



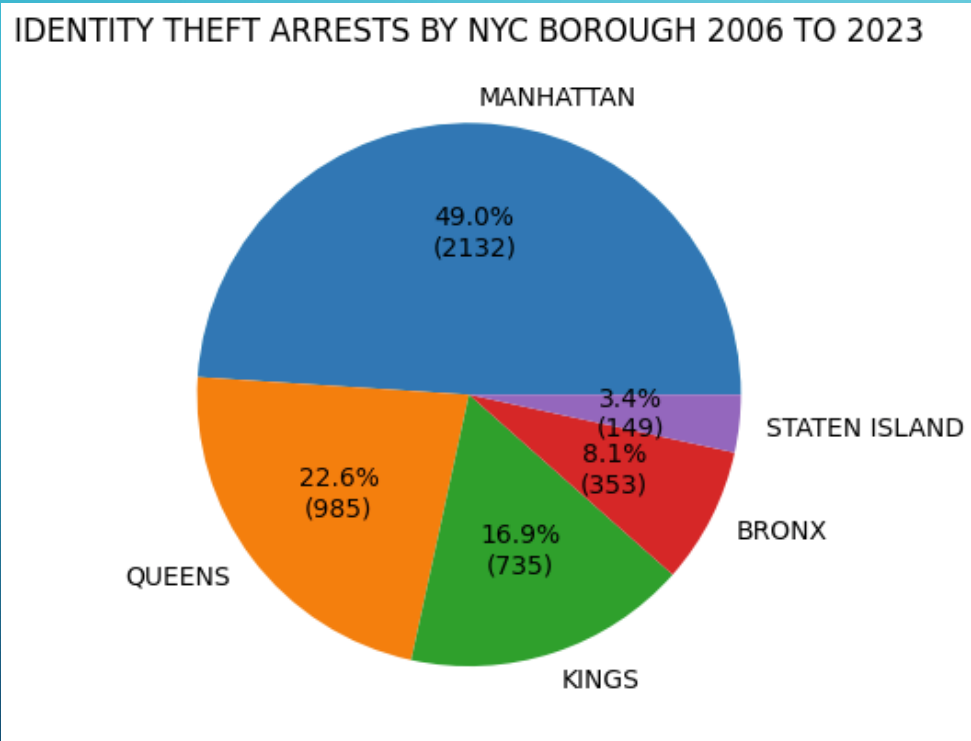
DATA 201 CAPSTONE

ERIK RIGBY

DATA 201 CAPSTONE

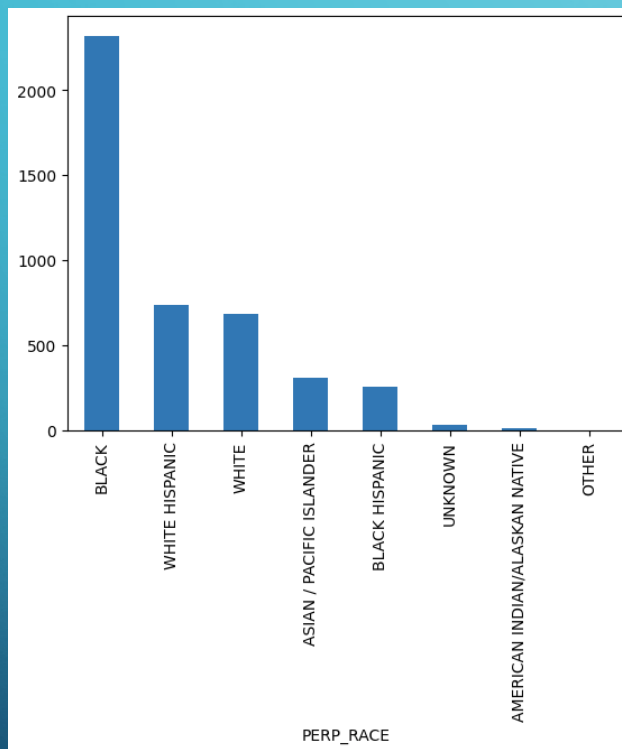
- The focus is on Identity Theft in New York City using historical NYPD data from 2006 to April 2023. Does this data align with the Federal Trade Commission's ("FTC") national or metropolitan regional average?
- The dataset is 5.5 million rows and 19 columns.
- My interest is two-fold; I am a Certified Fraud Examiner ("CFE") and I have had to file my own fraud claims, specifically, in the summer of 2020, when someone filed a fraudulent unemployment claim in my name.

