Auburn University
Masters of Public Administration
POLI 7000: Research Methods for Public and Nonprofit Organizations
Dr. Soren Jordan
April 29, 2024

The Fight for Fairness

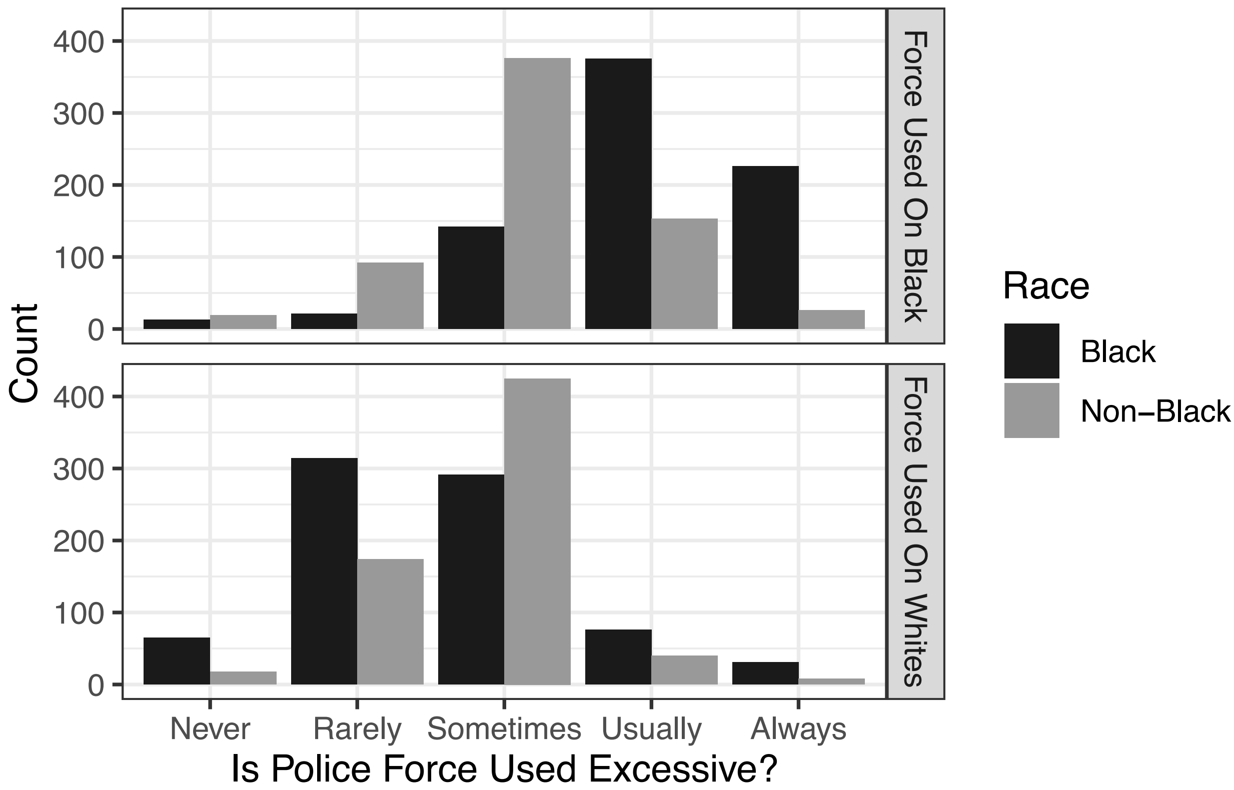
Fixing the Interwoven Racial Bias in America's Police Force

Prepared by

Abbey Crank and NPO Team

Opinion

Do the Police Use Excessive Force Based on Race?

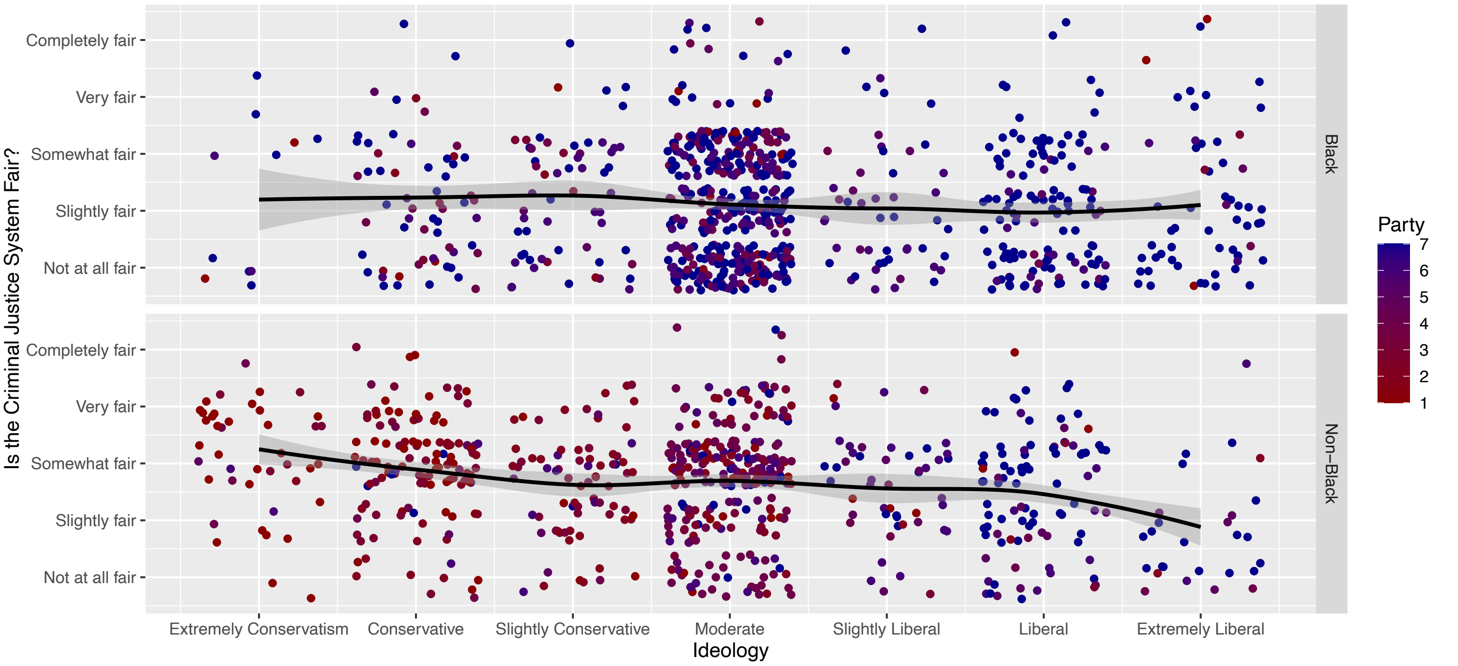


Researchers Jefferson, Neuner, and Pasek asked a group of 1,430 respondents about their opinions of police brutality/use of excessive police force in correspondence with racial bias. The sample was split 50% Black and 50% white. When looking at the graph, one can conclude that the pattern of “Force Used On Black” has most responses fall between the Sometimes, Usually, and Always categories. It didn’t matter if the respondent was white or Black, it seems most are in agreement that police use excessive force on Black Americans at least some of the time. It’s interesting to see that the main difference is that most white respondent responded with “Sometimes” while the Black respondents were more inclined to put “Usually.”

When it comes to “Force Used On Whites,” it’s almost flipped from the responses to their Black counterparts. Instead of the respondents’ answers trending towards “Sometimes” and upwards, respondents trend from “Sometimes” all the way down to “Never.” It didn’t drastically change whether the person was white or Black. It shows that, no matter the race, all respondents conclude that more excessive force is used on Black citizens compared to white when it comes to police interaction. It also proves that the force on white can be seen as excessive “some” of the time rather than never. So, while this seems to suggest that racial bias is at play, this may also conclude that the police force is seen as a threat rather than help. In both graphs “Never” is basically inexistent, and the respondents don’t see interaction with police officers as a 100% safe interaction.

It's also important to note that this was taken in 2016, before the second wave of the Black Lives Matter movement. Police brutality was covered in the news, but the nation-wide outrage hasn’t happened yet. Brown’s death, which is something we examine later, ignited the first wave of BLM 6 years before the murder of George Floyd According to Southern Oregon University researcher Travis Campbell, BLM protests reduced police-involved homicides/lethal force by 13%, and that many major cities had protests about excessive force after Brown’s death (Campbell, 2021). So, with these results taken 2 years after the incident, it's important to note that the outrage people felt years ago might be lessened.

Does One's Ideology Alter Their View of the Justice System?



A critical piece of a person's identity is their ideology. The way they view the world represents their core values and empathy. In this chart, the same respondents were asked if the U.S. justice system was fair, and their answers were compared by their ideology, race, and political party association. The results suggest that were many moderates within the sample. What's interesting is how Black moderates view the justice system more negatively than their white counterparts. White respondents, no matter their ideology, see the justice system as more fair than Black respondents across all ideology. The more liberal non-Black respondents got, it obviously got more negative, but that trend doesn't fully dip until one considered themselves liberal. It's important to note that the more conservative a non-Black respondent got, the more likely they were to see the system are fair. For Black, the trend line suggests, no matter the ideology/party, the opinion of the justice system is pretty consistent. Most Black respondents see the justice system as slightly fair or somewhat fair.

How Do Personal Characteristics Affect One's Opinion?

The team decided to pinpoint how certain characteristics of respondents come together to create their overall opinion. The information used was the respondents' race, party/political ideology, whether they lived in the suburbs, gender, income, education, and their age.

Opinion 1: Police use more force than necessary on Black citizens

Characteristics put side-by-side with the opinion of excessive police force on Black Americans. The force was rated on a 1-to-5 scale ranging from "Never" to "Always." Out of all defining characteristics, the one that affected one's view on excessive force against Black citizens is if the respondent was black. Race increased one's opinion by 0.7 on the 1-to-5 scale, making it almost a 20% jump in believing there is unneeded force. Other variables that had impact were party and ideology, meaning the more democratic/liberal the respondent was, the more they moved up the scale. Age also had a small influence was age; the older someone got, the more they thought there was use of unnecessary force. The variables we didn't find a connection were gender, income bracket, education, and suburban life.

Opinion 2: Police use more force than necessary on white citizens

Along with force on Black citizens, researchers also looked at police force used on white citizens. Just like the one above, the force was rated on the 1-to-5 scale ranging from "Never" to "Always." The characteristic that affected the overall answer to this was the respondent's ideology. The strongest effect was found in the respondent's ideology, meaning the more liberal a respondent was the more they believed excessive police force was used on white citizens. The change is still small, with a maximum effect of +0.26 on the 1-to-5 scale. There were two negative effects found. First, the more educated they were, the less likely they were to think that excessive force was used. Second, if the respondent was Black, they were less likely to think excessive force was used. The variables we didn't find a connection with were gender, party affiliation, suburban life, income bracket, and age.

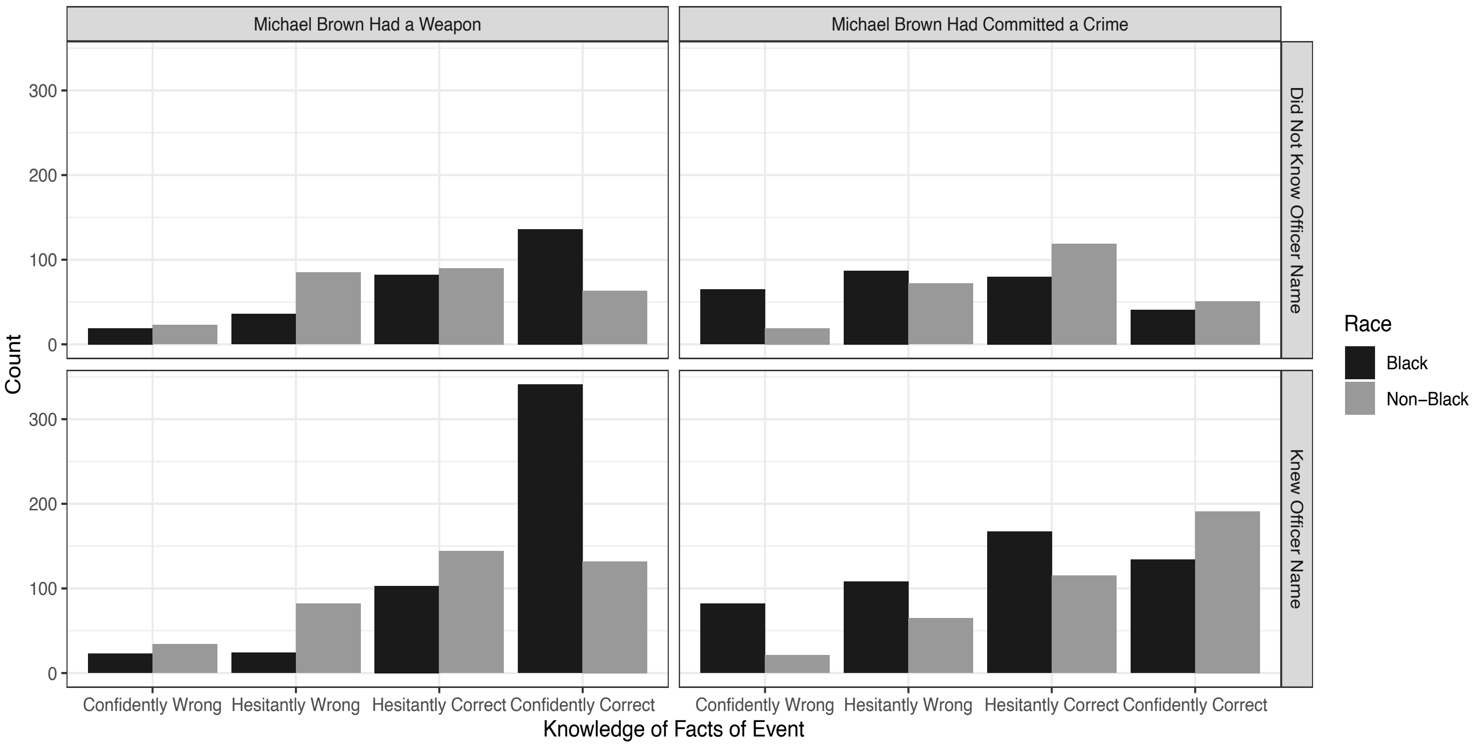
Opinion 3: The Criminal Justice System is fair for all Americans

Third, respondents were asked if they thought the criminal justice system was fair. The fairness was measured on a 1-to-5 scale, ranging from "Not at all fair" to "Completely fair." The characteristic that affected the overall answer the most was the respondent's race. If the person was Black, they were -0.5 less likely to think the system is fair, which is half a point on the only 5-point scale. Ideology followed closely behind with affecting the scale by -0.48 if the respondent was "extremely liberal." The last found variable that negatively affected responses was if the respondent was female. The variables that weren't seeming to have an confirmed effect were party affiliation, suburban life, income bracket, education, and age.

Opinion 4: The police treat Black and white citizens differently

Lastly, researchers wanted to see if people thought there was a difference in overall treatment depending on race. The scale for treatment was 1-to-7, ranging from "Police treat Blacks much better" to "Police treat whites much better." Out of all variables, race affected the response most significantly. A Black respondent's answer affected the scale by +0.96, meaning they are more likely to think white people are treated better. All other factors that caused the respondent to lean toward white favorability were party affiliation, ideology, education, and age. None of the confirmed, correlation variables had a negative impact. The variables that weren't confirmed to have an effect were suburban life, gender, and income bracket.

How People Responded to Michael Brown's Case



The last barplot was created to show the respondent's knowledge on Michael Brown's case. The plots were separated into 4 sections interlaced within each other. On top, the data is separated by knowing if Brown having a weapon and if Brown committed a crime. The horizontal separator is if the respondent knew the name of the officer who killed Brown. An important indicator is that most of the Black respondents who knew the officer's name were confidently correct about Brown not having a weapon. In the same ballpark, most Black respondents that knew the officer's name were correct, whether hesitantly or confidently, about Brown committing a crime. In most areas, it seems like those who knew the officer's name were more likely to know the key details about the case, no matter if it was Black or white respondents. For those who didn't know the officer's name, many weren't confidently to answer the questions either way. When comparing the bar plots of those that knew the name versus not, they look similar towards each other, especially in the "hesitant" categories. The main difference is that it seems most of the population knew the officer's name. Even though this happened in 2014, it speaks volumes that two years later many respondents still knew the officer's name and the details of the case. It shows the impact the murder had on many respondent's opinions about police brutality and racial bias.

Next, the team pinpointed how certain characteristics of respondents come together to create an overall opinion and knowledge of Brown's case. The attributes highlighted in the models below are the respondent's race, party affiliation, ideology, if they lived in a suburban area, gender, if they knew the officer's name, if they knew the officer wasn't indicted, education, income, and age.

Response 1: Did the respondent know if Brown have a weapon?

The first model focused on if the respondent knew that Brown didn't have a weapon. This is a 1-to-4 scale, ranging from being confidently incorrect to confidently correct. The variable that had the most impact was the respondent's race. If the answerer was Black, they were 0.41 more likely to correctly guess that Brown had a weapon. The other variables with a positive effect are the respondent's party affiliation (the more democratic, the more likely they were correct) and if they knew the name of the officer. Variables that weren't confirmed to have an effect are ideology, suburban living, gender, if they knew the officer was indicted, income, education, and age.

Response 2: Did the respondent know if Brown committed a crime?

The second model focused on if the respondent knew Brown committed a crime. This was a 1-to-4 scale, ranging from being confidently incorrect and confidently correct. The variable that had the most impact was if the respondent knew the name of the officer. If they got the name correct, they were almost +0.4 more likely to know Brown committed a crime. The two other components that increase the likelihood of knowing was their income bracket and if they are older in age. There were also many different characteristics that caused a respondent more likely to be wrong: if they were Black, have left-leaning ideologies, democratic political affiliation, and if they were female. Characteristics that weren't confirmed to have an effect are suburban living, if they knew the officer was indicted, and education.

Response 3: What made the respondent invested in Brown's case?

A respondent's investment in Brown's case was measured on a 1-to-5 scale, ranging from "Not at all" to "A great deal." Investment was most affected by the race of the respondent. If they were Black, they'd increase on the scale by +0.7, almost a whole point making up 20% of the scale. The other characteristics that can increase one's investment are democratic party affiliations, left-leaning ideology, if they knew Brown didn't have a weapon, and the more income the respondent had. Characteristics that weren't confirmed to have an effect on one's investment are suburban living, gender, knowing if a crime was committed, educational background, and age.

Response 4: Did the respondent think racial bias was involved?

Lastly, researchers asked if respondents thought racial bias played a role in Brown's murder. Answers were collected on a 1-to-5 scale, ranging "No role at all" to "An enormous role." The characteristic that affected this opinion the most was race. If the respondent was black, they were 0.86+ more likely to think racial bias played a role. Other aspects that make people think racial bias is involved is if they lean towards the Democratic party, have left-leaning ideologies, knew Brown didn't have a weapon, their education, and their age. The only attribute that decreases the belief in racial bias is if the respondent knew Brown committed a crime, which makes sense as a reason to believe racial bias wasn't involved. Characteristics that weren't confirmed to have an effect on one's opinion are suburban life, gender, and income bracket.

Next Steps: Building Emotional Investment

From Jefferson, Neuner, and Pasek's research, one can conclude Americans, especially Black citizens, are fearful when interacting with police officers. To build emotional investment within the community, it's crucial to build programs within the area as a way to bring awareness and how to's when interacting with officers. As a journalist who specialized in local community relations, using this information as a way to convince police departments to create community liaisons within their community. In July 2021, I had the privilege to attend Opelika Police Department's first meeting with elected community members as a way to build a bond between the department and the community. When liaisons met with the police chief and important captains in the force, it was a time for the two sides to understand how to work together (Crank, 2021). Using this information, our nonprofit could use this to catapult a similar task force within our area. It's a step in the right direction, and through this partnership, we could provide training for police interaction with citizens. We could meet with the police force for monthly trainings on implicit racial bias and how to safely protect everyone while on the job. Building emotional investment in any issue requires on thing: for people to care. People care through human connection, personal hardships, and passion. All we need to do is start the ball rolling, which is exactly what connecting community leaders and the police force. It gives humanity to both sides of the divided spectrum. I'm not saying the problems will immediately be fixed, but this strong foundation will provide the future with a brighter, stable relationship among the community and government/police officials.

References

Campbell, T. (2021). Black Lives Matter's Effect on Police Lethal Use of Force. *Journal of Urban Economics*, *141*, 103587. https://doi.org/10.1016/j.jue.2023.103587

Crank, A. (2021, July 9). *OPD strives for change*. The Observer. https://opelikaobserver.com/2021/07/07/opd-strives-for-change/

Jefferson, H., Neuner, F. G., & Pasek, J. (2020). Seeing blue in black and white: Race and perceptions of officer-involved shootings. *Perspectives on Politics*, *19*(4), 1165–1183. https://doi.org/10.1017/s1537592720003618

R-script

```
library(tidyverse)

#load in dataset
data = read.csv(file.choose())

#1. First ggplot: barplot of the beliefs of whether the use of force is excessive, faceted by whether the use of force is
#against whites or against Blacks, and filled by whether the respondent was Black or non-Black.

data$black.fac <- NA
data$white.force.fac <- NA
data$black.force.fac <- NA

data$black.fac[data$black == 0] <- "Non-Black"
data$black.fac[data$black == 1] <- "Black"

data$white.force.fac[data$forceonwhite == 1] <- "Never"
data$white.force.fac[data$forceonwhite == 2] <- "Rarely"
data$white.force.fac[data$forceonwhite == 3] <- "Sometimes"
data$white.force.fac[data$forceonwhite == 4] <- "Usually"
data$white.force.fac[data$forceonwhite == 5] <- "Always"

data$black.force.fac[data$forceonblack == 1] <- "Never"
data$black.force.fac[data$forceonblack == 2] <- "Rarely"
data$black.force.fac[data$forceonblack == 3] <- "Sometimes"
data$black.force.fac[data$forceonblack == 4] <- "Usually"
data$black.force.fac[data$forceonblack == 5] <- "Always"

data$black.force.fac = factor(data$black.force.fac, levels = c("Never", "Rarely", "Sometimes",
"Usually", "Always"))
data$white.force.fac = factor(data$white.force.fac, levels = c("Never", "Rarely", "Sometimes",
"Usually", "Always"))

group.names <- c(black.force.fac = "Force Used On Black", white.force.fac = "Force Used On
Whites")

data.plot <- data %>%
  pivot_longer(cols = c(white.force.fac, black.force.fac), names_to = "Group", values_to = "Excessive")
%>%
  select(black.fac, Excessive, Group) %>%
  na.omit()

summary(data.plot)

length(data$black.fac)
summary(data$black)
length(data$black)
ggplot(data.plot, aes(x = Excessive, fill = as.factor(black.fac))) + geom_bar(stat = "count", position =
"dodge") + scale_fill_manual(values = c("grey10", "grey60")) + facet_grid(rows = vars(Group), labeller
= labeller(Group = group.names)) + labs(x = "Is Police Force Used Excessive?", y = "Count", fill =
"Race") + theme_bw()
```

```
#2. Second ggplot: jittered plot with a smoothed regression line, showing the relationship between ideology and whether the
#criminal justice system is fair. Points colored by party, and panels faceted by whether the respondent was Black or non-Black
#colors must be changed so that Republicans are the traditional red, and Democrats are traditional blue

data$ideology.fac <- NA
data$crimfair.fac <- NA
data$fac.party <- NA

data$ideology.fac[data$ideology == 1] <- "Extremely Conservative"
data$ideology.fac[data$ideology == 2] <- "Conservative"
data$ideology.fac[data$ideology == 3] <- "Slightly Conservative"
data$ideology.fac[data$ideology == 4] <- "Moderate"
data$ideology.fac[data$ideology == 5] <- "Slightly liberal"
data$ideology.fac[data$ideology == 6] <- "Liberal"
data$ideology.fac[data$ideology == 7] <- "Extremely Liberal"

data$crimfair.fac[data$crimfair == 1] <- "Not at all fair"
data$crimfair.fac[data$crimfair == 2] <- "Slightly fair"
data$crimfair.fac[data$crimfair == 3] <- "Somewhat fair"
data$crimfair.fac[data$crimfair == 4] <- "Very fair"
data$crimfair.fac[data$crimfair == 5] <- "Completely fair"

data$fac.party[data$party == 1] <- "Strong Republican"
data$fac.party[data$party == 2] <- "Republican"
data$fac.party[data$party == 3] <- "Independent, closer to Republican"
data$fac.party[data$party == 4] <- "Independent"
data$fac.party[data$party == 5] <- "Independent, closer to Democrat"
data$fac.party[data$party == 6] <- "Democrat"
data$fac.party[data$party == 7] <- "Strong Democrat"

data.plot2 <- data %>%
  select(black.fac, party, ideology, crimfair) %>%
  na.omit()

summary(data.plot2)

ggplot(data.plot2, aes(ideology, crimfair, color = party)) + geom_jitter() + facet_grid(row =
vars(black.fac)) + labs(x= "Ideology", y = "Is the Criminal Justice System Fair?", color = "Party") +
geom_smooth(method = "loess", se = TRUE, color = "black") + scale_color_gradient(low = "darkred",
high = "darkblue") + scale_x_continuous(breaks = c(1:7), labels = c("Extremely Conservatism",
"Conservative", "Slightly Conservative", "Moderate", "Slightly Liberal", "Liberal", "Extremely Liberal"))
+ scale_y_continuous(breaks = c(1:5), labels = c("Not at all fair", "Slightly fair", "Somewhat fair", "Very
fair", "Completely fair"))

#3. Investigate the appropriate bivariate test of the relationship between policetreat and party
#choices: tubular analysis, difference of means, or Pearson's r

cor.test(data$policetreat, data$party)

#a). justify your choice of bivariate test
# I chose to do a Pearson's R (correlation) to see if one's political party affiliation could affect how they see police treatment across races (or if they are even correlated at all).
```

#b). interpret the statistical significance
When it comes to the statistical significance of this bivariate test, the p-value is less than 0.05 (2.2e-16), meaning it is statistically significant and reject the null hypothesis.

#c). interpret the substantive significance
The test explains that the correlation between a respondent's party affiliation and their thoughts on police having racial bias when interacting with citizens is a 0.4508169 (0.5 if rounded up). This makes the relationship a moderatley strong positive correlation, which I would say is somewhat substantive, but not extremely strong.

#4. Investigate the appropriate bivariate test of the relationship between white and forceonwhite as well as white and forceonblack
#first bivariate test: white and forceonwhite

```
t.test(forceonwhite ~ white, data = data)
```

#a). justify your choice of bivariate test
I decided to do both of these bivariate tests as a difference of means because white and black are categorical variables while force on white and force on black could be seen as continuous.

#b). interpret the statistical significance
#The p-value of this test is a 0.000388, which is less than 0.05. This means we can reject the null hypothesis.

#c). interpret the substantive significance
#The mean of group 0 (Non-white respondents) is 2.610304 while the mean of group 1 (white respondents) is 2.759124. This is only a 0.15 difference on the scale for the forceonwhite's scale of 1 to 5, which in my opinion isn't substantive at all.

#second bivariate test: white and forceonblack

```
t.test(forceonblack ~ white, data = data)
```

#a). justify your choice of bivariate test
I decided to do both of these bivariate tests as a difference of means because white and black are categorical variables while force on white and force on black could be seen as continuous.

#b). interpret the statistical significance
The p-value of this given test is less than 0.05 (2.2e-16). This means we can reject the null hypothesis and conclude it's statistically significant.

#c). interpret the substantive significance
The mean of group 0 (Non-white respondents) was a 3.997358 while the mean of group 1 (white respondents) is 3.145773. The distance between the two numbers is around a 0.85 difference for the two groups. I think this is substantivey significant because this is almost a full point different on the 1-to-5 scale.

#5. estimate the regression of forceonblack (Y) on black, suburban and female.

```
model1 <- lm(forceonblack ~ black + suburban + female, data = data)
summary(model1)
```

#a). interpret the α. Is it useful? What kind of respondent does it represent?

The α is 3.05380 on a scale from 1-to-5. This basically represents that those who aren't Black, from rural or urban areas, and are male lie at around 3 on the scale, meaning they overall think the force on Black citizens is "sometimes" excessive force.

#b). interpret the R2 from each regression
#The R2 is 0.2263, meaning there is a 22.6% variance in forceonblack that is explained by all X variables within this model (explained by if the respondent was black, from a suburban area, and female).

#c). interpret the RMSE from each regression
The RMSE is 0.2279. In my opinion, I don't this is super drastic because in most cases it won't raise or lower the scales to be a different number. I think it's just a small error when it's just a 0.23 error on the 1-to-5 scale.

#5 part 2: estimate the regression of forceonwhite (Y) on black, suburban and female

```
model2 <- lm(forceonwhite ~ black + suburban + female, data = data)
summary(model2)
```

#a). interpret the α. Is it useful? What kind of respondent does it represent?
#The α for this regression is 2.82751 on a scale from 1-to-5. Just like the regression above, it represents those that aren't Black, from rural/urban areas, and are male. I think it is important because it showcases how the population we are wanting to understand differs from their opposite. The number 2.8 which is "rarely" on if police use more force than necessary on white citizens.

#b). interpret the R2 from each regression
#The R2 is 009.318, meaning there is a 9.3% variance in forceonwhite (Y) that is explained by all X variables within this model (explained by if the respondent was black, from the suburbs, and female).

#c). interpret the RMSE from each regression
#The RMSE is 0.8065. This is a huge jump in my opinion because it could skew any result by almost 1 full point on the 1-to-5 scale. If a number was originally a 3.2, meaning "sometimes," the error could cause the overall result to say "usually." I think there is a huge difference between someone thinking something happens "sometimes" and "usually"

#6 estimate four more regressions. four dependent variables: forceonblack, forceonwhite, crimfair, and policetreat
#independent variables in all 4: black, party, ideology, suburban, female, income, education, and age

#regression 1: forceonblack (Y)

```
model3 <- lm(forceonblack ~ black + party + ideology + suburban + female + income + education +
age, data = data)
summary(model3)
```

#a). interpret the ˆα. is it substantively useful?
#The ˆα is 2.241875. Personally, I don't see this as useful in any way because there's no way for someone to have no party, 0 ideology, 0 income, 0 education, and be 0 years old.

#b). interpret the R2
#The R2 is 0.2826, meaning there is a 28.3% variance in forceonblack that is explained by all the X variables in the regression.

#c). interpret the RMSE
#The RMSE is a 0.7939 (basically a 0.8). This is a huge range for error seeing how the scale is still a 1-to-5 and can move an answer into the wrong section.

#d). interpret the statistical significance of each X variable
#Out of all the variables, the ones that are seen as statistically significant are black, party, ideology, and age. All of these variables have a p-value less than 0.05.
#The variables over 0.05, meaning they aren't statistically significant, are suburban, female, income, and education.

#e). interpret the ˆ β on black
#The ˆ β of black is 0.711668. This means if the respondent was black, they are responding +0.7 more than the intercept. This basically means being Black makes the responded think there is more unnecessary force from police on Black citizens.

#f). interpret the effect of shifting from “conservative” to “liberal” on ideology
#The effect of shifting from "conservative" to "liberal" on ideology is a 0.178556, which honestly isn't that much of an effect on the 1-to-5 scale. Moving from conservative to liberal is 4 spots apart on the 1-to-7 scale.

#g). interpret the effect of 25-unit increase in age.
#If someone aged by 25 years, it changes the variable's β to 0.14395. Honestly, for 25 years, that is not a huge increase at all on the 1-to-5 scale.

```
summary(data$age)
summary(data$income)
summary(data$education)
```

#h). generate a prediction (Y) for a non-Black, Independent, moderate, suburban, male, with average income, average education, and average age

```
coef(model3)["(Intercept)"] + (coef(model3)["black"] *0) + (coef(model3)["party"] * 4) +
(coef(model3)["ideology"] * 4) + (coef(model3)["suburban"] * 1) + (coef(model3)["female"] * 0) +
(coef(model3)["income"] * 2.619) + (coef(model3)["education"] * 3.34) + (coef(model3)["age"] * 50.03)
```

#The prediction is a 3.037861, which means this type of person falls on "sometimes" when it comes to excessive police force on Black citizens.

#regression 2: forceonwhite (Y)

```
model4 <- lm(forceonwhite ~ black + party + ideology + suburban + female + income + education +
age, data = data)
summary(model4)
```

#a). interpret the ˆα. is it substantively useful?
\# The ˆα is 2.751219, and just like the one above, I don't think it's substantively useful because many of the variables (the same as the ones in the regression above) can't be 0 and make sense.

#b). interpret the R2
#The R2 is 0.01459, meaning there is a 1.5% variance in forceonwhite that is explained by all the X variables in the regression. This is basically inexistent to me.

#c). interpret the RMSE
#The RMSE is 0.8011, which is a lot of room for error on the 1-to-5 scale.

#d). interpret the statistical significance of each X variable
#The variables that are statistically significant in this regression are black, ideology, and education.
#The variables that are more than 0.05 and not statistically significant are party, suburban, female, income, and age.
#Personally, the fact that there are more variables that are insigificant than significant makes me feel like this regression isn't accurately measuring everything.

#e). interpret the ˆ β on black
#The β on black is -0.160773, which means Black respondents are less likely to think that unnesssecary police force is used on white citizens than the intercept.

#f). interpret the effect of shifting from “conservative” to “liberal” on ideology
#Shifting from conservative to liberal on ideology is causes the effect to be a +0.171988, and I don't think that is much of a shift on the 1-to-5 scale.

#g). interpret the effect of 25-unit increase in age.
#A 25-unit increase in age is 0.0472 on the 1-to-5 scale. Personally, I don't think this matters because it isn't statistically significant.

#h). generate a prediction (Y) for a non-Black, Independent, moderate, suburban, male, with average income, average education, and average age

```
coef(model4)["(Intercept)"] + (coef(model4)["black"] *0) + (coef(model4)["party"] * 4) + (coef(model4)["ideology"] * 4) + (coef(model4)["suburban"] * 1) + (coef(model4)["female"] * 0) + (coef(model4)["income"] * 2.619) + (coef(model4)["education"] * 3.34) + (coef(model4)["age"] * 50.03)
```

#The prediction for this specific type of person is a 2.807726, meaning this person is predicted to fall under "rarely" and close to "sometimes" in thinking that white citizens get unnecessary force from police.

#regression 3: crimfair

```
model5 <- lm(crimfair ~ black + party + ideology + suburban + female + income + education + age, data = data)
summary(model5)
```

#a). interpret the ˆα. is it substantively useful?
#The α for this regression is 3.2202654. Just like the regressions above, I don't think it's substantively useful because many of the variables (the same as the ones in the regression above) can't be 0 and make sense.

#b). interpret the R2
#The R2 is 0.1011, meaning there is a 10.11% variance in crimfair that is explained by the X variables.

#c). interpret the RMSE
#The RMSE is 0.9872. I think this is huge room for error because that number, which is basically 1, is about 20% of the 1-to-5 scale. It could really skew results.

#d). interpret the statistical significance of each X variable
#The X variables that had a p-value of less than 0.05 (statistically significant) are black, ideology, and female.
#The X variables that aren't statistically significant are party, suburban, income, education, and age.

#e). interpret the ˆ β on black

#The β on black is -0.4910625. This means that if the respondent is black, they are less likely to think the criminal justice system is fair compared to those that aren't black.

#f). interpret the effect of shifting from “conservative” to “liberal” on ideology
#Shifting from conservative to liberal has a -0.3234644 effect on the crimfair scale. This is the second biggest shift so far out of all regressions.

#g). interpret the effect of 25-unit increase in age.
#The effect of a 25-year increase makes the respondent 0.0092325 less likely to think the criminal justice system is fair. I'm not sure if this is even useful, though, because it wasn't statistically significsnt to begin with.

#h). generate a prediction (Y) for a non-Black, Independent, moderate, suburban, male, with average income, average education, and average age

coef(model5)["(Intercept)"] + (coef(model5)["black"] *0) + (coef(model5)["party"] * 4) + (coef(model5)["ideology"] * 4) + (coef(model5)["suburban"] * 1) + (coef(model5)["female"] * 0) + (coef(model5)["income"] * 2.619) + (coef(model5)["education"] * 3.34) + (coef(model5)["age"] * 50.03)

#This predicted person is a 2.789841 on the crimfair 1-to-5 scale, meaning that they think the criminal justice system is "slightly fair" and it's closely coming towards "somewhat fair."

#regression 4: policetreat

model6 <- lm(policetreat ~ black + party + ideology + suburban + female + income + education + age, data = data)
summary(model6)

#a). interpret the ˆα. is it substantively useful?
#The α for this regression is 3.217014, or "police treat Blacks a little better." Just like the one above, I don't think it's substantively useful because many of the variables can't be 0 and make sense.

#b). interpret the R2
#The R2 is 0.3207 meaning 32.07% of the variance in policetreat are explained by all of the Xs in the model.

#c). interpret the RMSE
#The RMSE is 0.3246 which honeslty isn't that significant on a scale of 1-to-7 where it's a gradual change from police treat Blacks much better to police treat whites much better.

#d). interpret the statistical significance of each X variable
#The X variables that are statistically significant (have a p-value less than 0.05) are black, party, ideology, education, and age.
#The X variables that are not statisticaly significant (p-value of more than 0.05) are suburban, female, and income

#e). interpret the ˆ β on black
#The β on black is 0.964547, meaning if the respondent is Black they move almost a full point on the 1-to-7 scale. This means they are more in the ballpark of "police treat both the same" and "police treat whites a little better."

#f). interpret the effect of shifting from “conservative” to “liberal” on ideology
#The effect of shifting from conservative to liberal on ideology is a 0.43872, which is almost half a point. This is the biggest jump out of all regressions when moving from "conservative" to "liberal"

```r
#g). interpret the effect of 25-unit increase in age.
#A 25-year age jump would cause the β to be a 0.11685. I don't think this is that important because the age variable is not statistically significant.
#It's interesting how each regression has such a low change when it comes to age, somewhat suggesting that age doesn't have the biggest effect of changing opinions.

#h). generate a prediction (Y) for a non-Black, Independent, moderate, suburban, male, with average income, average education, and average age

coef(model6)["(Intercept)"] + (coef(model6)["black"] *0) + (coef(model6)["party"] * 4) +
(coef(model6)["ideology"] * 4) + (coef(model6)["suburban"] * 1) + (coef(model6)["female"] * 0) +
(coef(model6)["income"] * 2.619) + (coef(model6)["education"] * 3.34) + (coef(model6)["age"] * 50.03)

#This prediction is 4.843122 on the policetreat's 1-to-7 scale, which is "police treat both the same," very closely to entering "police treat whites a little better.

#7. third ggplot: barplot on whether the respondent had the correct knowledge about whether Michael Brown had committed a crime or had a weapon.
#facet (in rows) based on whether the respondent knew the name of the officer
#facet (in columns) based on whether the question is about the commission of a crime or whether he had a weapon
#fill by whether the respondent is Black or non-Black
#match the colors, axis labels, and control of the labels of the x-axis to represent the category labels rather than numbers

data$had.weapon.fac <- NA
data$had.weapon.fac[data$hadweapon == 1] <- "Confidently Wrong"
data$had.weapon.fac[data$hadweapon == 2] <- "Hesitantly Wrong"
data$had.weapon.fac[data$hadweapon == 3] <- "Hesitantly Correct"
data$had.weapon.fac[data$hadweapon == 4] <- "Confidently Correct"

data.plot3 <- data %>%
  pivot_longer(cols = c(hadweapon, hadcrime), names_to = "Question", values_to = "Values") %>%
  select(black.fac, nameofficer, Question, Values) %>%
  na.omit()

data.plot3$values.fac <- NA
data.plot3$values.fac[data.plot3$Values == 1] <- "Confidently Wrong"
data.plot3$values.fac[data.plot3$Values == 2] <- "Hesitantly Wrong"
data.plot3$values.fac[data.plot3$Values == 3] <- "Hesitantly Correct"
data.plot3$values.fac[data.plot3$Values == 4] <- "Confidently Correct"

data.plot3$nameofficer.fac <- NA
data.plot3$nameofficer.fac[data.plot3$nameofficer == 0] <- "Did Not Know Officer Name"
data.plot3$nameofficer.fac[data.plot3$nameofficer == 1] <- "Knew Officer Name"

data.plot3$Question.fac <- NA
data.plot3$Question.fac[data.plot3$Question == "hadcrime"] <- "Michael Brown Had Committed a Crime"
data.plot3$Question.fac[data.plot3$Question == "hadweapon"] <- "Michael Brown Had a Weapon"

head(data.plot3)
```

data.plot3$values.fac = factor(data.plot3$values.fac, levels <- c("Confidently Wrong", "Hesitantly Wrong", "Hesitantly Correct", "Confidently Correct"))

ggplot(data.plot3, aes(values.fac, fill = as.factor(black.fac))) + geom_bar(stat = "count", position = "dodge") + facet_grid(row = vars(nameofficer.fac), col = vars(Question.fac)) + scale_fill_manual(values = c("grey10", "grey60")) + labs(x = "Knowledge of Facts of Event", y = "Count", fill = "Race") + theme_bw()

#8. estimate 2 regressions: 2 dependent variables (Y) hadweapon and hadcrime.
#independent variables (X) black, party, ideology, suburban, female, nameofficer, officerindicted, income, education, and age

#regression 1: hadweapon

model7 <- lm(hadweapon ~ black + party + ideology + suburban + female + nameofficer + officerindicted + income + education + age, data = data)
summary(model7)

#a). interpret the ˆα. is it substantively useful?
#The ˆα is a 2.300899 on hadweapon's 1-to-4 scale. I still don't think its substantively useful because a respondent can't have 0 income, 0 education, and be 0 years old.

#b). interpret the R2.
#The R2 is 0.1327, meaning 13.3% of the variance in hadweapon is explained by the X variables in the regression.

#c). interpret the RMSE
#The RMSE is 0.8725, which honestly is a lot. That's almost an entire point on the 1-to-4 scale, meaning it can greatly skew results.

#d). interpret the statistical significance of each X variable.
#The X variables that have a p-value of less than 0.05, proving to be statistically significant, are black, party, and nameofficer.
#The variables that have more that a 0.05 p-value, proving not to be statistically significant, are ideology, suburban, female, officerindicted, income, education, and age.

#e). interpret the ˆ β on black.
#The β of black is 0.414203, meaning if the respondent is black their response is shifted to be more likely to believe that Michael didn't have a weapon, whether it's "probably" or "confidently."

#f) interpret the maximum effect of party.
#The maximum effect of party (the respondent is a strong Democrat) in this regression is 0.364026, which isn't even half a point on the 1-to-4 scale. I wouldn't say it's that important. Seeing how the p-value is statistically significant, it can be concluded that it's reliable to trust.

#g). interpret the maximum effect of ideology.
#The maximum effect of ideology (the respondent is extremely Liberal) is 0.205848, which isn't that super strong effect on the 1-to-4 scale. I'm not sure if the numbers are too trustworthy, though, because the variable isn't statistically significant.

#h). interpret the ˆ β on nameofficer.

#The β on name officer is 0.238886, meaning if the respondent picked the correct name of the officer that shot Michael Brown, they are more likely to guess the respondent did not have weapon, or at least don't confidently believe Brown had a weapon.

#i). which variable is the most substantively significant? justify your answer.
#I would say the respondent beung black has the most subsrantive significance. Even when multiplying the maximum effect of all variables, black being a scale from 0-to-1 had the most effect on if the respondent thought Brown had a weapon.

#regression 2: hadcrime

model8 <- lm(hadcrime ~ black + party + ideology + suburban + female + nameofficer + officerindicted + income + education + age, data = data)
summary(model8)

#a). interpret the ˆα. is it substantively useful?
#The α is 2.892061 meaning when all X variables are 0, the respondent is a 2.89 on the 1-to-4 hadcrime scale. This isn't substantively useful because some X variables don't equal 0 (party, ideology, income, education, and age).

#b). interpret the R2.
#The R2 is 0.1081, meaning 10.81% of the variance in hadcrime is explained by all the X variables within the model.

#c). interpret the RMSE
#The RMSE is 0.9536, and on the 1-to-4 scale, it could greatly skew results by saying someone that was incorrect about Brown not committing a crime but could fall into the "probably correct" category.

#d). interpret the statistical significance of each X variable.
#The X variables in the regression that have a p-value of less than 0.05 making them statistically significant are black, party, ideology, female, nameofficer, income, and age.
#The variables that have a p-value more than 0.05 making them not statistically significant are suburban, officerindicted, and education.
#This regression is the one that has the most statistically significant variables so far.

#e). interpret the ˆ β on black.
#The β on black is -0.320422 which means if the respondent was Black they are more likely to be incorrect and think Brown didn't commit a crime.

#f) interpret the maximum effect of party.
#The maximum effect of party (the respondent is strong Democrat) in this regression is 0.217236 on the 1-to-4 scale. It's definitely a small push, but it's not something crazy significant. I think this is also an accurate reading because the p-value of party is less than 0.05.

#g). interpret the maximum effect of ideology.
#The maximum effect of ideology (the respondent is extremely Liberal) in this regression is -0.27045 on the 1-to-4 scale. Just like the one above, it's a small push toward believing a crime hadn't been committed (which wasn't correct). I think this is also an accurate readinf because the p-value for ideology is less than 0.05.

#h). interpret the ˆ β on nameofficer.
#The β of nameofficer is 0.383642, meaning if that if the respondent chose the correct officer name, it increases the likelihood by 0.4 that the respondent was correct about Brown committing a crime.

#i). which variable is the most substantively significant? justify your answer.
#The variable that's most substantively significant would have to be nameofficer with black following not too far behind. While both are on a 0-to-1 scale, just moving to 1 causes such a big jump compared to the other X variables.

#j). what does it mean for the sign on party to flip across the regressions?
#For the first model, it's saying that many people who are moving up the party scale (becoming more liberal) thought that Brown didn't have a weapon and that he didn't commit a crime.
#It makes sense to me that more liberal respondents would lean towards Brown not having a weapon and not committing a crime, especially in cases where police brutality/murder of a citizen is involved.
#side note: this dataset is super interesting to me, especially because this was before BLM.

#9. estimate two regressions: two dependent variables (Y) invested and racerole
#independent variables (X): black, party, ideology, suburban, female, hadweapon, hadcrime, income, education, and age

#regression 1: invested

model9 <- lm(invested ~ black + party + ideology + suburban + female + hadweapon + hadcrime + income + education + age, data = data)
summary (model9)

#a). interpret the ˆα. is it substantively useful?
#The α of this regression is 0.568602. This isn't even a point on the scale, seeing how it starts at 1. Just like the others, many of these scales don't have a 0 and make no sense being held at 0, like party/ideology, hadweapon, hadcrime, income, education, and age

#b). interpret the R2.
#The R2 is 0.1799, meaning 17.99% (basically 18%) of the variance in invested is explained by all the X variables within the model.

#c). interpret the RMSE.
#The RMSE is 1.086 on a scale of 1-to-5. I think this is an RMSE to be concerned about because it's basically say every observation could be off by one full point. 1.086 is about 21% of the overall scale.

#d). interpret the statistical significance of each X variable.
#The variables that have a p-value less than 0.05, meaning they're statistically significant are black, party, ideology, hadweapon, and income.
#The variables that have a p-value of more than 0.05 are suburban, female, hadcrime, education, and age.

#e). interpret the ˆ β on black.
#The β on black is 0.737239. This means if the respondent is black, they are invested almost one whole point more about the Michael Brown case.

#f) interpret the maximum effect of party.
#The maximum effect party (the respondent being a strong Democrat) can have on the respondent being invested is 0.379878. This is not too big of a jump for the most part in the invested's 1-to-5 scale, especially when black has an effect of 0.7. It is helpful that the p-value for this number is less than 0.05, meaning it is still statistically significant.

#g). interpret the maximum effect of ideology.

#The maximum effect of ideology (the respondent is extremely liberal) is 0.35484 on the 1-to-5 invested scale. This maximum effect is not too far off from party's maximum. Just like party, this variable's p-value is statistically significant, meaning we can confirm there is some relationship.

#h). interpret the maximum effect of hadweapon.
#The maximum effect hadweapon can have on invested is 0.435105, which is the second-highest effect out of all the variables. It's also statistically significant meaning there is a relationship between invested and hadweapon.

#i). interpret the maximum effect of hadcrime.
#The maximum effect of hadcrime on invested is 0.16847. I'm not sure this matters though because the variable isn't statistically significant and we cannot reject the null hypothesis.

#regression 2: racerole

model10 <- lm(racerole ~ black + party + ideology + suburban + female + hadweapon + hadcrime + income + education + age, data = data)
summary(model10)

#side note: racerole scale is 1-to-5

#a). interpret the $\hat{\alpha}$. is it substantively useful?
#The α of this regression is 0.838166. Just like the other, this isn't useful for 2 reason. First, the α is less than 1, which is the beginning of the scale. Second, it's impossible for all of the X variables to be 0 because they don't have a 0 on the scale (or it makes no sense for them to be 0).

#b). interpret the R2.
#The R2 is 0.3411, meaning 34% of the variance in racerole is explained by all the X variables within the model.

#c). interpret the RMSE.
#The RMSE is 1.105 on the 1-to-5 race role scale. I would say this is important, as it can completely skew every result.

#d). interpret the statistical significance of each X variable.
#The X variables in this regression with a p-value of less than 0.05 are black, party, ideology, hadweapon, hadcrime, education, and age.
#The X variables that have a p-value over 0.05, meaning they aren't statistically significant, are suburban, female, and income.

#e). interpret the $\hat{\beta}$ on black.
#The β on black is 0.857568, which is honestly a huge jump for the 1-to-5 racerole scale. It basically means if the respondent is Black they move up the 1-to-5 scale basically saying that race had a part in the Brown case. This is almost a full point aka almost 20% of the scale.

#f) interpret the maximum effect of party.
#The maximum effect of party (the respondent is a strong Democrat) on racerole is 0.68442, which can affect a respondent by more than half a point on the 1-to-5 scale. The more Democratic the respondent is, the more likely they are to think there was a racial bias involved in the Brown case. It's important to note the p-value is less than 0.05 on this variable, meaning we can reject the null hypothesis.

#g). interpret the maximum effect of ideology.
#The maximum effect of ideology (the respondent is extremely liberal) on racerole is 0.744372, which is the second-highest following behind black. This is pretty important because it takes up

more than half a point on the 1-to-5 scale. The p-value is also less than 0.05 meaning the relationship is statistically significant.

#h). interpret the maximum effect of hadweapon.
#The maximum effect of hadweapon (which represents the respondent confidently and correctly guessed the Brown did not have the weapon) on racerole is 0.690858, being the third highest following behind ideology. It's also important to not it is statistically significant because the p-value is less than 0.05 and we can reject the null hypothesis.

#i). interpret the maximum effect of hadcrime.
#The maximum effect of hadcrime (which represents the respondent confidently and correctly guessed Brown committed a crime) on racerole is -0.383901, which is semi-meaningful but it's not a huge increase (not even half a point). The p-value for this variable is statistically significnt meaning we can reject the null hypothesis.

#j). Predictions

#Pred. 1/2: A non-Black, conservative, Republican, suburban, male, with average income,
#high school education, and average age, who is confidently wrong that
#Michael Brown had weapon, and confidently wrong that Michael Brown had not committed a crime.
#Prediction w/ model 9 (invested)
coef(model9)["(Intercept)"] + (coef(model9)["black"] * 0) + (coef(model9)["ideology"] * 2) + (coef(model9)["party"] * 2) + (coef(model9)["suburban"] * 1) + (coef(model9)["female"] * 0) + (coef(model9)["income"] * 2.619) + (coef(model9)["education"] * 2) + (coef(model9)["age"] * 50.03) + (coef(model9)["hadweapon"] * 1) + (coef(model9)["hadcrime"] * 1)
#1.457446

#Prediction w model 10 (racerole)
coef(model10)["(Intercept)"] + (coef(model10)["black"] * 0) + (coef(model10)["ideology"] * 2) + (coef(model10)["party"] * 2) + (coef(model10)["suburban"] * 1) + (coef(model10)["female"] * 0) + (coef(model10)["income"] * 2.619) + (coef(model10)["education"] * 2) + (coef(model10)["age"] * 50.03) + (coef(model10)["hadweapon"] * 1) + (coef(model10)["hadcrime"] * 1)
#1.949055

#Pred. 3/4: A non-Black, conservative, Republican, suburban, male, with average income,
#high school education, and average age, who is confidently correct that Michael Brown did not have weapon,
#and confidently correct that Michael Brown had committed a crime.
#Prediction w/ model 9 (invested)
coef(model9)["(Intercept)"] + (coef(model9)["black"] * 0) + (coef(model9)["ideology"] * 2) + (coef(model9)["party"] * 2) + (coef(model9)["suburban"] * 1) + (coef(model9)["female"] * 0) + (coef(model9)["income"] * 2.619) + (coef(model9)["education"] * 2) + (coef(model9)["age"] * 50.03) + (coef(model9)["hadweapon"] * 4) + (coef(model9)["hadcrime"] * 4)
#2.009398

#Prediction w/ model 10 (racerole)
coef(model10)["(Intercept)"] + (coef(model10)["black"] * 0) + (coef(model10)["ideology"] * 2) + (coef(model10)["party"] * 2) + (coef(model10)["suburban"] * 1) + (coef(model10)["female"] * 0) + (coef(model10)["income"] * 2.619) + (coef(model10)["education"] * 2) + (coef(model10)["age"] * 50.03) + (coef(model10)["hadweapon"] * 4) + (coef(model10)["hadcrime"] * 4)
#2.256013

#Pred. 5/6: A non-Black, moderate, Independent, suburban, male, with average income,
#high school education, and average age, who is confidently wrong that Michael

```
#Brown had weapon, and confidently wrong that Michael Brown had not committed a crime.
#Prediction w/ model 9 (invested)
coef(model9)["(Intercept)"] + (coef(model9)["black"] * 0) + (coef(model9)["ideology"] * 4) +
(coef(model9)["party"] * 4) + (coef(model9)["suburban"] * 1) + (coef(model9)["female"] * 0) +
(coef(model9)["income"] * 2.619) + (coef(model9)["education"] * 2) + (coef(model9)["age"] * 50.03) +
(coef(model9)["hadweapon"] * 1) + (coef(model9)["hadcrime"] * 1)
#1.702352

#Prediction w/ model 10 (racerole)
coef(model10)["(Intercept)"] + (coef(model10)["black"] * 0) + (coef(model10)["ideology"] * 4) +
(coef(model10)["party"] * 4) + (coef(model10)["suburban"] * 1) + (coef(model10)["female"] * 0) +
(coef(model10)["income"] * 2.619) + (coef(model10)["education"] * 2) + (coef(model10)["age"] * 50.03) +
(coef(model10)["hadweapon"] * 1) + (coef(model10)["hadcrime"] * 1)
#2.425317

#Pred. 7/8:A non-Black, moderate, Independent, suburban, male, with average income,
#high school education, and average age, who is confidently correct that Michael
#Brown did not have weapon, and confidently correct that Michael Brown had committed a crime.
#Prediction w/ model 9 (invested)
coef(model9)["(Intercept)"] + (coef(model9)["black"] * 0) + (coef(model9)["ideology"] * 4) +
(coef(model9)["party"] * 4) + (coef(model9)["suburban"] * 1) + (coef(model9)["female"] * 0) +
(coef(model9)["income"] * 2.619) + (coef(model9)["education"] * 2) + (coef(model9)["age"] * 50.03) +
(coef(model9)["hadweapon"] * 4) + (coef(model9)["hadcrime"] * 4)
#2.254304

#Prediction w/ model 10 (racerole)
coef(model10)["(Intercept)"] + (coef(model10)["black"] * 0) + (coef(model10)["ideology"] * 4) +
(coef(model10)["party"] * 4) + (coef(model10)["suburban"] * 1) + (coef(model10)["female"] * 0) +
(coef(model10)["income"] * 2.619) + (coef(model10)["education"] * 2) + (coef(model10)["age"] * 50.03) +
(coef(model10)["hadweapon"] * 4) + (coef(model10)["hadcrime"] * 4)
#2.732276

#Pred. 9/10: A non-Black, liberal, Democrat, suburban, male, with average income, high
#school education, and average age, who is confidently wrong that Michael
#Brown had weapon, and confidently wrong that Michael Brown had not committed a crime
#Prediction w/ model 9 (invested)
coef(model9)["(Intercept)"] + (coef(model9)["black"] * 0) + (coef(model9)["ideology"] * 6) +
(coef(model9)["party"] * 6) + (coef(model9)["suburban"] * 1) + (coef(model9)["female"] * 0) +
(coef(model9)["income"] * 2.619) + (coef(model9)["education"] * 2) + (coef(model9)["age"] * 50.03) +
(coef(model9)["hadweapon"] * 1) + (coef(model9)["hadcrime"] * 1)
#1.947258

#Prediction w/ model 10 (racerole)
coef(model10)["(Intercept)"] + (coef(model10)["black"] * 0) + (coef(model10)["ideology"] * 6) +
(coef(model10)["party"] * 6) + (coef(model10)["suburban"] * 1) + (coef(model10)["female"] * 0) +
(coef(model10)["income"] * 2.619) + (coef(model10)["education"] * 2) + (coef(model10)["age"] * 50.03) +
(coef(model10)["hadweapon"] * 1) + (coef(model10)["hadcrime"] * 1)
#2.90158

#Pred. 11/12: A non-Black, liberal, Democrat, suburban, male, with average income, high
#school education, and average age, who is confidently correct that Michael
#Brown did not have weapon, and confidently correct that Michael Brown had committed a crime.
#Prediction w/ model 9 (invested)
coef(model9)["(Intercept)"] + (coef(model9)["black"] * 0) + (coef(model9)["ideology"] * 6) +
(coef(model9)["party"] * 6) + (coef(model9)["suburban"] * 1) + (coef(model9)["female"] * 0) +
```

