澳門大學
UNIVERSIDADE DE MACAU
UNIVERSITY OF MACAU

CISC3015 - Data and Information Visualization (2024-1 001)

PAIR 05: PROJECT PRESENTATION

Tailine J. S. Nonato
Zhong Wanyi

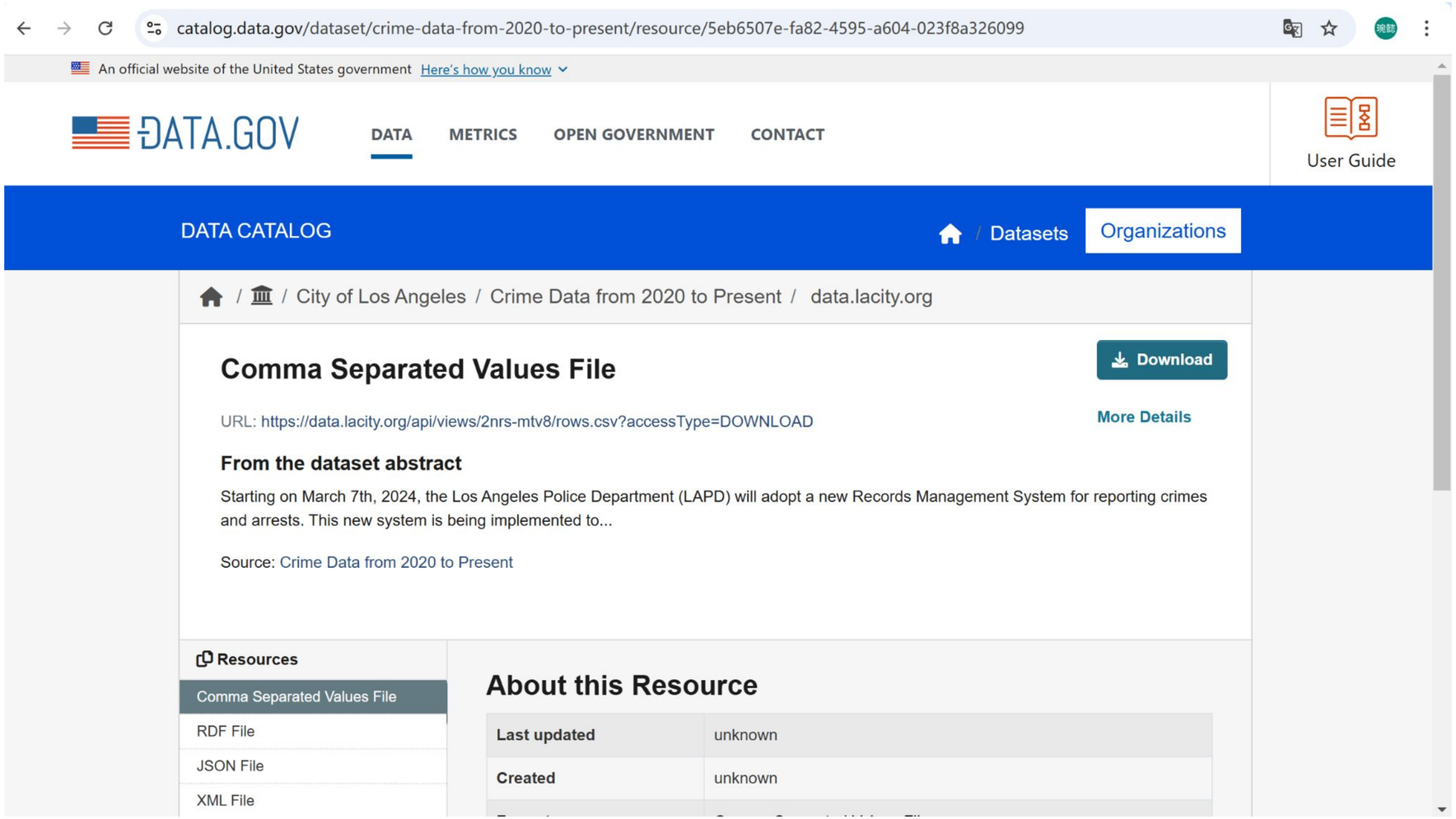
DOMAIN CATEGORY

Crimes;
Population safety.

REGION

LOS ANGELES

ABOUT THE DATA



OVERVIEW OF THE DATASET

[1]:

	DR_NO	Date Rptd	DATE OCC	TIME OCC	AREA	AREA NAME	Rpt Dist No	Part 1-2	Crm Cd	Crm Cd Desc	...	Status	Status Desc	Crm Cd 1	Crm Cd 2	Crm Cd 3	Crm Cd 4	LOCATION	Cro Stre
0	190326475	03/01/2020 12:00:00 AM	03/01/2020 12:00:00 AM	2130	7	Wilshire	784	1	510	VEHICLE - STOLEN	...	AA	Adult Arrest	510.0	998.0	NaN	NaN	1900 S LONGWOOD AV	N
1	200106753	02/09/2020 12:00:00 AM	02/08/2020 12:00:00 AM	1800	1	Central	182	1	330	BURGLARY FROM VEHICLE	...	IC	Invest Cont	330.0	998.0	NaN	NaN	1000 S FLOWER ST	N
2	200320258	11/11/2020 12:00:00 AM	11/04/2020 12:00:00 AM	1700	3	Southwest	356	1	480	BIKE - STOLEN	...	IC	Invest Cont	480.0	NaN	NaN	NaN	1400 W 37TH ST	N
3	200907217	05/10/2023 12:00:00 AM	03/10/2020 12:00:00 AM	2037	9	Van Nuys	964	1	343	SHOPLIFTING-GRAND THEFT ($950.01 & OVER)	...	IC	Invest Cont	343.0	NaN	NaN	NaN	14000 RIVERSIDE DR	N
4	220614831	08/18/2022 12:00:00 AM	08/17/2020 12:00:00 AM	1200	6	Hollywood	666	2	354	THEFT OF IDENTITY	...	IC	Invest Cont	354.0	NaN	NaN	NaN	1900 TRANSIENT	N

5 rows × 28 columns

```
:  crimes.shape
```

```
:  (722906, 29)
```

[70]:

```
# dataset variables
crimes.columns
```

[70]:

```
Index(['DR_NO', 'Date Rptd', 'DATE OCC', 'TIME OCC', 'AREA', 'AREA NAME',
       'Rpt Dist No', 'Part 1-2', 'Crm Cd', 'Crm Cd Desc', 'Mocodes',
       'Vict Age', 'Vict Sex', 'Vict Descent', 'Premis Cd', 'Premis Desc',
       'Weapon Used Cd', 'Weapon Desc', 'Status', 'Status Desc', 'Crm Cd 1',
       'Crm Cd 2', 'Crm Cd 3', 'Crm Cd 4', 'LOCATION', 'Cross Street', 'LAT',
       'LON'],
      dtype='object')
```

Huge dataset. So it was required to clean and filter the data.

MAIN QUESTION

What are the factors that explain the number of crimes by category in Los Angeles from 2020 to 2024?

Does anything change when looking in detail into criminal homicides?

RELATED REFERENCES

Bureau of Justice Statistics
Special Report

Age Patterns of Victim of Serious Violent Crin

By Craig A. Perkins
BJS Statistician

Vulnerability to violent crime victimization varies across the age spectrum. The victimization rate increases through the teenage years, crests at around age 20, and steadily decreases through the remaining years. This pattern, with some exceptions, exists across all race, sex, and ethnic groups.

