



19

IVA STORIES F20

NYPD SHOOTING DATA PROJECT:

INTRA-RACIAL GUN VIOLENCE

ABSTRACT

- Compare rates of fatal and non-fatal gun violence between different racial groups in New York City: (Asian / Pacific Islander, Black, Black Hispanic, White, and White Hispanic)
- Majority of perpetrators reported to be Black, but shooting incidents are more likely to happen between individuals of the same race



ABOUT THE DATASET

- Data taken from the NYPD Shooting Incident (Historic) dataset.
- Contains every shooting incident from 2006 through 2023.



INITIAL OVERVIEW OF DATA

Raw dataset before cleaning:

- 28652 shooting incidents described over 21 variables
- Landing page shows descriptions of all 21 variables, but not all are needed for our analysis!



Column (Variable) Name	Description
INCIDENT_KEY	Randomly generated persistent ID for each arrest
OCCUR_DATE	Exact date of the shooting incident
OCCUR_TIME	Exact time of the shooting incident
BORO	Borough where the shooting incident occurred
LOC_OF_OCCUR_DESC	(no description available)
PRECINCT	Precinct where the shooting incident occurred
JURISDICTION_CODE	Jurisdiction where the shooting incident occurred. Jurisdiction codes: 0 = (Patrol), 1 = (Transit), and 2 = (Housing) represent NYPD whilst codes 3 and more represent non NYPD jurisdictions
LOC_CLASSFCTN_DESC	(no description available)
LOCATION_DESC	Location of the shooting incident
STATISTICAL_MURDER_FLAG	Shooting resulted in the victim's death which would be counted as a murder
PERP_AGE_GROUP	Perpetrator's age within a category
PERP_SEX	Perpetrator's sex description
PERP_RACE	Perpetrator's race description
VIC_AGE_GROUP	Victim's age within a category
VIC_SEX	Victim's sex description
VIC_RACE	Victim's race description
X_COORD_CD	Midblock X-coordinate for New York State Plane Coordinate System, Long Island Zone, NAD 83, units feet (FIPS 3104)
Y_COORD_CD	Midblock Y-coordinate for New York State Plane Coordinate System, Long Island Zone, NAD 83, units feet (FIPS 3104)
Latitude	Latitude coordinate for Global Coordinate System, WGS 1984, decimal degrees (EPSG 4326)
Longitude	Longitude coordinate for Global Coordinate System, WGS 1984, decimal degrees (EPSG 4326)
Lon_Lat	Longitude and Latitude Coordinates for mapping

CLEAN AND TRANSFORM DATA

- We removed the variables we did not need for our analysis leaving us with just four variables:
 - **OCCUR_DATE** (date the shooting incident occurred)
 - **STATISTICAL_MURDER_FLAG** (where "TRUE" indicates the shooting incident was fatal and "FALSE" indicates that it was not),
 - **PERP_RACE** (race of the perpetrator)
 - **VIC_RACE** (race of the victim)

Column (Variable) Name	Description
OCCUR_DATE	Exact date of the shooting incident
STATISTICAL_MURDER_FLAG	Shooting resulted in the victim's death which would be counted as a murder
PERP_RACE	Perpetrator's race description
VIC_RACE	Victim's race description

CLEAN AND TRANSFORM DATA

- Remove rows with missing data and invalid data entered as strings like "(null)" or "UNKNOWN"
- Remove data where **PERP_RACE** and **VIC_RACE** are reported as “AMERICAN INDIAN / ALASKAN NATIVE” due to small sample sizes (≤ 5) for each these categories
- Set **STATISTICAL_MURDER_FLAG** to a numeric object and replace "TRUE" with 1 and "FALSE" with 0.

```
##      OCCUR_DATE      STATISTICAL_MURDER_FLAG PERP_RACE
##  Min.   :2006-01-01   Min.   :0.0000          Length:16223
##  1st Qu.:2008-10-20   1st Qu.:0.0000          Class :character
##  Median :2012-11-22   Median :0.0000          Mode   :character
##  Mean    :2013-12-10   Mean    :0.2185
##  3rd Qu.:2019-04-19   3rd Qu.:0.0000
##  Max.    :2023-12-29   Max.    :1.0000

##      VIC_RACE
##  Length:16223
##  Class :character
##  Mode  :character
```

INITIAL VISUALIZATIONS

- Before we begin our formal analysis, we create two initial visualizations of our data.



INITIAL VISUALIZATIONS:

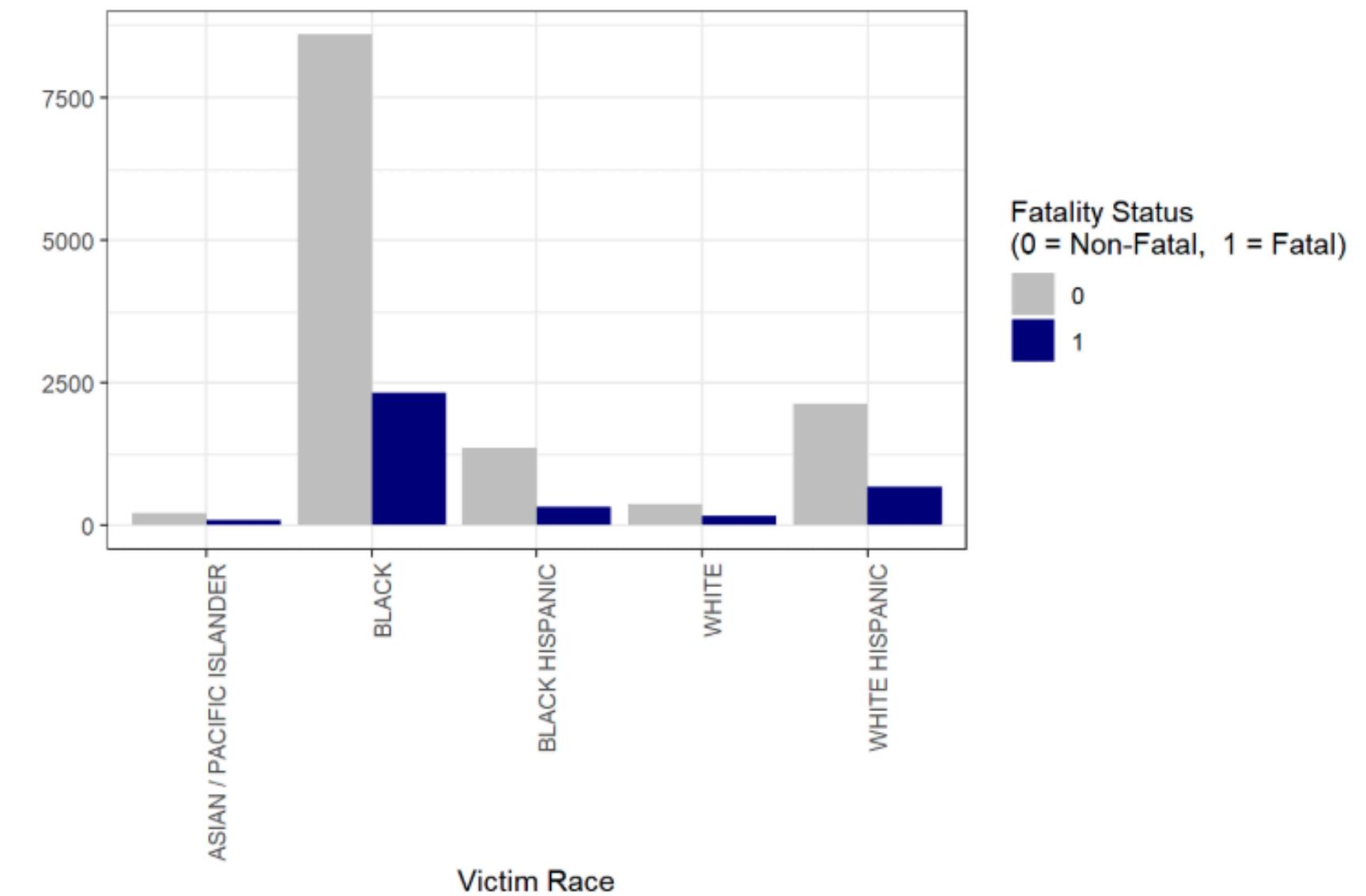
VICTIMS BY RACE

Victims by Race

- Most incidents non-fatal
- Black victims account for the majority of both non-fatal and fatal shootings.

```
## # A tibble: 10 x 4
##   VIC_RACE      STATISTICAL_MURDER_FLAG count    rate
##   <chr>          <dbl> <int>    <dbl>
## 1 ASIAN / PACIFIC ISLANDER     0     217  0.0134
## 2 ASIAN / PACIFIC ISLANDER     1      83  0.0051
## 3 BLACK             0    8597  0.530
## 4 BLACK             1    2318  0.143
## 5 BLACK HISPANIC        0   1359  0.0838
## 6 BLACK HISPANIC        1    322  0.0198
## 7 WHITE            0    371  0.0229
## 8 WHITE            1    150  0.0092
## 9 WHITE HISPANIC       0   2134  0.132
## 10 WHITE HISPANIC      1    672  0.0414
```

