The Relationship Between Education Level, Race, and Opinion on the Black Condition.

Robert James

Paper Assignment #4

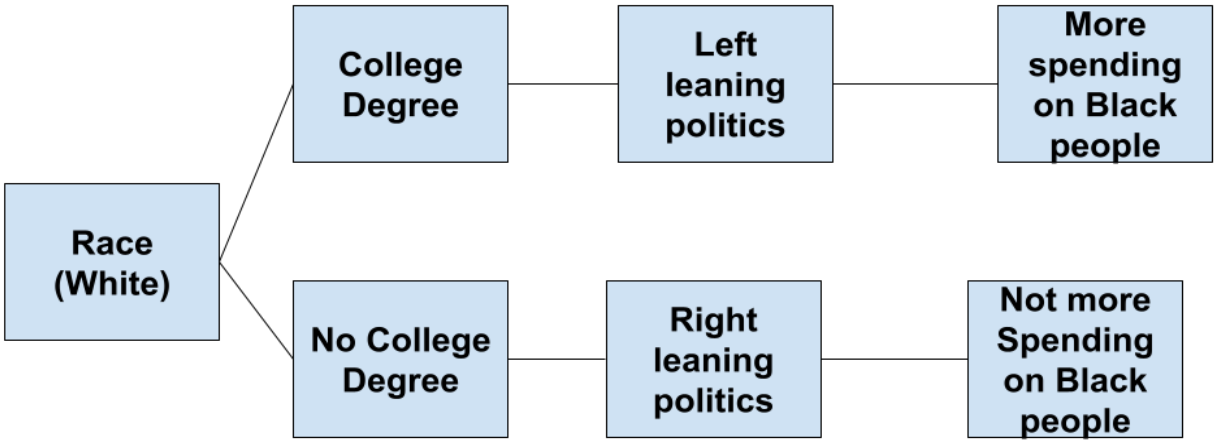
Data 202 - Fall 2023, Section 1

The aim of this paper is to explore the relationship between education level, political affiliation, and opinion on the question ". are we spending too much, too little, or about the right amount on) Improving the conditions of Blacks" of White American respondents. I chose to examine White respondents because they are the largest population of the US at 58.9% (US Census, 2020) and would be less biased than Black respondents as it is not about them. Additionally, knowing the stances on social justice issues among such a large population is important when applying ideas and sharing your message.

This exploration is based on data from the 2018 GSS survey. The variables used in the analysis were about race, education level, political affiliation, and opinion on government spending on Black people. In the beginning of the analysis the "race" variable was filtered to only select "white". Party affiliation, (partyid) was explored with the answers filtered to remove independent and other and change democrat and republican leaning answered to left and right respectively. The degree option was changed to two variables: no college degree and college degree. Lastly, the responses to the question "are we spending too much, too little, or about the right amount on improving the conditions of Blacks?" were to "More spending" and "Not more spending". I chose to simplify the responses to see the relationship between the binary options.

My hypothesis is that college educated and left leaning white Americans will be more likely to agree that we should be spending more on Black people, while non college educated and right leaning people will not agree that we should spend more on Black people. College educated people are more likely to hold liberal views (Pew Research Center, 2016), which being in favor of more spending on Black people is. Additionally, those that are college educated may have learned about social justice issues through classes or from interactions in a possibly more

diverse setting. This could influence their politics to be more left leaning which in turn would make them more in favor of more spending.



