

Stats 101 Assignment Report:

CS6603

Nehmya Negash
nnegash6@gatech.edu

1 STATS 101 REPORT

1.1 Step 1 - Select one of the datasets available

death-in-custody.csv

1.2 Step 2 - Explore the data

Dataset: Death in Custody

Number of Observations: 8,373

Number of Variables: 16

Regulated Domain in Law: Housing and Public Accommodation (Fair Housing Act, Civil Rights Act)

	Protected Class	Law
gender	Sex	Civil Rights Act of 1964, Equal Pay Act of 1963
age	Age	Age Discrimination Act (less directly applicable here but age is a demographic) / Eighth Amendment (cruel and unusual punishment, relevant to conditions of confinement which may differ by age)
race	Race/Ethnicity	Civil Rights Act of 1964, Fourteenth Amendment (Equal Protection)

1.3 Step 3 - Determine the relationships between variables

1.3.1 Race vs Manner of Death

	White	Black	Hispanic	Other
Homicide Justified	368	248	575	81
Homicide Willful	100	46	71	16
Suicide	419	94	263	80
Natural	2142	1348	1211	234
Accidental	265	142	247	29
Other	146	94	127	27

Table 1—Race vs Manner of Death

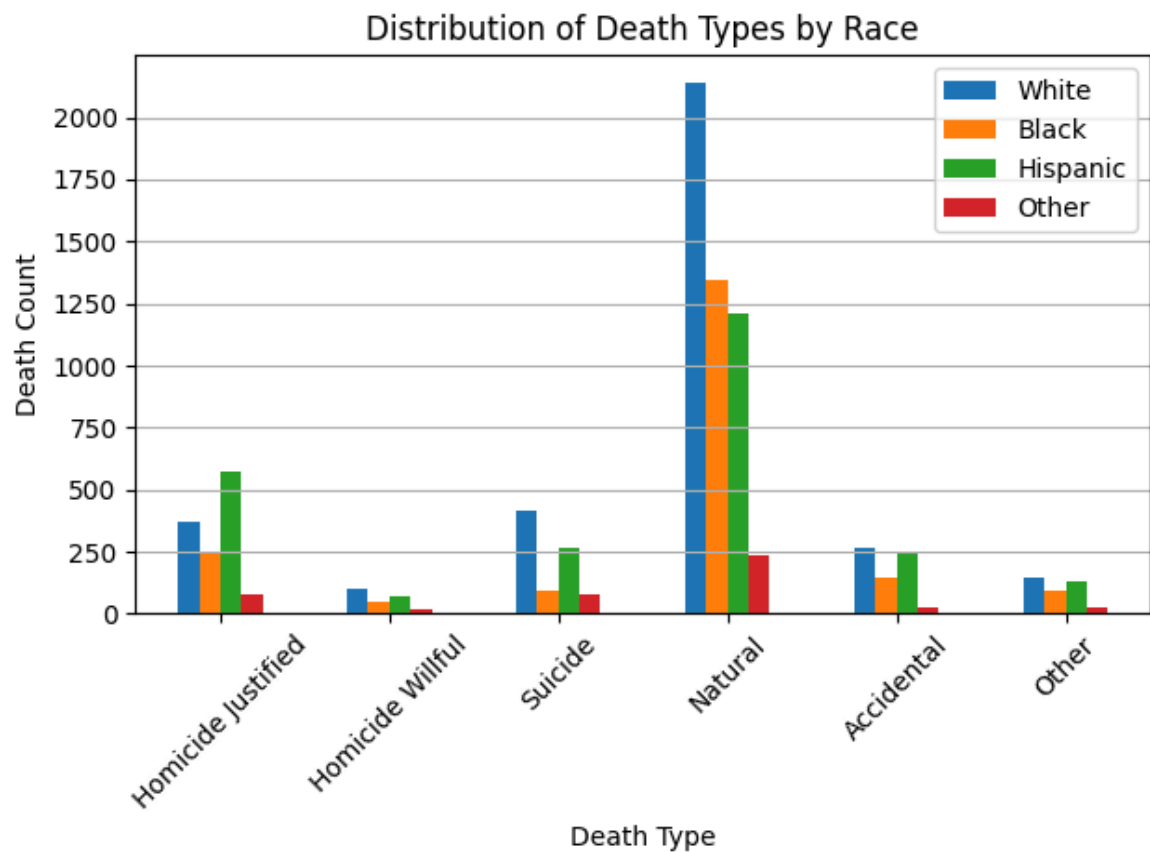


Figure 1—Race vs Manner of Death

1.3.2 Gender vs Manner of Death

	Male	Female
Homicide Justified	1223	49
Homicide Willful	230	3
Suicide	798	58
Natural	4654	281
Accidental	628	55
Other	369	25