For 1992-94, the rate of serious violent crime ranged from 37 per 1,000 persons age 12 to 14, to 50 per 1,000 persons age 18 to 21, to 3 violent crimes per 1,000 persons age 65 or older. Crime rates for individuals age 18 to 21 were

Highlights

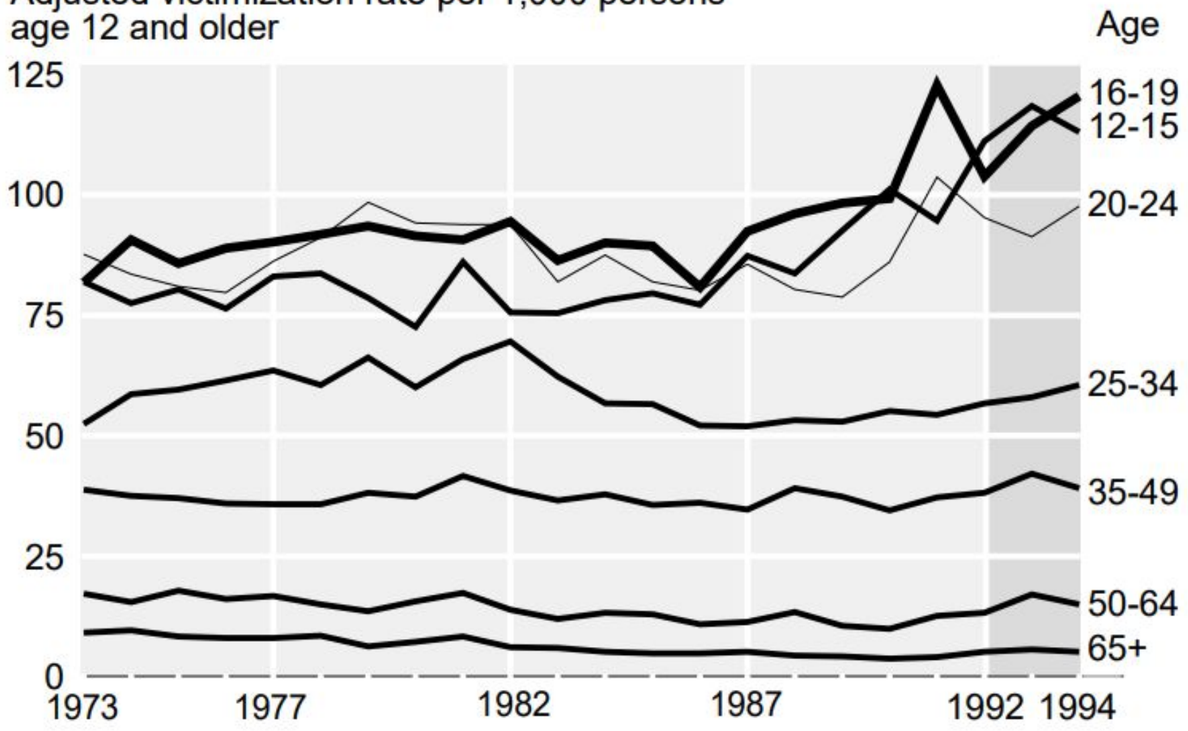
- Persons age 12 to 24 comprised:
  22% of the population,
  35% of murder victims, and
  49% of serious violent crime victims.
- Persons age 25 to 49 constituted:
  47% of the population,
  53% of murder victims, and
  44% of serious violent crime victims.
- Persons age 50 or older made up:
  30% of the population,
  12% of murder victims, and
  7% of serious violent crime victims.

- Persons age
  likely to experie
  crime, and blac
  were the most
  72 victimi
  50 victimi
  Hispanics
  46 victimi

- More than 52
  sault victims w
  25.

- Almost 1 in 10 murder victims age 18 to 21 were black.

Violent crime rates by age
Adjusted victimization rate per 1,000 persons age 12 and older



PURPOSE OF OUR PROJECT

1. **For Individual**
   (1) Advance knowledge in fields like data analysis, sociology, psychology related to deviant behaviour.
2. **For whole society**
   (1) Identify high-risk gender and age groups that may need more support services or protection programs to protect their safety in the future.

WHY CRIMINAL HOMICIDE?

After cleaning the data, we find that the number of criminal homicide is 1,538 extracting from the whole 722,906 crimes.

Crm Cd Desc	Vict Age
CHILD NEGLECT (SEE 300 W.I.C.)	8.15393
CRM AGNST CHLD (13 OR UNDER) (14-15 & SUSP 10 YRS OLDER)	10.6564
CHILD ABANDONMENT	11
CHILD ABUSE (PHYSICAL) - SIMPLE ASSAULT	11.5292
CHILD ABUSE (PHYSICAL) - AGGRAVATED ASSAULT	11.7922
CHILD ANNOYING (17YRS & UNDER)	13.3268
LEWD/LASCIVIOUS ACTS WITH CHILD	13.5904
CONTRIBUTING	15.2683
SEX,UNLAWFUL(INC MUTUAL CONSENT, PENETRATION W/ FRGN OBJ	16.0679
HUMAN TRAFFICKING - COMMERCIAL SEX ACTS	17.4586
DRUGS, TO A MINOR	17.8333
INCITING A RIOT	22
KIDNAPPING - GRAND ATTEMPT	22.2796
HUMAN TRAFFICKING - INVOLUNTARY SERVITUDE	22.7064
PIMPING	22.9609
INCEST (SEXUAL ACTS BETWEEN BLOOD RELATIVES)	24.8333
CHILD PORNOGRAPHY	25.4803
ORAL COPULATION	25.677
SEXUAL PENETRATION W/FOREIGN OBJECT	27.1282
PANDERING	27.5429
RAPE, FORCIBLE	28.1692
SODOMY/SEXUAL CONTACT B/W PENIS OF ONE PERS TO ANUS OTH	29.0885
KIDNAPPING	29.6366
REPLICA FIREARMS(SALE,DISPLAY,MANUFACTURE OR DISTRIBUTE)	30.2222
BIGAMY	30.25
INDECENT EXPOSURE	35.5879
LEWD CONDUCT	35.7419
BIKE - STOLEN	35.9169
STALKING	36.1429
RECKLESS DRIVING	36.1702
DISRUPT SCHOOL	36.3333
THREATENING PHONE CALLS/LETTERS	36.4201
MANSLAUGHTER, NEGLIGENT	36.625
SHOPLIFTING - ATTEMPT	36.6774
LYNCHING - ATTEMPTED	37
LETTERS, LEWD - TELEPHONE CALLS, LEWD	37.1193
CRIMINAL HOMICIDE	37.3687
ROBBERY	37.4131
BATTERY ON A FIREFIGHTER	37.4226
TILL TAP - GRAND THEFT ($950.01 & OVER)	37.8889
FIREARMS RESTRAINING ORDER (FIREARMS RO)	38
DRUNK ROLL	38.0323
ASSAULT WITH DEADLY WEAPON, AGGRAVATED ASSAULT	38.1582
BURGLARY FROM VEHICLE	38.2459
BRANDISH WEAPON	38.7482
FIREARMS EMERGENCY PROTECTIVE ORDER (FIREARMS EPO)	39
ATTEMPTED ROBBERY	39.2869
WEAPONS POSSESSION/BOMBING	39.4167
DISCHARGE FIREARMS/SHOTS FIRED	39.4784
THROWING OBJECT AT MOVING VEHICLE	39.4791
OTHER MISCELLANEOUS CRIME	39.5106
CRUELTY TO ANIMALS	39.9831
THEFT PLAIN - PETTY ($950 & UNDER)	40.1192
CRIMINAL THREATS - NO WEAPON DISPLAYED	40.2312
TILL TAP - PETTY ($950 & UNDER)	40.25
OTHER ASSAULT	40.3065
TELEPHONE PROPERTY - DAMAGE	40.3333
EMBEZZLEMENT, PETTY THEFT ($950 & UNDER)	47.871
PICKPOCKET, ATTEMPT	48
ILLEGAL DUMPING	49.2143
TRAIN WRECKING	50
GRAND THEFT / INSURANCE FRAUD	51
GRAND THEFT / AUTO REPAIR	52.3333
DOCUMENT FORGERY / STOLEN FELONY	52.6182
LYNCHING	53
BLOCKING DOOR INDUCTION CENTER	54.6667
DRUNK ROLL - ATTEMPT	58
DISHONEST EMPLOYEE ATTEMPTED THEFT	60
THEFT, COIN MACHINE - GRAND ($950.01 & OVER)	65