Shooting Incidents by Victim Race and Fatality Status



INITIAL VISUALIZATIONS: PERPETRATORS BY RACE

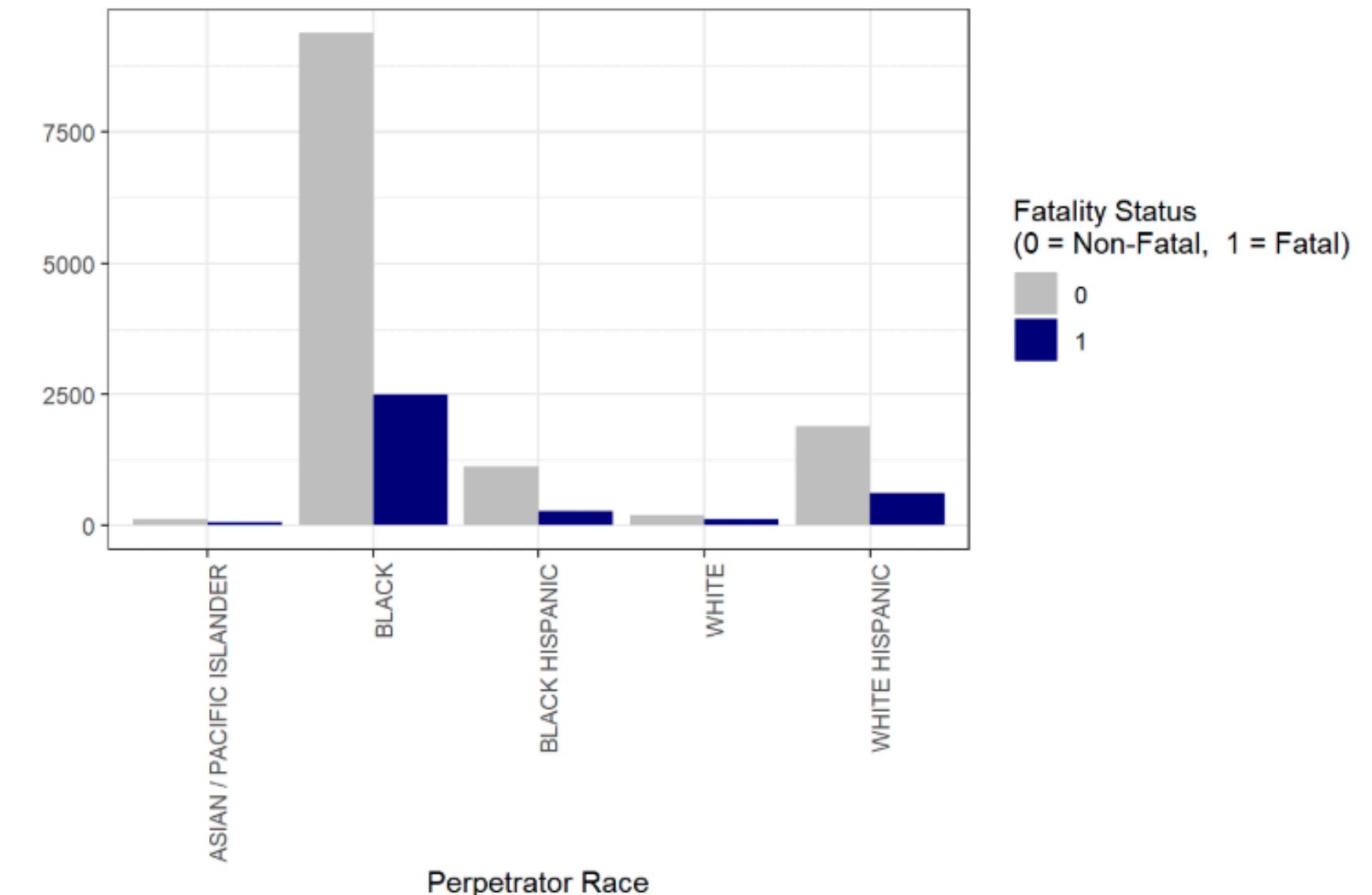
Perpetrator by Race

- Most incidents non-fatal
- Black perpetrators account for the majority of both non-fatal and fatal shootings.

The rates marked in the blue box will be the expected values for our chi-squared analysis!

## # A tibble: 10 × 4	PERP_RACE	STATISTICAL_MURDER_FLAG	count	rate
##	<chr>	<dbl>	<int>	<dbl>
## 1	ASIAN / PACIFIC ISLANDER	0	118	0.0073
## 2	ASIAN / PACIFIC ISLANDER	1	51	0.0031
## 3	BLACK	0	9384	0.578
## 4	BLACK	1	2496	0.154
## 5	BLACK HISPANIC	0	1108	0.0683
## 6	BLACK HISPANIC	1	278	0.0171
## 7	WHITE	0	186	0.0115
## 8	WHITE	1	111	0.0068
## 9	WHITE HISPANIC	0	1882	0.116
## 10	WHITE HISPANIC	1	615	0.0379

Shooting Incidents by Perpetrator Race and Fatality Status



HYPOTHESIS AND ANALYSIS:

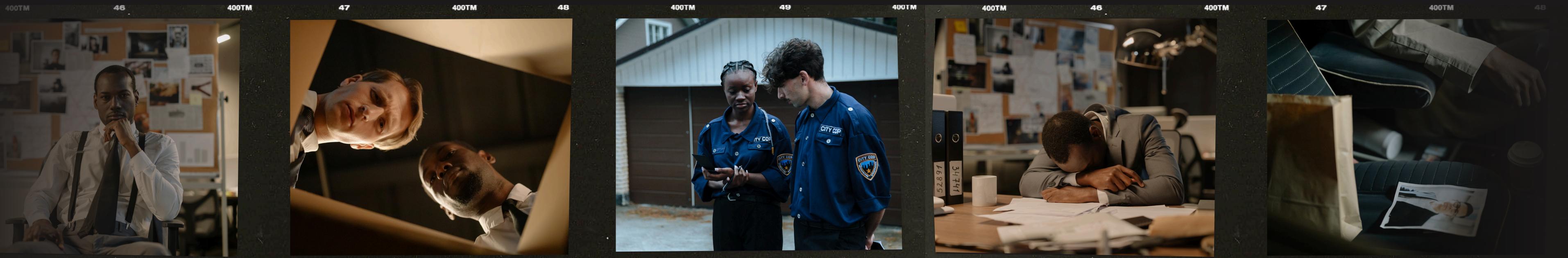
THE EXPECTED RATES FOR PERPETRATOR BY RACE ARE
NOT
REFLECTED ACROSS ALL VICTIM RACES.

- Chi-squared goodness of fit test
 - Expected Values: Rates of perpetrators by race for whole dataset
 - Observed Values: Rates of perpetrators by race for each sample by victim race
- Significance level set to $\alpha = 0.05$.



