

Racial Profiling 101: How to judge a book by its cover

Emina Belko & Malak Matar



Main question of the presentation

Are *certain races*
associated with
certain crimes?

Data collection and preparation

Source: FBI
Arrests by Race and Ethnicity
[12,212 agencies; 2018
estimated population
247,752,415]

[Link ↗](#)



Data collection and preparation

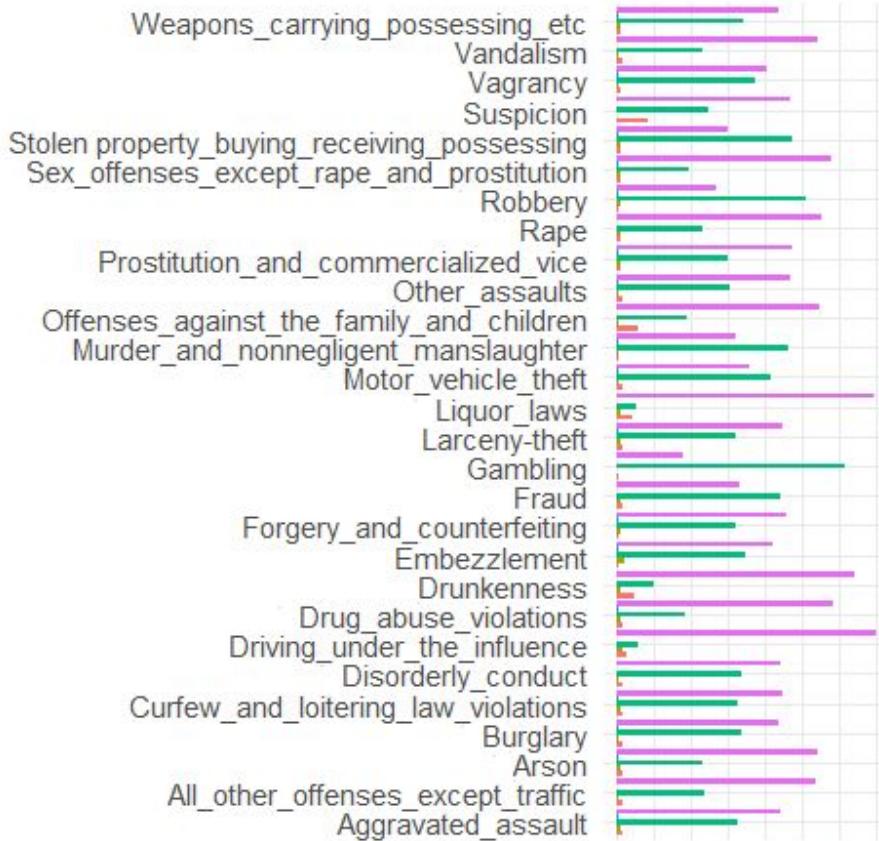
The data: Under 18 and Over 18
categorized into 5 races and 28 criminal offenses

A	B	C	D	E	F	G
Offense_charged	Total	Race_White	Race_Black_or_Af	Race_American	Race_Asian	Race_Native
TOTAL	7167658.00	4983627.00	1925785.00	154084.00	86897.00	17265.00
Count_Murder_and_nonnegligent_manslaughter	8264.00	3673.00	4377.00	99.00	91.00	24.00
Count_Rape	15706.00	10677.00	4492.00	234.00	259.00	44.00
Count_Robbery	53707.00	24602.00	27789.00	604.00	516.00	196.00
Count_Aggravated_assault	277097.00	172900.00	91786.00	6368.00	4810.00	1233.00
Count_Burglary	117843.00	82398.00	32588.00	1310.00	1267.00	280.00
Count_Larceny-theft	600408.00	409000.00	173055.00	10844.00	6323.00	1186.00
Count_Motor_vehicle_theft	57913.00	39536.00	16525.00	939.00	734.00	179.00
Count_Arson	5608.00	4031.00	1356.00	111.00	84.00	26.00
Count_Other_assaults	701782.00	457339.00	216686.00	14165.00	9516.00	2076.00
Count_Forgery_and_counterfeiting	36945.00	24695.00	11323.00	329.00	537.00	61.00
Count_Fraud	86111.00	57125.00	26443.00	1352.00	1052.00	139.00
Count_Embzelement	10739.00	6694.00	3767.00	118.00	147.00	13.00
Count_Stolen_property_buying_receiving_posses	62938.00	41581.00	19538.00	805.00	724.00	290.00
Count_Vandalism	112041.00	75758.00	32275.00	2453.00	1355.00	200.00
Count_Weapons_carrying_possessing/etc.	113492.00	61773.00	49214.00	960.00	1206.00	339.00
Count_Prostitution_and_commercialized_vice	23303.00	12810.00	9034.00	92.00	1307.00	60.00
Count_Sex_offenses_except_rape_and_prostitution	29572.00	21276.00	7042.00	526.00	667.00	61.00
Count_Drug_abuse_violations	1167082.00	821984.00	317809.00	12791.00	12443.00	2055.00
Count_Gambling	2331.00	970.00	1095.00	10.00	205.00	51.00
Count_Offenses_against_the_family_and_children	61910.00	41697.00	17957.00	1714.00	492.00	50.00
Count_Driving_under_the_influence	732646.00	594407.00	108425.00	13017.00	14259.00	2538.00
Count_Liquor_laws	108923.00	83666.00	17483.00	6014.00	1632.00	128.00
Count_Drunkenness	249011.00	191053.00	37465.00	17274.00	2793.00	426.00
Count_Disorderly_conduct	205095.00	134516.00	59775.00	8903.00	1640.00	261.00

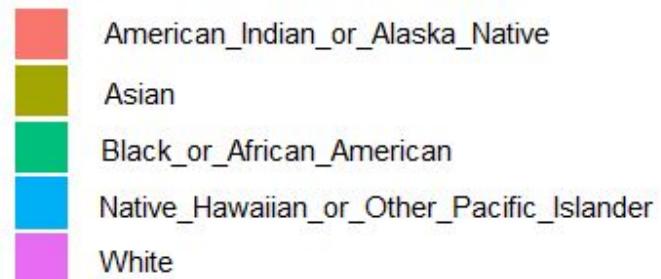
A	B	C	D	E	F	G
Offense_charged	Total	Race_White	Race_Black_or_Af	Race_American	Race_Asian	Race_Native_H
TOTAL	543242.00	336027.00	189596.00	10346.00	5840.00	1433.00
Count_Murder_and_nonnegligent_manslaughte	693.00	280.00	401.00	6.00	3.00	3.00
Count_Rape	3070.00	2117.00	884.00	33.00	30.00	6.00
Count_Robbery	13082.00	4423.00	8398.00	72.00	125.00	64.00
Count_Aggravated_assault	20943.00	11627.00	8607.00	368.00	268.00	73.00
Count_Burglary	16699.00	9183.00	7029.00	280.00	155.00	52.00
Count_Larceny-theft	69575.00	39193.00	28031.00	1143.00	1001.00	207.00
Count_Motor_vehicle_theft	11089.00	4976.00	5780.00	212.00	84.00	37.00
Count_Arson	1338.00	907.00	384.00	26.00	17.00	4.00
Count_Other_assaults	93005.00	54686.00	35674.00	1546.00	832.00	267.00
Count_Forgery_and_counterfeiting	779.00	445.00	314.00	6.00	11.00	3.00
Count_Fraud	3499.00	1447.00	1944.00	67.00	36.00	5.00
Count_Embzelement	435.00	229.00	188.00	4.00	12.00	2.00
Count_Stolen_property_buying_receiving_posse	6936.00	2598.00	4123.00	76.00	95.00	44.00
Count_Vandalism	22753.00	15418.00	6612.00	496.00	197.00	30.00
Count_Weapons_carrying_possessing/etc.	12840.00	6983.00	5501.00	134.00	184.00	38.00
Count_Prostitution_and_commercialized_vice	199.00	118.00	75.00	3.00	2.00	1.00
Count_Sex_offenses_except_rape_and_prostitutu	5585.00	4062.00	1361.00	61.00	84.00	17.00
Count_Drug_abuse_violations	67096.00	49311.00	15304.00	1357.00	902.00	222.00
Count_Gambling	134.00	30.00	103.00	1.00	0.00	0.00
Count_Offenses_against_the_family_and_childr	2447.00	1674.00	573.00	181.00	18.00	1.00
Count_Driving_under_the_influence	3598.00	3512.00	278.00	133.00	64.00	11.00
Count_Liquor_laws	19530.00	17021.00	1260.00	959.00	244.00	46.00
Count_Drunkenness	2479.00	1989.00	316.00	138.00	32.00	4.00
Count_Disorderly_conduct	43621.00	24017.00	18417.00	867.00	286.00	34.00

Proportion of Each Race within Each Offense Category

Offense Charged

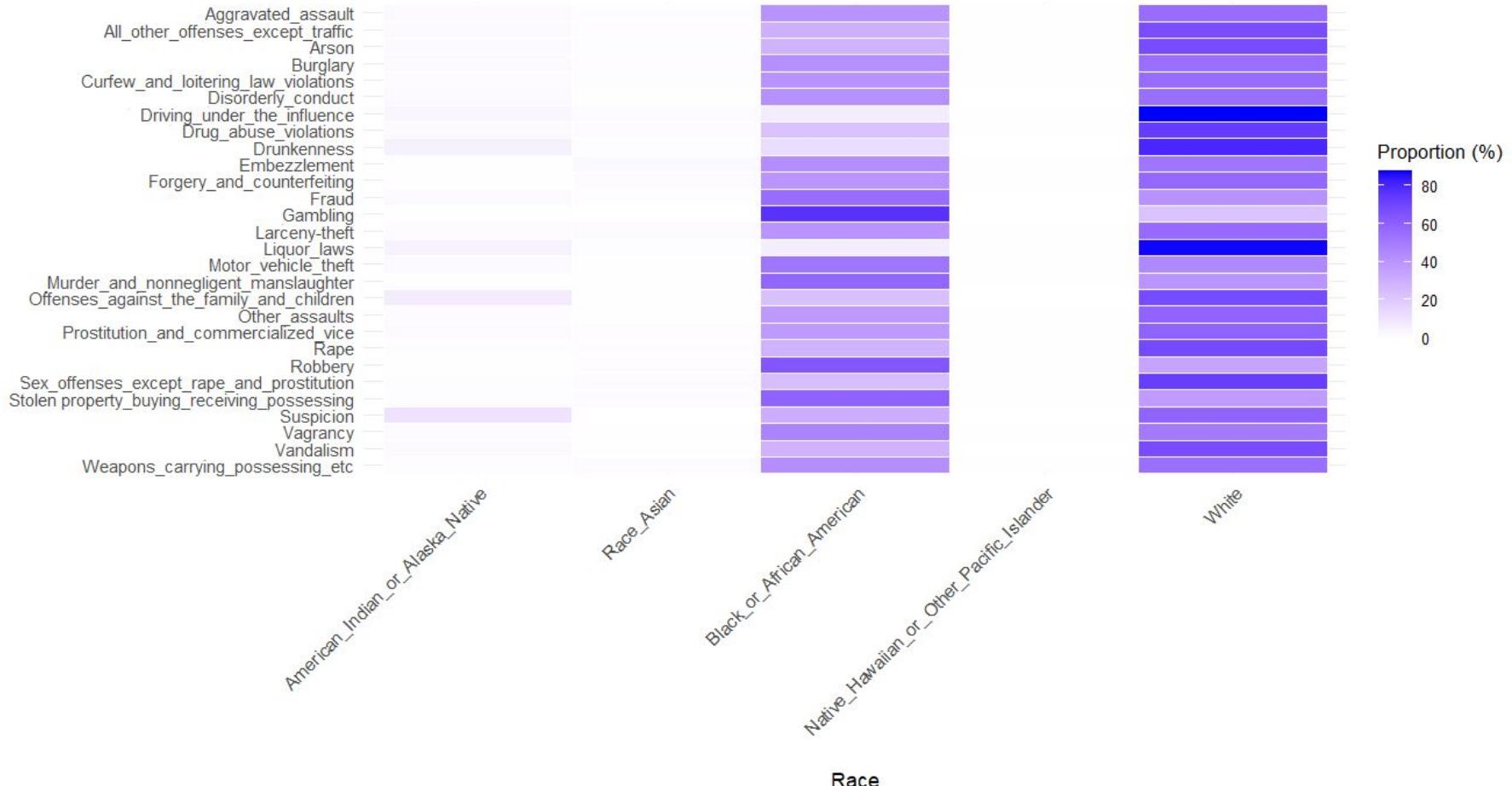


Race



Heatmap of Proportion of Each Race within Each Offense Category

Offense Charged



Initial Analysis

Out of all juveniles arrested for murder:

- White - 40.4%
- Black or African American - 57.9%

Out of all juveniles arrested for violent crimes:

- White juveniles - 48.8%
- Black or African American - 48.4%.

