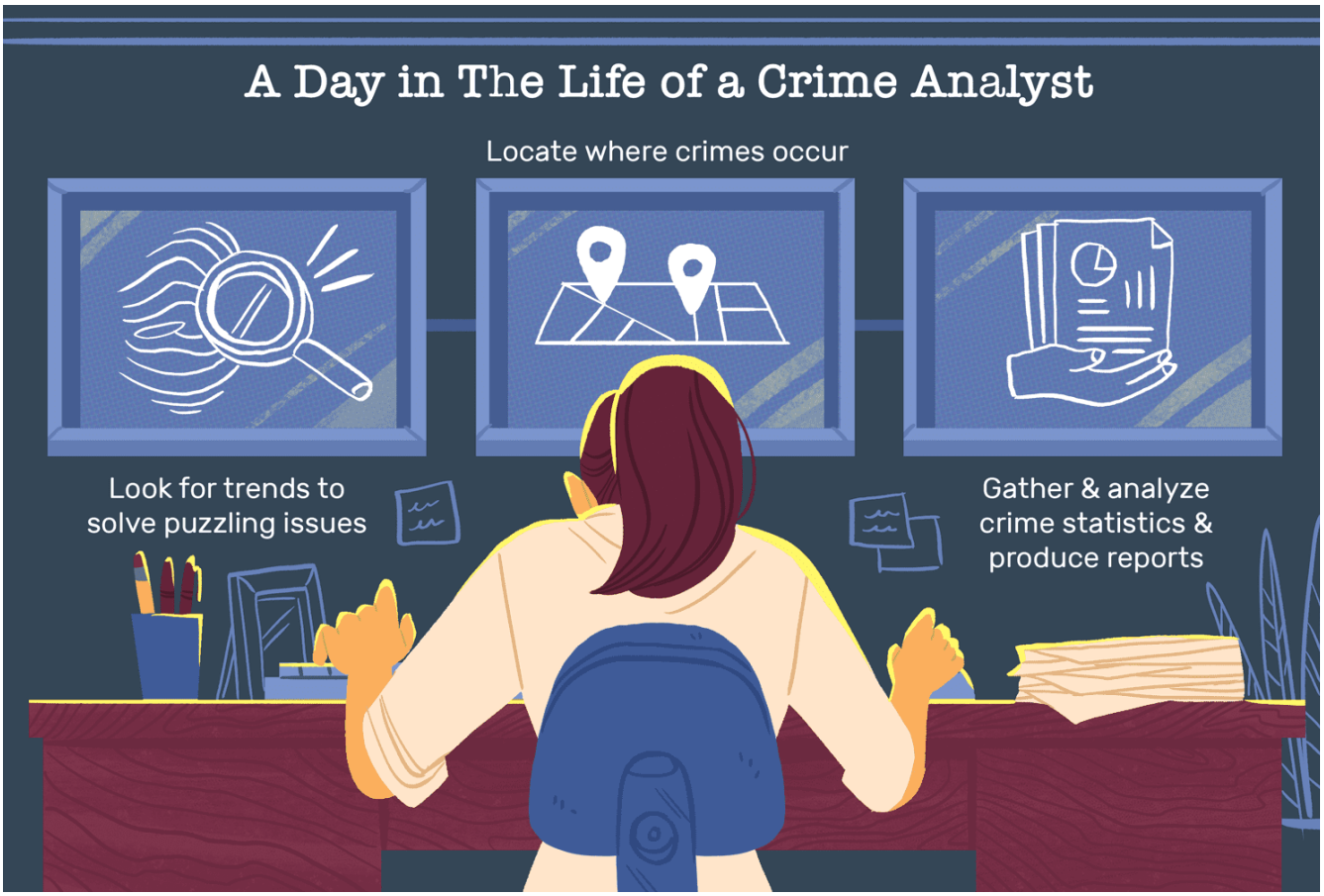


CRIME ANALYSIS OF NYC 2017-2019

Done By Bibi Nahida Shaik

Dab 06

Table of Contents

1 INTRODUCTION

Crime against women remains a significant social issue, affecting the safety, wellbeing, and freedom of half the population. Understanding the patterns and dynamics of these crimes is essential for creating effective prevention strategies. This project focuses on analyzing crime complaints involving female victims in New York City from 2017 to 2019 to identify trends, hotspots, and actionable insights that can help law enforcement and policymakers better protect women.