PRIMARY FINDINGS

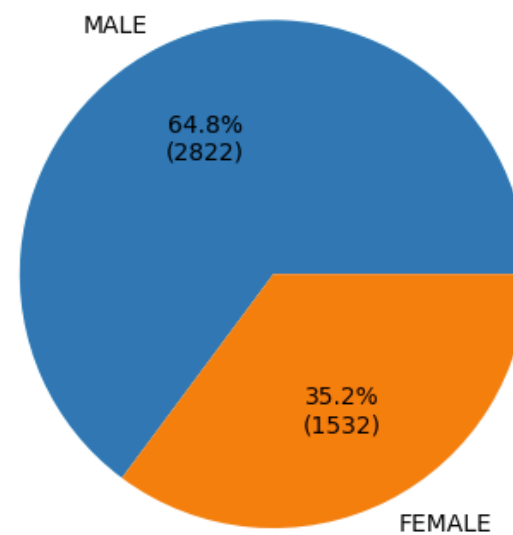


- Almost half of the Identity Theft Arrests in New York City occurred in the Borough of Manhattan.
- Black males were arrested at a rate of almost three times that of other races.
- While other arrest categories continued to be reported by the NYPD through April 2023, Identity Theft no longer appeared after December 2018.

NYC IDENTITY THEFT ARRESTS BY RACE AND GENDER



NEW YORK CITY IDENTITY THEFT ARRESTS BY GENDER 2006 TO 2023



FTC 2019 NY METRO ID THEFT AVERAGE 225/100K

FTC CONSUMER SENTINEL NETWORK

Published November 1, 2023
(data as of September 30, 2023)

New York-Newark-Jersey City, NY-NJ-PA Metro Area
Year: 2019

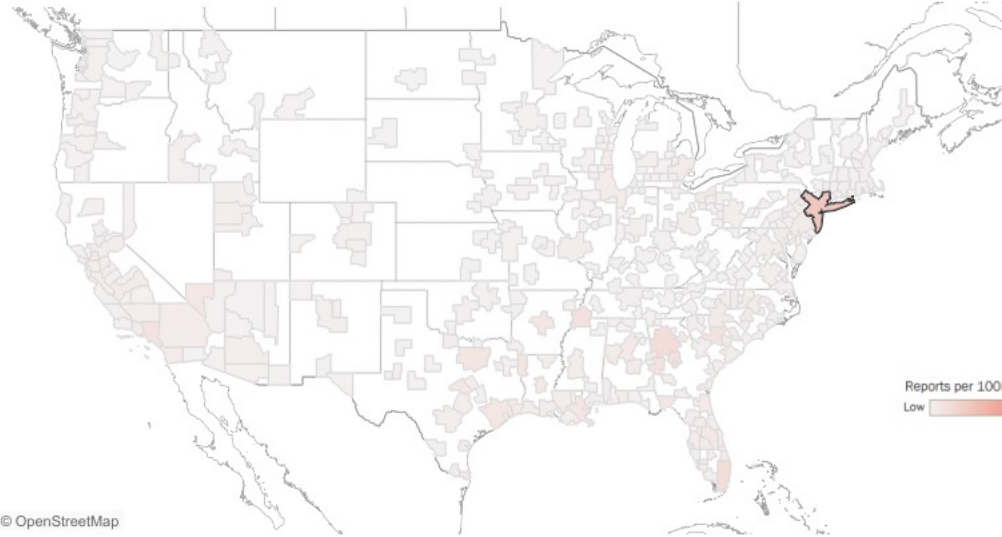
View
Map

Year
2019

Quarter
All



© 2024 Mapbox © OpenStreetMap



Reports per 100K Population
Low High

Theft Type

Credit Card Fraud	24,128
Other Identity Theft	11,263
Phone or Utilities Fraud	5,930
Loan or Lease Fraud	5,655
Bank Fraud	3,984
Employment or Tax-Related Fraud	2,122
Government Documents or Benefits Fraud	1,383

44,962

Total Identity Theft Reports

225

Reports per 100K Population

Metropolitan Areas are defined by the Office of Management and Budget, and population estimates are based on 2021 U.S. Census figures. Metropolitan Areas with a population of 100,000 or more are identified (identified Wyoming Metropolitan Areas have a population of 99,000 or more).

FEDERAL TRADE COMMISSION · ftc.gov/exploredata

SUMMARY

- NYPD Identity Theft reporting ended in 2018 and did not mirror what was being reported by the FTC. The data appeared flawed or was captured elsewhere.
- Libraries: Pandas, NumPy, Matplotlib
- Data Cleaning: Dropped unnecessary column; Replaced null/NaN data; Assigned DateTime value to column
- Wrangling: New data frames; Renamed column headers; Filtered by string contains
- Visualizations: Bar and Pie Charts

REFERENCES

- <https://data.cityofnewyork.us/Public-Safety/NYPD-Arrests-Data-Historic-/8h9b-rp9u>
- <https://public.tableau.com/app/profile/federal.trade.commission/viz/FraudandIDTheftMaps/FraudbyMetroArea>
- https://www.ic3.gov/Media/PDF/AnnualReport/2022_IC3Report.pdf
- <https://identitytheft.org/statistics/>