Table 2—Gender vs Manner of Death

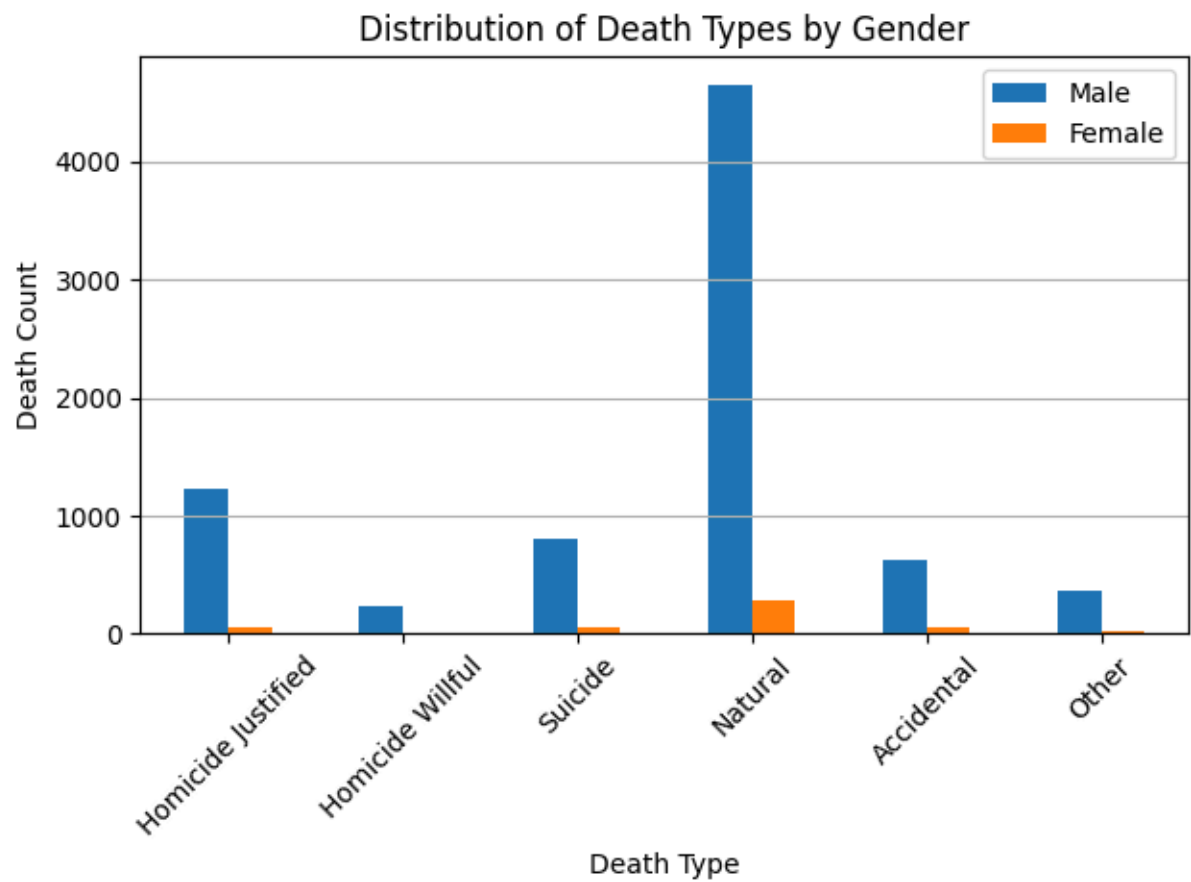


Figure 2—Gender vs Manner of Death

1.3.3 Age vs Manner of Death

	0-17	18-34	35-59	60+
Homicide Justified	34	694	501	42
Homicide Willful	0	63	140	30
Suicide	5	335	474	42
Natural	2	257	2828	1848
Accidental	6	243	397	37
Other	5	147	197	45

Table 3—Age vs Manner of Death

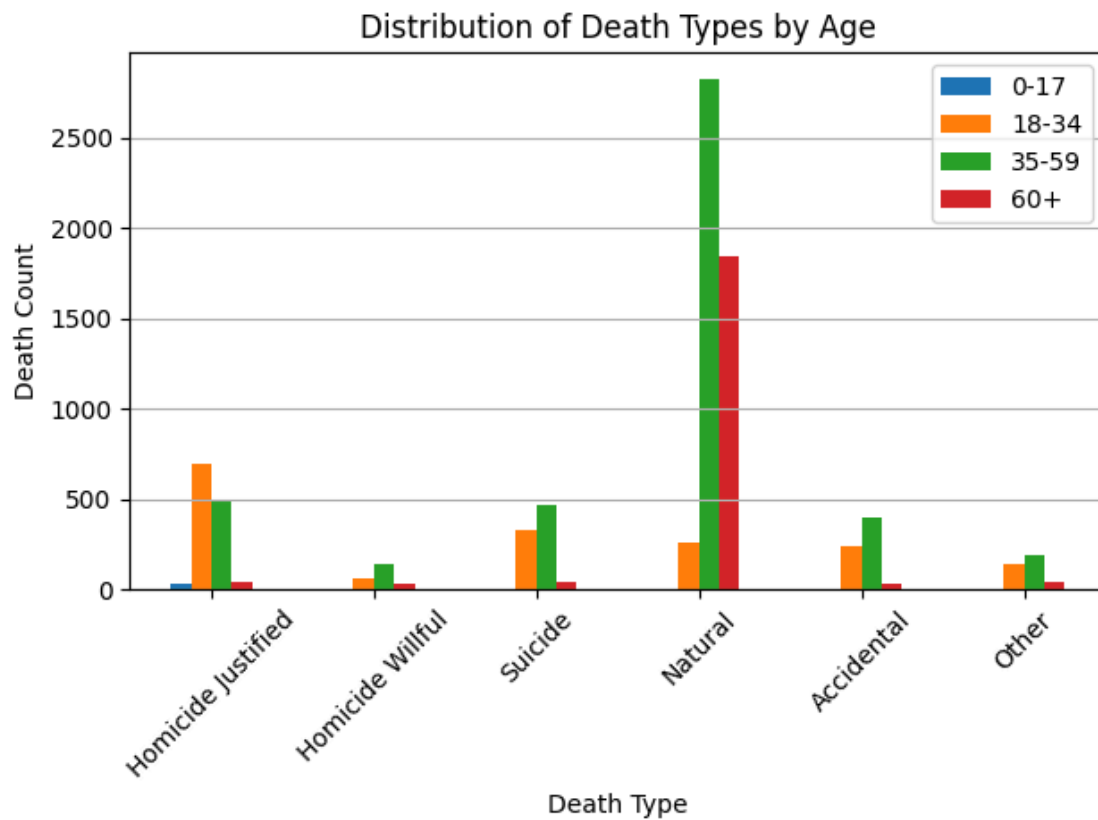


Figure 3—Age vs Manner of Death

1.3.4 Age vs Custody Status

	0-17	18-34	35-59	60+
Booked	2	311	752	169
Sentenced	8	488	2937	1762
Process of Arrest	41	827	681	52
Other	1	113	167	61

Table 4— Age vs Custody Status

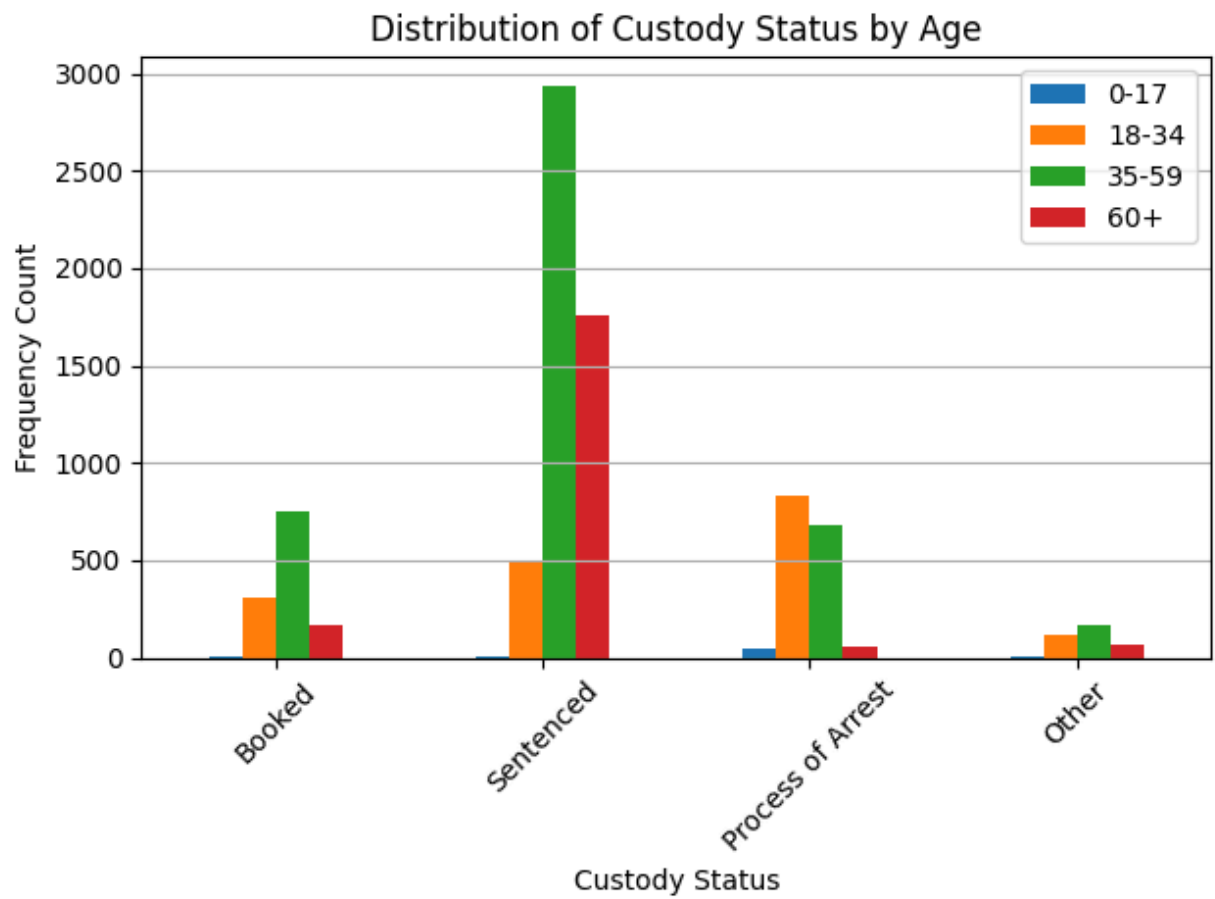


Figure 4— Age vs Custody Status

1.3.5 Gender vs Custody Status

	Male	Female
Booked	1096	138
Sentenced	4961	234
Process of Arrest	1535	67
Other	310	32

Table 5—Gender vs Custody Status

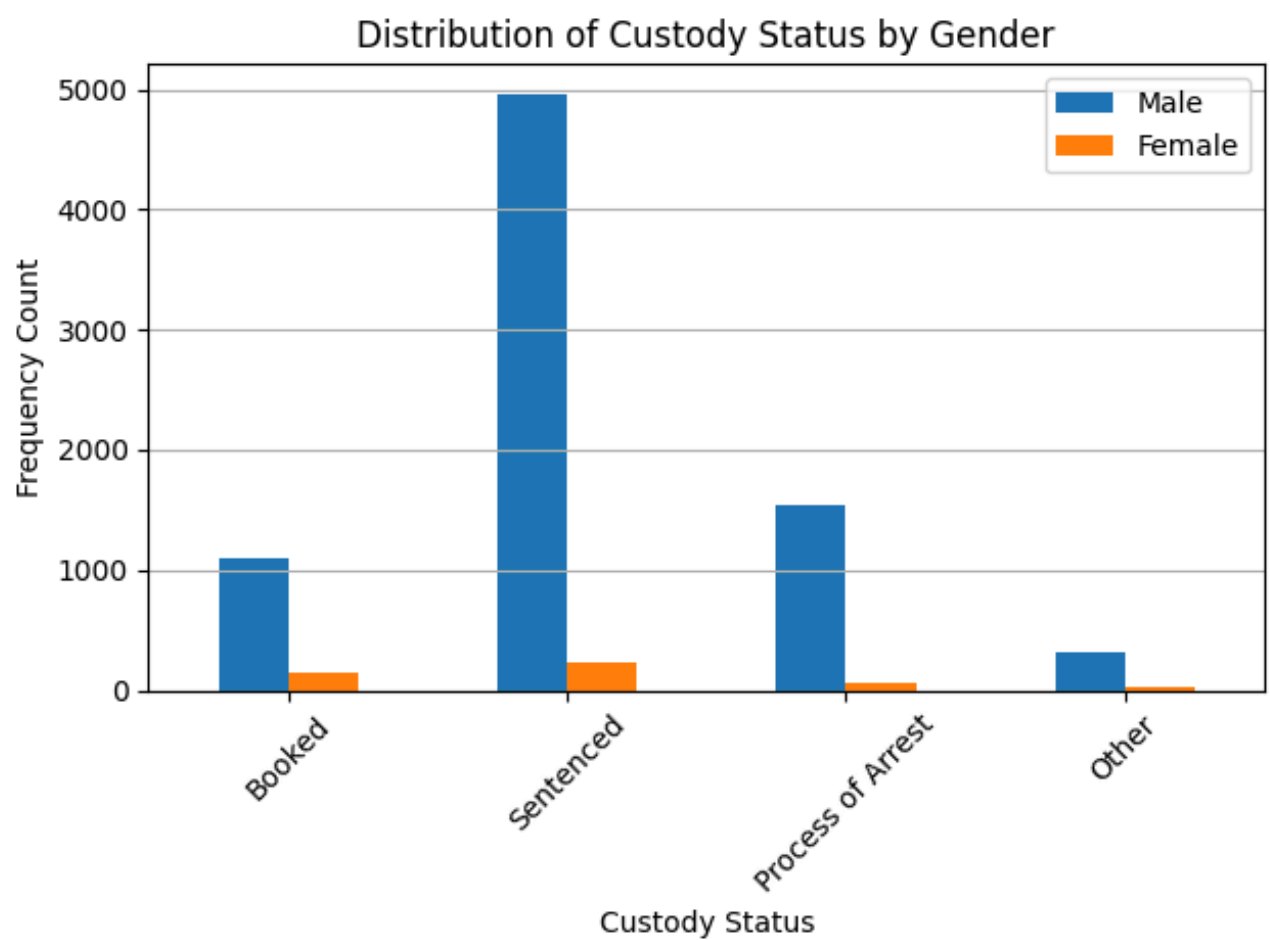


Figure 5—Gender vs Custody Status

1.3.6 Gender vs Custody Status

	White	Black	Hispanic	Other
Booked	563	254	340	77
Sentenced	2249	1316	1365	265
Process of Arrest	491	323	679	109
Other	137	79	110	16

Table 6—Race vs Custody Status

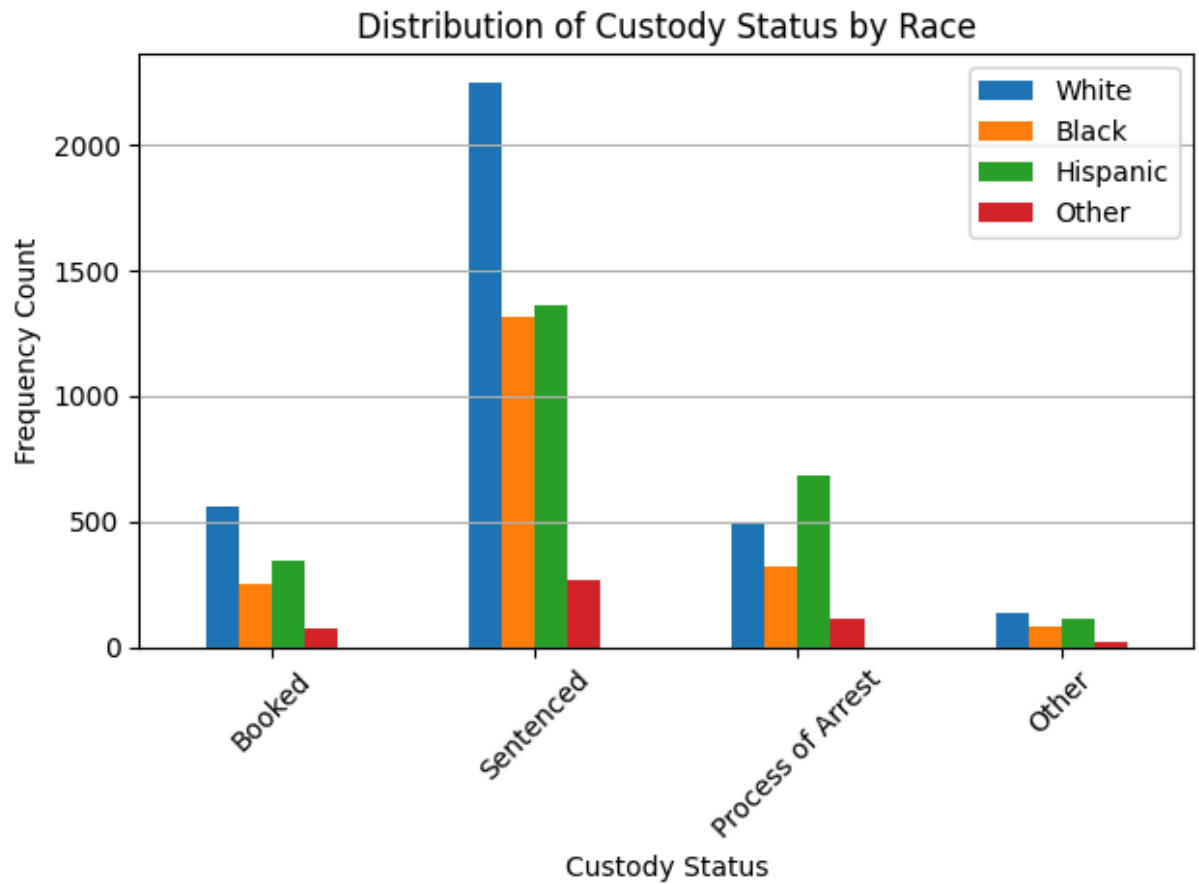


Figure 6—Race vs Custody Status

1.4 Step 4 - Show how to manipulate with data

1.4.1 Fair Hypothesis

As seen from the graph, no race is at a higher risk of dying to homicide.
[Manipulations: Increased scale to 10 and normalized ratios by race population size to always be less than 1; No label on the Y-Axis]

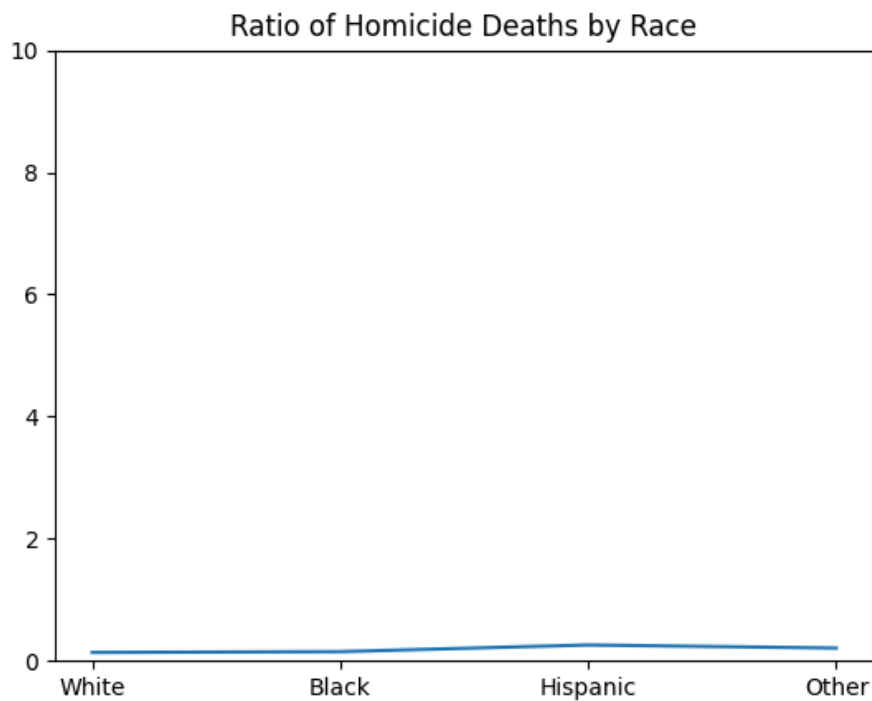


Figure 7—Race vs Homicide rate misleading

1.4.1 Bias Hypothesis

As seen from the graph, Hispanics are at a higher risk of dying via homicide compared to other races. [Manipulations: Data supports hypothesis so no need for manipulations]

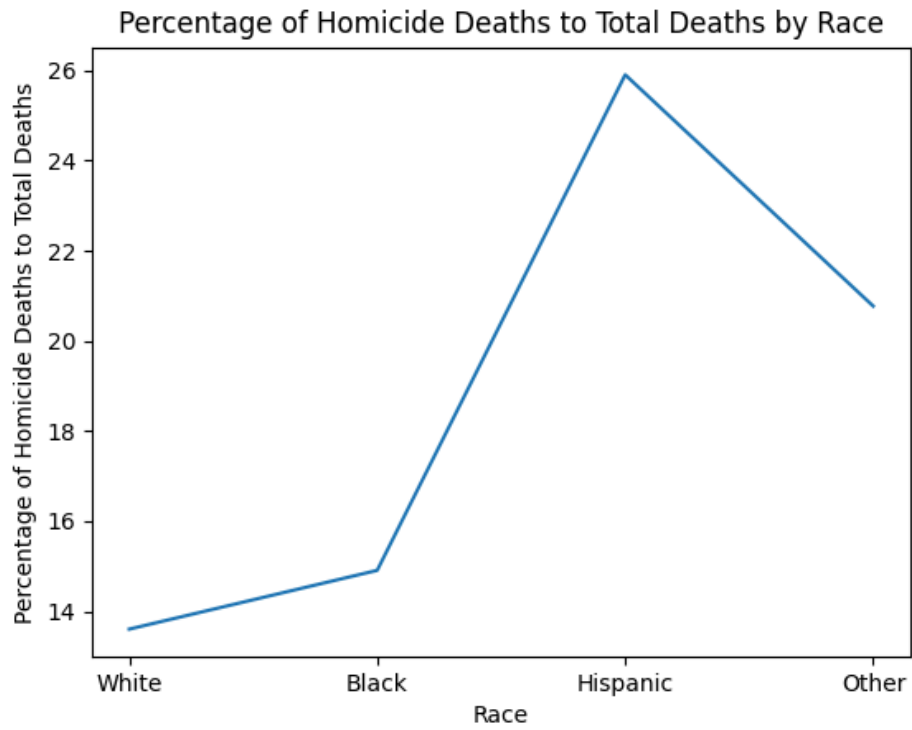


Figure8—Race vs Homicide rate

1.5 Step 5

Protected Class Variable (Race)	Mean	Median	Mode
Original	2.2 (Natural)	2 (Natural)	2 (Natural)
Reduced	2.1 (Natural)	2 (Natural)	2 (Natural)
Difference	No Difference	No Difference	No Difference

Table 7—Statistics of original vs reduced data set

1.6 Step 6: 50% sampled Race vs Manner of Death

	White	Black	Hispanic	Other
Homicide Justified	167	134	286	40
Homicide Willful	55	21	32	7
Suicide	206	53	133	43
Natural	1068	658	619	105
Accidental	139	78	130	14
Other	76	44	63	15

Table 8—Race vs Custody Status

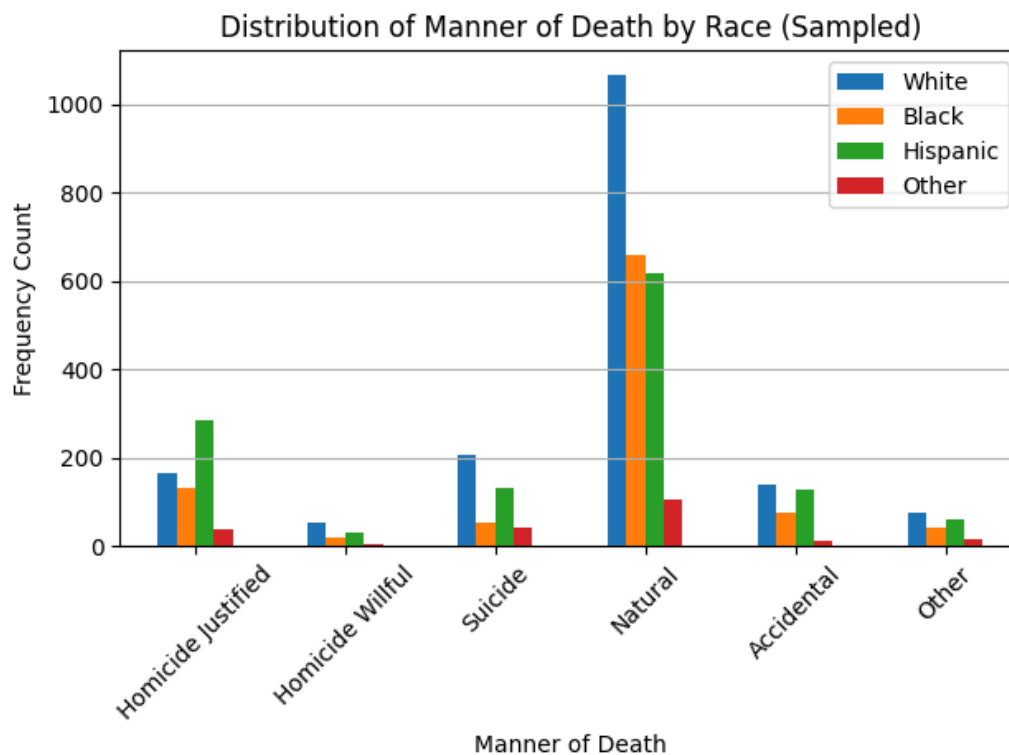


Figure 9—Race vs Custody Status

There is no meaningful difference observed, likely due to the large size of the dataset. A smaller sample might introduce more variability. Since the results remained consistent, no group would be either advantaged or disadvantaged based on a sampled dataset.