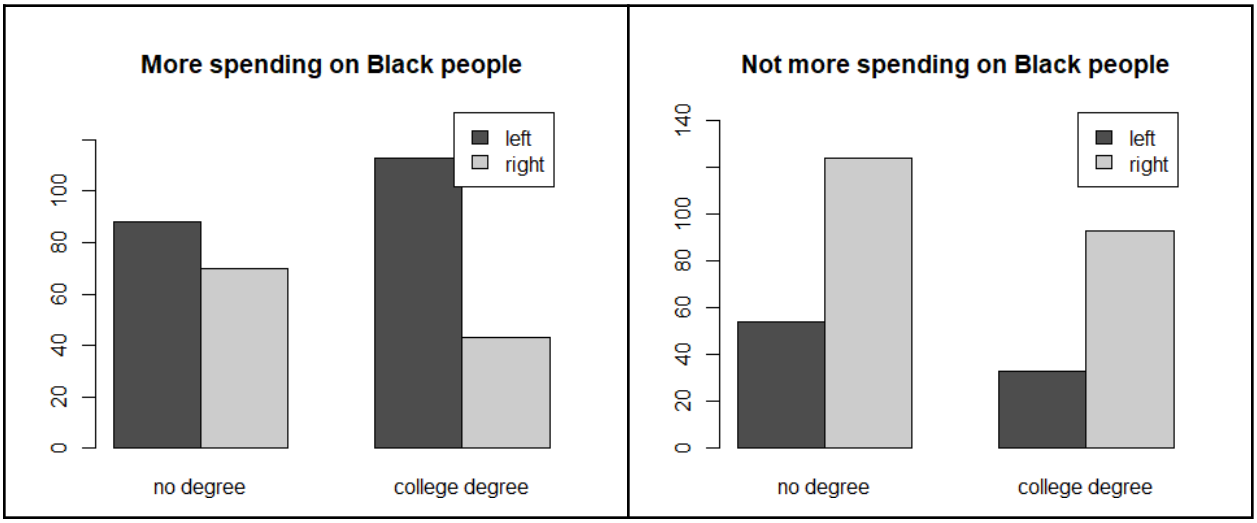
With all of the variables converted to completely binary options I created a table and charts to visualize the difference.

> table
, , = more spending

	no degree	college degree
left	88	113
right	70	43

, , = not more spending

	no degree	college degree
left	54	33
right	124	93




```
> cat("Left - More Spending: ", percentage_left_more_spending, "%\n")
Left - More Spending:  69.79167 %
> cat("Left - Not More Spending: ", percentage_left_not_more_spending, "%\n")
Left - Not More Spending:  30.20833 %
> cat("Right - More Spending: ", percentage_right_more_spending, "%\n")
Right - More Spending:  34.24242 %
> cat("Right - Not More Spending: ", percentage_right_not_more_spending, "%\n")
Right - Not More Spending:  65.75758 %

> cat("College Degree - More Spending: ", perc_col_more_spend, "%\n")
College Degree - More Spending:  55.31915 %
> cat("College Degree - Not More Spending: ", perc_col_not_more_spend, "%\n")
College Degree - Not More Spending:  44.68085 %
> cat("No Degree - More Spending: ", perc_no_deg_more_spend, "%\n")
No Degree - More Spending:  47.02381 %
> cat("No Degree - Not More Spending: ", perc_no_deg_not_more_spend, "%\n")
No Degree - Not More Spending:  52.97619 %
```

Nearly 70% of those on the left are for more spending while nearly 67% of those on the right are not for more spending. 55% of those with a degree are for more spending whereas 53% of those

without a degree are not. These results lean towards agreeing with my hypothesis though it seems that political affiliation is a stronger influence than education level. This makes sense as one's political beliefs inform their opinion on the subject.

```
> cat("Left - College Degree: ", perc_left_col_deg, "%\n")
Left - College Degree:  50.69444 %
> cat("Left - No Degree: ", perc_left_no_deg, "%\n")
Left - No Degree:  49.30556 %
> cat("Right - College Degree: ", perc_right_col_deg, "%\n")
Right - College Degree:  41.21212 %
> cat("Right - No Degree: ", perc_right_no_deg, "%\n")
Right - No Degree:  58.78788 %
```

The relationship between college degree and political affiliation shows that The left is split pretty evenly on having a degree while the right leans more at 58% not having a degree.

I chose a logistic regression model as the variables are all categorical and I wanted to see the relationship between “degree” and “partyid” on the “natrace” question.

```
> logmod <- glm(natrace ~ degree + partyid, family = binomial, data = dfc)
> summary(logmod)

Call:
glm(formula = natrace ~ degree + partyid, family = binomial,
    data = dfc)

Coefficients:
                    Estimate Std. Error z value Pr(>|z|)
(Intercept)          -0.7251     0.1535  -4.723 2.33e-06 ***
degreecollege degree -0.2266     0.1734  -1.307    0.191
partyidright          1.4729     0.1735   8.489  < 2e-16 ***
```

```
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 856.57  on 617  degrees of freedom
Residual deviance: 775.30  on 615  degrees of freedom
AIC: 781.3

Number of Fisher Scoring iterations: 4
```

The model shows that those with a college degree are more likely to be for more spending though the p-value is .20 so it is not statistically significant. The partyid variable is far more of an influence as it shows that those on the right are more likely to not be for more spending and has a very low p-value.

The null hypothesis can be rejected as there is a clear relationship. The opinion of white Americans on spending on Black people is related to education level and political leaning though politics is a far greater influence. Those on the left are more likely to be in favor of more spending compared to those on the right. Education level does not make as much of a difference but the majority of those with a degree were in favor of more spending.