MEAN AND MEDIAN AGE OF THE VICTIM

[12]:
```
# mean age of the victims
overall_mean = homicides['Vict Age'].mean()
mean_victim_ages_by_gender = homicides.groupby(['Crm Cd Desc', 'Vict Sex'])['Vict Age'].mean().reset_index()

mean_victim_ages_by_gender['Overall Mean Victim Age'] = overall_mean

print(mean_victim_ages_by_gender.sort_values(by=['Crm Cd Desc', 'Vict Sex']).to_markdown(index=False))
```

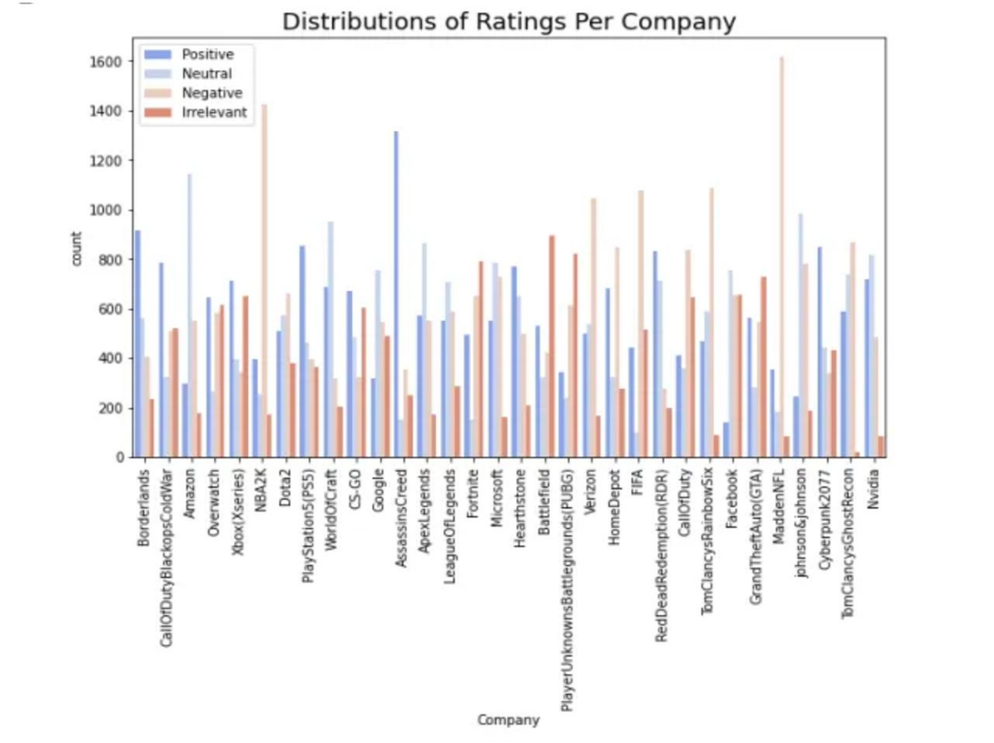
Crm Cd Desc	Vict Sex	Vict Age	Overall Mean Victim Age
CRIMINAL HOMICIDE	F	38.7379	37.3687
CRIMINAL HOMICIDE	M	37.1569	37.3687

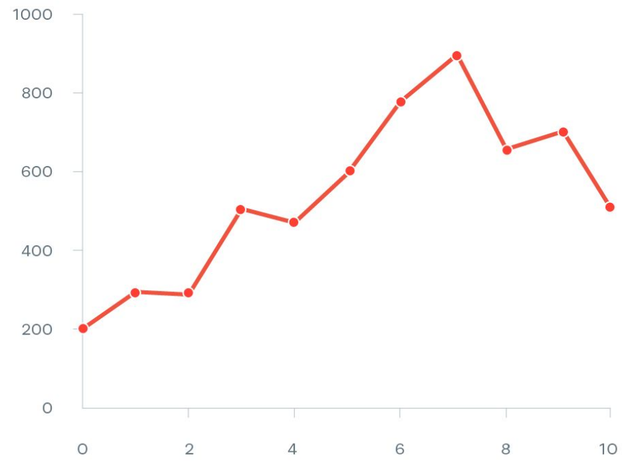
[21]:
```
# median age of the victims
overall_median = homicides['Vict Age'].median()
median_victim_ages_by_gender = homicides.groupby(['Crm Cd Desc', 'Vict Sex'])['Vict Age'].median().reset_index()

median_victim_ages_by_gender['Overall Median Victim Age'] = overall_median

print(median_victim_ages_by_gender.sort_values(by=['Crm Cd Desc', 'Vict Sex']).to_markdown(index=False))
```

Crm Cd Desc	Vict Sex	Vict Age	Overall Median Victim Age
CRIMINAL HOMICIDE	F	35	34.5
CRIMINAL HOMICIDE	M	34	34.5

EXPLORATORY DATA ANALYSIS

Countplot and line plots as our main types of data visualization.





CORRELATION COEFFICIENT

[15]:
```
# correlation between the crime being a homicide and the age of the victims
correlation_age = crimes['Is Homicide'].corr(crimes['Vict Age'])
print(f"Correlation between crime being a homicide and victim age: {correlation_age}")
```

```
Correlation between crime being a homicide and victim age: -0.007096097740188022
```

[16]:
```
# correlation between the crime being a homicide and the gender of the victims
correlation_gender = crimes['Is Homicide'].corr(crimes['Vict Sex Encoded'])
print(f"Correlation between crime being a homicide and victim gender: {correlation_gender}")
```

```
Correlation between crime being a homicide and victim gender: 0.03307853721475881
```

Value is too small and close to zero!

