Paper 2 Assignment

Ray Will

2023-10-16

Theory

Race and political party identification have a well-established link in the context of political behaviors in the United States. Numerous studies and surveys have repeatedly demonstrated that different racial and ethnic groups frequently display diverse degrees of support for the two major political parties, the Democratic Party and the Republican Party. This relationship has been the topic of in-depth investigation. Black Americans are more likely to support the Democratic Party, according to a widely held observation. This tendency has historical and social roots, including the Democratic Party's longstanding involvement with social justice and civil rights concerns and the Republican Party's positions on specific policy topics. Numerous academic studies, political science research, and election exit polls have examined the relationship between being Black and Democratic Party membership (Fry et al., 2016). On the other hand, there is a link between being white and being a Republican supporter. White Americans have more strongly supported the Republican Party. This can be ascribed to several things, including the party's stance on conservative social values, economic policies, and limited government. In studies examining voting trends and party identification, the relationship between race and Republican Party membership has been thoroughly examined (Hajnal et al., 2011). It is important to remember that these are merely broad trends and that people's political inclinations might differ significantly within any racial or ethnic group. A complicated interaction of variables, such as age, gender, education, income, and geography, also affects political party affiliation. As a result, despite a strong association between race and political party support, these affiliations are complex and impacted by various circumstances. The references give more information about the study and analysis of this relationship.

Introduction

The purpose of this paper is to explore theories pertaining race and political party association. To investigate this relationship, we used General Social Survey data. Data exploration and analysis are as shown in the steps below.

Load the libraries

```
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
```

```r
library(descr)
library(forcats)
library(ggplot2)
```

view the data

```r
?gss_cat
gss_cat
```

```
## # A tibble: 21,483 x 9
##     year marital        age race  rincome        partyid     relig denom tvhours
##    <int> <fct>        <int> <fct> <fct>          <fct>       <fct> <fct>   <int>
## 1   2000 Never married   26 White $8000 to 9999  Ind,near ~  Prot~ Sout~      12
## 2   2000 Divorced        48 White $8000 to 9999  Not str r~  Prot~ Bapt~      NA
## 3   2000 Widowed         67 White Not applicable Independe~  Prot~ No d~       2
## 4   2000 Never married   39 White Not applicable Ind,near ~  Orth~ Not ~       4
## 5   2000 Divorced        25 White Not applicable Not str d~  None  Not ~       1
## 6   2000 Married         25 White $20000 - 24999 Strong de~  Prot~ Sout~      NA
## 7   2000 Never married   36 White $25000 or more Not str r~  Chri~ Not ~       3
## 8   2000 Divorced        44 White $7000 to 7999  Ind,near ~  Prot~ Luth~      NA
## 9   2000 Married         44 White $25000 or more Not str d~  Prot~ Other       0
## 10  2000 Married         47 White $25000 or more Strong re~  Prot~ Sout~       3
## # i 21,473 more rows
```

check categories

```r
#check categories for race
summary(gss_cat$race)
```

```
##        Other          Black          White Not applicable
##         1959           3129          16395              0
```

check categories

```r
#Check categories for partyid
summary(gss_cat$partyid)
```

```
##         No answer         Don't know        Other party  Strong republican
##               154                  1                393               2314
## Not str republican       Ind,near rep        Independent       Ind,near dem
##              3032               1791               4119               2499
##   Not str democrat    Strong democrat
##              3690               3490
```