2 PROBLEM STATEMENT

Despite various safety initiatives, crimes against women in NYC remain concerning. There is a need for a data-driven analysis to identify when and where these crimes are most frequent, the nature of the offenses, and potential victim demographics.

3 OBJECTIVES

- Identify time-based, location-based, and demographic patterns in crimes against women.
- Highlight high-risk boroughs and precincts.
- Understand the most common types of offenses.
- Provide actionable recommendations for prevention efforts

4 TARGET AUDIENCE

- NYC Police Department (NYPD)
- City policymakers and safety planners
- Non-profit organizations focusing on women's safety
- General Public.

5 DATA OVERVIEW

1.1 Dataset:

- **Source**: NYC Open Data – NYPD Complaint Data Historic
- **Dataset Description**:

  Complaint-level records of all crime complaints reported to the NYPD.
- **Data Dictionary**:

Column Name	Data Type	Description
CMPLNT_NUM	int64	Unique identifier for each crime complaint.
CMPLNT_FR_DT	object (date)	Date the crime allegedly occurred.
CMPLNT_FR_TM	object (time)	Time the crime allegedly occurred.
CMPLNT_TO_DT	object (date)	End date of the crime occurrence.
CMPLNT_TO_TM	object (time)	End time of the crime occurrence.
ADDR_PCT_CD	float64	Precinct code where the complaint was filed.
RPT_DT	object (date)	Date the complaint was officially reported.
KY_CD	int64	Numeric offense classification code.
OFNS_DESC	object	Description of the offense (e.g., ROBBERY, ASSAULT).
PD_CD	float64	NYPD internal classification code.
PD_DESC	object	NYPD internal offense description.
CRM_ATPT_CPTD_CD	object	Indicates whether the crime was completed (COMPLETED) or attempted (ATTEMPTED).
LAW_CAT_CD	object	Legal category: FELONY, MISDEMEANOR, or VIOLATION.
BORO_NM	object	Borough where the incident occurred (e.g., MANHATTAN, BROOKLYN).
LOC_OF_OCCUR_DESC	object	Location description relative to the building (e.g., FRONT OF, INSIDE).
PREM_TYP_DESC	object	Type of premises where the crime occurred (e.g., RESIDENCE, STREET).
JURIS_DESC	object	Agency responsible for the jurisdiction (usually NYPD).
JURISDICTION_CODE	float64	Numeric code for the jurisdiction.
HOUSING_PSA	object	Public Service Area code for NYCHA developments.
X_COORD_CD	float64	X coordinate (easting) of the crime location (NYC spatial reference).

Y_COORD_CD	float64	Y coordinate (northing) of the crime location (NYC spatial reference).
SUSP_AGE_GROUP	object	Age group of the suspect (e.g., <18, 18-24, 25-44, 45-64, 65+).
SUSP_RACE	object	Race of the suspect.
SUSP_SEX	object	Gender of the suspect.
TRANSIT_DISTRICT	float64	Code of the Transit District (if crime occurred in the transit system).
Latitude	float64	Latitude coordinate of the crime location.
Longitude	float64	Longitude coordinate of the crime location.
Lat_Lon	object	Combined latitude and longitude field.
PATROL_BORO	object	NYPD patrol borough responsible for policing the area.
STATION_NAME	object	Name of the transit station if crime occurred there.
VIC_AGE_GROUP	object	Age group of the victim.
VIC_RACE	object	Race of the victim.
VIC_SEX	object	Gender of the victim.

6 DATA HANDELING

- **Timeframe**: Records from January 1, 2017, to December 31, 2019.
- **Handling Missing Values**:
  - Dropped rows where critical fields like BORO_NM or OFNS_DESC were missing.
  - Imputed missing age groups where appropriate using 'UNKNOWN' tag.
  - Replaced nan values from overall dataset to unknown.
  - Renamed offences into comprehensible strings.

- **Data Transformation**:
  - Extracted Year, Month, Weekday from CMPLNT_FR_DT.(Using Python and Power BI)
  - Created a crime_type_category based on OFNS_DESC grouping.

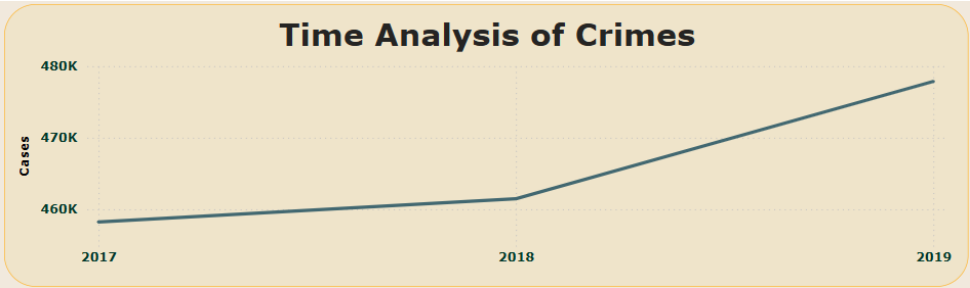
- **Preprocessing Techniques**:
  - Standardized offense descriptions (removing extra spaces, fixing typos)(Using Power Query).
    Aggregated crimes by borough, precinct, and time.

7 ANALYSIS AND FINDINGS

1.2 Overall Crime Analysis (All Victims)

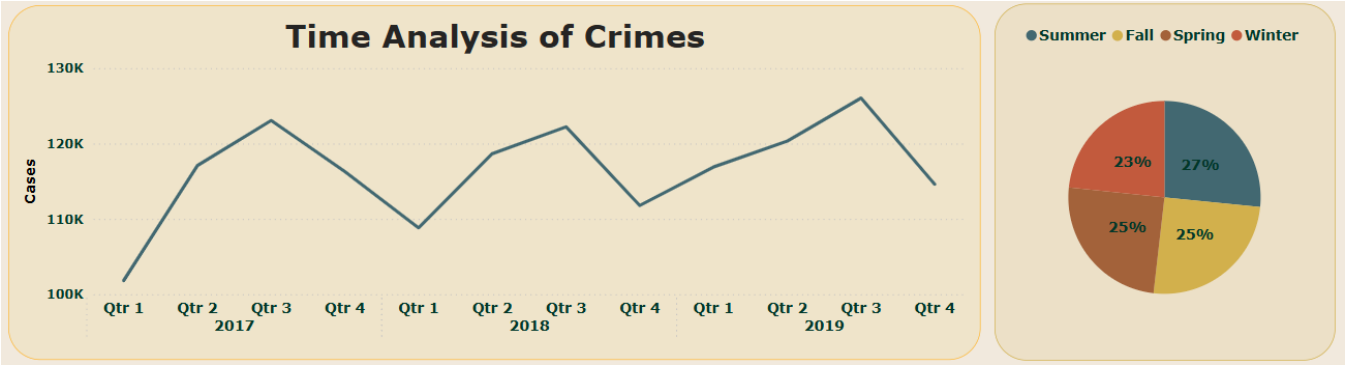
1.2.1 Temporal Analysis

- **Yearly Trends**:



The overall crime rate dropped significantly in 2017. Which can be explained by 2017 election where Trump was elected, and his focus was to drop the crime rate and indeed the strategies were effective as in that year the crime rate was lowest since 1950s. In addition to that, at the start of 2017, the employment rate increased, there were many community programs which eventually decreased the number of criminals hence the crime rate dropped. But the crime rate increased notably increase after in the coming years.