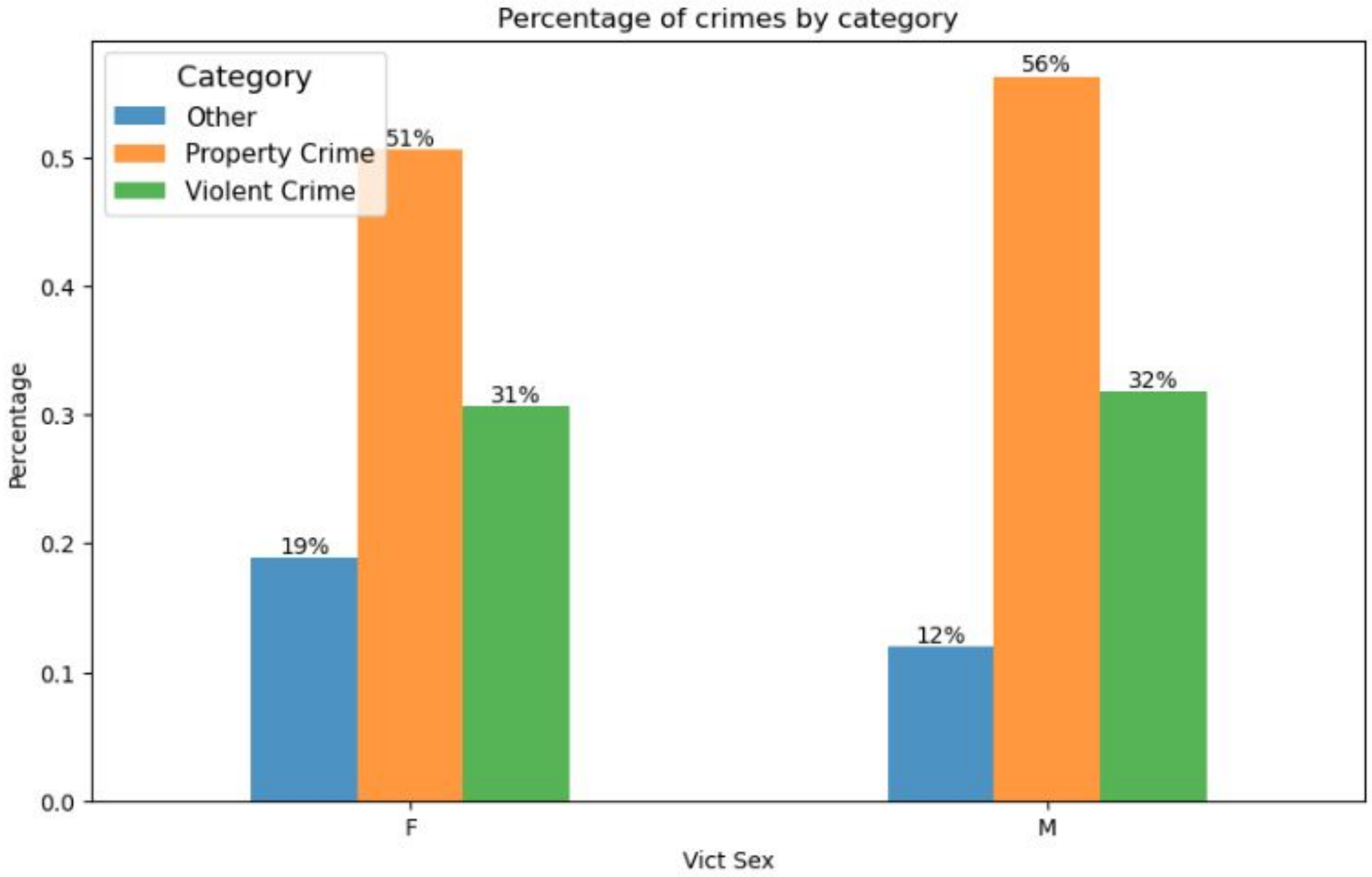
FIRST APPROACH

There is a total of 138 different crimes in the dataset. In order to smoothly work with the data, we categorized based on the wording used to describe the crime.

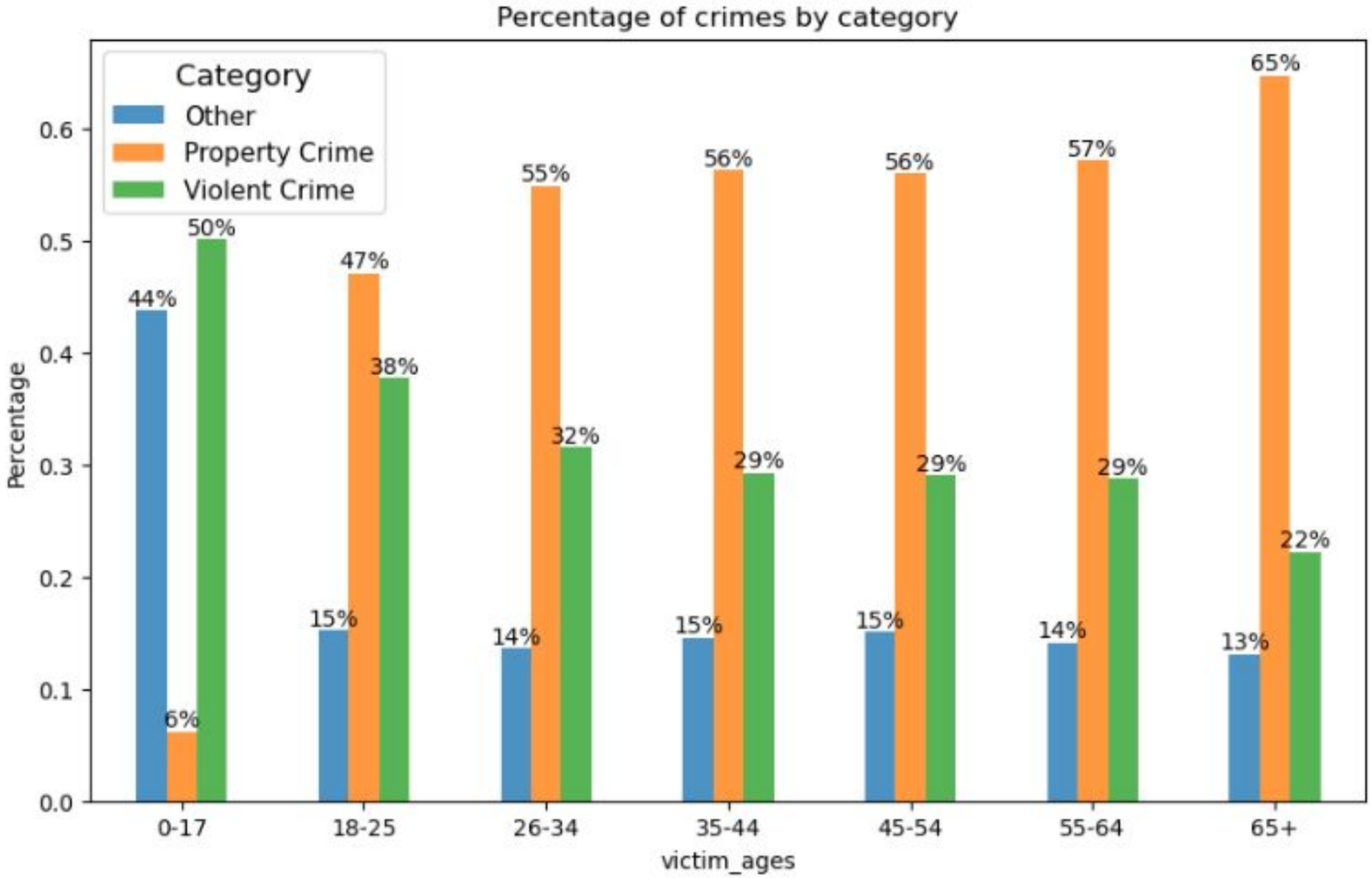
```
category
Property Crime    381063
Violent Crime     222237
Other             109424
Name: count, dtype: int64
```

```python
# define broader categories for crimes
def categorize_crime(crime_description):
    if'ASSAULT' in crime_description or 'ROBBERY' in crime_description or 'HOMICIDE' in crime_description:
        return 'Violent Crime'
    elif 'BURGLARY' in crime_description or 'THEFT' in crime_description or 'VANDALISM' in crime_description or 'STOLEN' in crime_description:
        return 'Property Crime'
    else:
        return 'Other'
```