ANALYSIS:

- For victims of all races, Black perpetrators were reported to account for the majority of both non-fatal and fatal shooting incidents.
 - **BUT...**analysis shows observed rates of Black perpetrators across victim races are actually lower than expected
- Data suggests rates of victims being shot by someone of their own race are higher than expected



ASIAN / PACIFIC ISLANDER VICTIMS

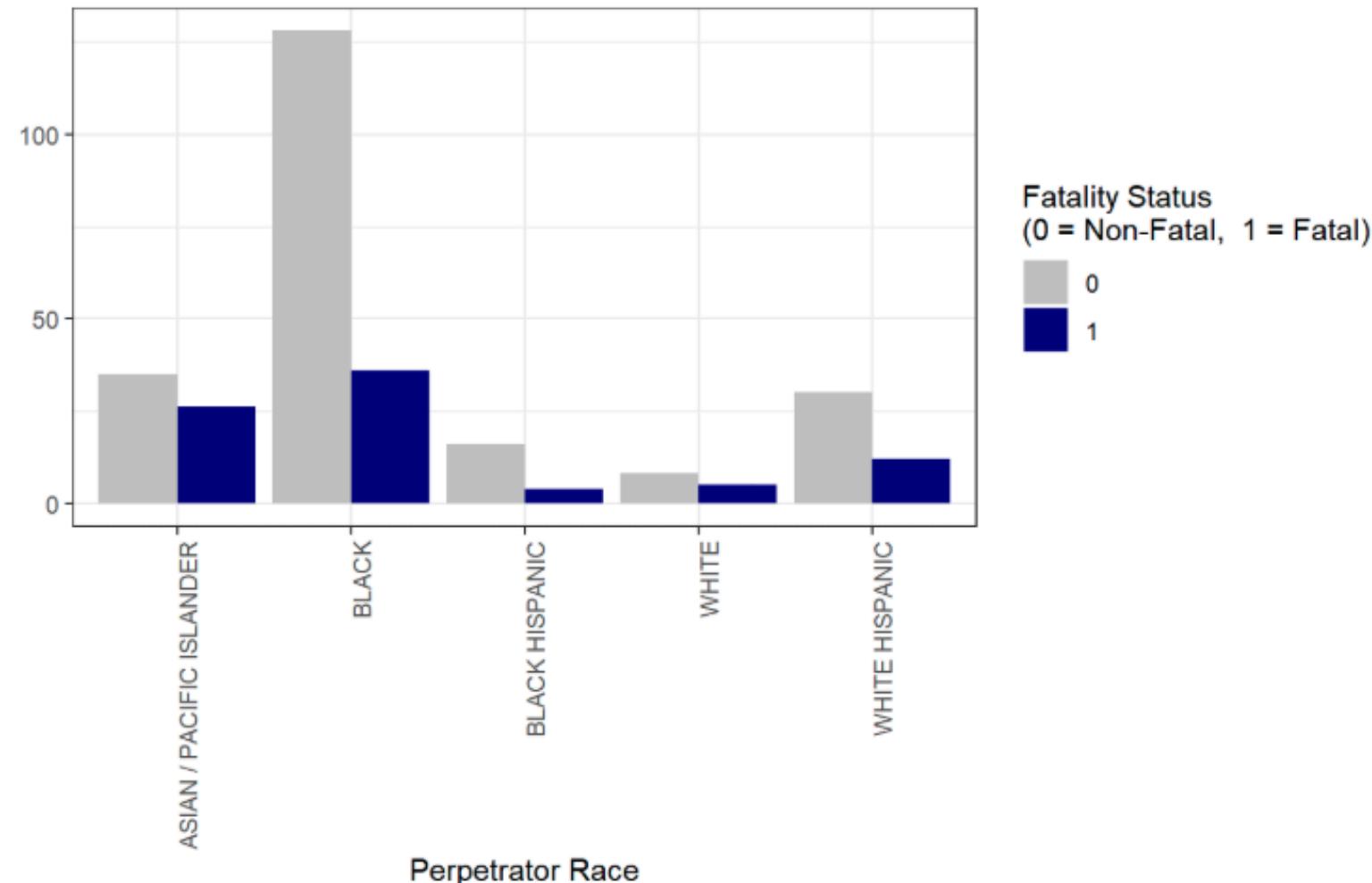
- P-value of 0.9126 suggests the city-wide rates of perpetrators by race do not reflect what we see among Asian / Pacific Islander victims
- Observed rates for both fatal and non-fatal shooting incidents due to Asian / Pacific Islander perpetrators significantly higher than expected

```

## # A tibble: 10 × 4
##   PERP_RACE   STATISTICAL_MURDER_FLAG count percentage
##   <chr>          <dbl> <int>      <dbl>
## 1 ASIAN / PACIFIC ISLANDER     0     35    0.117
## 2 ASIAN / PACIFIC ISLANDER     1     26    0.0867
## 3 BLACK             0    128    0.427
## 4 BLACK             1     36    0.12
## 5 BLACK HISPANIC       0     16    0.0533
## 6 BLACK HISPANIC       1      4    0.0133
## 7 WHITE            0      8    0.0267
## 8 WHITE            1      5    0.0167
## 9 WHITE HISPANIC       0     30    0.1
## 10 WHITE HISPANIC      1     12    0.04

```

Shooting Incidents for Asian / Pacific Islander Victims
by Perpetrator Race and Fatality Status

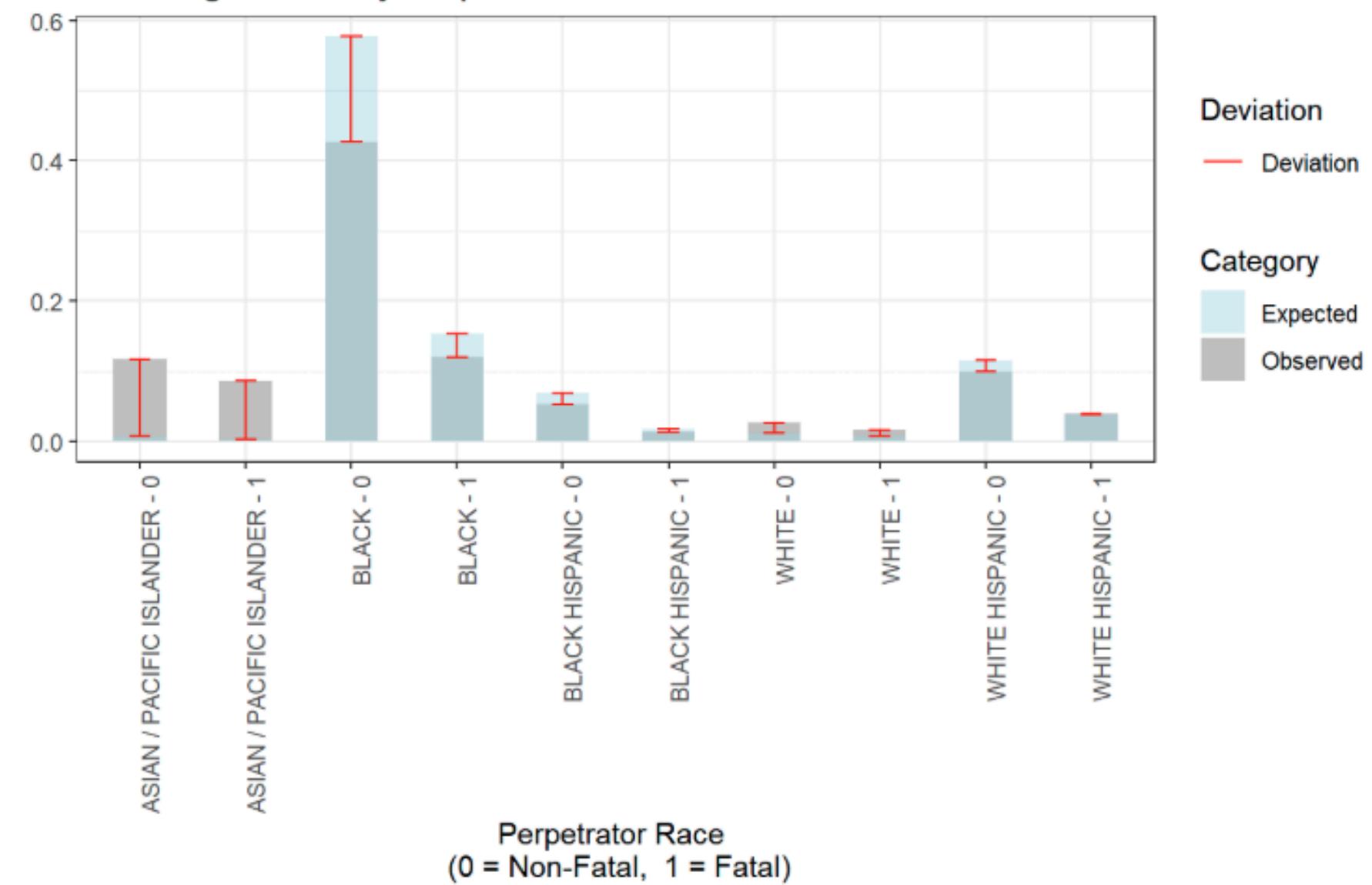


```

##           PERP_RACE Observed Expected deviation abs(deviation)
## 1 ASIAN / PACIFIC ISLANDER - 0 0.1167 0.0073 0.1094 0.1094
## 2 ASIAN / PACIFIC ISLANDER - 1 0.0867 0.0031 0.0836 0.0836
## 3 BLACK - 0 0.4267 0.5784 -0.1517 0.1517
## 4 BLACK - 1 0.1200 0.1535 -0.0335 0.0335
## 5 BLACK HISPANIC - 0 0.0533 0.0683 -0.0150 0.0150
## 6 BLACK HISPANIC - 1 0.0133 0.0171 -0.0038 0.0038
## 7 WHITE - 0 0.0267 0.0115 0.0152 0.0152
## 8 WHITE - 1 0.0167 0.0068 0.0099 0.0099
## 9 WHITE HISPANIC - 0 0.1000 0.1160 -0.0160 0.0160
## 10 WHITE HISPANIC - 1 0.0400 0.0379 0.0021 0.0021