- **Quarterly Trends**:

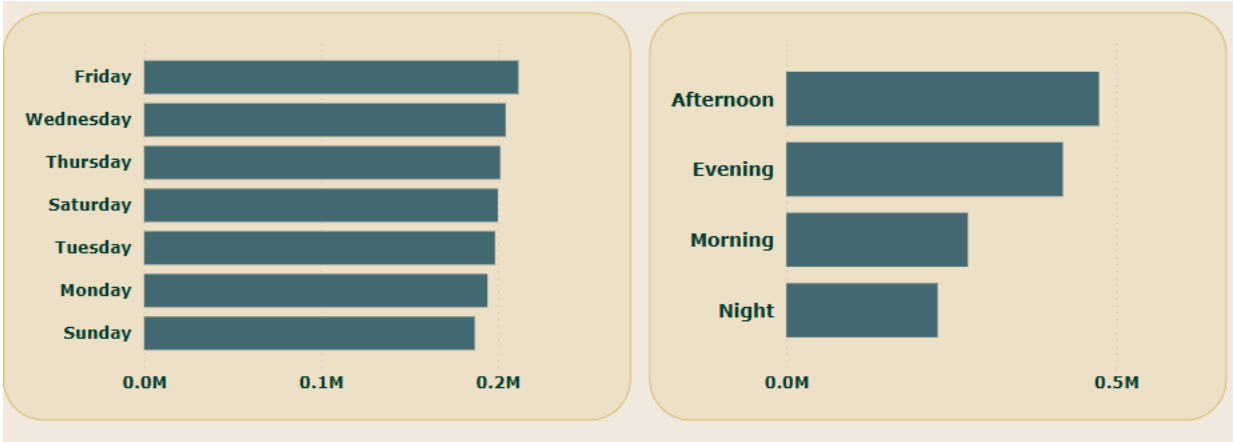


- Quarter one of year 2017, we steep is at its lowest as it was mentioned in the yearly analysis crime rate where the lowest
- Higher crimes during the summer (June–August). Time when people prefer to go out with friends and family. While it can also be noted crimes drop around quarter four, which may be due to cold weather, increased policing for Thanksgiving, and Christmas.

- **Quarters by crime:**

- When looking at the crime rate by the crime it could be noted that some crimes do not follow the same pattern, in fact they are increasing, which is really concerning.

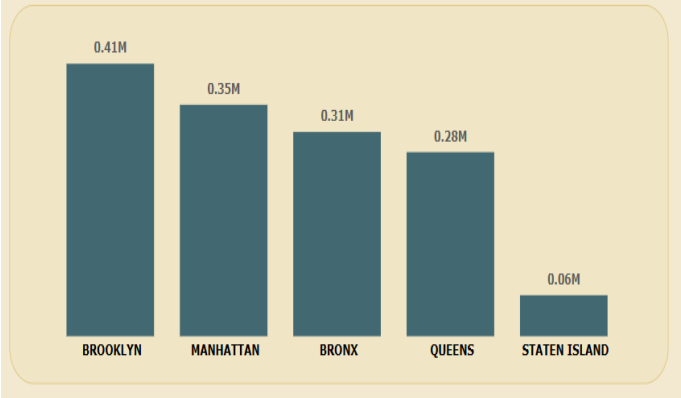
- **Day of the Week**:

