Uruguayan Crimes

July 19, 2025

1 Uruguayan reported crimes

1.1 What is this about?

Using the dataset from the Ministerio del Interior we will try to achieve some analysis about this things: - The anual growth of crimes on Uruguay, per department and in general. - The trend on crime types. - Does the neighbor with the most crimes reported got better over the year? And viceversa?

The dataset can be found here It has over 2 million rows being extremely helpful for my learning. What each column has can be found here

1.2 Data Insight

From the data that was retrieved I decided to only get some of these columns for this view.
After some cleaning on the data I got this results:
Columns:
- crime : Crime type - date : Date of the crime - year : Year of the crime - month : Month of the crime - dept : Department of Uruguay where the crime took place - barrio : Neighborhood on Montevideo where the crime took place.

Total: 2296511 crimes reported.

Starting date: 2013-01-01

Last date: 2025-03-31

[1]:
```
sql_query = 'SELECT * FROM delitos_uruguay.clean_reports'

df = pd.read_sql(sql_query,engine)
```

Now that we have correctly upload the data to this file lets start by reviewing what we want to do.

Starting from the first one, we want to retrieve the anual growth of crimes on Uruguay, filter this by department and in general.
I know we can retrieve the data from SQL but my main objective is to learn to do it with Pandas.

We want to group the data by years and then we can use it to split it between departments, see the trend and see if it gets better or worse overtime.
2025 was excluded since there is data only up to march 31.

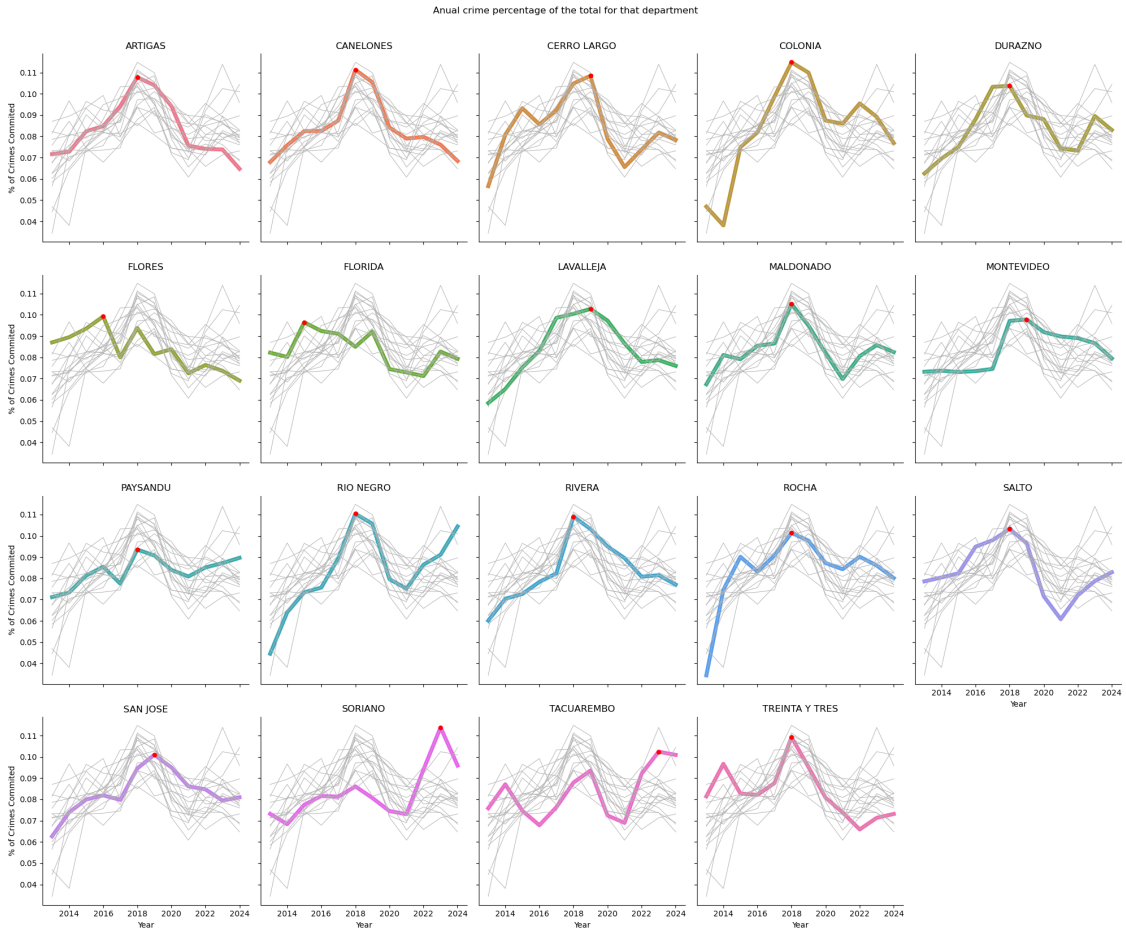
df_crimes_year_dept : Contains yearly crimes splitted by departments.
df_crimes_year : Contains all the crimes accumulated per year.