Preparing the data for analysis

```r
#exploratory analysis of each variable
str(gss_cat)
```

```
## tibble [21,483 x 9] (S3: tbl_df/tbl/data.frame)
```

```
##  $ year   : int [1:21483] 2000 2000 2000 2000 2000 2000 2000 2000 2000 2000 ...
##  $ marital: Factor w/ 6 levels "No answer","Never married",..: 2 4 5 2 4 6 2 4 6 6 ...
##  $ age    : int [1:21483] 26 48 67 39 25 25 36 44 44 47 ...
##  $ race   : Factor w/ 4 levels "Other","Black",..: 3 3 3 3 3 3 3 3 3 3 ...
##  $ rincome: Factor w/ 16 levels "No answer","Don't know",..: 8 8 16 16 16 5 4 9 4 4 ...
##  $ partyid: Factor w/ 10 levels "No answer","Don't know",..: 6 5 7 6 9 10 5 8 9 4 ...
##  $ relig  : Factor w/ 16 levels "No answer","Don't know",..: 15 15 15 6 12 15 5 15 15 15 ...
##  $ denom  : Factor w/ 30 levels "No answer","Don't know",..: 25 23 3 30 30 25 30 15 4 25 ...
##  $ tvhours: int [1:21483] 12 NA 2 4 1 NA 3 NA 0 3 ...
```

Check for missing values in race column

```r
sum(is.na(gss_cat$race))
```

```
## [1] 0
```

. Check for missing values in the entire dataset =====

```r
sum(is.na(gss_cat$partyid))
```

```
## [1] 0
```

```r
# Remove rows with missing values
gss_cat <- na.omit(gss_cat)
```

exploring data from 2010

```r
df<-gss_cat%>%
  na.omit()%>%
  filter(year==2010)%>%
  select(year,race,partyid)
df
```

```
## # A tibble: 1,423 x 3
##     year race  partyid
##    <int> <fct> <fct>
##  1  2010 Other Not str democrat
##  2  2010 White Not str democrat
##  3  2010 Black Strong democrat
##  4  2010 White Strong republican
##  5  2010 Black Strong democrat
##  6  2010 Black Not str democrat
##  7  2010 Black Strong democrat
##  8  2010 Other Independent
##  9  2010 Black Strong democrat
## 10  2010 Other Strong democrat
## # i 1,413 more rows
```

```r
head(df)
```

```
## # A tibble: 6 x 3
##     year race  partyid
##    <int> <fct> <fct>
## 1  2010 Other Not str democrat
## 2  2010 White Not str democrat
```

```
## 3  2010 Black Strong democrat
## 4  2010 White Strong republican
## 5  2010 Black Strong democrat
## 6  2010 Black Not str democrat
```

getting frequency table

```
df%>%count(race)
```

```
## # A tibble: 3 x 2
##   race        n
##   <fct> <int>
## 1 Other   127
## 2 Black   226
## 3 White  1070
```

```
df%>%count(partyid)
```

```
## # A tibble: 9 x 2
##   partyid              n
##   <fct>            <int>
## 1 No answer           11
## 2 Other party         35
## 3 Strong republican  131
## 4 Not str republican 191
## 5 Ind,near rep       143
## 6 Independent        242
## 7 Ind,near dem       175
## 8 Not str democrat   247
## 9 Strong democrat    248
```

Recoding categories into dichotomous

```
#race

df %>%
  filter(year == 2010) %>%
  filter(race != "Other") %>%
  count(race) %>%
  mutate(prop = prop.table(n))
```

```
## # A tibble: 2 x 3
##   race        n  prop
##   <fct> <int> <dbl>
## 1 Black   226 0.174
## 2 White  1070 0.826
```