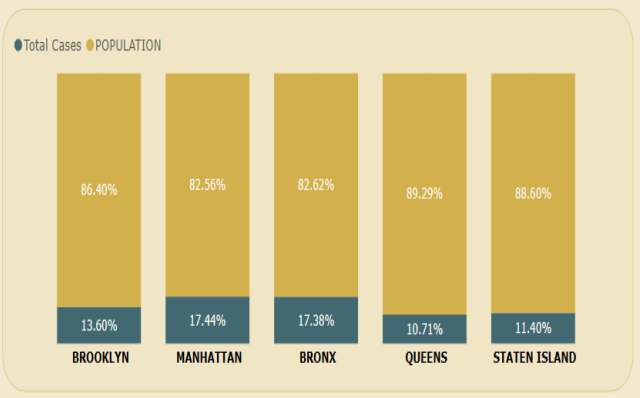


- Crimes peak on weekdays, specifically mid-week when people are busy, at break, pick up children from school. It may be because of it being mid-day when people are frustrated hungry hence easily starting any conflict or pickpocketing.
- **Time of Day**:
- Most crimes occurred between afternoon and evening, when there is a lot of foot traffic.

1.2.2 Location Analysis

- **By Borough**:



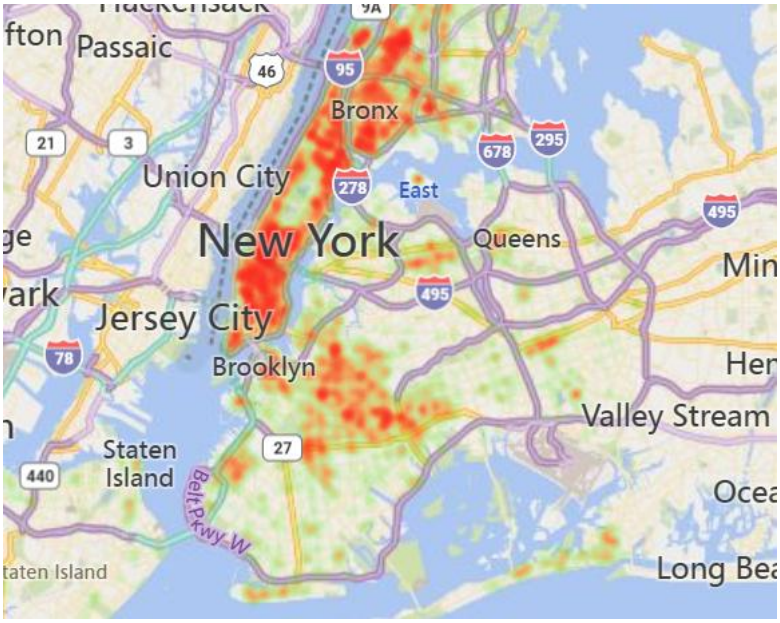


Brooklyn had the highest overall crime reports, followed by The Bronx and Manhattan. But!!

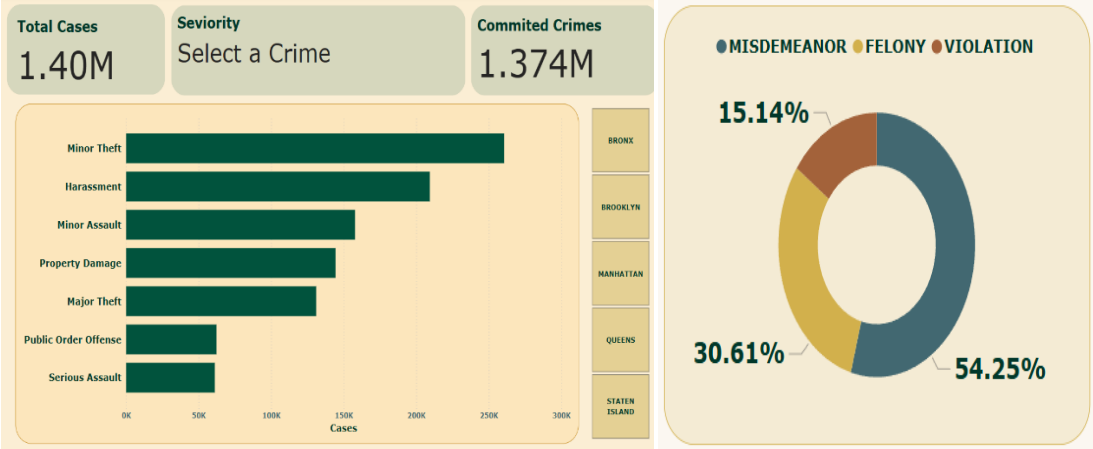
After comparing the number of reported crimes with the population it can be noted that yes Manhattan is still the top with most registered cases, but the next most populated borough Queens is 5th on the scale. That shows the crime rate in Queens in relatively lower the other boroughs.

- **By heatmap**:

Observing the following heatmap. It can be concluded that northern Queens and southern Brooklyn deem safer than other areas. While Manhattan is not safe at all.

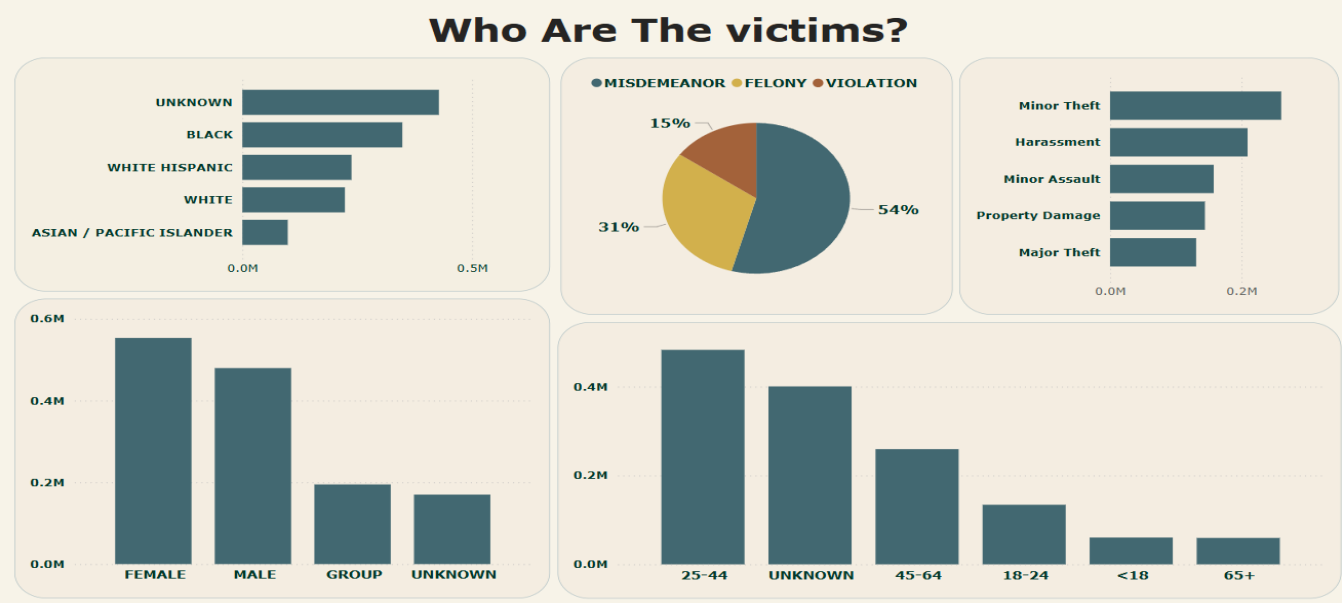


1.2.3 Offense Type Analysis



- **Top Reported Offenses**:
  - Minor Theft (Theft of property worth less than 1000$)
  - Harassment (Stalking, Following, Physical contact without consent)
  - Minor Assault (Injuring a third party while intending to harm someone else)
- **Felonies vs. Misdemeanors**:
  - Felonies made up around 40% of crimes, while misdemeanors dominated (55%).