(coef(model9)["income"] * 2.619) + (coef(model9)["education"] * 2) + (coef(model9)["age"] * 50.03) + (coef(model9)["hadweapon"] * 4) + (coef(model9)["hadcrime"] * 4)
#2.49921

#Prediction w/ model 10 (racerole)
coef(model10)["(Intercept)"] + (coef(model10)["black"] * 0) + (coef(model10)["ideology"] * 6) + (coef(model10)["party"] * 6) + (coef(model10)["suburban"] * 1) + (coef(model10)["female"] * 0) + (coef(model10)["income"] * 2.619) + (coef(model10)["education"] * 2) + (coef(model10)["age"] * 50.03) + (coef(model10)["hadweapon"] * 4) + (coef(model10)["hadcrime"] * 4)
#3.208539

#10. look across all of the regressions. which dependent variable Y do we explain best? justify your answer.
#There are three regression that stands out the most as explaining the dependent variable the best is model 10 (racerole). It has 10 X variables and 7 were proven to be statistically significant, meaning there is an indication that all 7 have a relationship with the intercept (racerole). The overall p-value is also less than 0.05, meaning it is also statistically significant.
#The RMSE is pretty high on this one, but none of the models have a super low RMSE, and no other model (besides model 8), have 7 out of 10 variables correct. The more variables you have to describe the relationships, the more detailed the observation could be.

#11. where does party have the biggest effect, accounting for the scale of Y
#Party has the biggest effect on Y in model 10 (racerole) as it takes up almost a full point on a 1-to-5 scale. The only one i saw as a contender is model 6 (policetreat) as the maximum effect was higher, but the scale for this one is 1-to-7 which means it doesn't really cover as much grounds.

#12. do not estimate a model, but explain what it would mean to include an interaction between party and age. What would this interaction indicate?
#If I were to include and interaction between party and age in a regression, this basically means I'm trying to see if the relationship between age and party together could cause a difference in the outcome of both variables (whether positive or negative). It's basically seeing how the value of one X would change the other if it was different on the scale.

R-Console

```
> library(tidyverse)
> #load in dataset
> data = read.csv(file.choose())
> data$black.fac <- NA
> data$white.force.fac <- NA
> data$black.force.fac <- NA
>
> data$black.fac[data$black == 0] <- "Non-Black"
> data$black.fac[data$black == 1] <- "Black"
>
> data$white.force.fac[data$forceonwhite == 1] <- "Never"
> data$white.force.fac[data$forceonwhite == 2] <- "Rarely"
> data$white.force.fac[data$forceonwhite == 3] <- "Sometimes"
> data$white.force.fac[data$forceonwhite == 4] <- "Usually"
> data$white.force.fac[data$forceonwhite == 5] <- "Always"
>
> data$black.force.fac[data$forceonblack == 1] <- "Never"
> data$black.force.fac[data$forceonblack == 2] <- "Rarely"
> data$black.force.fac[data$forceonblack == 3] <- "Sometimes"
```

```
> data$black.force.fac[data$forceonblack == 4] <- "Usually"
> data$black.force.fac[data$forceonblack == 5] <- "Always"
>
> data$black.force.fac = factor(data$black.force.fac, levels = c("Never", "Rarely", "Sometimes",
"Usually", "Always"))
> data$white.force.fac = factor(data$white.force.fac, levels = c("Never", "Rarely", "Sometimes",
"Usually", "Always"))
>
> group.names <- c(black.force.fac = "Force Used On Black", white.force.fac = "Force Used On
Whites")
> data.plot <- data %>%
+   pivot_longer(cols = c(white.force.fac, black.force.fac), names_to = "Group", values_to =
"Excessive") %>%
+   select(black.fac, Excessive, Group) %>%
+   na.omit()
>
> summary(data.plot)
  black.fac            Excessive       Group
 Length:2885        Never    : 115   Length:2885
 Class :character   Rarely   : 601   Class :character
 Mode  :character   Sometimes:1234   Mode  :character
                    Usually  : 644
                    Always   : 291
>
> length(data$black.fac)
[1] 1565
> summary(data$black)
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.    NA's
 0.0000  0.0000  0.0000  0.4994  1.0000  1.0000       1
> length(data$black)
[1] 1565
> ggplot(data.plot, aes(x = Excessive, fill = as.factor(black.fac))) + geom_bar(stat = "count", position =
"dodge") + scale_fill_manual(values = c("grey10", "grey60")) + facet_grid(rows = vars(Group), labeller
= labeller(Group = group.names)) + labs(x = "Is Police Force Used Excessive?", y = "Count", fill =
"Race") + theme_bw()
> data$ideology.fac <- NA
> data$crimfair.fac <- NA
> data$fac.party <- NA
>
> data$ideology.fac[data$ideology == 1] <- "Extremely Conservative"
> data$ideology.fac[data$ideology == 2] <- "Conservative"
> data$ideology.fac[data$ideology == 3] <- "Slightly Conservative"
> data$ideology.fac[data$ideology == 4] <- "Moderate"
> data$ideology.fac[data$ideology == 5] <- "Slightly liberal"
> data$ideology.fac[data$ideology == 6] <- "Liberal"
> data$ideology.fac[data$ideology == 7] <- "Extremely Liberal"
>
> data$crimfair.fac[data$crimfair == 1] <- "Not at all fair"
> data$crimfair.fac[data$crimfair == 2] <- "Slightly fair"
> data$crimfair.fac[data$crimfair == 3] <- "Somewhat fair"
> data$crimfair.fac[data$crimfair == 4] <- "Very fair"
> data$crimfair.fac[data$crimfair == 5] <- "Completely fair"
>
> data$fac.party[data$party == 1] <- "Strong Republican"
> data$fac.party[data$party == 2] <- "Republican"
```

```
> data$fac.party[data$party == 3] <- "Independent, closer to Republican"
> data$fac.party[data$party == 4] <- "Independent"
> data$fac.party[data$party == 5] <- "Independent, closer to Democrat"
> data$fac.party[data$party == 6] <- "Democrat"
> data$fac.party[data$party == 7] <- "Strong Democrat"
> data.plot2 <- data %>%
+   select(black.fac, party, ideology, crimfair) %>%
+   na.omit()
>
> summary(data.plot2)
  black.fac            party          ideology      crimfair
 Length:1443        Min.   :1.000   Min.   :1.0   Min.   :1.000
 Class :character   1st Qu.:4.000   1st Qu.:3.0   1st Qu.:1.000
 Mode  :character   Median :6.000   Median :4.0   Median :2.000
                    Mean   :5.072   Mean   :4.1   Mean   :2.376
                    3rd Qu.:7.000   3rd Qu.:5.0   3rd Qu.:3.000
                    Max.   :7.000   Max.   :7.0   Max.   :5.000
> ggplot(data.plot2, aes(ideology, crimfair, color = party)) + geom_jitter() + facet_grid(row =
vars(black.fac)) + labs(x= "Ideology", y = "Is the Criminal Justice System Fair?", color = "Party") +
geom_smooth(method = "loess", se = TRUE, color = "black") + scale_color_gradient(low = "darkred",
high = "darkblue") + scale_x_continuous(breaks = c(1:7), labels = c("Extremely Conservatism",
"Conservative", "Slightly Conservative", "Moderate", "Slightly Liberal", "Liberal", "Extremely Liberal"))
+ scale_y_continuous(breaks = c(1:5), labels = c("Not at all fair", "Slightly fair", "Somewhat fair", "Very
fair", "Completely fair"))
`geom_smooth()` using formula = 'y ~ x'
Warning messages:
1: In simpleLoess(y, x, w, span, degree = degree, parametric = parametric,  :
  pseudoinverse used at 5
2: In simpleLoess(y, x, w, span, degree = degree, parametric = parametric,  :
  neighborhood radius 1
3: In simpleLoess(y, x, w, span, degree = degree, parametric = parametric,  :
  reciprocal condition number  0
4: In predLoess(object$y, object$x, newx = if (is.null(newdata)) object$x else if
(is.data.frame(newdata)) as.matrix(model.frame(delete.response(terms(object)),  :
  pseudoinverse used at 5
5: In predLoess(object$y, object$x, newx = if (is.null(newdata)) object$x else if
(is.data.frame(newdata)) as.matrix(model.frame(delete.response(terms(object)),  :
  neighborhood radius 1
6: In predLoess(object$y, object$x, newx = if (is.null(newdata)) object$x else if
(is.data.frame(newdata)) as.matrix(model.frame(delete.response(terms(object)),  :
  reciprocal condition number  0
> cor.test(data$policetreat, data$party)

	Pearson's product-moment correlation

data:  data$policetreat and data$party
t = 19.152, df = 1438, p-value < 2.2e-16
alternative hypothesis: true correlation is not equal to 0
95 percent confidence interval:
 0.4086768 0.4910390
sample estimates:
      cor
0.4508169

> t.test(forceonwhite ~ white, data = data)
```

```
	Welch Two Sample t-test

data:  forceonwhite by white
t = -3.557, df = 1361.5, p-value = 0.000388
alternative hypothesis: true difference in means between group 0 and group 1 is not equal to 0
95 percent confidence interval:
 -0.23089581 -0.06674471
sample estimates:
mean in group 0 mean in group 1 
       2.610304        2.759124 

> t.test(forceonblack ~ white, data = data)

	Welch Two Sample t-test

data:  forceonblack by white
t = 19.356, df = 1438.5, p-value < 2.2e-16
alternative hypothesis: true difference in means between group 0 and group 1 is not equal to 0
95 percent confidence interval:
 0.7652816 0.9378892
sample estimates:
mean in group 0 mean in group 1 
       3.997358        3.145773 

> model1 <- lm(forceonblack ~ black + suburban + female, data = data)
> summary(model1)

Call:
lm(formula = forceonblack ~ black + suburban + female, data = data)

Residuals:
     Min       1Q   Median       3Q      Max 
-3.04529 -0.16439 -0.04529  0.83561  1.98398 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept)  3.05380    0.05283  57.803   <2e-16 ***
black        0.88090    0.04424  19.911   <2e-16 ***
suburban    -0.03778    0.04412  -0.856   0.3919    
female       0.11059    0.04958   2.231   0.0259 *  
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.8257 on 1436 degrees of freedom
  (125 observations deleted due to missingness)
Multiple R-squared:  0.2279,	Adjusted R-squared:  0.2263 
F-statistic: 141.3 on 3 and 1436 DF,  p-value: < 2.2e-16

> model2 <- lm(forceonwhite ~ black + suburban + female, data = data)
> summary(model2)

Call:
lm(formula = forceonwhite ~ black + suburban + female, data = data)
```

```
Residuals:
    Min      1Q  Median      3Q     Max 
-1.8275 -0.6033  0.2338  0.3967  2.4251 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept)  2.82751    0.05160  54.794  < 2e-16 ***
black       -0.16291    0.04321  -3.770  0.00017 ***
suburban    -0.02837    0.04309  -0.658  0.51041    
female      -0.06132    0.04842  -1.266  0.20558    
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.8065 on 1436 degrees of freedom
  (125 observations deleted due to missingness)
Multiple R-squared:  0.01138,	Adjusted R-squared:  0.009318 
F-statistic: 5.511 on 3 and 1436 DF,  p-value: 0.0009187

> model3 <- lm(forceonblack ~ black + party + ideology + suburban + female + income + education + age, data = data)
> summary(model3)

Call:
lm(formula = forceonblack ~ black + party + ideology + suburban + 
    female + income + education + age, data = data)

Residuals:
    Min      1Q  Median      3Q     Max 
-3.1733 -0.3279 -0.0006  0.5900  2.3238 

Coefficients:
             Estimate Std. Error t value Pr(>|t|)    
(Intercept)  2.241875   0.139916  16.023  < 2e-16 ***
black        0.711668   0.051009  13.952  < 2e-16 ***
party        0.086465   0.013776   6.277  4.6e-10 ***
ideology     0.044639   0.016810   2.656 0.008007 ** 
suburban    -0.024877   0.043579  -0.571 0.568192    
female       0.082840   0.048483   1.709 0.087735 .  
income      -0.022161   0.017229  -1.286 0.198579    
education    0.019878   0.022944   0.866 0.386433    
age          0.005758   0.001571   3.665 0.000257 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.7939 on 1409 degrees of freedom
  (147 observations deleted due to missingness)
Multiple R-squared:  0.2866,	Adjusted R-squared:  0.2826 
F-statistic: 70.76 on 8 and 1409 DF,  p-value: < 2.2e-16

> coef(model3)["(Intercept)"] + (coef(model3)["black"] *0) + (coef(model3)["party"] * 4) + (coef(model3)["ideology"] * 4) + (coef(model3)["suburban"] * 1) + (coef(model3)["female"] * 0) + (coef(model3)["income"] * 2.619) + (coef(model3)["education"] * 3.34) + (coef(model3)["age"] * 50.03)
(Intercept) 
   3.037861 
```

```
> model4 <- lm(forceonwhite ~ black + party + ideology + suburban + female + income + education +
age, data = data)
> summary(model4)

Call:
lm(formula = forceonwhite ~ black + party + ideology + suburban +
    female + income + education + age, data = data)

Residuals:
    Min      1Q  Median      3Q     Max
-1.9156 -0.6257  0.1948  0.3601  2.5681

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept)  2.751219   0.141180  19.487  < 2e-16 ***
black       -0.160773   0.051469  -3.124  0.00182 **
party       -0.007879   0.013900  -0.567  0.57092
ideology     0.042997   0.016962   2.535  0.01135 *
suburban    -0.025344   0.043972  -0.576  0.56447
female      -0.063566   0.048921  -1.299  0.19403
income       0.002571   0.017385   0.148  0.88246
education   -0.047843   0.023152  -2.067  0.03896 *
age          0.001888   0.001585   1.191  0.23399
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.8011 on 1409 degrees of freedom
  (147 observations deleted due to missingness)
Multiple R-squared:  0.02015,	Adjusted R-squared:  0.01459
F-statistic: 3.622 on 8 and 1409 DF,  p-value: 0.0003518

> coef(model4)["(Intercept)"] + (coef(model4)["black"] *0) + (coef(model4)["party"] * 4) +
(coef(model4)["ideology"] * 4) + (coef(model4)["suburban"] * 1) + (coef(model4)["female"] * 0) +
(coef(model4)["income"] * 2.619) + (coef(model4)["education"] * 3.34) + (coef(model4)["age"] * 50.03)
(Intercept)
   2.807726
> model5 <- lm(crimfair ~ black + party + ideology + suburban + female + income + education + age,
data = data)
> summary(model5)

Call:
lm(formula = crimfair ~ black + party + ideology + suburban +
    female + income + education + age, data = data)

Residuals:
    Min      1Q  Median      3Q     Max
-2.1968 -0.8765  0.0464  0.7772  3.1648

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)  3.2202654  0.1739301  18.515  < 2e-16 ***
black       -0.4910625  0.0632921  -7.759 1.64e-14 ***
party       -0.0154598  0.0171161  -0.903 0.366558
ideology    -0.0808661  0.0208840  -3.872 0.000113 ***
suburban     0.0249434  0.0541315   0.461 0.645019
```

```
female      -0.1444507  0.0602740  -2.397 0.016679 *
income       0.0372210  0.0213962   1.740 0.082145 .
education   -0.0446320  0.0285157  -1.565 0.117766
age         -0.0003693  0.0019500  -0.189 0.849829
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.9872 on 1412 degrees of freedom
  (144 observations deleted due to missingness)
Multiple R-squared:  0.1061,Adjusted R-squared:  0.1011
F-statistic: 20.96 on 8 and 1412 DF,  p-value: < 2.2e-16

> coef(model5)["(Intercept)"] + (coef(model5)["black"] *0) + (coef(model5)["party"] * 4) +
(coef(model5)["ideology"] * 4) + (coef(model5)["suburban"] * 1) + (coef(model5)["female"] * 0) +
(coef(model5)["income"] * 2.619) + (coef(model5)["education"] * 3.34) + (coef(model5)["age"] * 50.03)
(Intercept)
   2.789841
> model6 <- lm(policetreat ~ black + party + ideology + suburban + female + income + education +
age, data = data)
> summary(model6)

Call:
lm(formula = policetreat ~ black + party + ideology + suburban +
    female + income + education + age, data = data)

Residuals:
    Min      1Q  Median      3Q     Max
-5.2414 -0.6598  0.0904  0.7497  2.8980

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept)  3.217014   0.185410  17.351  < 2e-16 ***
black        0.964547   0.067594  14.270  < 2e-16 ***
party        0.128426   0.018255   7.035 3.10e-12 ***
ideology     0.109680   0.022275   4.924 9.49e-07 ***
suburban    -0.006498   0.057748  -0.113 0.910430
female       0.068453   0.064247   1.065 0.286846
income       0.022090   0.022832   0.968 0.333454