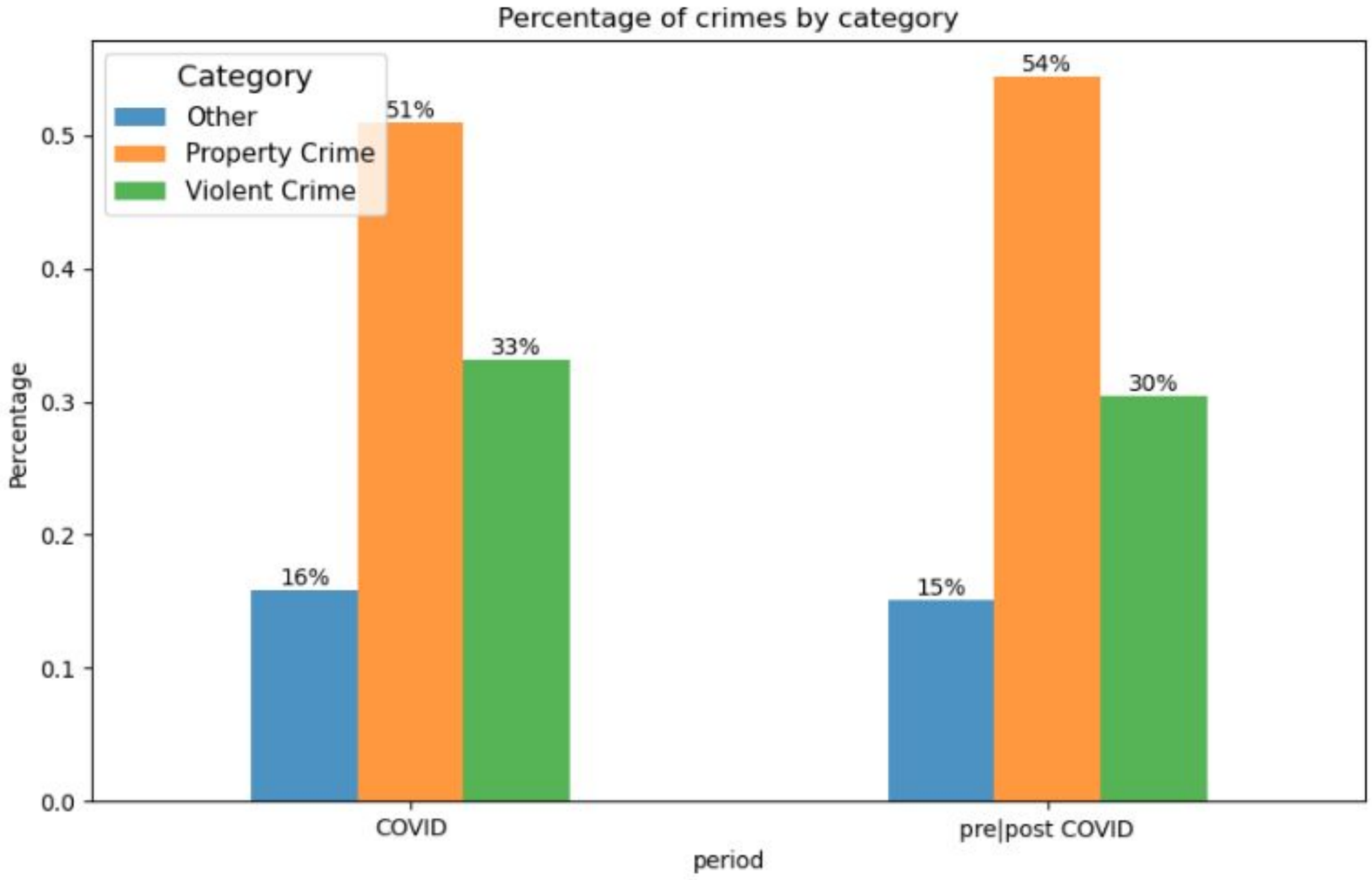
CRIMES BY VICTIM'S SEX



CRIMES BY VICTIM'S AGE

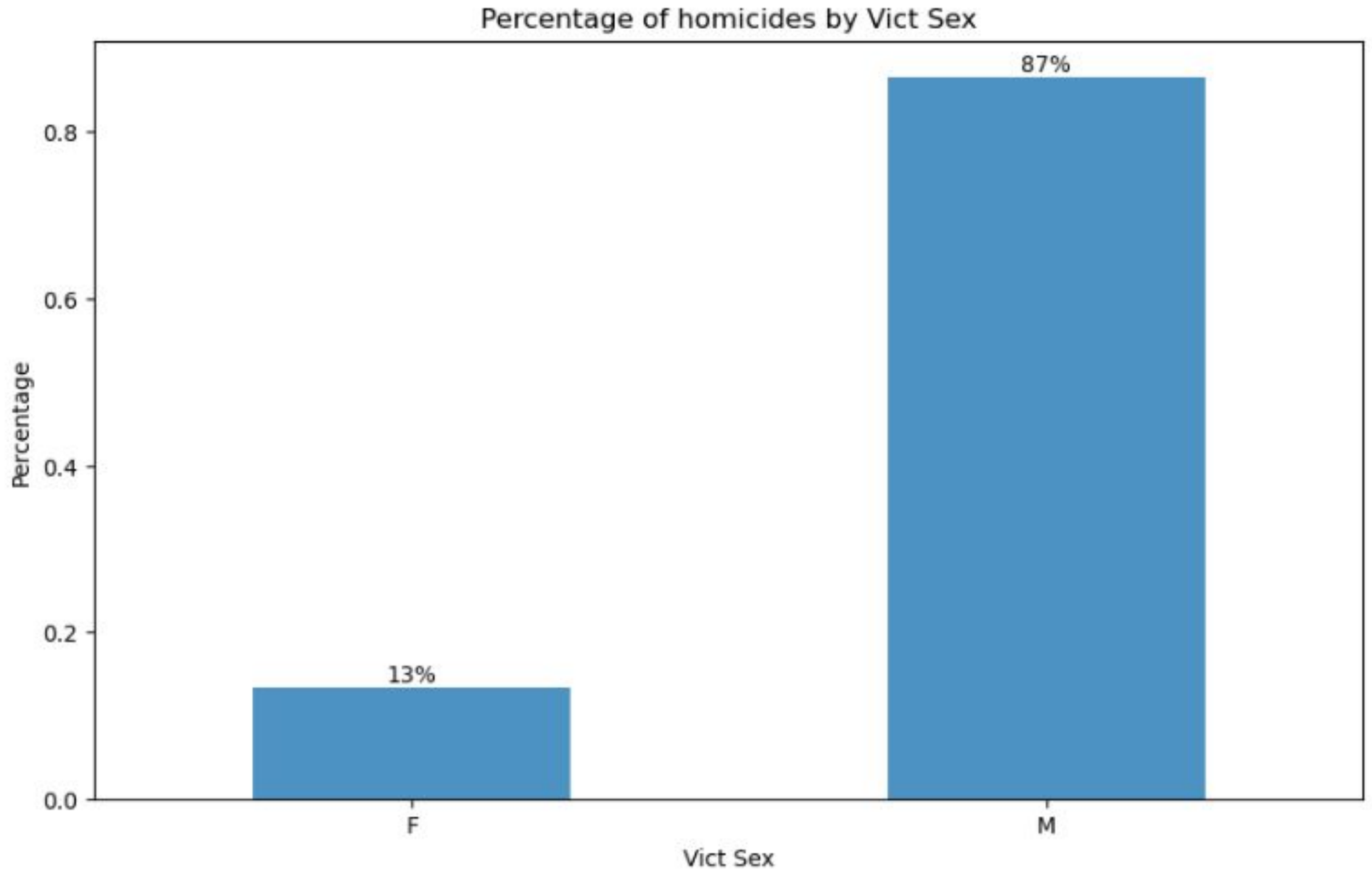


CRIMES BY PERIOD

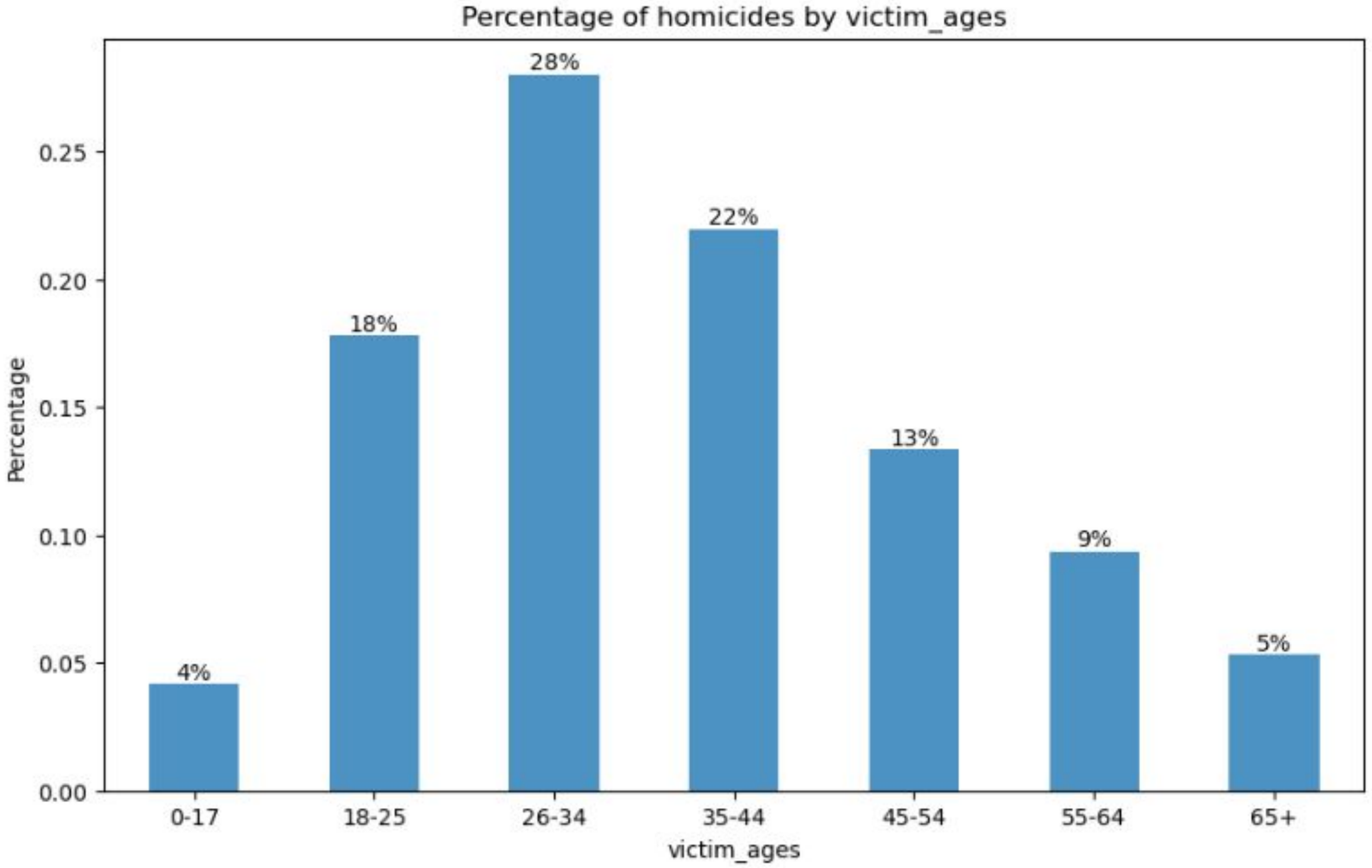


HOMICIDES BY VICTIM'S SEX

CISC3015 - Data and Information Visualization (2024-1 001)

