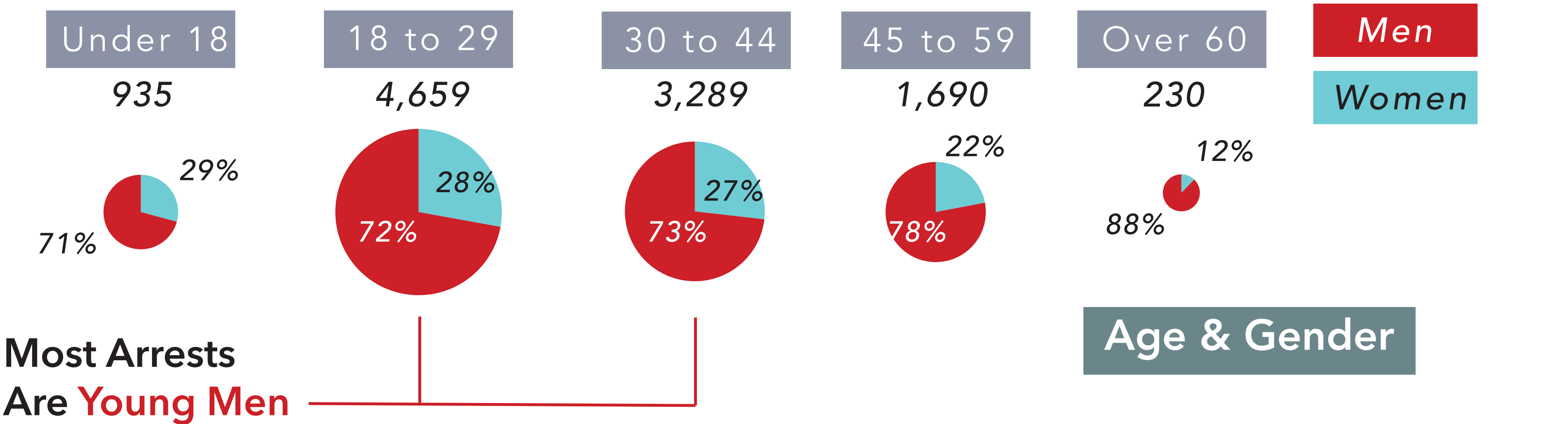
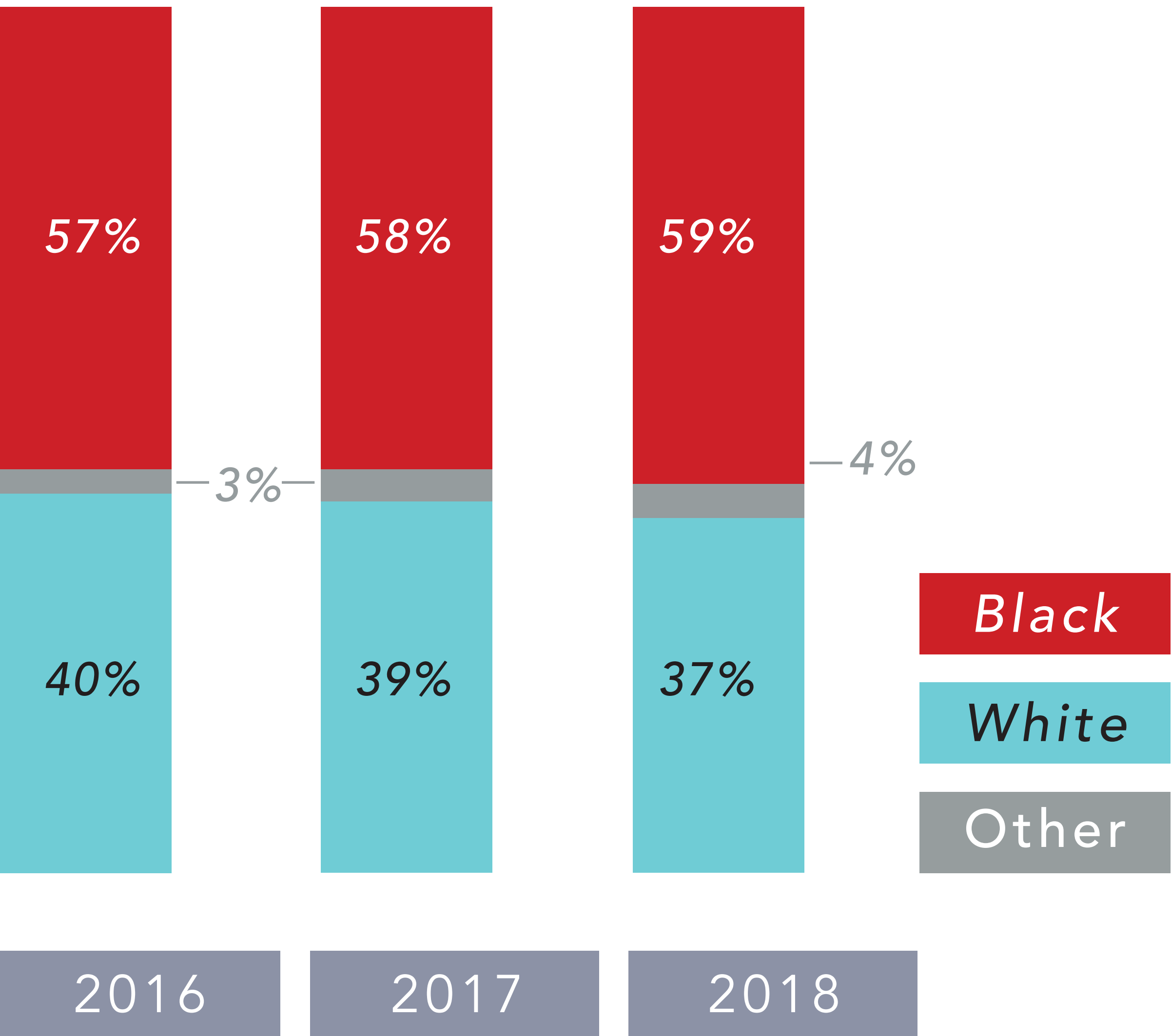