References

Pew Research Center. (2016, April 26). *A Wider Ideological Gap Between More and Less Educated Adults*. Pew Research Center - U.S. Politics & Policy. https://www.pewresearch.org/politics/2016/04/26/a-wider-ideological-gap-between-more-and-less-educated-adults/

The United States Census Bureau. (n.d.). *U.S. Census Bureau QuickFacts: United States*. Census Bureau QuickFacts. https://www.census.gov/quickfacts/fact/table/US/PST045222

Code:

```
## Name: Robert James
## Assignment: Paper #4
## Date: 12/08/2023
## Purpose: Explore the Relationship Education levels, Race, age, and Opinion on spending on Black people
# install packages
install.packages("remotes", repos = "http://cran.us.r-project.org")
install.packages("tidyverse", repos = "http://cran.us.r-project.org")
install.packages("tidyr", repos = "http://cran.us.r-project.org")
install.packages("survey", repos = "http://cran.us.r-project.org")
install.packages("srvyr", repos = "http://cran.us.r-project.org")
install.packages("forcats", repos = "http://cran.us.r-project.org")
install.packages("statsr", repos = "http://cran.us.r-project.org")
# load gssr package
remotes::install_github("kjhealy/gssr")



# load libraries
library(gssr)
library(critstats)
library(descr)
library(Hmisc)
library(dplyr)
library(tidyr)
library(ggplot2)
library(haven)
library(tibble)
library(survey)
library(srvyr)
library(forcats)
library(statsr)
library(dplyr)



# load the data
data(gss_all)
data(gss_doc)
data(gss_dict)
```

```
gss_dict

gss_all %>% # check for variable in each year
  gss_which_years(c("degree", "partyid", "natrace", "age", "race", "wtssall")) %>%
  print(n = Inf)
gss18 <- gss_all %>% filter(year == 2018) %>% filter(race == 1) # Select ONLY for the year
2018 and White respondents

df <- gss18 %>%
  select(c("degree", "partyid", "natrace", "race", "wtssall")) %>%
  drop_na()
sapply(df, function(x) sum(is.na(x))) # check that all NA was dropped

cat_vars = c("degree", "natrace", "partyid")
wt_vars = c("wtssall")

# use variable lists to mutate the variables
dfc <- df %>% # create a clean data set
  mutate(wtssall = as.numeric(wtssall)) %>%
  mutate(across(all_of(cat_vars), forcats::as_factor))
dfc

dfc %>% # count # of values in race (should be only white)
  group_by(race) %>%
  summarize(count = n())

dfc %>%  # count # of values in degree
  group_by(degree) %>%
  summarize(count = n())

dfc %>%  # count # of values in party
  group_by(partyid) %>%
  summarize(count = n())

dfc %>%  # count # of values in natrace
  group_by(natrace) %>%
  summarize(count = n())
dfc <- dfc %>% ## drop out non left or right leaning responses
  filter(partyid != "independent (neither, no response)",
         partyid != "other party",) %>%
```

```
  droplevels()

dfc %>%  # count # of values in party
  group_by(partyid) %>%
  summarize(count = n())

dfc <- dfc %>% # Make the only political affiliations left and right
  mutate(partyid = fct_recode(partyid,
                              "left" = "strong democrat",
                              "left" = "not very strong democrat",
                              "left" = "independent, close to democrat",
                              "right" = "independent, close to republican",
                              "right" = "not very strong republican",
                              "right" = "strong republican"))
dfc %>%  # count # of values in party
  group_by(partyid) %>%
  summarize(count = n())

dfc %>%  # count # of values in degree
  group_by(degree) %>%
  summarize(count = n())

dfc <- dfc %>% # Make the only education levels no college degree and college degree
  mutate(degree = fct_recode(degree,
                             "no degree" = "less than high school",
                             "no degree" = "high school",
                             "college degree" = "associate/junior college",
                             "college degree" = "independent, close to republican",
                             "college degree" = "bachelor's",
                             "college degree" = "graduate"))
dfc %>%  # count # of values in degree
  group_by(degree) %>%
  summarize(count = n())

dfc %>%  # count # of values in natrace
  group_by(natrace) %>%
  summarize(count = n())
dfc <- dfc %>% # Make the only options for natrace "more spending" "not more spending"
  mutate(natrace = fct_recode(natrace,
                              "more spending" = "too little",
```

```
                    "not more spending" = "about right",
                    "not more spending" = "too much"))
dfc %>%  # count # of values in natrace
  group_by(natrace) %>%
  summarize(count = n())

table <- table(dfc$partyid, dfc$degree, dfc$natrace) # make a table of binary variables

table

# split the table more or not more spending
table_more <- table[, , "more spending"]
table_not_more <- table[, , "not more spending"]

# Set up the margins (bottom, left, top, right)
par(mfrow=c(1,1), mar=c(2, 3, 3, 2))

# Barplot for 'more spending'
barplot(table_more, legend.text = rownames(table_more), main="More spending on Black
people",
        beside=TRUE, ylim=c(0, max(table_more)*1.2),
        col=c("grey30", "grey80"))

# Barplot for 'not more spending'
barplot(table_not_more, legend.text = rownames(table_not_more), main="Not more spending on
Black people",
        beside=TRUE, ylim=c(0, max(table_not_more)*1.2),
        col=c("grey30", "grey80"))

# Calculate sums for the left
left_more_spend <- sum(table["left", , "more spending"])
left_not_more_spend <- sum(table["left", , "not more spending"])
total_left <- left_more_spend + left_not_more_spend

# Calculate percentages for the left
perc_left_more_spend <- (left_more_spend / total_left) * 100
perc_left_not_more_spend <- (left_not_more_spend / total_left) * 100

# Calculate sums for the right
right_more_spend <- sum(table["right", , "more spending"])
```

```
right_not_more_spend <- sum(table["right", , "not more spending"])
total_right <- right_more_spend + right_not_more_spend

# Calculate percentages for the right
perc_right_more_spend <- (right_more_spend / total_right) * 100
perc_right_not_more_spend <- (right_not_more_spend / total_right) * 100

# percentages for political affiliation
cat("Left - More Spend: ", perc_left_more_spend, "%\n")
cat("Left - Not More Spend: ", perc_left_not_more_spend, "%\n")
cat("Right - More Spend: ", perc_right_more_spend, "%\n")
cat("Right - Not More Spend: ", perc_right_not_more_spend, "%\n")

# Calculate sums for college degree
col_more_spend <- sum(table[, "college degree", "more spending"])
col_not_more_spend <- sum(table[, "college degree", "not more spending"])
total_col <- col_more_spend + col_not_more_spend

# Calculate percentages for college degree
perc_col_more_spend <- (col_more_spend / total_col) * 100
perc_col_not_more_spend <- (col_not_more_spend / total_col) * 100

# Calculate sums for no degree
no_deg_more_spend <- sum(table[, "no degree", "more spending"])
no_deg_not_more_spend <- sum(table[, "no degree", "not more spending"])
total_no_deg <- no_deg_more_spend + no_deg_not_more_spend

# Calculate percentages for no degree
perc_no_deg_more_spend <- (no_deg_more_spend / total_no_deg) * 100
perc_no_deg_not_more_spend <- (no_deg_not_more_spend / total_no_deg) * 100

# percentages for education level
cat("Col Degree - More Spend: ", perc_col_more_spend, "%\n")
cat("Col Degree - Not More Spend: ", perc_col_not_more_spend, "%\n")
cat("No Degree - More Spend: ", perc_no_deg_more_spend, "%\n")
cat("No Degree - Not More Spend: ", perc_no_deg_not_more_spend, "%\n")



# Calculate sums for left
```

```
left_col_deg <- sum(table["left", "college degree", ])
left_no_deg <- sum(table["left", "no degree", ])
total_left <- left_col_deg + left_no_deg

# Calculate percentages for left
perc_left_col_deg <- (left_col_deg / total_left) * 100
perc_left_no_deg <- (left_no_deg / total_left) * 100

# Calculate sums for right
right_col_deg <- sum(table["right", "college degree", ])
right_no_deg <- sum(table["right", "no degree", ])
total_right <- right_col_deg + right_no_deg

# Calculate percentages for right
perc_right_col_deg <- (right_col_deg / total_right) * 100
perc_right_no_deg <- (right_no_deg / total_right) * 100

# Percentages of degrees and politics
cat("Left - College Degree: ", perc_left_col_deg, "%\n")
cat("Left - No Degree: ", perc_left_no_deg, "%\n")
cat("Right - College Degree: ", perc_right_col_deg, "%\n")
cat("Right - No Degree: ", perc_right_no_deg, "%\n")

logmod <- glm(natrace ~ degree + partyid, family = binomial, data = dfc)
summary(logmod)
```