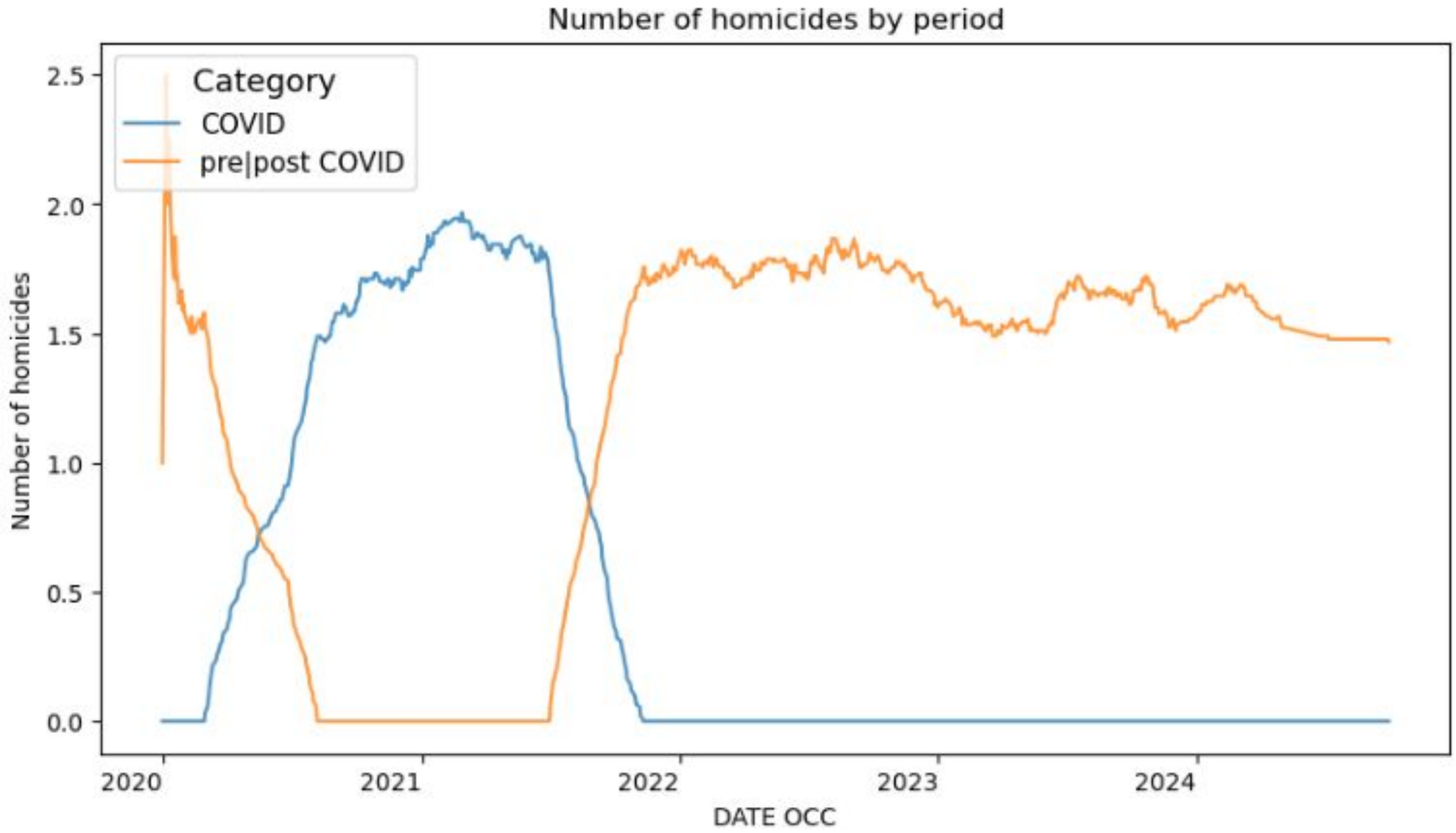


HOMICIDES BY VICTIM'S AGE

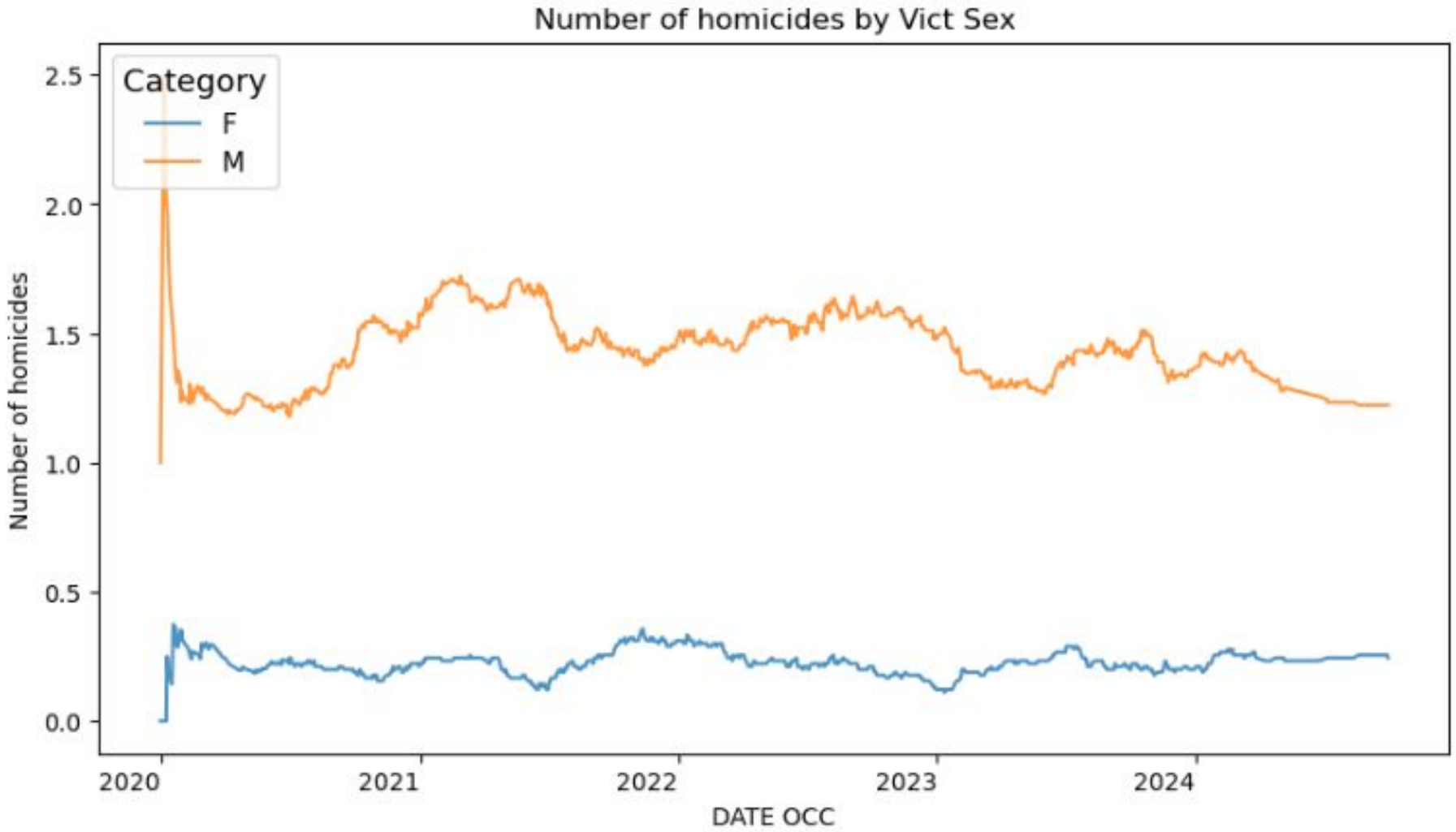


the age group of 26-34 has the highest number of victims ! The chart is shown as a bad normal distribution with lower normality.