```

Observed vs. Expected Rates for Asian / Pacific Islander Shooting Victims by Perpetrator Race



BLACK VICTIMS

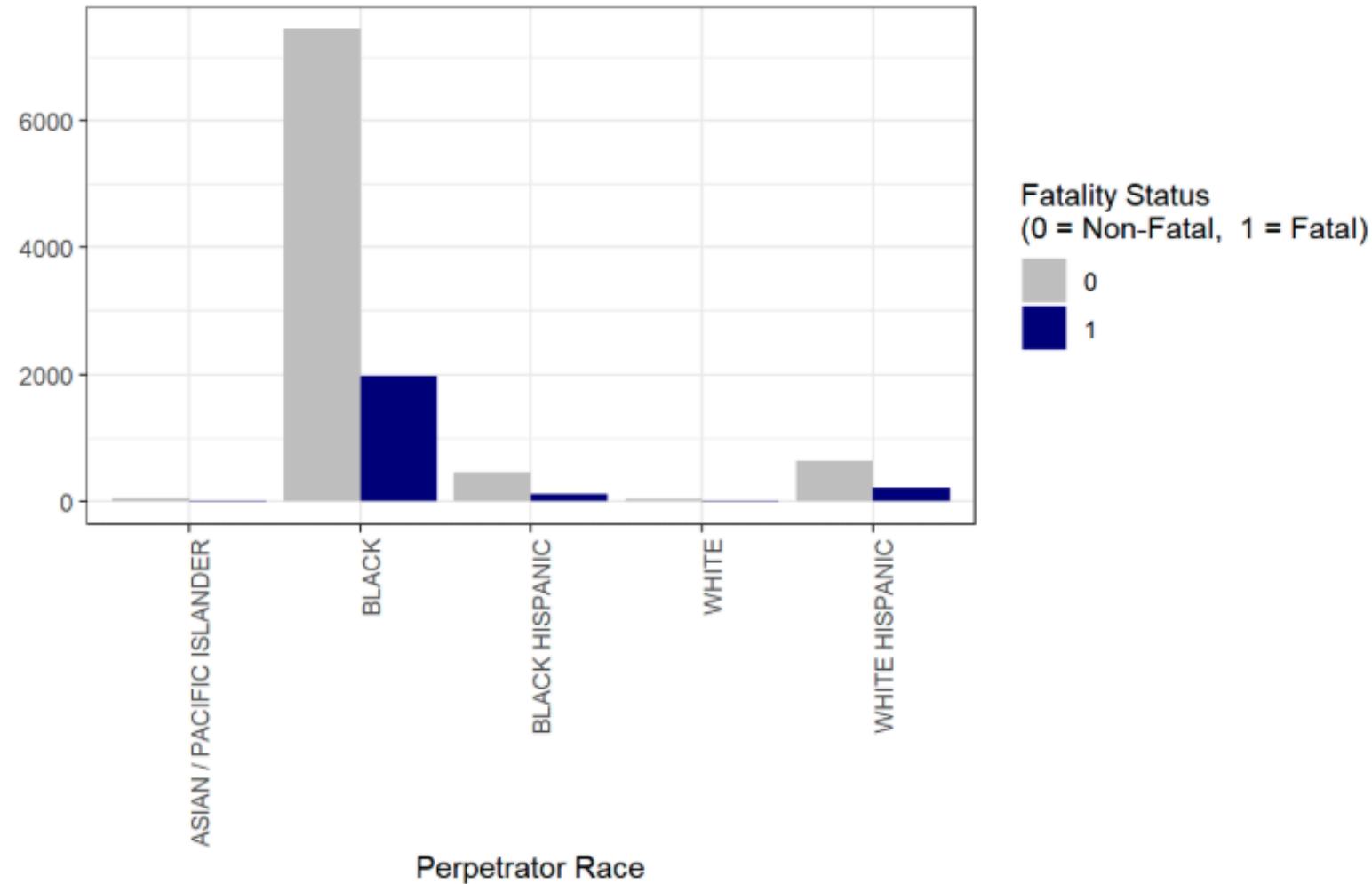
- P-value of 0.9999 suggests the city-wide rates of perpetrators by race do not reflect what we see among Black victims
- Observed rate for non-fatal shooting incidents due to Black perpetrators significantly higher than expected

```

## # A tibble: 10 × 4
##   PERP_RACE    STATISTICAL_MURDER_FLAG count percentage
##   <chr>          <dbl> <int>      <dbl>
## 1 ASIAN / PACIFIC ISLANDER        0     44      0.004
## 2 ASIAN / PACIFIC ISLANDER        1     12      0.0011
## 3 BLACK                         0    7435      0.681
## 4 BLACK                         1    1976      0.181
## 5 BLACK HISPANIC                 0     450      0.0412
## 6 BLACK HISPANIC                 1     111      0.0102
## 7 WHITE                         0     33      0.003
## 8 WHITE                         1      9      0.0008
## 9 WHITE HISPANIC                 0     635      0.0582
## 10 WHITE HISPANIC                1    210      0.0192

```

Shooting Incidents for Black Victims
by Perpetrator Race and Fatality Status

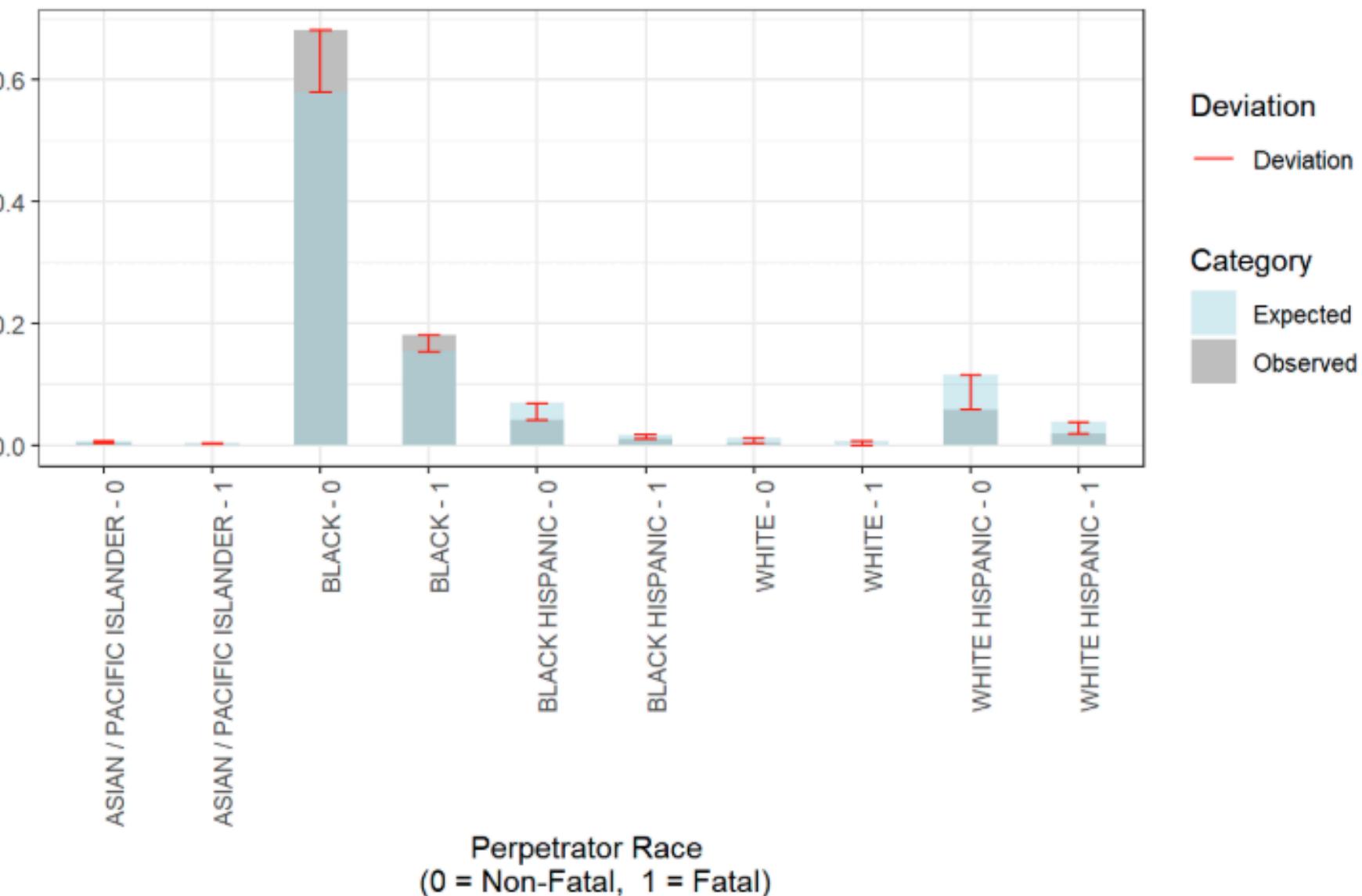


```

##           PERP_RACE Observed Expected deviation abs(deviation)
## 1 ASIAN / PACIFIC ISLANDER - 0  0.0040  0.0073 -0.0033  0.0033
## 2 ASIAN / PACIFIC ISLANDER - 1  0.0011  0.0031 -0.0020  0.0020
## 3 BLACK - 0  0.6812  0.5784  0.1028  0.1028
## 4 BLACK - 1  0.1810  0.1535  0.0275  0.0275
## 5 BLACK HISPANIC - 0  0.0412  0.0683 -0.0271  0.0271
## 6 BLACK HISPANIC - 1  0.0102  0.0171 -0.0069  0.0069
## 7 WHITE - 0  0.0030  0.0115 -0.0085  0.0085
## 8 WHITE - 1  0.0008  0.0068 -0.0060  0.0060
## 9 WHITE HISPANIC - 0  0.0582  0.1160 -0.0578  0.0578
## 10 WHITE HISPANIC - 1  0.0192  0.0379 -0.0187  0.0187