education    0.116309   0.030405   3.825 0.000136 ***
age          0.004674   0.002082   2.245 0.024924 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1.052 on 1409 degrees of freedom
  (147 observations deleted due to missingness)
Multiple R-squared:  0.3246,	Adjusted R-squared:  0.3207
F-statistic: 84.63 on 8 and 1409 DF,  p-value: < 2.2e-16

> coef(model6)["(Intercept)"] + (coef(model6)["black"] *0) + (coef(model6)["party"] * 4) +
(coef(model6)["ideology"] * 4) + (coef(model6)["suburban"] * 1) + (coef(model6)["female"] * 0) +
(coef(model6)["income"] * 2.619) + (coef(model6)["education"] * 3.34) + (coef(model6)["age"] * 50.03)
(Intercept)
   4.843122
> data$had.weapon.fac <- NA
> data$had.weapon.fac[data$hadweapon == 1] <- "Confidently Wrong"
```

```
> data$had.weapon.fac[data$hadweapon == 2] <- "Hesitantly Wrong"
> data$had.weapon.fac[data$hadweapon == 3] <- "Hesitantly Correct"
> data$had.weapon.fac[data$hadweapon == 4] <- "Confidently Correct"
>
> data.plot3 <- data %>%
+   pivot_longer(cols = c(hadweapon, hadcrime), names_to = "Question", values_to = "Values") %>%
+   select(black.fac, nameofficer, Question, Values) %>%
+   na.omit()
>
> data.plot3$values.fac <- NA
> data.plot3$values.fac[data.plot3$Values == 1] <- "Confidently Wrong"
> data.plot3$values.fac[data.plot3$Values == 2] <- "Hesitantly Wrong"
> data.plot3$values.fac[data.plot3$Values == 3] <- "Hesitantly Correct"
> data.plot3$values.fac[data.plot3$Values == 4] <- "Confidently Correct"
>
> data.plot3$nameofficer.fac <- NA
> data.plot3$nameofficer.fac[data.plot3$nameofficer == 0] <- "Did Not Know Officer Name"
> data.plot3$nameofficer.fac[data.plot3$nameofficer == 1] <- "Knew Officer Name"
>
> data.plot3$Question.fac <- NA
> data.plot3$Question.fac[data.plot3$Question == "hadcrime"] <- "Michael Brown Had Committed a
Crime"
> data.plot3$Question.fac[data.plot3$Question == "hadweapon"] <- "Michael Brown Had a Weapon"
> head(data.plot3)
# A tibble: 6 × 7
  black.fac nameofficer Question  Values values.fac          nameofficer.fac   Question.fac
  <chr>           <int> <chr>      <int> <chr>               <chr>             <chr>
1 Non-Black           1 hadweapon      4 Confidently Correct Knew Officer Name Michael Brown Had a
Weapon
2 Non-Black           1 hadcrime       4 Confidently Correct Knew Officer Name Michael Brown Had
Committed a Crime
3 Black               1 hadweapon      1 Confidently Wrong   Knew Officer Name Michael Brown Had a
Weapon
4 Black               1 hadcrime       3 Hesitantly Correct  Knew Officer Name Michael Brown Had
Committed a Crime
5 Non-Black           1 hadweapon      1 Confidently Wrong   Knew Officer Name Michael Brown Had a
Weapon
6 Non-Black           1 hadcrime       4 Confidently Correct Knew Officer Name Michael Brown Had
Committed a Crime
> data.plot3$values.fac = factor(data.plot3$values.fac, levels <- c("Confidently Wrong", "Hesitantly
Wrong", "Hesitantly Correct", "Confidently Correct"))
> ggplot(data.plot3, aes(values.fac, fill = as.factor(black.fac))) + geom_bar(stat = "count", position =
"dodge") + facet_grid(row = vars(nameofficer.fac), col = vars(Question.fac)) +
scale_fill_manual(values = c("grey10", "grey60")) + labs(x = "Knowledge of Facts of Event", y =
"Count", fill = "Race") + theme_bw()
> model7 <- lm(hadweapon ~ black + party + ideology + suburban + female + nameofficer +
officerindicted + income + education + age, data = data)
> summary(model7)

Call:
lm(formula = hadweapon ~ black + party + ideology + suburban +
    female + nameofficer + officerindicted + income + education +
    age, data = data)

Residuals:
```

```
    Min      1Q  Median      3Q     Max 
-2.7002 -0.5256  0.2741  0.5894  1.6258 

Coefficients:
                 Estimate Std. Error t value Pr(>|t|)    
(Intercept)      2.300899   0.162098  14.195  < 2e-16 ***
black            0.414203   0.056703   7.305 4.68e-13 ***
party            0.060671   0.015257   3.976 7.36e-05 ***
ideology         0.034308   0.018640   1.841   0.0659 .  
suburban        -0.084711   0.048360  -1.752   0.0801 .  
female          -0.019689   0.053676  -0.367   0.7138    
nameofficer      0.238886   0.049407   4.835 1.48e-06 ***
officerindicted  0.075395   0.073750   1.022   0.3068    
income           0.026237   0.019190   1.367   0.1718    
education        0.025006   0.025653   0.975   0.3298    
age             -0.002104   0.001752  -1.201   0.2299    
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.8725 on 1382 degrees of freedom
  (172 observations deleted due to missingness)
Multiple R-squared:  0.139,	Adjusted R-squared:  0.1327 
F-statistic:  22.3 on 10 and 1382 DF,  p-value: < 2.2e-16

> model8 <- lm(hadcrime ~ black + party + ideology + suburban + female + nameofficer + officerindicted + income + education + age, data = data)
> summary(model8)

Call:
lm(formula = hadcrime ~ black + party + ideology + suburban + 
    female + nameofficer + officerindicted + income + education + 
    age, data = data)

Residuals:
    Min      1Q  Median      3Q     Max 
-2.3340 -0.6886  0.1735  0.6968  1.9322 

Coefficients:
                 Estimate Std. Error t value Pr(>|t|)    
(Intercept)      2.892061   0.177163  16.324  < 2e-16 ***
black           -0.320422   0.061983  -5.169 2.69e-07 ***
party           -0.036206   0.016676  -2.171   0.0301 *  
ideology        -0.045075   0.020372  -2.213   0.0271 *  
suburban        -0.012223   0.052881  -0.231   0.8172    
female          -0.147140   0.058669  -2.508   0.0123 *  
nameofficer      0.383642   0.053988   7.106 1.91e-12 ***
officerindicted  0.029879   0.080811   0.370   0.7116    
income           0.050460   0.020973   2.406   0.0163 *  
education       -0.011545   0.028050  -0.412   0.6807    
age              0.003949   0.001915   2.062   0.0394 *  
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.9536 on 1382 degrees of freedom
  (172 observations deleted due to missingness)
```

```
Multiple R-squared:  0.1145, Adjusted R-squared:  0.1081
F-statistic: 17.88 on 10 and 1382 DF,  p-value: < 2.2e-16

> model9 <- lm(invested ~ black + party + ideology + suburban + female + hadweapon + hadcrime +
income + education + age, data = data)
> summary (model9)

Call:
lm(formula = invested ~ black + party + ideology + suburban +
    female + hadweapon + hadcrime + income + education + age,
    data = data)

Residuals:
    Min      1Q  Median      3Q     Max
-2.2964 -0.8683 -0.0325  0.7958  3.4146

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept)  0.568602   0.234885   2.421 0.015615 *
black        0.737239   0.071877  10.257  < 2e-16 ***
party        0.063313   0.019016   3.329 0.000893 ***
ideology     0.059140   0.023163   2.553 0.010781 *
suburban     0.043569   0.060080   0.725 0.468460
female      -0.051881   0.066756  -0.777 0.437184
hadweapon    0.145035   0.033671   4.307 1.77e-05 ***
hadcrime     0.038949   0.030546   1.275 0.202485
income       0.084321   0.023898   3.528 0.000432 ***
education    0.016108   0.031633   0.509 0.610674
age          0.003265   0.002165   1.508 0.131783
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1.086 on 1390 degrees of freedom
  (164 observations deleted due to missingness)
Multiple R-squared:  0.1857,	Adjusted R-squared:  0.1799
F-statistic: 31.71 on 10 and 1390 DF,  p-value: < 2.2e-16

> model10 <- lm(racerole ~ black + party + ideology + suburban + female + hadweapon + hadcrime +
income + education + age, data = data)
> summary(model10)

Call:
lm(formula = racerole ~ black + party + ideology + suburban +
    female + hadweapon + hadcrime + income + education + age,
    data = data)

Residuals:
    Min      1Q  Median      3Q     Max
-3.5687 -0.7939  0.0722  0.7912  3.0890

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept)  0.838166   0.261023   3.211 0.001359 **
black        0.857568   0.081455  10.528  < 2e-16 ***
party        0.114070   0.021247   5.369 9.57e-08 ***
```

```
ideology     0.124062   0.025885   4.793 1.86e-06 ***
suburban     0.011966   0.067040   0.178 0.858367
female       0.026905   0.074701   0.360 0.718788
hadweapon    0.230286   0.037926   6.072 1.71e-09 ***
hadcrime    -0.127967   0.034246  -3.737 0.000196 ***
income       0.006265   0.026357   0.238 0.812166
education    0.070877   0.035236   2.012 0.044503 *
age          0.007239   0.002394   3.024 0.002549 **
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 1.105 on 1160 degrees of freedom
  (394 observations deleted due to missingness)
Multiple R-squared:  0.3467,	Adjusted R-squared:  0.3411
F-statistic: 61.56 on 10 and 1160 DF,  p-value: < 2.2e-16

> coef(model9)["(Intercept)"] + (coef(model9)["black"] * 0) + (coef(model9)["ideology"] * 2) +
(coef(model9)["party"] * 2) + (coef(model9)["suburban"] * 1) + (coef(model9)["female"] * 0) +
(coef(model9)["income"] * 2.619) + (coef(model9)["education"] * 2) + (coef(model9)["age"] * 50.03) +
(coef(model9)["hadweapon"] * 1) + (coef(model9)["hadcrime"] * 1)
(Intercept)
   1.457446
> #1.457446
> coef(model10)["(Intercept)"] + (coef(model10)["black"] * 0) + (coef(model10)["ideology"] * 2) +
(coef(model10)["party"] * 2) + (coef(model10)["suburban"] * 1) + (coef(model10)["female"] * 0) +
(coef(model10)["income"] * 2.619) + (coef(model10)["education"] * 2) + (coef(model10)["age"] * 50.03) +
(coef(model10)["hadweapon"] * 1) + (coef(model10)["hadcrime"] * 1)
(Intercept)
   1.949055
> #1.949055
> coef(model9)["(Intercept)"] + (coef(model9)["black"] * 0) + (coef(model9)["ideology"] * 2) +
(coef(model9)["party"] * 2) + (coef(model9)["suburban"] * 1) + (coef(model9)["female"] * 0) +
(coef(model9)["income"] * 2.619) + (coef(model9)["education"] * 2) + (coef(model9)["age"] * 50.03) +
(coef(model9)["hadweapon"] * 4) + (coef(model9)["hadcrime"] * 4)
(Intercept)
   2.009398
> #2.009398
> coef(model10)["(Intercept)"] + (coef(model10)["black"] * 0) + (coef(model10)["ideology"] * 2) +
(coef(model10)["party"] * 2) + (coef(model10)["suburban"] * 1) + (coef(model10)["female"] * 0) +
(coef(model10)["income"] * 2.619) + (coef(model10)["education"] * 2) + (coef(model10)["age"] * 50.03) +
(coef(model10)["hadweapon"] * 4) + (coef(model10)["hadcrime"] * 4)
(Intercept)
   2.256013
> #2.256013
> coef(model9)["(Intercept)"] + (coef(model9)["black"] * 0) + (coef(model9)["ideology"] * 4) +
(coef(model9)["party"] * 4) + (coef(model9)["suburban"] * 1) + (coef(model9)["female"] * 0) +
(coef(model9)["income"] * 2.619) + (coef(model9)["education"] * 2) + (coef(model9)["age"] * 50.03) +
(coef(model9)["hadweapon"] * 1) + (coef(model9)["hadcrime"] * 1)
(Intercept)
   1.702352
> #1.702352
> coef(model10)["(Intercept)"] + (coef(model10)["black"] * 0) + (coef(model10)["ideology"] * 4) +
(coef(model10)["party"] * 4) + (coef(model10)["suburban"] * 1) + (coef(model10)["female"] * 0) +
(coef(model10)["income"] * 2.619) + (coef(model10)["education"] * 2) + (coef(model10)["age"] * 50.03) +
(coef(model10)["hadweapon"] * 1) + (coef(model10)["hadcrime"] * 1)
```

```
(Intercept)
   2.425317
> #2.425317
> coef(model9)["(Intercept)"] + (coef(model9)["black"] * 0) + (coef(model9)["ideology"] * 4) +
(coef(model9)["party"] * 4) + (coef(model9)["suburban"] * 1) + (coef(model9)["female"] * 0) +
(coef(model9)["income"] * 2.619) + (coef(model9)["education"] * 2) + (coef(model9)["age"] * 50.03) +
(coef(model9)["hadweapon"] * 4) + (coef(model9)["hadcrime"] * 4)
(Intercept)
   2.254304
> #2.254304
> coef(model10)["(Intercept)"] + (coef(model10)["black"] * 0) + (coef(model10)["ideology"] * 4) +
(coef(model10)["party"] * 4) + (coef(model10)["suburban"] * 1) + (coef(model10)["female"] * 0) +
(coef(model10)["income"] * 2.619) + (coef(model10)["education"] * 2) + (coef(model10)["age"] * 50.03) +
(coef(model10)["hadweapon"] * 4) + (coef(model10)["hadcrime"] * 4)
(Intercept)
   2.732276
> #2.732276
> coef(model9)["(Intercept)"] + (coef(model9)["black"] * 0) + (coef(model9)["ideology"] * 6) +
(coef(model9)["party"] * 6) + (coef(model9)["suburban"] * 1) + (coef(model9)["female"] * 0) +
(coef(model9)["income"] * 2.619) + (coef(model9)["education"] * 2) + (coef(model9)["age"] * 50.03) +
(coef(model9)["hadweapon"] * 1) + (coef(model9)["hadcrime"] * 1)
(Intercept)
   1.947258
> #1.947258
> coef(model10)["(Intercept)"] + (coef(model10)["black"] * 0) + (coef(model10)["ideology"] * 6) +
(coef(model10)["party"] * 6) + (coef(model10)["suburban"] * 1) + (coef(model10)["female"] * 0) +
(coef(model10)["income"] * 2.619) + (coef(model10)["education"] * 2) + (coef(model10)["age"] * 50.03) +
(coef(model10)["hadweapon"] * 1) + (coef(model10)["hadcrime"] * 1)
(Intercept)
    2.90158
> #2.90158
> coef(model9)["(Intercept)"] + (coef(model9)["black"] * 0) + (coef(model9)["ideology"] * 6) +
(coef(model9)["party"] * 6) + (coef(model9)["suburban"] * 1) + (coef(model9)["female"] * 0) +
(coef(model9)["income"] * 2.619) + (coef(model9)["education"] * 2) + (coef(model9)["age"] * 50.03) +
(coef(model9)["hadweapon"] * 4) + (coef(model9)["hadcrime"] * 4)
(Intercept)
    2.49921
> #2.49921
> coef(model10)["(Intercept)"] + (coef(model10)["black"] * 0) + (coef(model10)["ideology"] * 6) +
(coef(model10)["party"] * 6) + (coef(model10)["suburban"] * 1) + (coef(model10)["female"] * 0) +
(coef(model10)["income"] * 2.619) + (coef(model10)["education"] * 2) + (coef(model10)["age"] * 50.03) +
(coef(model10)["hadweapon"] * 4) + (coef(model10)["hadcrime"] * 4)
(Intercept)
   3.208539
> #3.208539
```