df_montevideo : Contains all crimes ocurred in Montevideo.
df_crime_type : Contains all crimes accumulated by crime, year and month.

Note: It seems that we cannot get a permanent state of grouped data like with dplyr on R, so we need to adapt.

1.3 Growth and crime rate

To answer out first question we need to create some visualizations, we will start with a facet grid to see the yearly crimes individually.



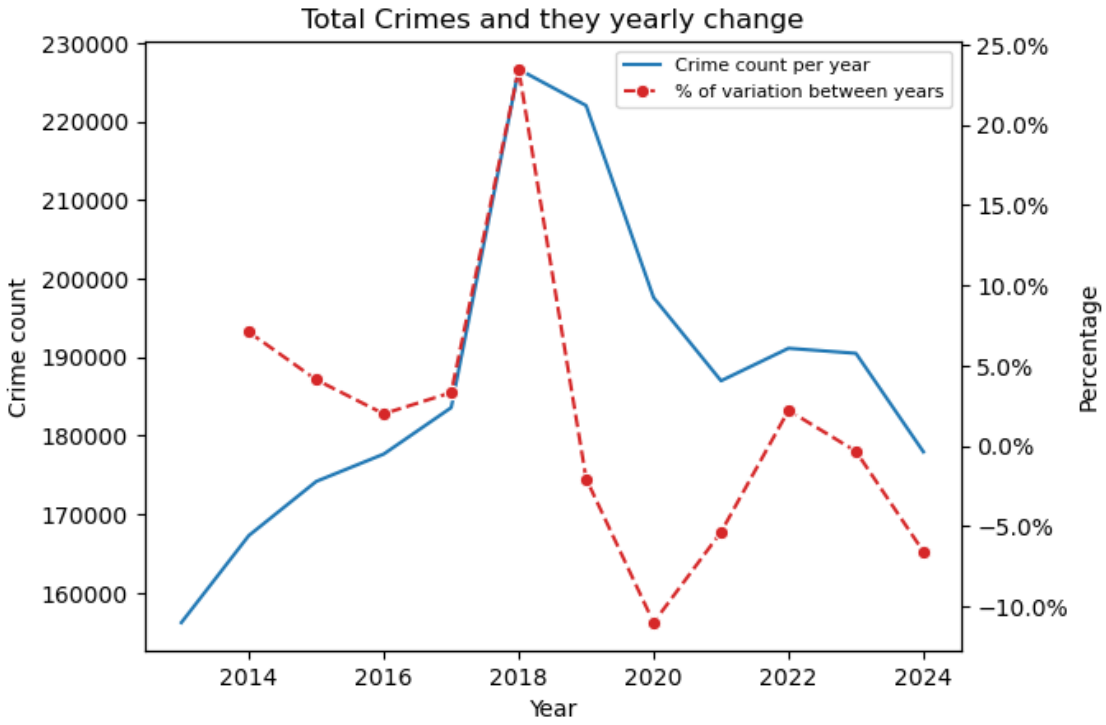
Main plot code:

```
g = sns.relplot(
    data = df_crimes_year_dept,
    x = 'year', y = 'percentage', col = 'dept', hue = 'dept',
    kind = 'line', linewidth = 5, col_wrap = 5, height = 4, legend=False
)
```

Within this graphs we can see something interesting.
In 17/19 departments there was a big spike on crime reports between 2018-2019.

As it may have been expected, on 2020 due to the global pandemic there was a decrease on all the departments.