```

Observed vs. Expected Rates for Black
Shooting Victims by Perpetrator Race



BLACK HISPANIC VICTIMS

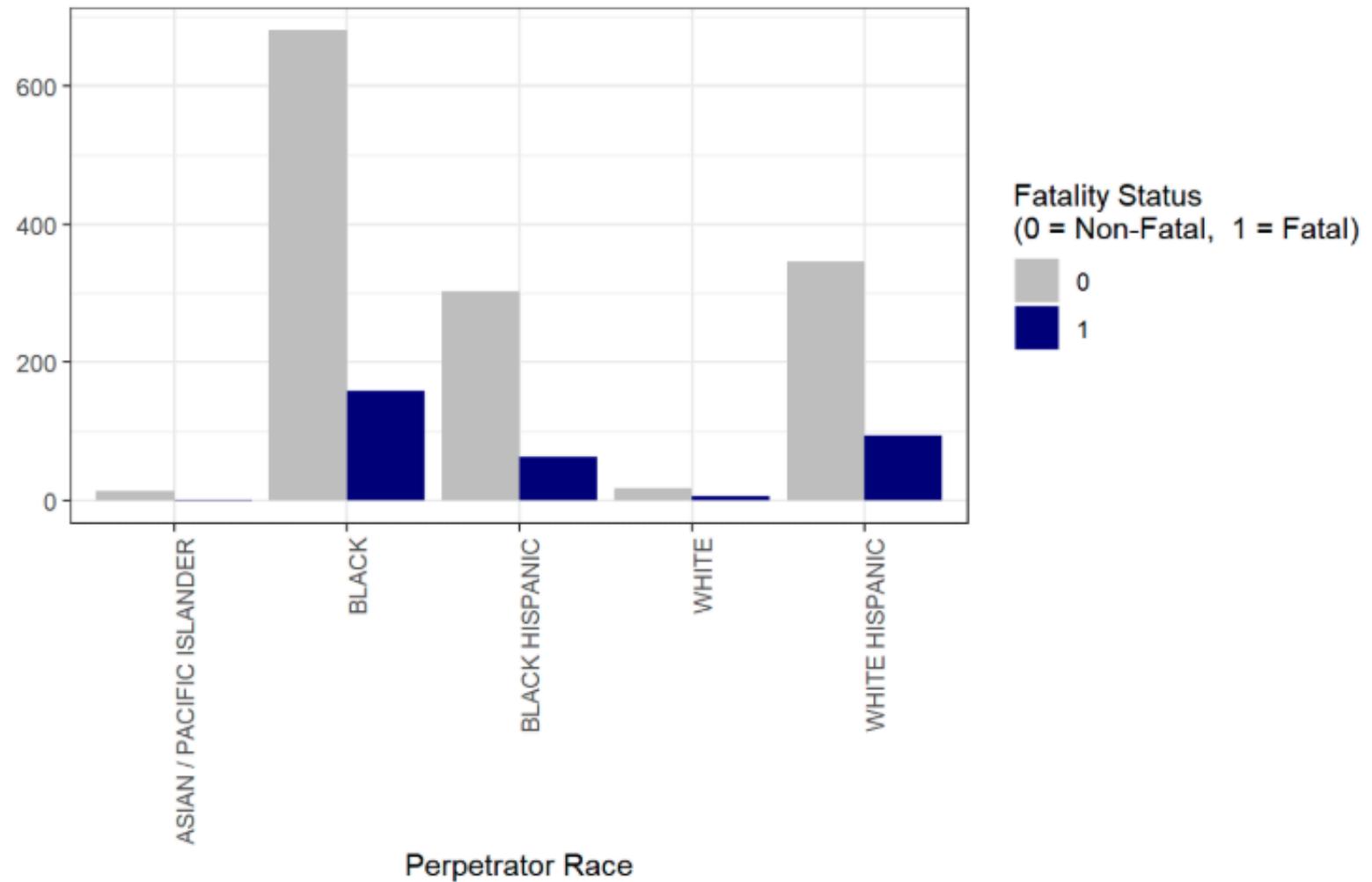
- P-value of 0.9999 suggests the city-wide rates of perpetrators by race do not reflect what we see among Black Hispanic victims
- Observed rates for both fatal and non-fatal shooting incidents due to Black Hispanic perpetrators significantly higher than expected

```

## # A tibble: 10 × 4
##   PERP_RACE STATISTICAL_MURDER_FLAG count percentage
##   <chr>      <dbl> <int>     <dbl>
## 1 ASIAN / PACIFIC ISLANDER 0       13     0.0077
## 2 ASIAN / PACIFIC ISLANDER 1       1      0.0006
## 3 BLACK        0       680     0.404
## 4 BLACK        1      159     0.0946
## 5 BLACK HISPANIC 0      302     0.180
## 6 BLACK HISPANIC 1      63      0.0375
## 7 WHITE         0      18      0.0107
## 8 WHITE         1       5      0.003
## 9 WHITE HISPANIC 0      346     0.206
## 10 WHITE HISPANIC 1      94      0.0559

```

Shooting Incidents for Black Hispanic Victims
by Perpetrator Race and Fatality Status

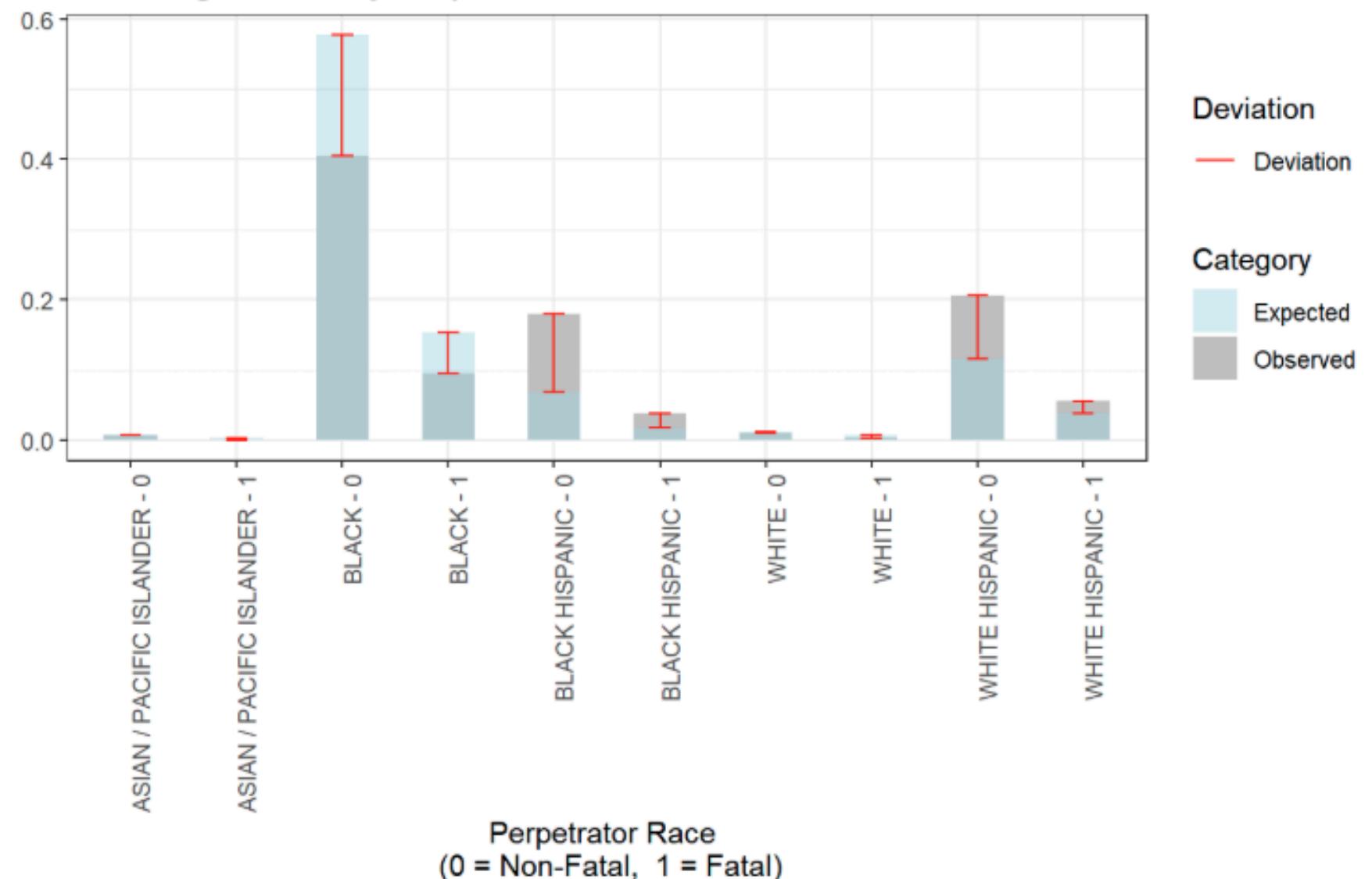


```

##          PERP_RACE Observed Expected deviation abs(deviation)
## 1 ASIAN / PACIFIC ISLANDER - 0 0.0077 0.0073 0.0004 0.0004
## 2 ASIAN / PACIFIC ISLANDER - 1 0.0006 0.0031 -0.0025 0.0025
## 3 BLACK        - 0 0.4045 0.5784 -0.1739 0.1739
## 4 BLACK        - 1 0.0946 0.1535 -0.0589 0.0589
## 5 BLACK HISPANIC - 0 0.1797 0.0683 0.1114 0.1114
## 6 BLACK HISPANIC - 1 0.0375 0.0171 0.0204 0.0204
## 7 WHITE         - 0 0.0107 0.0115 -0.0008 0.0008
## 8 WHITE         - 1 0.0030 0.0068 -0.0038 0.0038
## 9 WHITE HISPANIC - 0 0.2058 0.1160 0.0898 0.0898
## 10 WHITE HISPANIC - 1 0.0559 0.0379 0.0180 0.0180

```

Observed vs. Expected Rates for Black Hispanic Shooting Victims by Perpetrator Race



WHITE VICTIMS

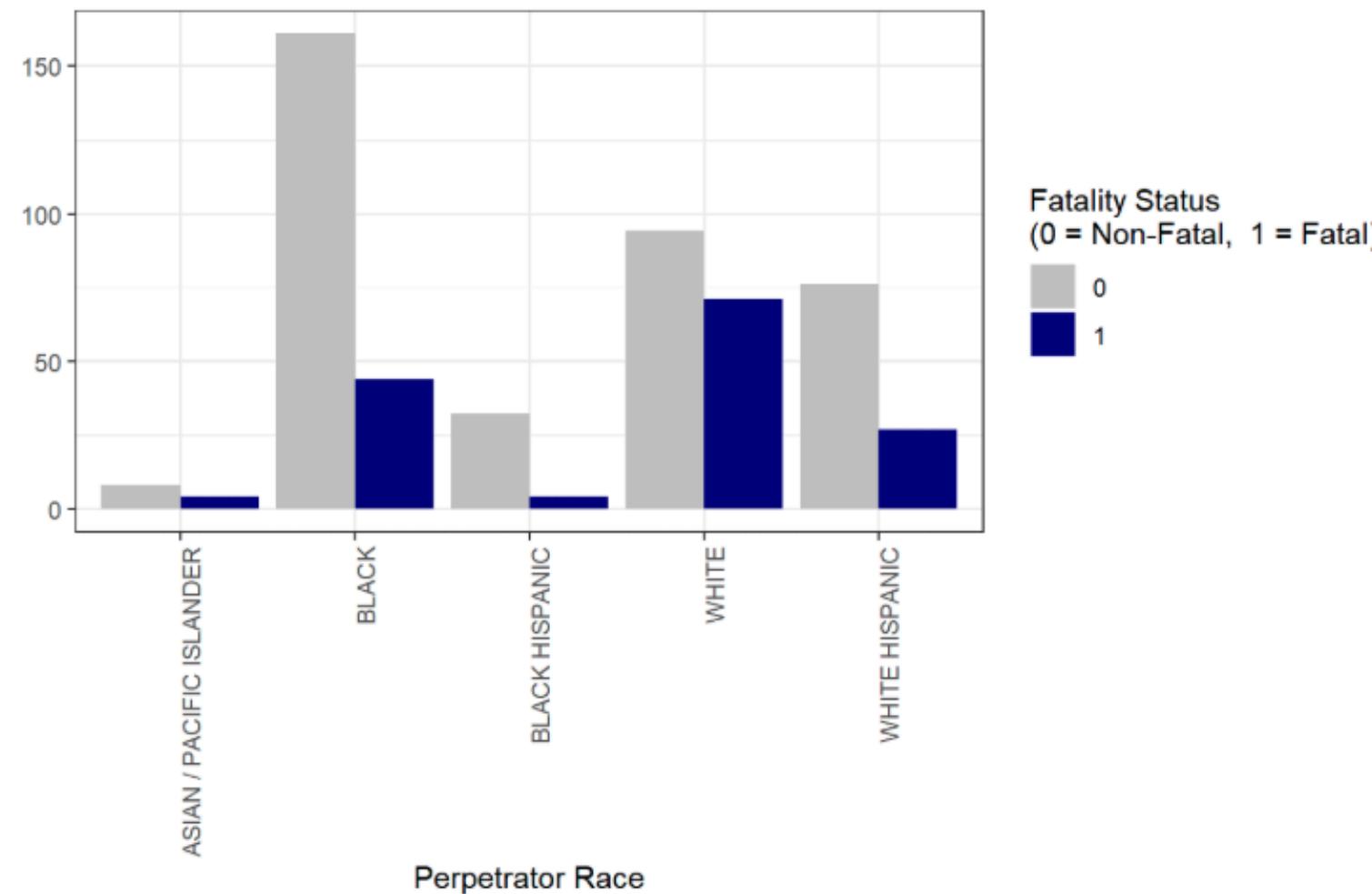
- P-value of 0.8229 suggests the city-wide rates of perpetrators by race do not reflect what we see among White victims
- Observed rates for both fatal and non-fatal shooting incidents due to White perpetrators significantly higher than expected

```

## # A tibble: 10 × 4
##   PERP_RACE   STATISTICAL_MURDER_FLAG count    rate
##   <chr>          <dbl> <int>    <dbl>
## 1 ASIAN / PACIFIC ISLANDER      0     8  0.0154
## 2 ASIAN / PACIFIC ISLANDER      1     4  0.0077
## 3 BLACK             0    161  0.309
## 4 BLACK             1     44  0.0845
## 5 BLACK HISPANIC        0    32  0.0614
## 6 BLACK HISPANIC        1     4  0.0077
## 7 WHITE            0    94  0.180
## 8 WHITE            1    71  0.136
## 9 WHITE HISPANIC       0    76  0.146
## 10 WHITE HISPANIC      1    27  0.0518

```

Shooting Incidents for White Victims
by Perpetrator Race and Fatality Status

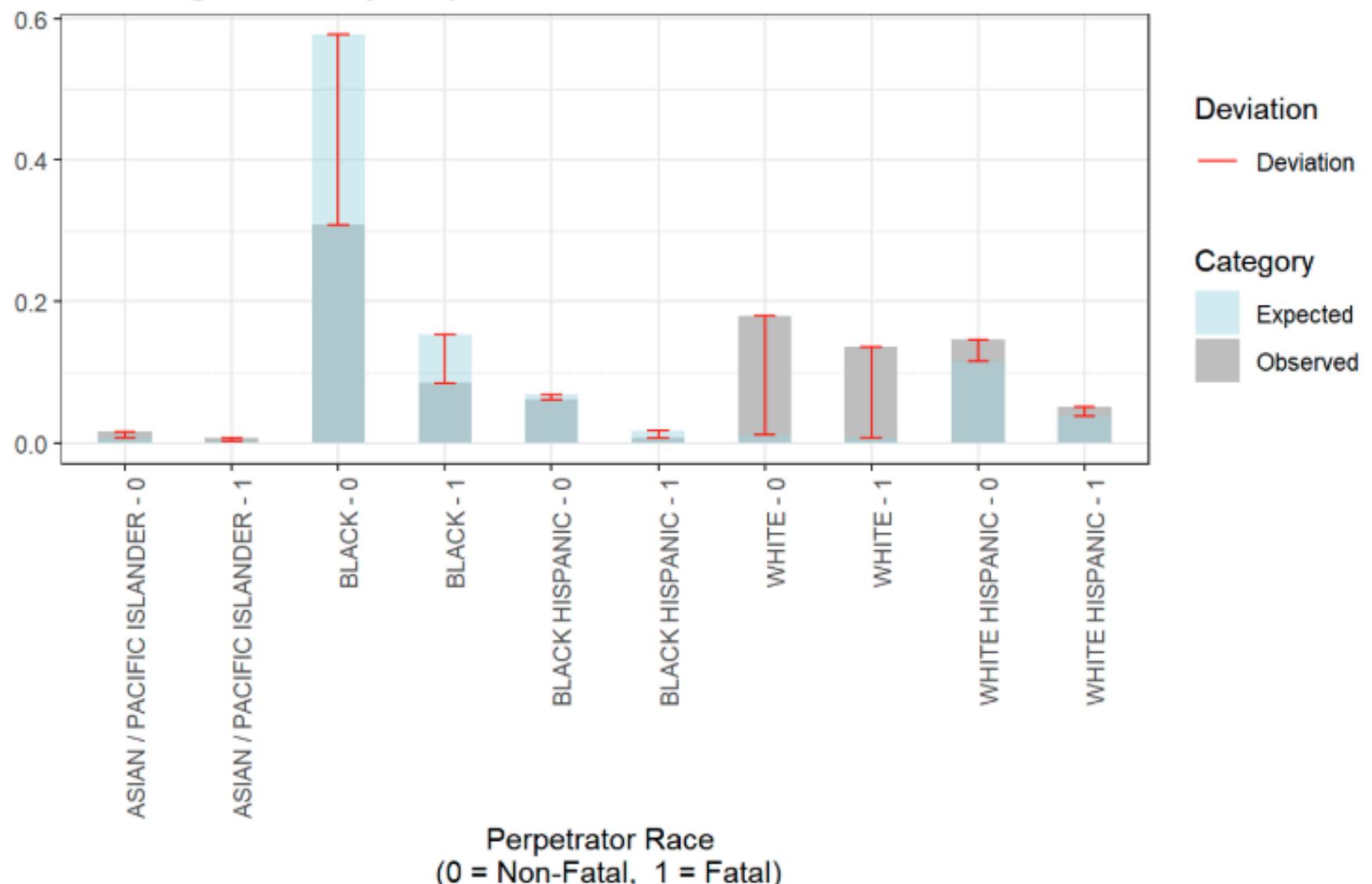


```

##           PERP_RACE Observed Expected deviation abs(deviation)
## 1 ASIAN / PACIFIC ISLANDER - 0 0.0154 0.0073 0.0081 0.0081
## 2 ASIAN / PACIFIC ISLANDER - 1 0.0077 0.0031 0.0046 0.0046
## 3 BLACK - 0 0.3090 0.5784 -0.2694 0.2694
## 4 BLACK - 1 0.0845 0.1535 -0.0690 0.0690
## 5 BLACK HISPANIC - 0 0.0614 0.0683 -0.0069 0.0069
## 6 BLACK HISPANIC - 1 0.0077 0.0171 -0.0094 0.0094
## 7 WHITE - 0 0.1804 0.0115 0.1689 0.1689
## 8 WHITE - 1 0.1363 0.0068 0.1295 0.1295
## 9 WHITE HISPANIC - 0 0.1459 0.1160 0.0299 0.0299
## 10 WHITE HISPANIC - 1 0.0518 0.0379 0.0139 0.0139

```

Observed vs. Expected Rates for White
Shooting Victims by Perpetrator Race



WHITE HISPANIC VICTIMS

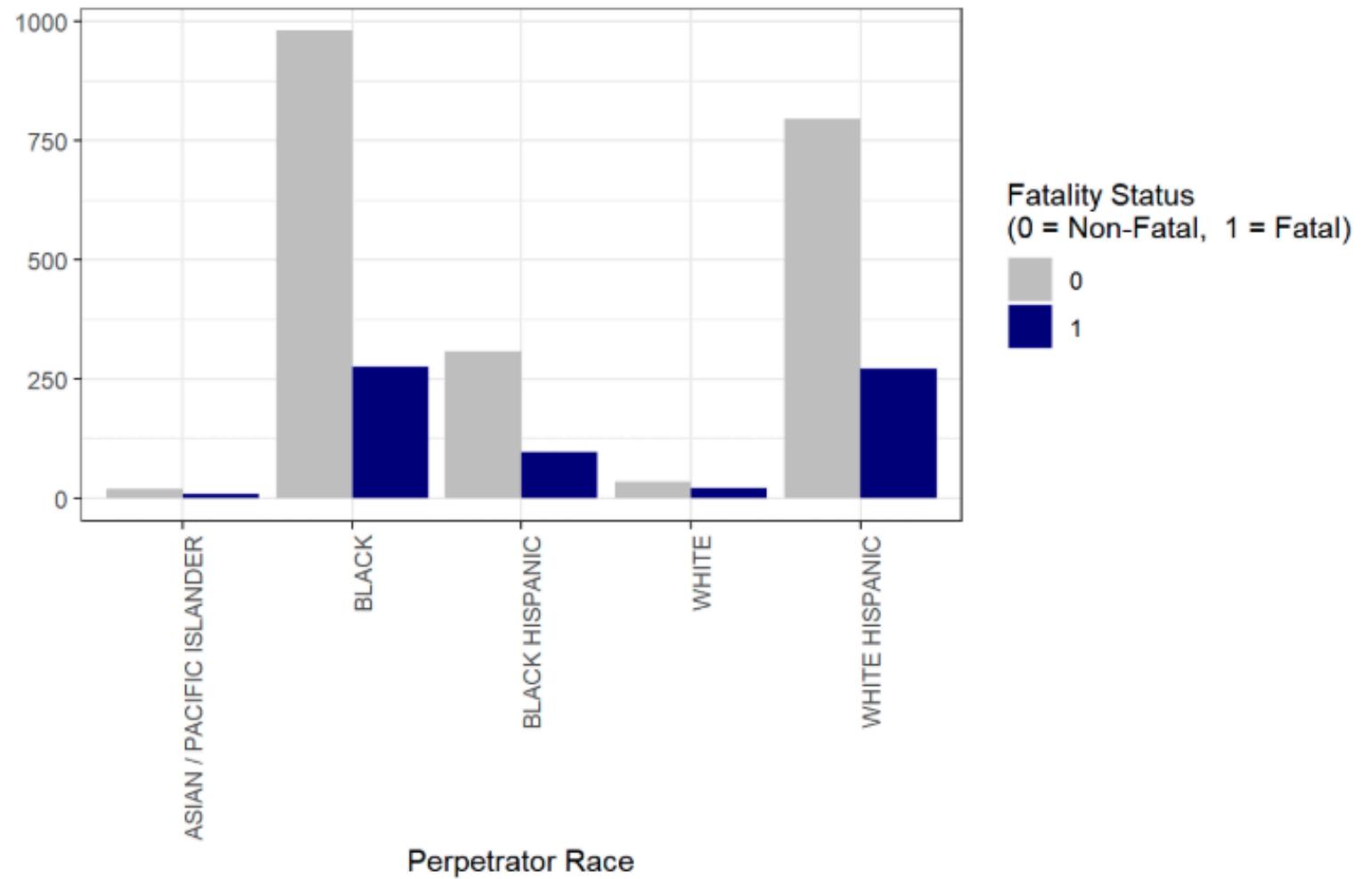
- P-value of 0.9999 suggests the city-wide rates of perpetrators by race do not reflect what we see among White Hispanic victims
- Observed rates for both fatal and non-fatal shooting incidents due to White Hispanic perpetrators significantly higher than expected

```

## # A tibble: 10 × 4
##   PERP_RACE    STATISTICAL_MURDER_FLAG count    rate
##   <chr>          <dbl> <int> <dbl>
## 1 ASIAN / PACIFIC ISLANDER      0     18  0.0064
## 2 ASIAN / PACIFIC ISLANDER      1      8  0.0029
## 3 BLACK             0    980  0.349
## 4 BLACK             1    275  0.098
## 5 BLACK HISPANIC        0    308  0.110
## 6 BLACK HISPANIC        1     96  0.0342
## 7 WHITE            0     33  0.0118
## 8 WHITE            1     21  0.0075
## 9 WHITE HISPANIC       0    795  0.283
## 10 WHITE HISPANIC       1    272  0.0969