The screenshot shows an RStudio interface with a script editor and a console. The script editor contains R code for categorizing crimes and calculating proportions. The console shows the results of running this code, including proportions for adults and juveniles arrested for murder and violent crimes across different racial groups.

```
8 # First we categorize the crimes
9 violent_crimes <- c("Count_Murder_and_nonnegligent_manslaughter", "Count_Rape",
10 "Count_Robbery", "Count_Aggravated_assault")
11 property_crimes <- c("Count_Burglary", "Count_Larceny-theft", "Count_Motor_vehicle",
12 "Count_Arson")
13
14
15 race_totals <- under18 %>%
16   summarise(across(starts_with("Race_"), sum))
17
18 # arrests for violent crimes by race
19 violent_arrests <- under18 %>%
20   filter(Offense_charged %in% violent_crimes) %>%
21   summarise(across(starts_with("Race_"), sum))
22
23 # property crimes by race
24 property_arrests <- under18 %>%
25   filter(Offense_charged %in% property_crimes) %>%
26   summarise(across(starts_with("Race_"), sum))
27
28 # Calculate proportions for violent crimes
29 violent_proportions <- violent_arrests / sum(violent_arrests) * 100
30
4:1 (Top Level) ⇣
```

R 4.3.3 · C:/Users/pc/Desktop/Year 2/Semester 2/DMA/PROJECT/archive/ ↗

Proportion of adults arrested for murder:
> print(murder_proportions)

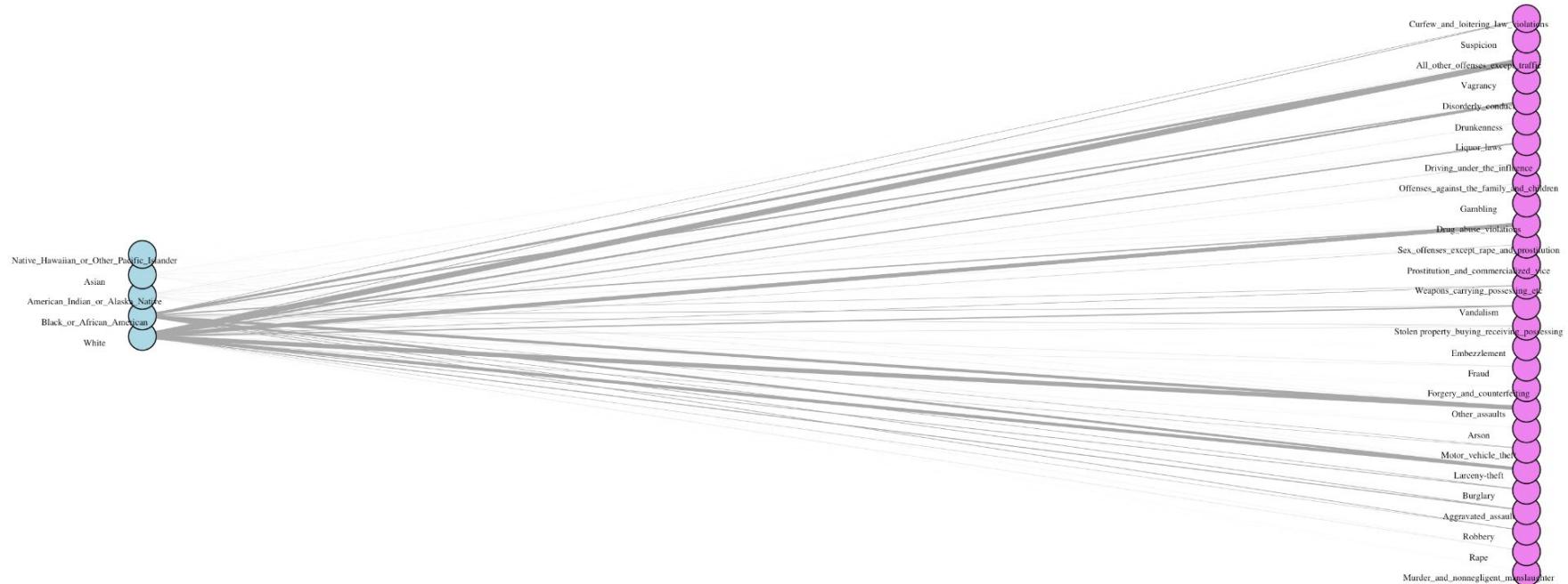
Race	Proportion
Race_White	40.40404
Race_Black_or_African_American	57.86436
Race_American_Indian_or_Alaska_Native	0.8658009
Race_Asian	0.4329004
Race_Native_Hawaiian_or_Other_Pacific_Islander	0.4329004

>
> cat("\nProportion of juveniles arrested for violent crimes:\n")