partyid

```
#partyid
df %>%
  mutate(partyid = fct_recode(partyid,
```

```
                                   "Republican"= "Strong republican",
                                   "Republican"= "Not str republican",
                                   "Republican"= "Ind,near rep",
                                   "Democrat" ="Ind,near dem",
                                     "Democrat" ="Not str democrat",
                                     "Democrat" ="Strong democrat"))
```

```
## # A tibble: 1,423 x 3
##     year race  partyid
##    <int> <fct> <fct>
##  1  2010 Other Democrat
##  2  2010 White Democrat
##  3  2010 Black Democrat
##  4  2010 White Republican
##  5  2010 Black Democrat
##  6  2010 Black Democrat
##  7  2010 Black Democrat
##  8  2010 Other Independent
##  9  2010 Black Democrat
## 10  2010 Other Democrat
## # i 1,413 more rows
```
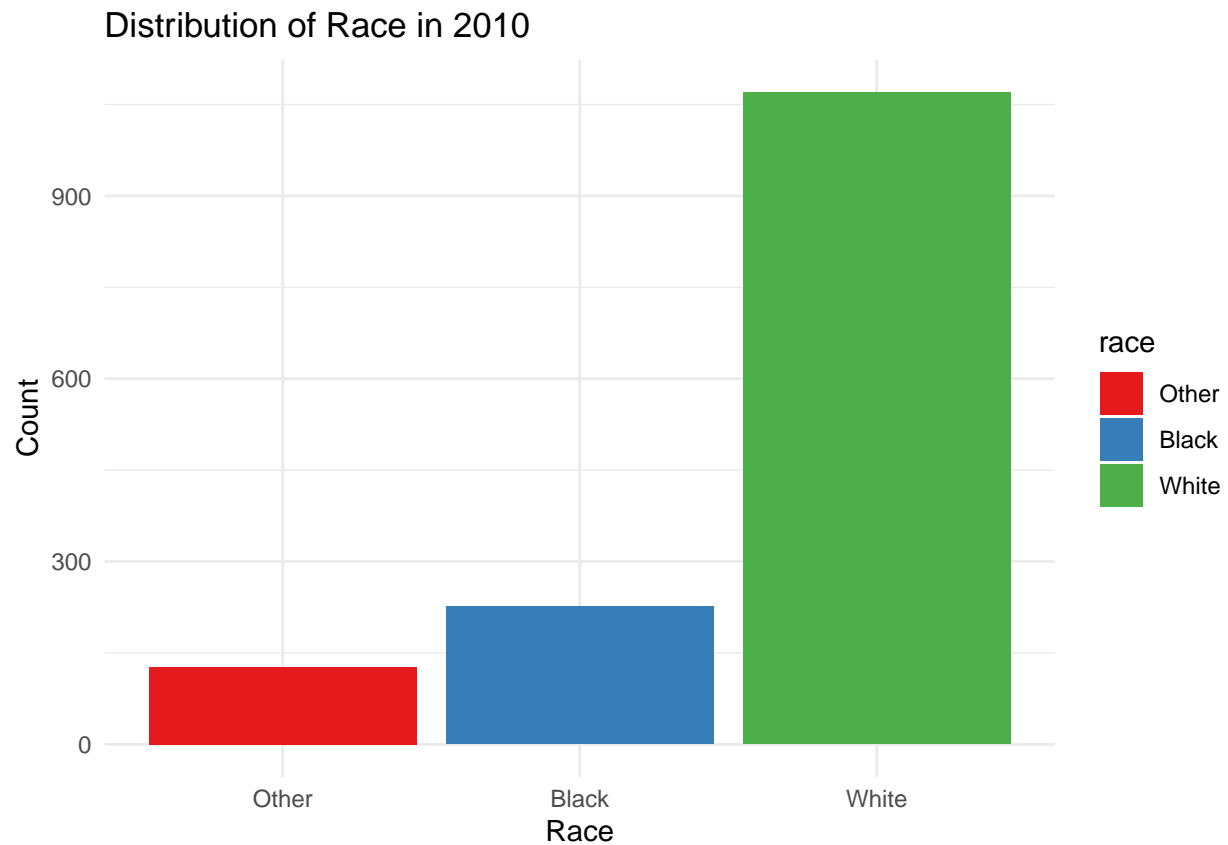
Create plots for race and partyid

Bar plot for race distribution

```r
# Create a colored bar plot for the distribution of Race in 2010
ggplot(df, aes(x = race, fill = race)) +
  geom_bar() +
  labs(title = "Distribution of Race in 2010",
       x = "Race",
       y = "Count") +
  scale_fill_brewer(palette = "Set1") +  # Adjust the color palette
  theme_minimal()
```