```

Shooting Incidents for White Hispanic Victims
by Perpetrator Race and Fatality Status

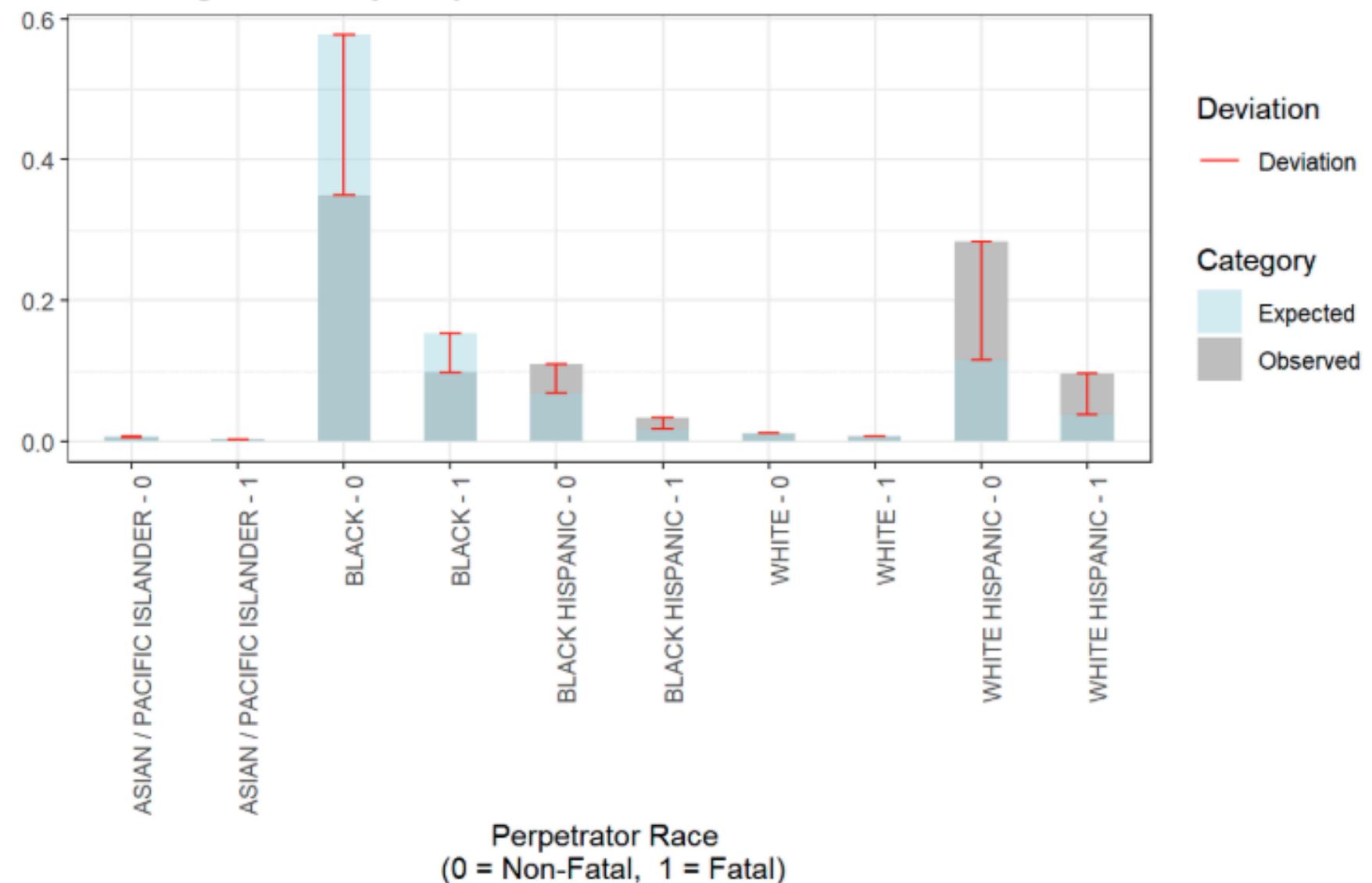


```

##           PERP_RACE Observed Expected deviation abs(deviation)
## 1 ASIAN / PACIFIC ISLANDER - 0  0.0064  0.0073 -0.0009   0.0009
## 2 ASIAN / PACIFIC ISLANDER - 1  0.0029  0.0031 -0.0002   0.0002
## 3 BLACK             - 0  0.3493  0.5784 -0.2291   0.2291
## 4 BLACK             - 1  0.0980  0.1535 -0.0555   0.0555
## 5 BLACK HISPANIC - 0  0.1098  0.0683  0.0415   0.0415
## 6 BLACK HISPANIC - 1  0.0342  0.0171  0.0171   0.0171
## 7 WHITE            - 0  0.0118  0.0115  0.0003   0.0003
## 8 WHITE            - 1  0.0075  0.0068  0.0007   0.0007
## 9 WHITE HISPANIC - 0  0.2833  0.1160  0.1673   0.1673
## 10 WHITE HISPANIC - 1  0.0969  0.0379  0.0590   0.0590

```

Observed vs. Expected Rates for White Hispanic
Shooting Victims by Perpetrator Race



Victim Race	Incident Type	Significant Deviations in			p-value
		Sample	Prepetrator Race		
Asian / Pacific Islander	Non-Fatal	-0.1517	Black	0.9126	
Asian / Pacific Islander	Non-Fatal	0.1094	Asian / Pacific Islander	0.9126	
Asian / Pacific Islander	Fatal	0.0836	Asian / Pacific Islander	0.9126	
Black	Non-Fatal	0.1028	Black	0.9999	
Black	Non-Fatal	-0.0578	White Hispanic	0.9999	
Black Hispanic	Non-Fatal	-0.1739	Black	0.9999	
Black Hispanic	Non-Fatal	0.1114	Black Hispanic	0.9999	
Black Hispanic	Non-Fatal	0.0898	White Hispanic	0.9999	
Black Hispanic	Fatal	-0.0589	Black	0.9999	
White	Non-Fatal	-0.2694	Black	0.8229	
White	Non-Fatal	0.1689	White	0.8229	
White	Fatal	0.1295	White	0.8229	
White	Fatal	-0.0690	Black	0.8229	
White Hispanic	Non-Fatal	-0.2291	Black	0.9999	
White Hispanic	Non-Fatal	0.1673	White Hispanic	0.9999	
White Hispanic	Fatal	0.0590	White Hispanic	0.9999	
White Hispanic	Fatal	-0.0555	Black	0.9999	

ADDITIONAL QUESTIONS

- Is this data reflective of New York City's population demographics?



CONCLUSIONS

- For both non-fatal and fatal shooting events, victims were more likely to be shot by perpetrators of their own race.
- All visualizations showed that the the majority of perpetrators were Black, but they may be over-represented in the current data.
- Overall rates of perpetrators by race did not fit neatly into what we saw with each sample when divided by victim race.
- Significantly lower p-value for White victims suggests that statistics may be under-reported for this group.



POSSIBLE SOURCES OF BIAS

- Inter-racial violence is more commonly portrayed in media, not the intra-racial violence seen in this study.
 - Communities where violence occurs may have higher levels of intra-racial interaction than inter-racial interaction.





FULL RMD FILE AT

[HTTPS://GITHUB.COM/APALOMER86/DTS...
RMD](https://github.com/apalomer86/dtsa_3301.NYPD_SHOOTING_DATA_PROJECT/blob/main/NYPD%20SHOOTING%20DATA%20PROJECT.RMD)