On the whole region we can not only see that max and big spike in reports from 2017 to 2018 being almost an increase of: 25% corresponding to previous year, I will research why this did happen. Also after the pandemic the crime rate seems to be normalized and seems to be decreasing since 2022.



Codes for blue and red graph

```
sns.lineplot(
    data = df_crimes_year,
    x='year', y='crime_count', ax=ax1,
    color = "tab:blue", label="Crime count per year", legend = None,
)
sns.lineplot(
    data =df_crimes_year,
    x='year', y='yearly_change', ax=ax2,
    color="tab:red", ls='--', label="% of variation between years",
    legend = None, marker = 'o'
)
```

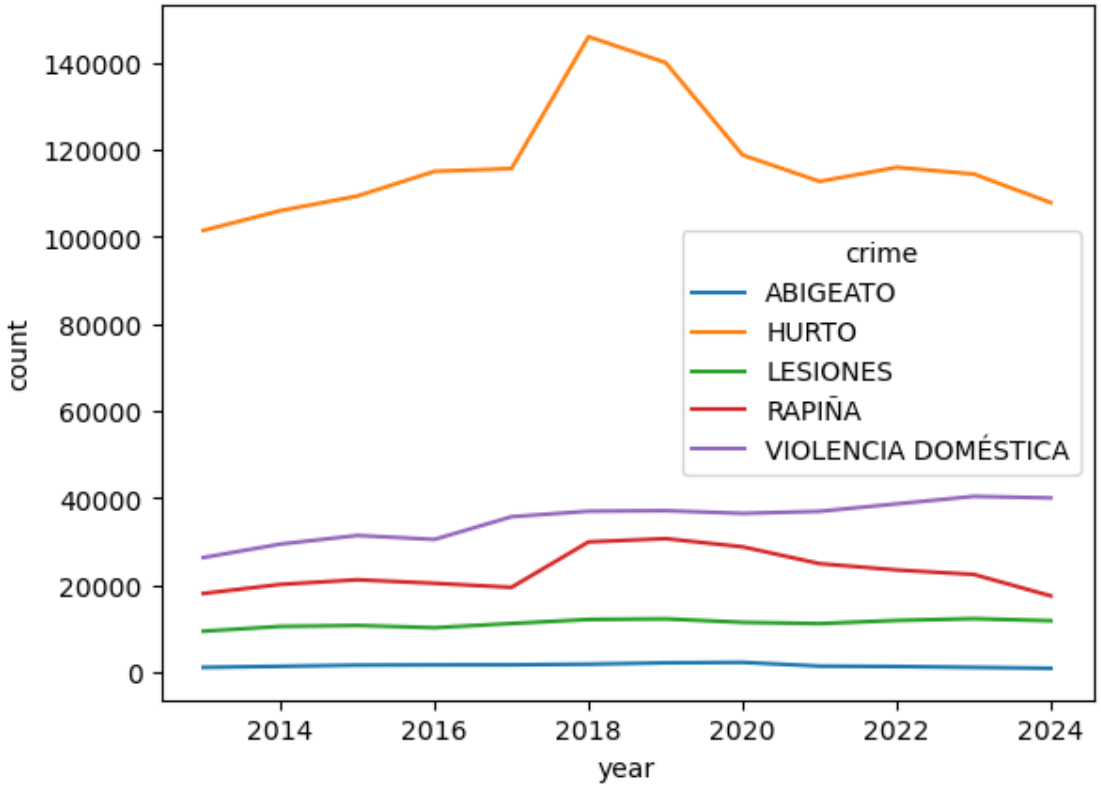
1.4 Crime Trends

We need to see which one is the most trending to happend and know they behavior, are they decreasing, increasing or keeping stable for each year.

From the metadata we found that there are this type of crimes: - Hurto : Steal - Rapiña : Mug - Violencia Domestica : Domestic violence - Lesiones : - Abigeato :

[10]:

year	2013	2018	2020	2024
crime				
ABIGEATO	1059	1794	2206	852
HURTO	101453	145973	118783	107845
LESIONES	9370	12069	11408	11780
RAPIÑA	18035	29885	28757	17480
VIOLENCIA DOMÉSTICA	26288	36946	36445	39990




```
sns.lineplot(
    data = df_yearly_crime_type,
    x='year', y='count', hue = 'crime',
)
```

Using the plot and looking on the resumed table where we choose 4 different years we can see some interesting things.