Pittsburgh Crime Stats



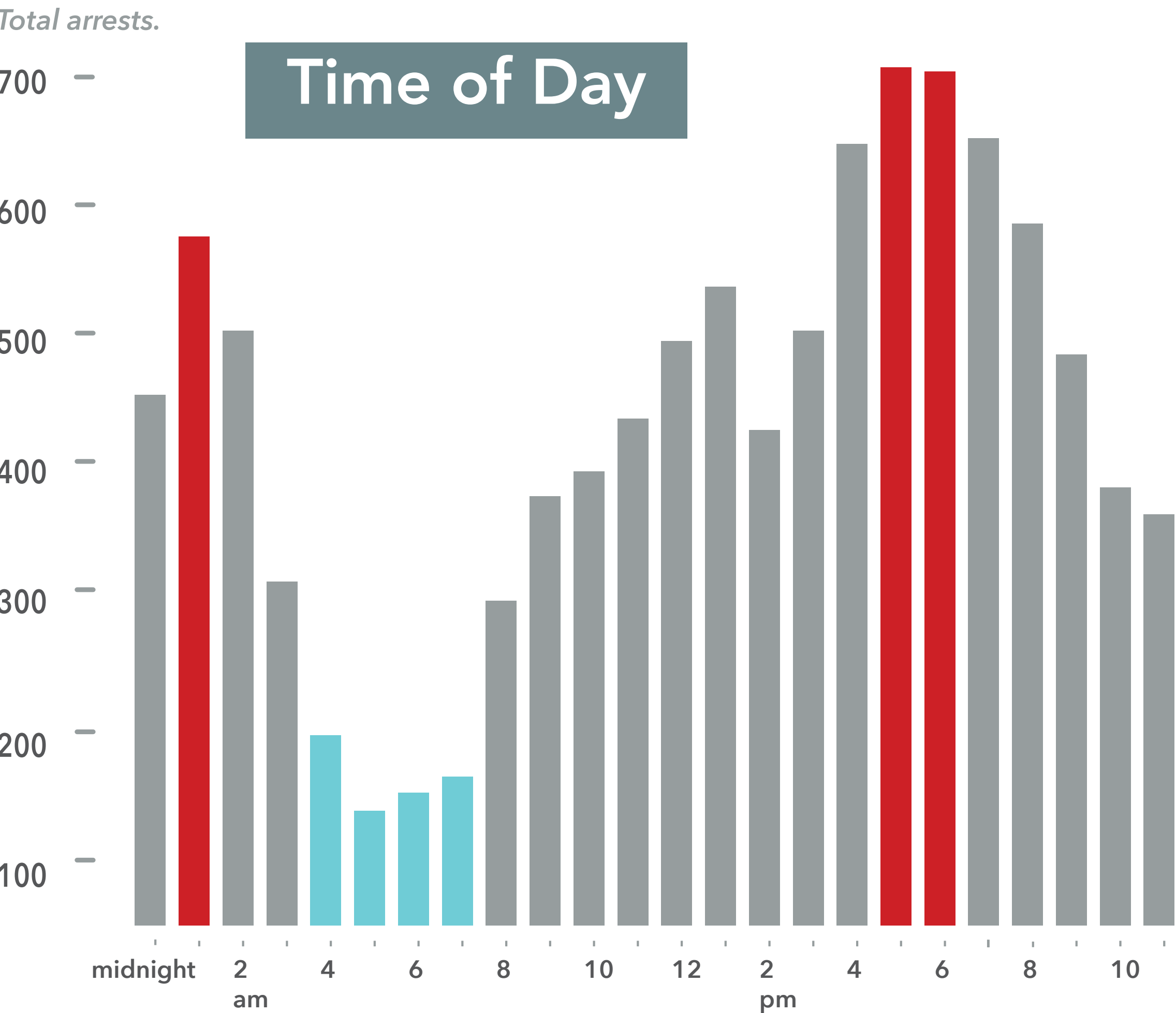
Racial Factors



* The arrest data was recorded sporadically for all years except 2017. To account for this, the arrest data was normalized to show percentage. Percentage was rounded to the nearest whole number.

The percentage of arrests shows little variation between years.*

Except for a few **high-crime** areas, most neighborhoods in Pittsburgh are **safe**.



Arrests have **two peaks** throughout the day: one after midnight and the other between five and seven pm.

Project Members:
Zinia Kamal,
John Casola,
Joe Hernandez,
Anthony Olivieri

Data Description

The final data set consisted of 10,803 rows and 10 columns. Incomplete cases, cases outside the Pittsburgh area, and cases where the gender was unidentified were dropped. Arrests where the race was neither white or black were categorized as "O" for "other", and age outliers were set to the median age.

Data Source

The data set was retrieved from Data.gov (<https://catalog.data.gov/dataset/pittsburgh-police-arrest-data>) the U.S. government's web-site for open data, tools, and resources. The R packages used were ggplot2, scales, ggmap, and dplyr. Tableau and Illustrator were also used for visualization.