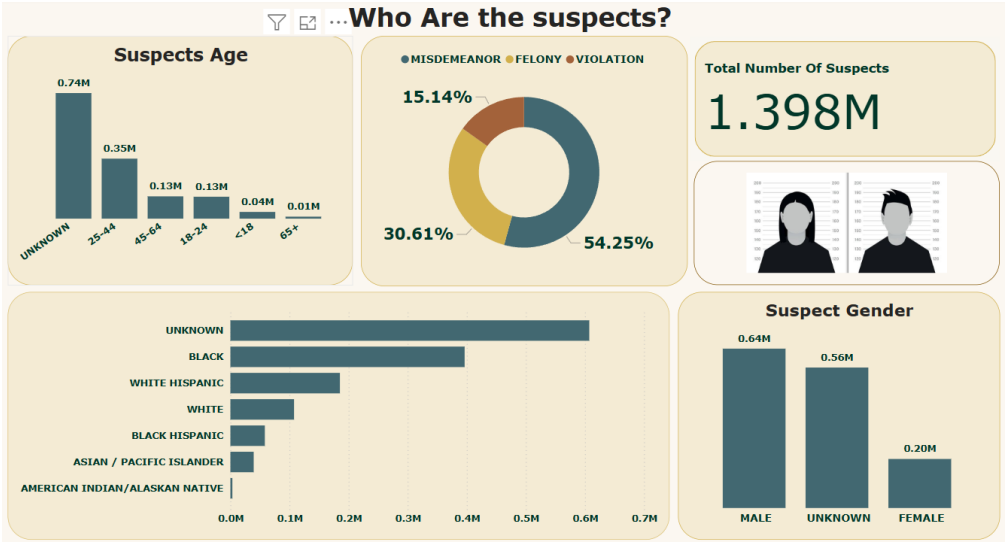
1.2.4 Victim Demographics



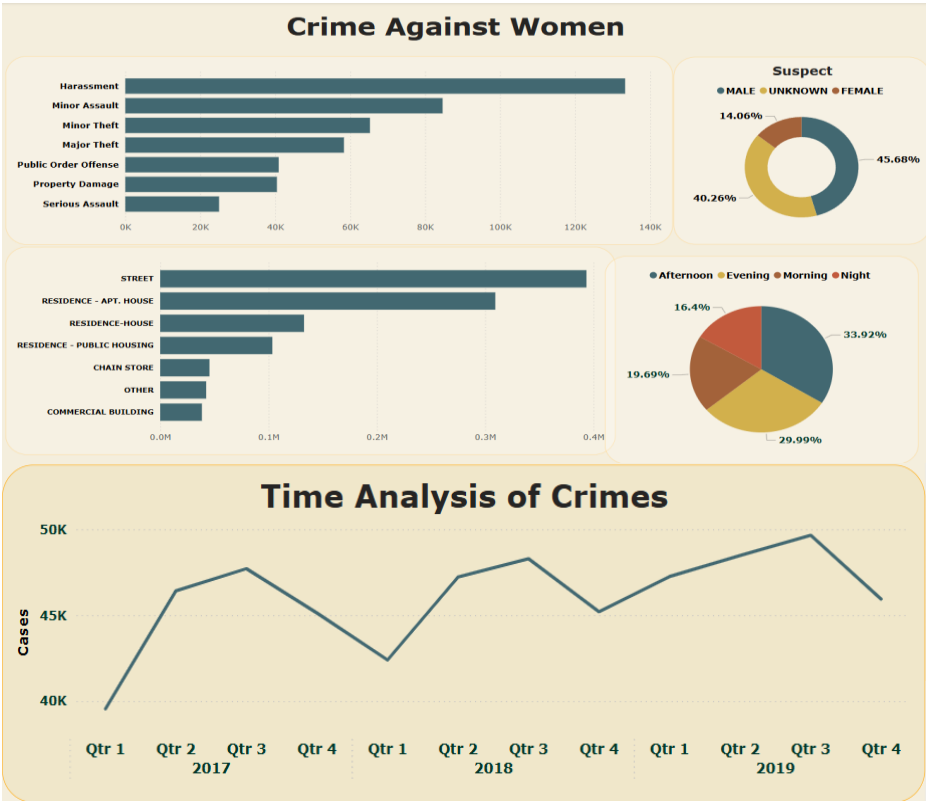
- **Age Groups**:
  - Most victims were aged 25–44.
- **Gender**:
  - More than 0.5 million of victims were female.
- **Race/Ethnicity**:
  - Black and Hispanic individuals were the most common victims after unknown(which includes groups and organizations).

1.2.5 Suspect Analysis

- **Age**:
  - Majority of suspects were aged 18–30.
- **Gender**:
  - 80–85% of suspects were male.
- **Race/Ethnicity**:
  - Predominantly Black and Hispanic.



1.3 Crime Analysis Focused on Female Victims



This visual breaks down crimes where women were the victims, revealing some concerning trends. Harassment stands out as the most reported offense, followed closely by minor assault and theft. Most of these incidents happen in public spaces, especially on the street and in apartment buildings, showing that women are often targeted in everyday environments where they should feel safe.

One of the most alarming findings is that in nearly half of the cases, the suspect is unknown. Where suspects are identified, the majority are male. This highlights not just the gendered nature of these crimes but also gaps in identification and accountability.

Looking at the time of day, evenings seem to be the most dangerous, though afternoons and nights also show significant numbers. This suggests that these crimes aren't limited to late hours but span most of the day, which is important for understanding risk patterns.

Over time, the data from 2017 to 2019 shows consistent rise and fall, with the highest number of cases recorded in mid to late 2019. This ongoing pattern underlines the need for sustained efforts in prevention, better public safety, and stronger support systems for women facing these threats in their daily lives.

1.4 Findings

- **Overall**: Brooklyn leads in total crimes; young adults (18–30) are the most frequent offenders and victims.
- **Female Victims**: Higher vulnerability during weekdays, day time hours, and in dense boroughs like Brooklyn and Bronx.
- **Data Gaps**: Lack of gang-related tags and suspect-victim relationship fields restrict root cause discovery.
- Females are being targeted in every environment. Even Spaces that should have been safest.

8 RECOMMENDATIONS

- Increase police presence in identified hotspots when foot traffic is high and summer.
- Launch public awareness and self-defense programs targeting women aged 18–44.
- Collaborate with community groups to support vulnerable neighborhoods.
- Conduct additional training for officers on handling assault and harassment cases sensitively.

The recommendation given does not address the root cause of the problem. A large portion of suspect data is unknown, and to properly address this issue, access to mafia- or gang-related data is necessary, as it may shed light on the motivations behind these crimes.

9 LIMITATIONS

- Missing or incorrect data entries could bias results.
- The key point to addressing high rates of crime is to understand the key factors which this dataset lacks. Such as, situation, or cause leading to the crime. Data regarding gang activities that has a role in increased crime rate.
- Victim or suspect race data is often incomplete.


THE END ?

References

- (NYPD), P. D. (2025, April 15). *NYC crime: NYC open data*. NYC crime | NYC Open Data. https://data.cityofnewyork.us/Public-Safety/NYC-crime/qb7u-rbmr/about_data
- Heward, S. (2025, February 6). *New York through the seasons*. Holiday Hypermarket. https://www.holidayhypermarket.co.uk/hype/when-to-visit-new-york-through-the-seasons
- Kaufman, E. (2017, December 28). *New York City on pace to record lowest murder tally in decades*. CNN. https://edition.cnn.com/2017/12/26/us/new-york-2017-crime-stats/index.html