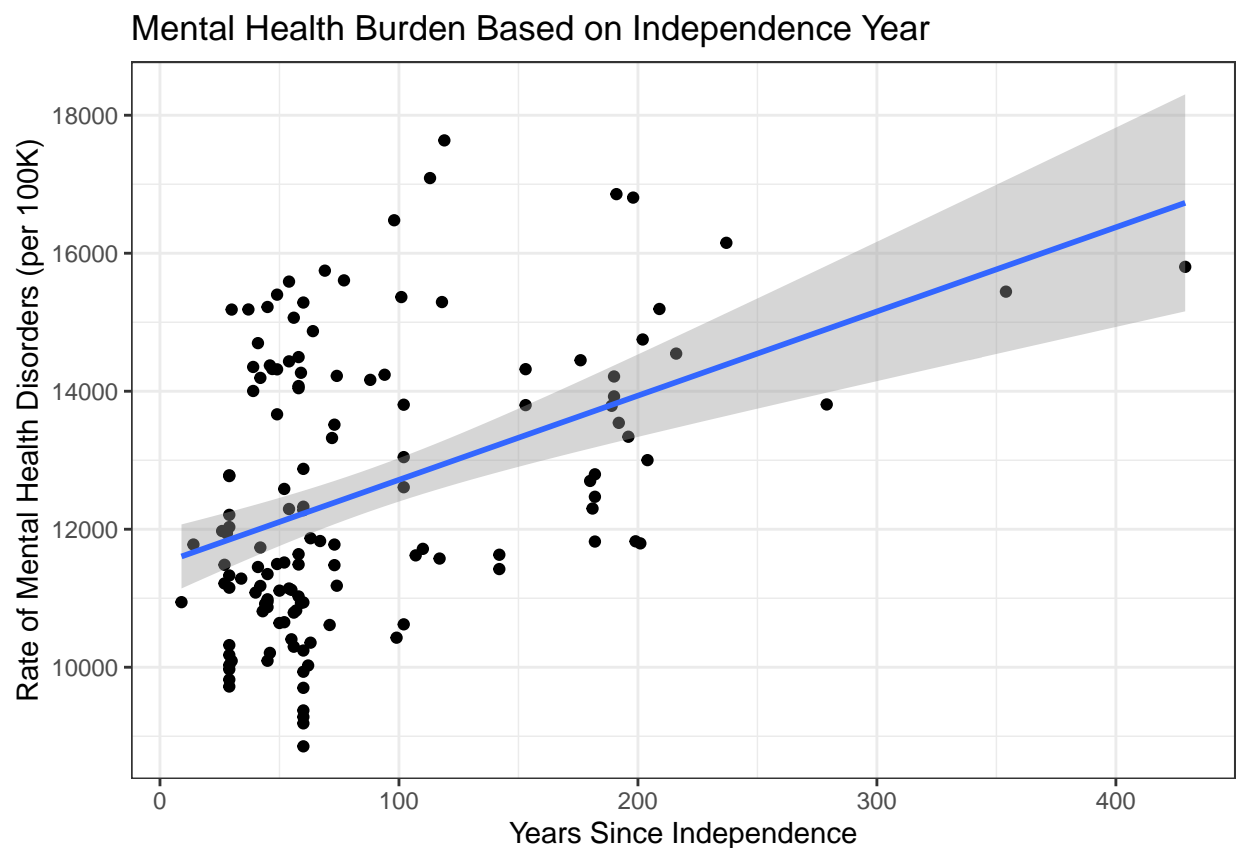
Brian

11/12/2020

Importing & Cleaning the Data

I need to merge two datasets but I don't have a common ID code - may have to go through and check names. For now, just using the names of countries and it's worked ok.

Data Analysis

There appears to be a positive relationship between the number of years since independence and the rate of mental health disorders in a country. This is surprising and contradicts my initial hypothesis, considering my theory that having more time since independence creates more room for the traumas and violence of colonization to heal, and thus results in less mental health problems over time.

Regression Table

Table 1: Mental Health Burden Based on Independence Year and Violent Independence

	Dependent variable:	
	Rate of Mental Health Disorders	
	(1)	(2)
Years Since Independence	12.201***	14.627***
	(2.274)	(2.610)
Violent Independence		−689.685*
		(374.372)
Constant	11,495.780***	11,530.210***
	(250.307)	(248.803)
Observations	136	136
R^2	0.177	0.197
Adjusted R^2	0.171	0.185
Residual Std. Error	1,817.429 (df = 134)	1,801.409 (df = 133)
F Statistic	28.801*** (df = 1; 134)	16.355*** (df = 2; 133)

Note:	*p<0.1; **p<0.05; ***p<0.01

For each year since independence, the rate of mental health disorders per 100K population goes up by about 12-15 cases. The adjusted-R squared values indicate that the linear model is able to explain about 18% of the data.

Interestingly, countries with a violent independence had a coefficient of -690, indicating that having a violent independence is correlated with a rate of 690 per 100K less mental health disorders today. This may potentially validate Frantz Fanon's theory about violence by the oppressor against the colonizer has the potential of being cathartic and restoring the humanity of the colonized, thus positively affecting mental health in the future.

Additional Ideas & Further Directions

- Analyze the role of violence and whether or not that is correlated with less mental health burdenn (maybe go back to earlier years)
- Subset the data based on different types of independence (decolonization in particular)
- Cross check mental health rates for countries that were colonizers (or not colonized) – maybe use that as a control, and then calculate the ATE.
- Use data for schizophrenia and particular mental health disorders
- Determine confidence intervals? How does that work for a line graph/ regression??
- Check interpretation of regression tables.

Code

```r
knitr::opts_chunk$set(echo = TRUE)
library(tidyverse)
library(stargazer)
# Read the data.
coldata <- read_csv("data/ICOW Colonial History 1.1/coldata110.csv") %>%
  select(Name, IndDate, ColRuler, IndViol) %>%
  rename(country = Name)

mentaldisorders <- read_csv("data/GBD Data/Mental Disorder_Prevalence_Percent_IHME-GBD_2019/Rate_Prevale
  select(location_name, val) %>%
  rename(country = location_name)


# Merge the datasets.
newdata <- inner_join(coldata, mentaldisorders, by = "country") %>%

  # Remove countries that did not have a colonial ruler.
  filter(ColRuler > 0) %>%
    # Should I take out other Western countries?
    # filter(country != "Canada",
         #   country != "Australia",
         #   country != "United States") %>%

  # Calculate the time since independence
  mutate(indep_date = as.integer(str_sub(IndDate, start = 1L, end = -3L)),
         years_indep = 2020 - indep_date)
# Plot years since independence to rate of mental health disorders
ggplot(newdata, aes(years_indep, val, label = country)) +
  geom_point() +
  geom_smooth(method = lm) +
  theme_bw() +
  labs(x = "Years Since Independence",
       y = "Rate of Mental Health Disorders (per 100K)",
       title = "Mental Health Burden Based on Independence Year")
  # geom_label(nudge_y = 50, size = 2)

# Fit linear regression models.
fit_indep_date <- lm(val ~ years_indep, newdata)
fit_indep_viol <- lm(val ~ years_indep + IndViol, newdata)

stargazer::stargazer(fit_indep_date, fit_indep_viol, title =
    "Mental Health Burden Based on Independence Year and Violent Independence",
                     covariate.labels =
                        c("Years Since Independence", "Violent Independence"),
                     dep.var.labels = "Rate of Mental Health Disorders",
                     header = FALSE)
```