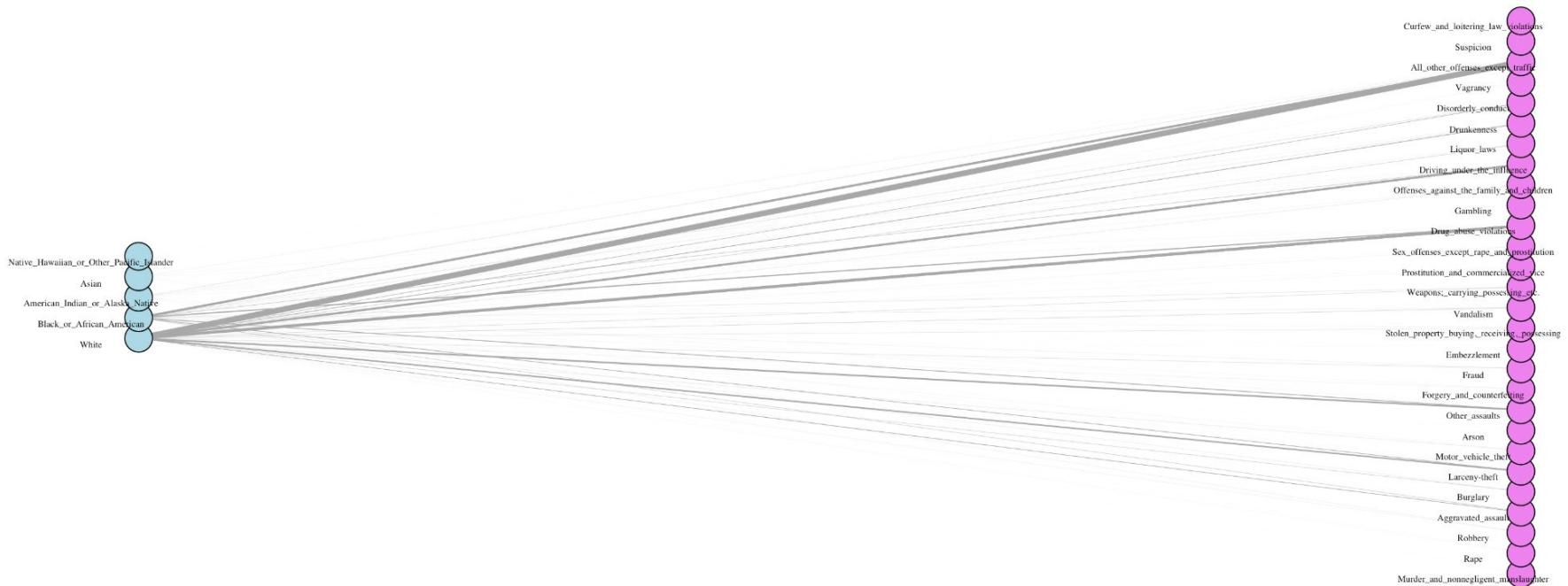
Proportion of juveniles arrested for violent crimes:
> print(juvenile_violent_proportions)

Race	Proportion
Race_White	48.81708
Race_Black_or_African_American	48.40161
Race_American_Indian_or_Alaska_Native	1.267598
Race_Asian	1.127342
Race_Native_Hawaiian_or_Other_Pacific_Islander	0.386366

Bipartite Graph - Under 18



Bipartite Graph - Over 18



Marginal proportions and conditional probability to study association

The marginal proportion of an event is the probability distribution that describes that single event only.

The conditional probability is a distribution that represents the likelihood of an event to occur given a particular outcome of another event.

In our case, the likelihood that a specific crime occurs given the race

Marginal proportions and conditional probability to study association

```
13
14 #data manipulation
15 over18_long <- over18 %>%
16   pivot_longer(cols = starts_with("Race_"), names_to = "Race", values_to = "Count")
17
18 conditional_prob <- over18_long %>%
19   group_by(Race) %>%
20   mutate(Total_Offenses = sum(Count),
21         Conditional_Probability = Count / Total_Offenses) %>%
22   ungroup()
23
24 offense_totals <- over18_long %>%
25   group_by(Offense_charged) %>%
26   summarise(Total_Offenses = sum(Count))
27 overall_total <- sum(offense_totals$Total_Offenses)
28
29 # marginal proportions
30 offense_totals <- offense_totals %>%
31   mutate(Marginal_Proportion = Total_Offenses / overall_total)
32
33 #have both the conditional probability and marginal proportions
34 comparison_table <- conditional_prob %>%
35   left_join(offense_totals, by = "Offense_charged") %>%
36   select(Offense_charged, Race, Count, Conditional_Probability, Marginal_Proportion)
37
38 # filter to keep only rows where the conditional probability is greater than the marginal proportion
39 filtered_table <- comparison_table %>%
40   filter(Conditional_Probability > Marginal_Proportion)
41
42 table_plot <- ggplot(data = filtered_table, aes(x = Offense_charged, y = Conditional_Probability, fill = Race)) +
43   geom_bar(stat = "identity", position = position_dodge())
44   geom_text(aes(label = round(Conditional_Probability, 3)), vjust = -1, position = position_dodge(3)) +
45   labs(title = "Races associated with each offense (Over 18)",
46       x = "Offense Charged",
47       y = "Conditional Probability") +
48   theme(axis.text.x = element_text(angle = 90, hjust = 1))
49
50 ggsave("conditional_probability_table_over_18.jpg", plot = table_plot, width = 20, height = 10)
51
52 print(filtered_table, n = nrow(filtered_table))
53
```

Marginal proportions and conditional probability to study association (under 18)

Source				
Console	Terminal	Background Jobs		
R 4.3.3 · C:/Users/pc/Desktop/Year 2/Semester 2/DMA/PROJECT/archive/				
Offense_charged	Race	Count	Conditional_Probability	Marginal_Proportion
<chr>	<chr>	<dbl>	<dbl>	<dbl>
1 Count_Murder_and_nonnegligent_manslaughter	Race_Black_or_African_American	401	0.00212	0.00128
2 Count_Murder_and_nonnegligent_manslaughter	Race_Native_Hawaiian_or_Other_Pacific_Islander	3	0.00209	0.00128
3 Count_Rape	Race_White	2117	0.00630	0.00565
4 Count_Robbery	Race_Black_or_African_American	8398	0.0443	0.0241
5 Count_Robbery	Race_Native_Hawaiian_or_Other_Pacific_Islander	64	0.0447	0.0241
6 Count_Aggravated_assault	Race_Black_or_African_American	8607	0.0454	0.0386
7 Count_Aggravated_assault	Race_Asian	268	0.0459	0.0386
8 Count_Aggravated_assault	Race_Native_Hawaiian_or_Other_Pacific_Islander	73	0.0509	0.0386
9 Count_Burglary	Race_Black_or_African_American	2029	0.0371	0.0307
10 Count_Burglary	Race_Native_Hawaiian_or_Other_Pacific_Islander	52	0.0363	0.0307
11 Count_Larceny_theft	Race_Black_or_African_American	28031	0.148	0.128
12 Count_Larceny_theft	Race_Asian	1001	0.171	0.128
13 Count_Larceny_theft	Race_Native_Hawaiian_or_Other_Pacific_Islander	207	0.144	0.128
14 Count_Motor_vehicle_theft	Race_Black_or_African_American	5780	0.0305	0.0204
15 Count_Motor_vehicle_theft	Race_American_Indian_or_Alaska_Native	212	0.0205	0.0204
16 Count_Motor_vehicle_theft	Race_Native_Hawaiian_or_Other_Pacific_Islander	37	0.0258	0.0204
17 Count_Arson	Race_White	907	0.00270	0.00246
18 Count_Arson	Race_American_Indian_or_Alaska_Native	26	0.00251	0.00246
19 Count_Arson	Race_Asian	17	0.00291	0.00246
20 Count_Arson	Race_Native_Hawaiian_or_Other_Pacific_Islander	4	0.00279	0.00246
21 Count_Other_assaults	Race_Black_or_African_American	35674	0.188	0.171
22 Count_Other_assaults	Race_Native_Hawaiian_or_Other_Pacific_Islander	267	0.186	0.171
23 Count_Forgery_and_counterfeiting	Race_Black_or_African_American	314	0.00166	0.00143
24 Count_Forgery_and_counterfeiting	Race_Asian	11	0.00188	0.00143
25 Count_Forgery_and_counterfeiting	Race_Native_Hawaiian_or_Other_Pacific_Islander	3	0.00209	0.00143
26 Count_Fraud	Race_Black_or_African_American	1944	0.0103	0.00644
27 Count_Fraud	Race_American_Indian_or_Alaska_Native	67	0.00648	0.00644
28 Count_Embezzlement	Race_Black_or_African_American	188	0.000992	0.000801
29 Count_Embezzlement	Race_Asian	12	0.00205	0.000801
30 Count_Embezzlement	Race_Native_Hawaiian_or_Other_Pacific_Islander	2	0.00140	0.000801
31 Count_Stolen_property_buying_receiving_possessing	Race_Black_or_African_American	4123	0.0217	0.0128
32 Count_Stolen_property_buying_receiving_possessing	Race_Asian	95	0.0163	0.0128
33 Count_Stolen_property_buying_receiving_possessing	Race_Native_Hawaiian_or_Other_Pacific_Islander	44	0.0307	0.0128
34 Count_Vandalism	Race_White	15418	0.0459	0.0419
35 Count_Vandalism	Race_American_Indian_or_Alaska_Native	496	0.0479	0.0419
36 Count_Weapons_carrying_possessing_etc	Race_Black_or_African_American	5501	0.0290	0.0236
37 Count_Weapons_carrying_possessing_etc	Race_Asian	184	0.0315	0.0236
38 Count_Weapons_carrying_possessing_etc	Race_Native_Hawaiian_or_Other_Pacific_Islander	38	0.0265	0.0236
39 Count_Prostitution_and_commercialized_vice	Race_Black_or_African_American	75	0.000396	0.000366
40 Count_Prostitution_and_commercialized_vice	Race_Native_Hawaiian_or_Other_Pacific_Islander	1	0.000000	0.000366

Marginal proportions and conditional probability to study association (over 18)