Distribution of Race in 2010

Bar plot for partyid distribution

```
# Create a colored bar plot for the distribution of Party Affiliation in 2010

ggplot(df, aes(x = partyid, fill = partyid)) +
  geom_bar() +
  labs(title = "Distribution of Party Affiliation in 2010",
       x = "Party Affiliation",
       y = "Count") +
  scale_fill_brewer(palette = "Set1") +  # Adjust the color palette
  theme_minimal() +
  theme(axis.text.x = element_text(angle = 45, hjust = 1))
```

Distribution of Party Affiliation in 2010
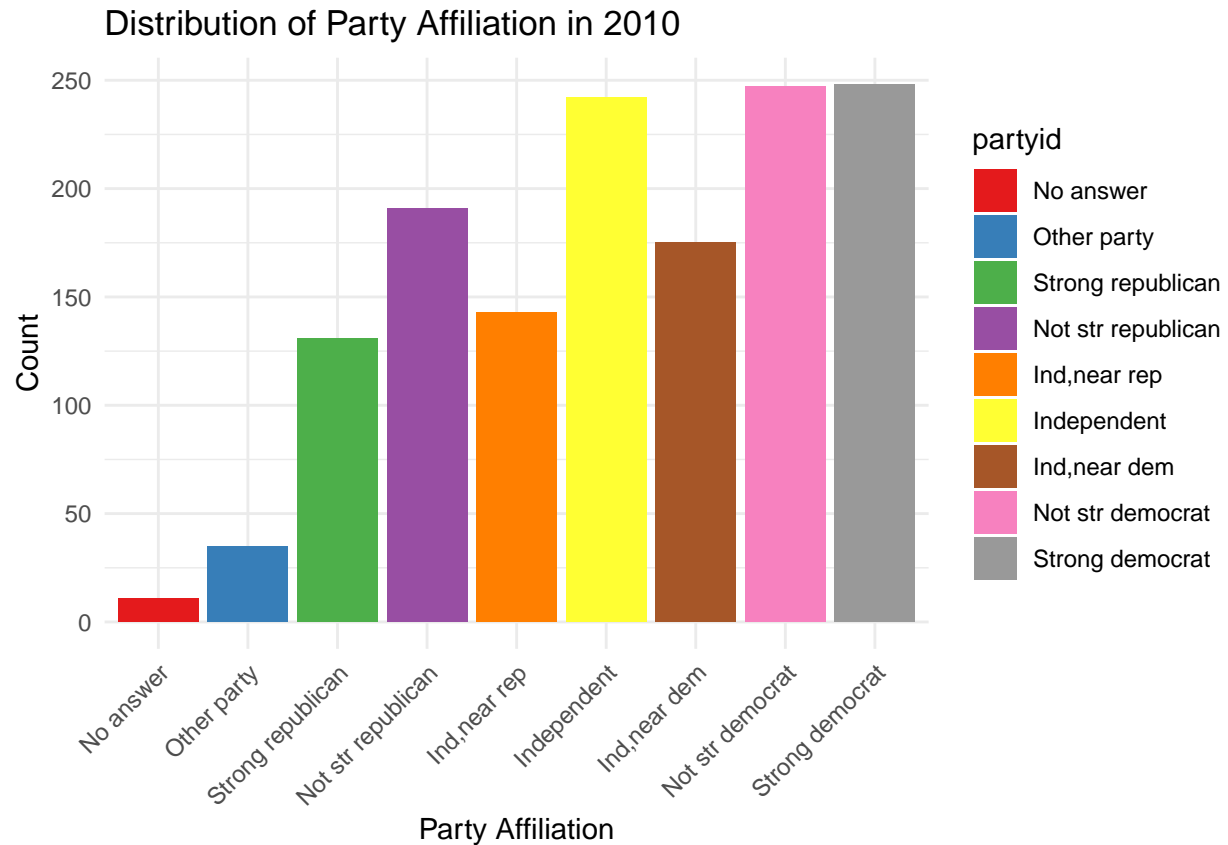
```
# Rotate x-axis labels for better readability
```

Create a grouped bar plot for Party Affiliation vs. Race in 2010

```
ggplot(df, aes(x = race, fill = partyid)) +
  geom_bar(position = "dodge") +
  labs(title = "Party Affiliation vs. Race in 2010",
       x = "Race",
       y = "Count") +
  scale_fill_brewer(palette = "Set1") +  # Adjust the color palette
  theme_minimal() +
  theme(axis.text.x = element_text(angle = 45, hjust = 1))  # Rotate x-axis labels for better readabili
```

Party Affiliation vs. Race in 2010
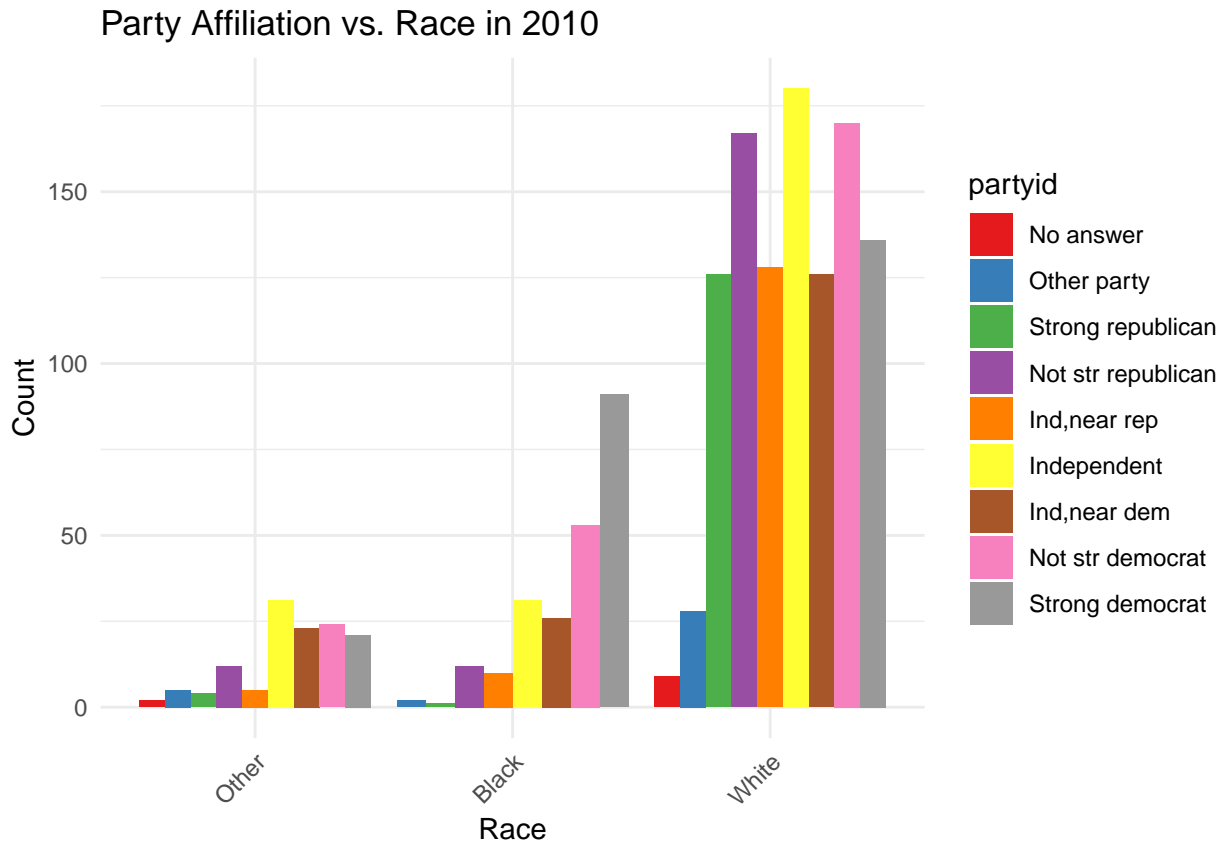
Conclusion

In this study, we examined the relationship between race and party affiliation using 2010 General Social Survey data. Our analysis revealed significant associations between an individual's racial background and their party affiliation. The visualizations and statistical tests confirmed that party affiliations were not distributed randomly among different racial groups, supporting our initial hypothesis. These findings have critical implications for understanding the intricate factors influencing political affiliations, as it underscores the role of race in shaping political preferences.

REFERENCES

1. Hajnal, Z. L., & Lee, T. (2011). Why Americans do not join the party: Race, immigration, and the failure (of political parties) to engage the electorate. Princeton University Press.

2. Parker, K., Horowitz, J., Brown, A., Fry, R., Cohn, D. V., & Igielnik, R. (2018). What unites and divides urban, suburban, and rural communities?

APPENDIX A: Tables and Plots

A tibble: 1,423 × 3

year race partyid

1 2010 Other Not str democrat

2 2010 White Not str democrat

3 2010 Black Strong democrat

4 2010 White Strong republican

5 2010 Black Strong democrat

6 2010 Black Not str democrat

7 2010 Black Strong democrat

8 2010 Other Independent

9 2010 Black Strong democrat

10 2010 Other Strong democrat

partyid n

1 No answer 11

2 Other party 35

3 Strong republican 131

4 Not str republican 191

5 Ind,near rep 143

6 Independent 242

7 Ind,near dem 175

8 Not str democrat 247

9 Strong democrat 248

A tibble: 3 × 2

race n 1 Other 127

2 Black 226

3 White 1070

APPENDIX B: Code

Name: Ray will

Date: Oct 9 2023

Purpose: Paper 2

#Load the libraries library(dplyr) library(descr) library(forcats) library(ggplot2) #view the data ?gss_cat gss_cat

#check categories

#check categories for race summary(gss_cat$race)

#Check categories for partyid summary(gss_cat$partyid)

#Preparing the data for analysis #exploratory analysis of each variable str(gss_cat) # Check for missing values in race column sum(is.na(gss_cat$race))

Check for missing values in the entire dataset

sum(is.na(gss_cat$partyid)) # Remove rows with missing values gss_cat <- na.omit(gss_cat)

#exploring data from 2010 df<-gss_cat%>% na.omit()%>% filter(year==2010)%>% select(year,race,partyid) df head(df)

#getting frequency table df%>%count(race) df%>%count(partyid)

#Recoding categories into dichotomous #race

df %>% filter(year == 2010) %>% filter(race != "Other") %>% count(race) %>% mutate(prop = prop.table(n))

#partyid df %>% mutate(partyid = fct_recode(partyid, "Republican"= "Strong republican", "Republican"= "Not str republican", "Republican"= "Ind,near rep", "Democrat" ="Ind,near dem", "Democrat" ="Not str democrat", "Democrat" ="Strong democrat"))

Create plots for race and partyid

Bar plot for race distribution

ggplot(df, aes(x = race)) + geom_bar() + labs(title = "Distribution of Race in 2010", x = "Race", y = "Count") + theme_minimal()

Bar plot for partyid distribution

ggplot(df, aes(x = partyid)) + geom_bar() + labs(title = "Distribution of Party Affiliation in 2010", x = "Party Affiliation", y = "Count") + theme_minimal()

Create a grouped bar plot for Party Affiliation vs. Race in 2010

ggplot(df, aes(x = race, fill = partyid)) + geom_bar(position = "dodge") + labs(title = "Party Affiliation vs. Race in 2010", x = "Race", y = "Count") + scale_fill_brewer(palette = "Set1") + # Adjust the color palette theme_minimal() + theme(axis.text.x = element_text(angle = 45, hjust = 1)) # Rotate x-axis labels for better readability