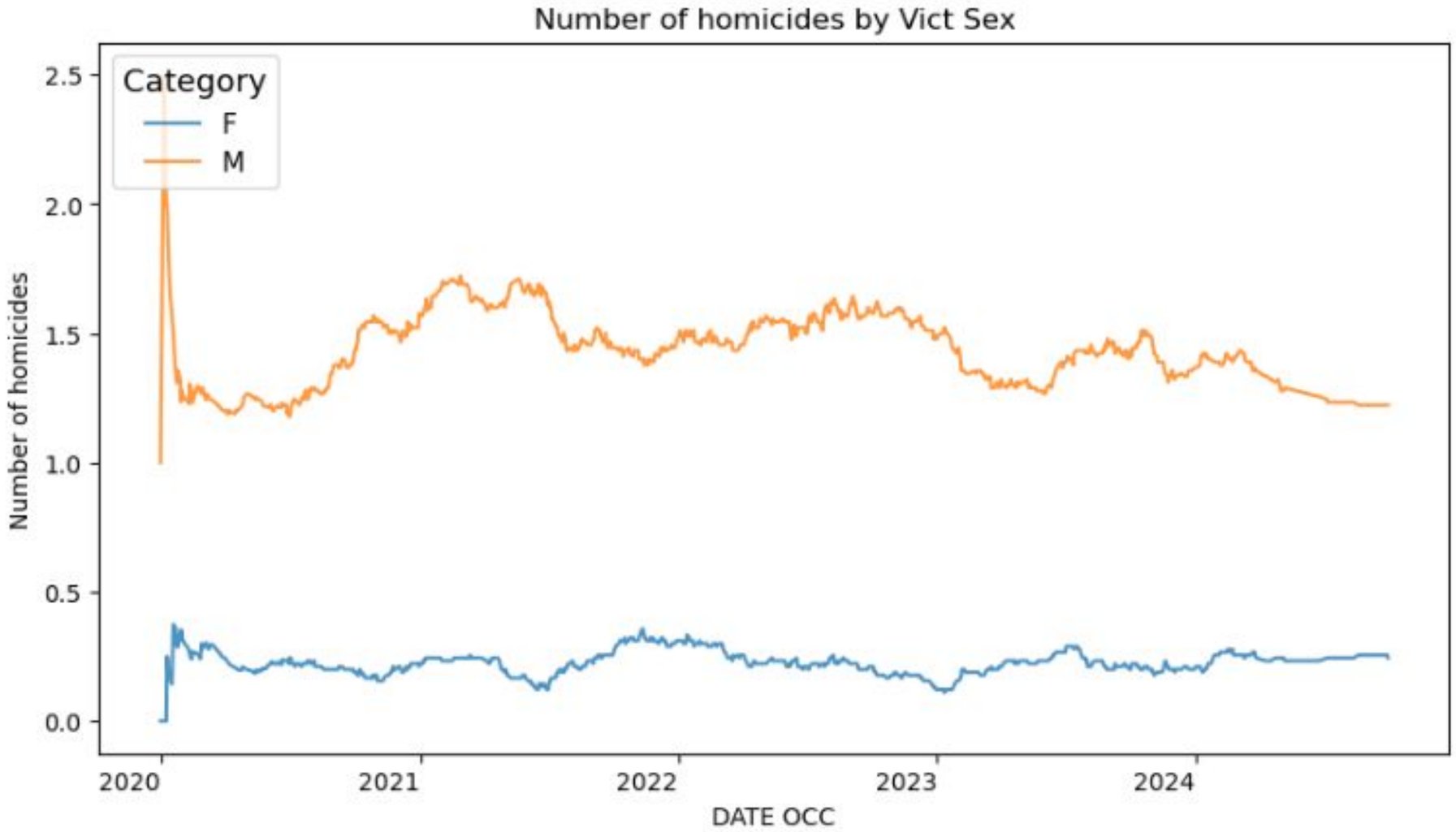
HOMICIDES OVER THE PERIODS



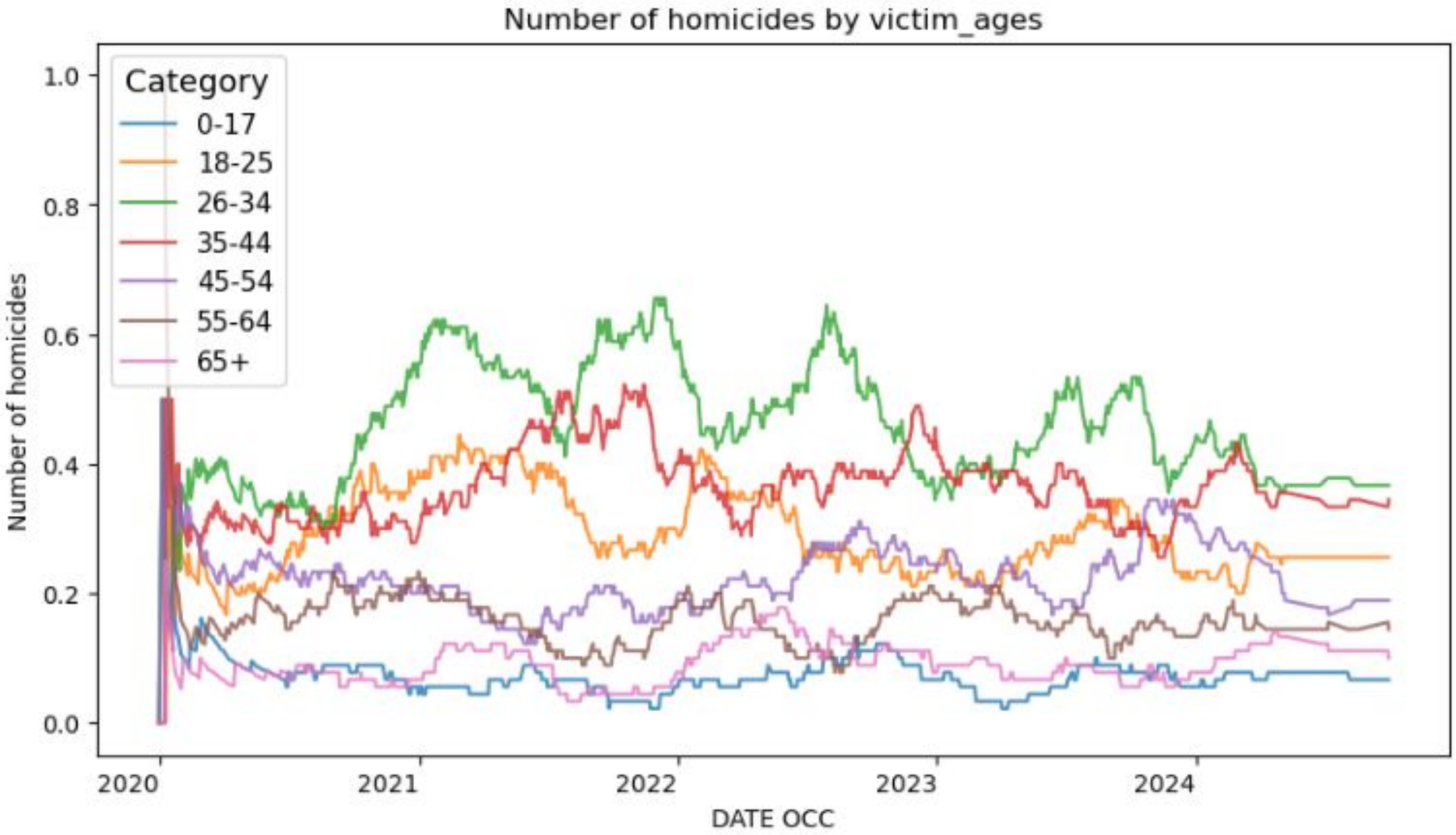
HOMICIDES OVER THE PERIODS



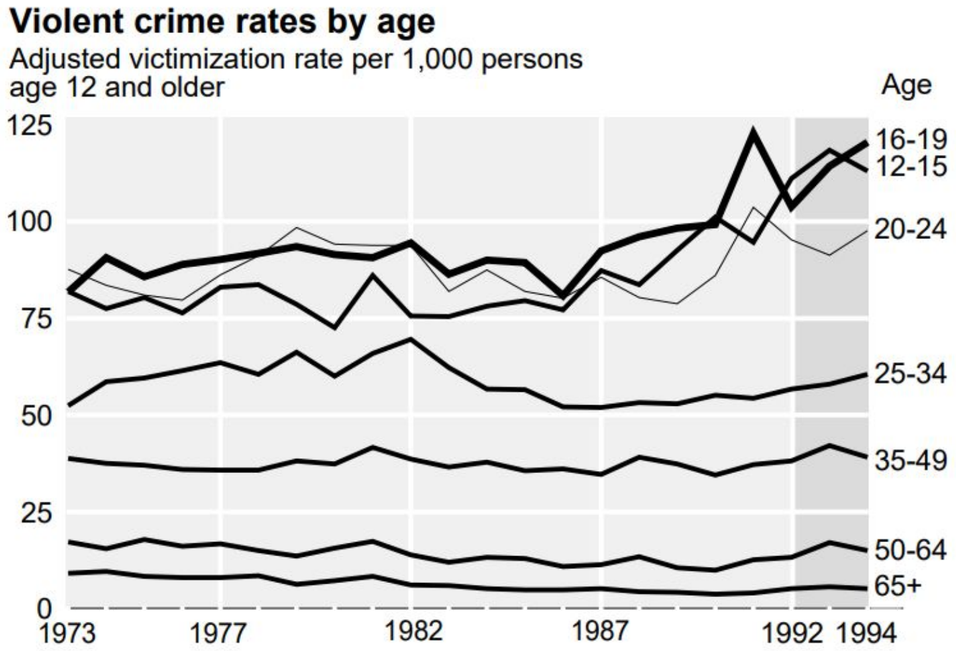
HOMICIDES OVER THE PERIODS

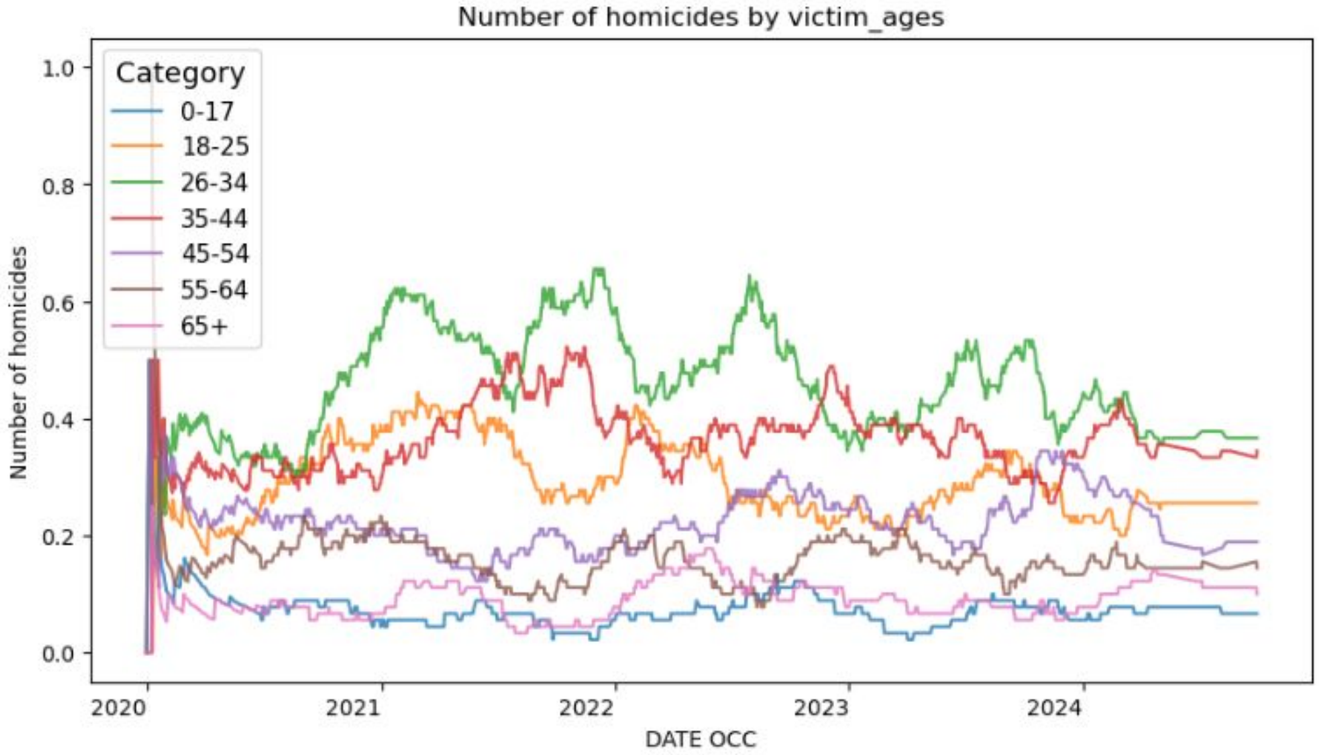


HOMICIDES OVER THE PERIODS



HOMICIDES (AGE OVERVIEW)





CONCLUSION

As much as we did find visual indication that there is some sort of relation between crime category and victim's age, and between homicide and victim's sex, it would be hard to affirm these relationships base only on that.

But we can know that **the age group of 26-34** has the highest number of victims, with a significantly higher number of **male victims** compared to female victims.

Hence, the high-risk gender is **males group** and age groups is people **aged 26-34** that may require more resource allocation for victim support services, and policies to enhance public safety.

Thank You for Your Listening !