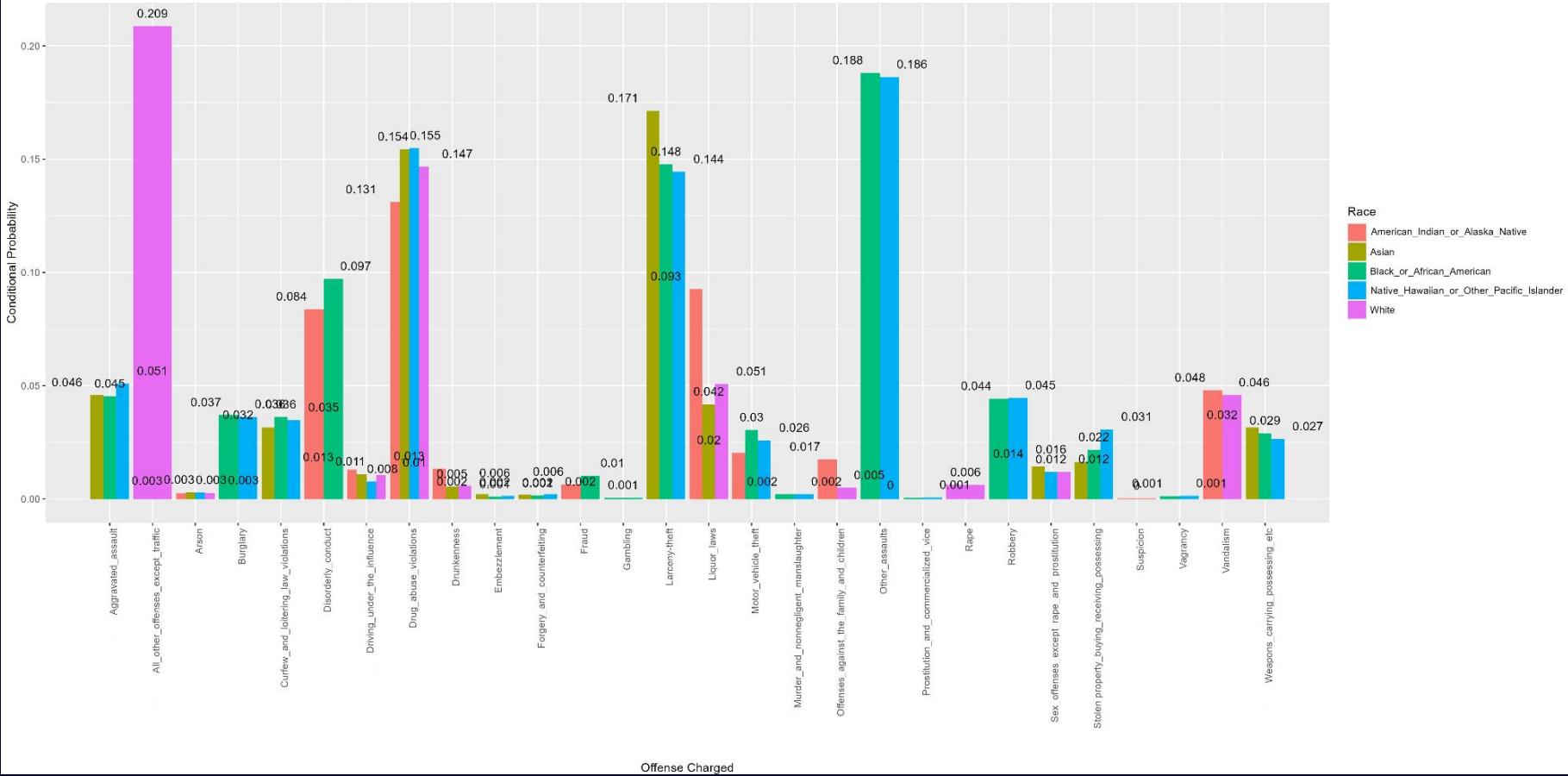
Source

Console Terminal × Background Jobs ×

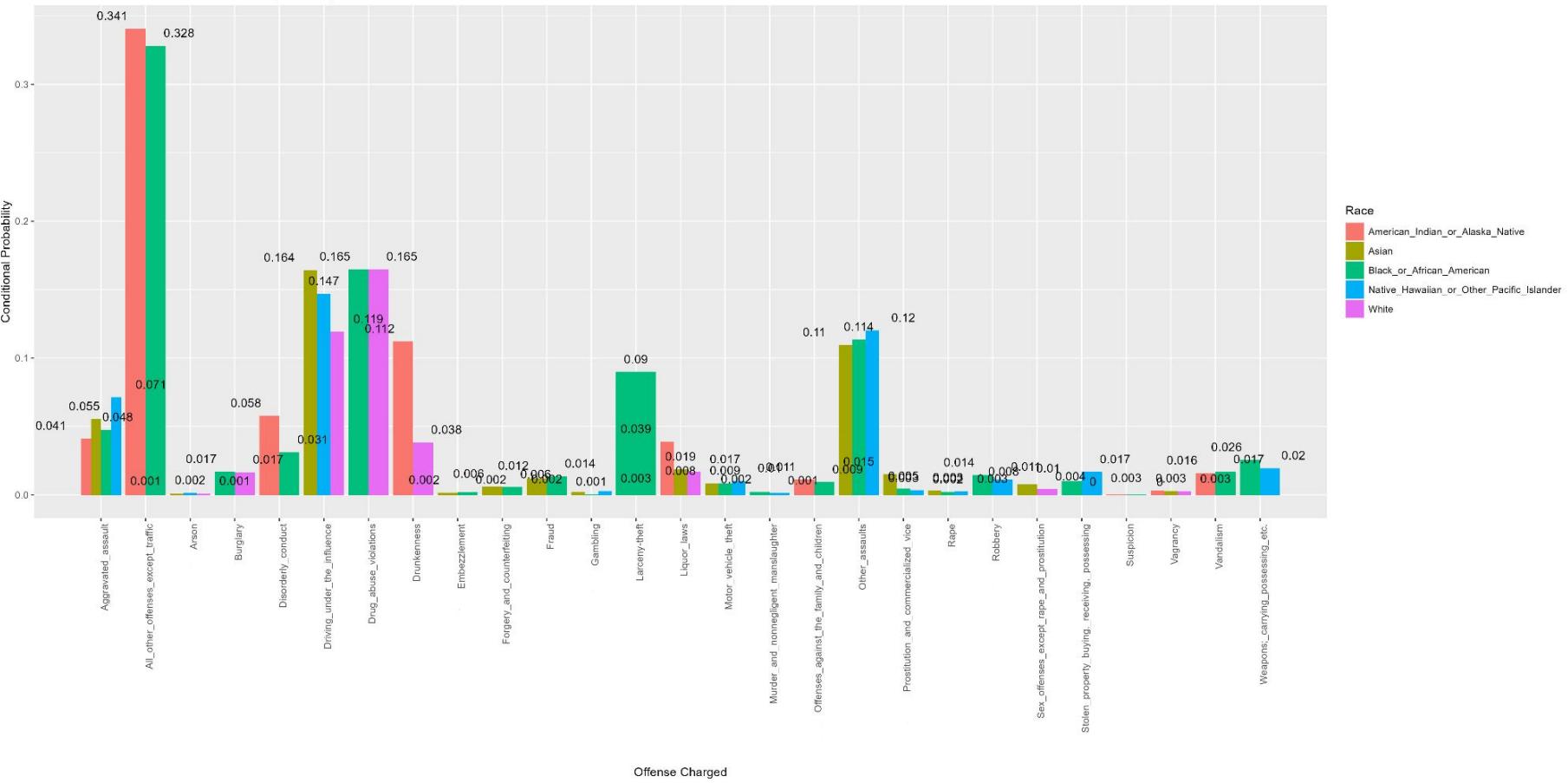
R 4.3.3 · C:/Users/pc/Desktop/Year 2/Semester 2/DMA/PROJECT/archive/

	Race	Count	Conditional_Probability	Marginal_Proportion
	<dbl>	<dbl>	<dbl>	<dbl>
1	Race_Black_or_African_American	4377	0.00227	0.00115
2	Race_Native_Hawaiian_or_Other_Pacific_Islander	24	0.00139	0.00115
3	Race_Black_or_African_American	4492	0.00233	0.00219
4	Race_Asian	259	0.00298	0.00219
5	Race_Native_Hawaiian_or_Other_Pacific_Islander	44	0.00255	0.00219
6	Race_Black_or_African_American	27789	0.0144	0.00749
7	Race_Native_Hawaiian_or_Other_Pacific_Islander	196	0.0114	0.00749
8	Race_Black_or_African_American	91786	0.0472	0.0387
9	Race_American_Indian_or_Alaska_Native	6368	0.0413	0.0387
10	Race_Asian	4810	0.0554	0.0387
11	Race_Native_Hawaiian_or_Other_Pacific_Islander	1233	0.0714	0.0387
12	Race_White	82398	0.0165	0.0164
13	Race_Black_or_African_American	32588	0.0169	0.0164
14	Race_Black_or_African_American	173055	0.0899	0.0838
15	Race_Black_or_African_American	16525	0.00858	0.00808
16	Race_Asian	734	0.00845	0.00808
17	Race_Native_Hawaiian_or_Other_Pacific_Islander	179	0.0104	0.00808
18	Race_White	4031	0.000809	0.000782
19	Race_Asian	84	0.000967	0.000782
20	Race_Native_Hawaiian_or_Other_Pacific_Islander	26	0.00151	0.000782
21	Race_Black_or_African_American	218686	0.114	0.0979
22	Race_Asian	9516	0.110	0.0979
23	Race_Native_Hawaiian_or_Other_Pacific_Islander	2076	0.120	0.0979
24	Race_Black_or_African_American	11323	0.00588	0.00515
25	Race_Asian	537	0.00618	0.00515
26	Race_Black_or_African_American	26443	0.0137	0.0120
27	Race_Asian	1052	0.0121	0.0120
28	Race_Black_or_African_American	3767	0.00196	0.00150
29	Race_Asian	147	0.00169	0.00150
30	Race_Black_or_African_American	19538	0.0101	0.00878
31	Race_Native_Hawaiian_or_Other_Pacific_Islander	290	0.0168	0.00878
32	Race_Black_or_African_American	32275	0.0168	0.0156
33	Race_American_Indian_or_Alaska_Native	2453	0.0159	0.0156
34	Race_Black_or_African_American	49214	0.0256	0.0158
35	Race_Native_Hawaiian_or_Other_Pacific_Islander	339	0.0196	0.0158
36	Race_Black_or_African_American	9034	0.00469	0.00325
37	Race_Asian	1307	0.0150	0.00325
38	Race_Native_Hawaiian_or_Other_Pacific_Islander	60	0.00348	0.00325
39	Race_White	21276	0.00427	0.00413
40	Race_Asian	667	0.00769	0.00413

Races associated with each offense (Under 18)



Races associated with each offense (Over 18)



Interesting observations

Racial Disparities

- "Driving under the influence" and "Drug abuse violations" → high conditional probabilities for Whites in both age groups
- "Weapons carrying" and "Prostitution and commercialized vice" → more associated with Black or African American individuals.



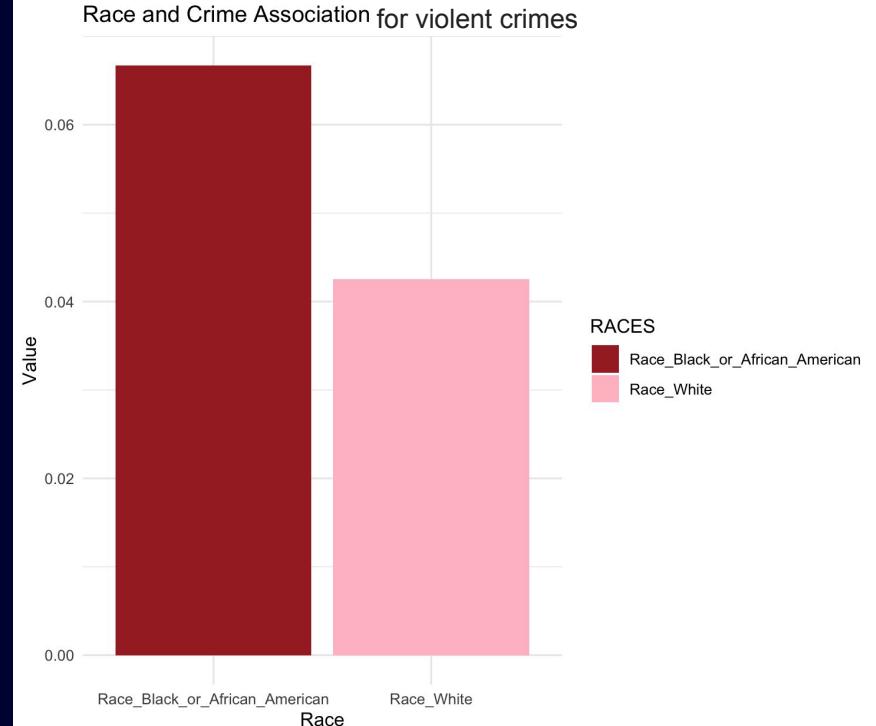
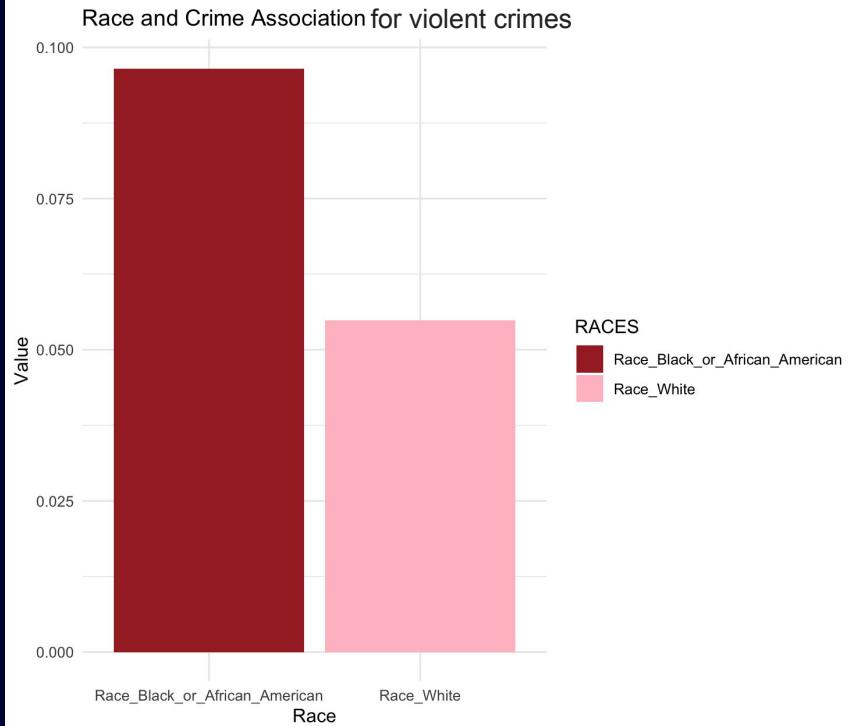
Number of crimes that each race was associated with:

Race	Under 18	Over 18
White	10	8
Black or African American	17	21
American Indian or Alaska Native	11	9
Asian	13	14
Native Hawaiian or other Pacific Islander	18	12

Number of crimes that each race was associated with:

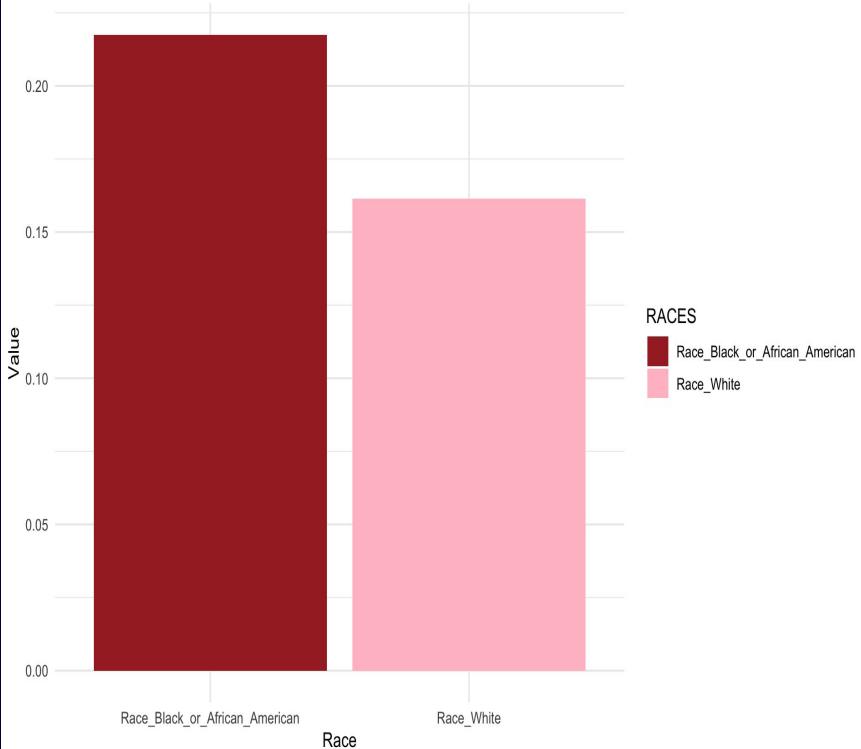
Race	Under 18	Over 18
White	10	8
Black or African American	17	21
American Indian or Alaska Native	11	9
Asian	13	14
Native Hawaiian or other Pacific Islander	18	12

Histogram representation

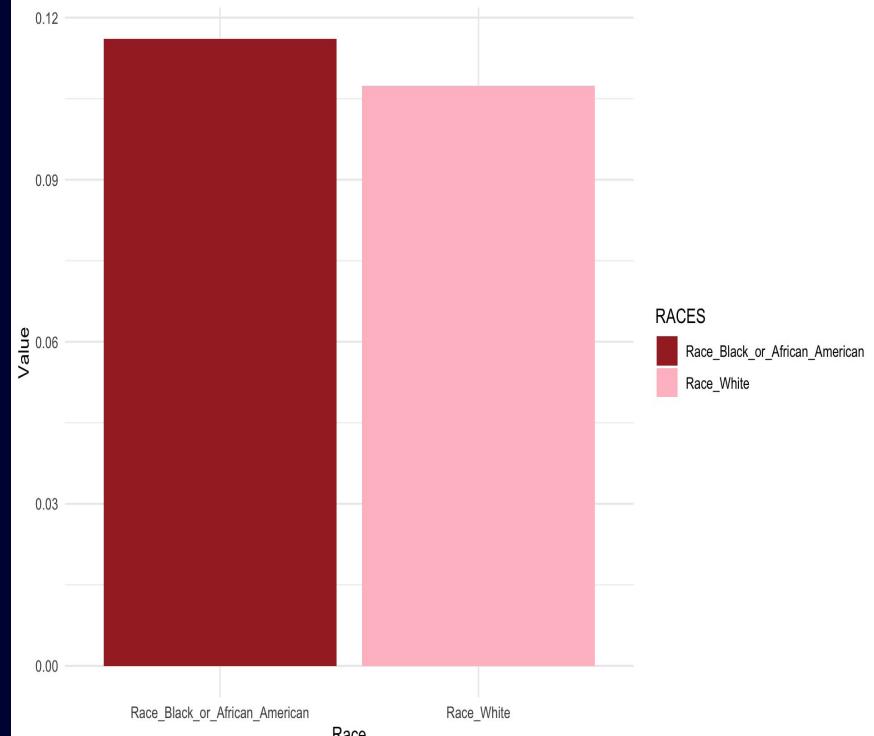


Histogram representation

Race and Crime Association for property crimes



Race and Crime Association for property crimes



Conclusions

Significant association between type of crime and offenders race could be found in the following cases:

Crimes committed	Under 18	Over 18
Violent Crimes	Black or African American	Black or African American
Property Crimes	Black or African American	Black or African American
Drugs abuse and intoxication Crimes	American Indian or Alaska Native	White

Conclusions

Yes, certain races are associated with certain types of crimes.

Certain crimes were associated with only one racial group:

- Age group - Under 18
 - Racial group: Black and African American → Gambling
 - Racial group: White → Rape
 - Racial group: American Indian or Alaska Native → Suspicion
- Age group - Over 18
 - Racial group: Black and African American → Larceny Theft

Thank you!