- Abigeato has decreased a lot since 2021, being on the all time low on 2024.
- Hurto has it's peak on 2018 as well as Rapiña did.
- Hurto is the most current type of crime.
- All the crimes are currently decreasing over the time except domestic violence.

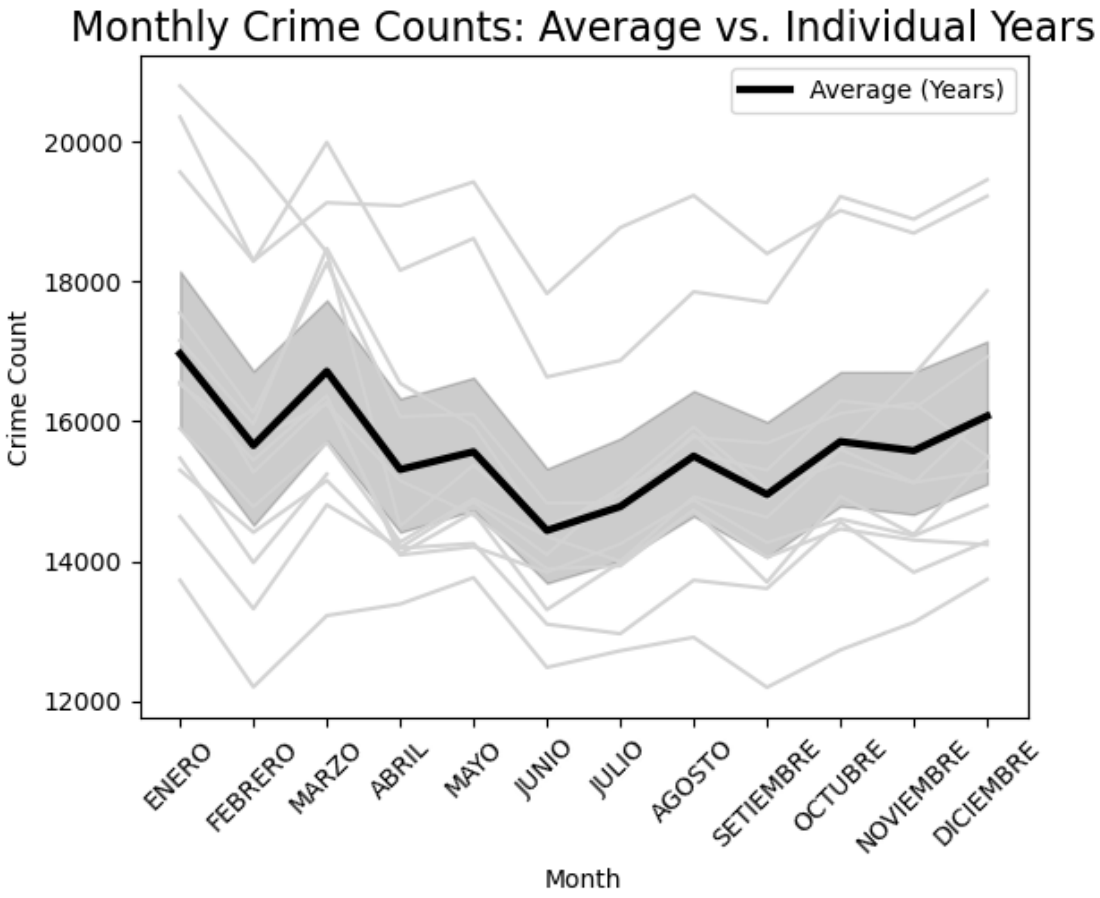
- Domestic violence seems to be increasing steady and slowly over the years, on 2023 it got it's max.

This is really concerning insights.

1.5 What month of the year tends to have more crimmes being commited?

On the following graph we can see that most of the crimes, on average, occur between January and March each year.
Then tend to decrease and re increse on the summer like months.

One would have thought that crimes tend to occur more on the winter nights than on summer.



Looking into our results, we can say and get to understand that most of the assaults and other crimes are commited during summer times with a gradual increase over the years for Domestic Violence.

1.6 Neighborhoods

We want to understand what happened with the top 5 neighborhoods on Montevideo with the most reports and last 5 ones.

On 2024, will they be better or worse?

[ ]: