

## **Escopo**

Este projeto se insere no campo de análise de crimes urbanos, utilizando dados de ocorrências criminais da cidade de Los Angeles, com o foco de criar uma visualização clara e organizada dos tipos de crimes e seus padrões. O estudo visa facilitar a compreensão e a análise dos dados por meio de gráficos e visualizações.

## **Desafio**

O desafio principal é representar visualmente a distribuição dos crimes com base em variáveis como o período do dia, gênero da vítima, tipo de crime e local, além de dados temporais como a data da ocorrência. Com um grande volume de dados, é necessário filtrar e tratar os dados para garantir que apenas as informações relevantes sejam consideradas.

## **Projeto de Visualização**

As seguintes etapas foram tomadas para gerar as representações visuais:

1. **Coleta e Tratamento dos Dados:** O dataset foi importado, os dados foram processados para remover colunas irrelevantes e valores inválidos. A coluna de data foi transformada para o formato datetime, e foi criada uma coluna adicional chamada 'YearMonth', que facilita a análise mensal dos crimes. Além disso, a idade das vítimas foi filtrada para valores realistas, e uma nova coluna foi criada para categorizar os crimes em períodos do dia.
2. **Criação de gráficos comparativos:** Foram gerados gráficos comparando os locais mais comuns, tipos, suas áreas e a distribuição de vítimas por sexo, além de analisar as ocorrências criminais por latitude e longitude.

3. **Visualização de crimes ao longo do tempo:** Gráfico que mostra a contagem de crimes por mês, permitindo observar tendências temporais.
4. **Análise por período do dia:** a análise das ocorrências de crimes durante diferentes períodos do dia (manhã, tarde, noite e madrugada).

## **Mapeamento**

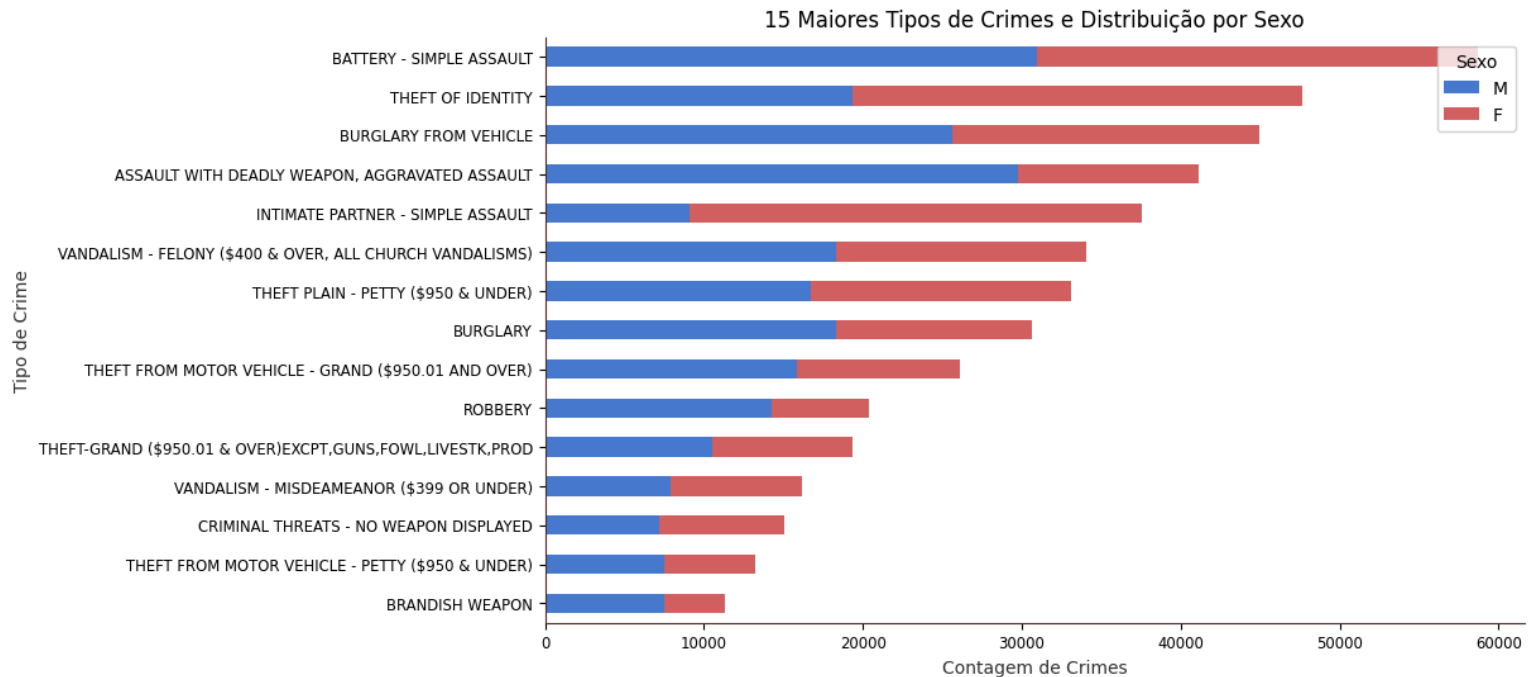
A codificação utilizada para o mapeamento dos dados de entrada incluiu as colunas:

- **CrimeType:** Descrição do tipo de crime cometido.
- **Vict Sex:** Sexo da vítima (F - Feminino, M - Masculino, X - Desconhecido).
- **DATE OCC:** Data em que o crime ocorreu (MM/DD/YYYY).
- **Vict Age:** Idade da vítima.
- **LAT:** Latitude geográfica da localização do crime.
- **LON:** Longitude geográfica da localização do crime.
- **YearMonth:** Mês e ano da ocorrência do crime, utilizado para contagem de crimes por mês.
- **Período:** Classificação da ocorrência do crime por períodos do dia.

## Representações Visuais (RVs)

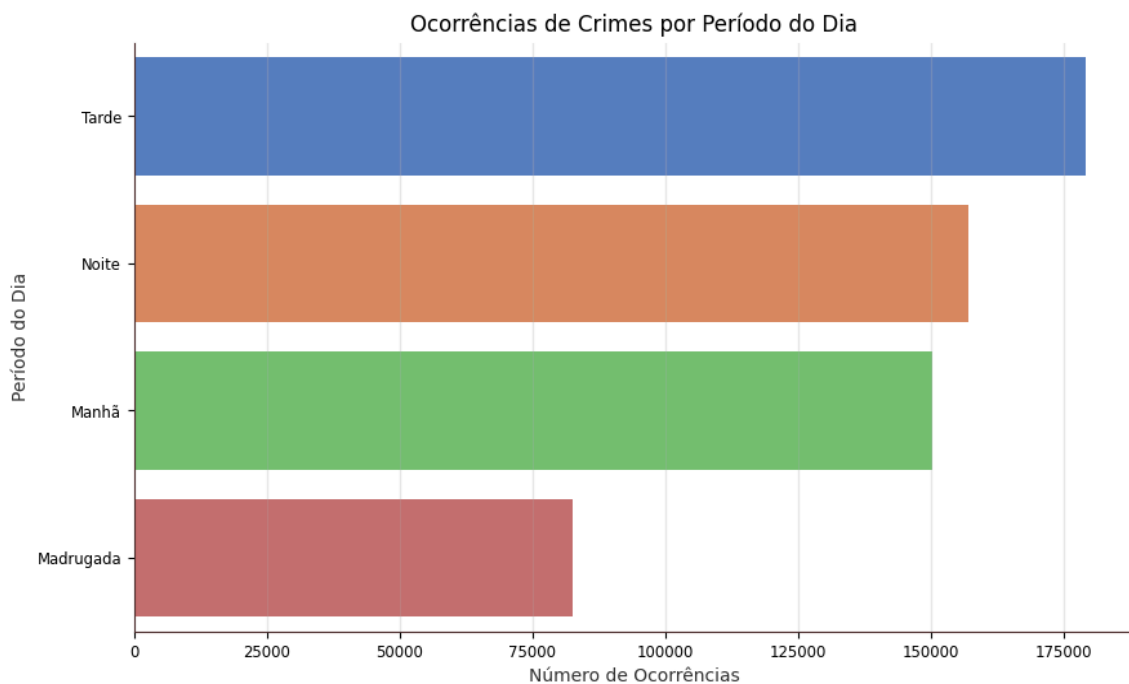
As representações visuais geradas incluem:

1. Gráfico de barras horizontal destacando os 15 maiores tipos de crimes e a distribuição por sexo (*Figura 1*).



*Figura 1 - Os 15 maiores tipos de crimes e a distribuição por sexo*

2. Gráfico de barras horizontal destacando as ocorrências de crimes por período do dia (*Figura 2*).



*Figura 2 - Ocorrências de crimes por período do dia.*

3. Mapa de calor destacando a densidade dos crimes em Los Angeles (Figura 3).

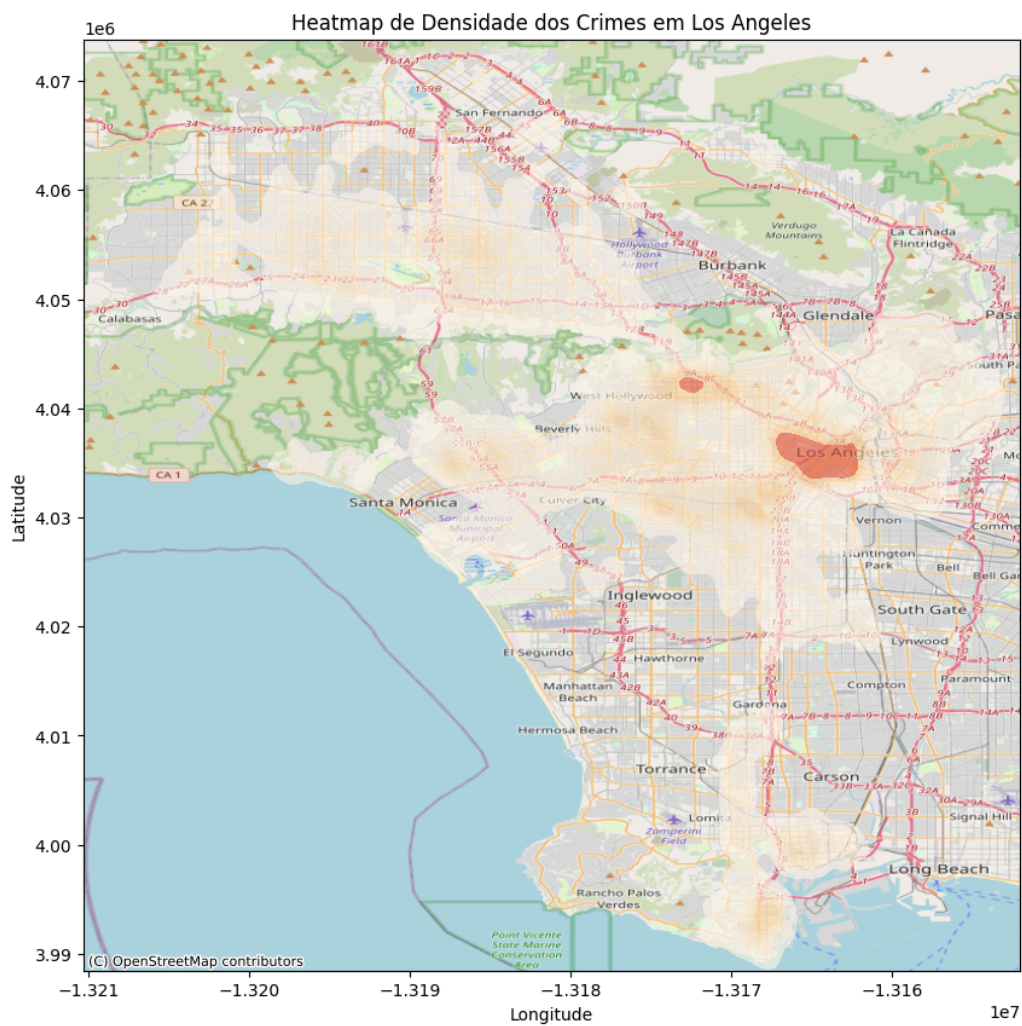


Figura 3 - Densidade dos crimes em Los Angeles.

4. Gráfico de linhas destacando a contagem de crimes por mês (Figura 4).

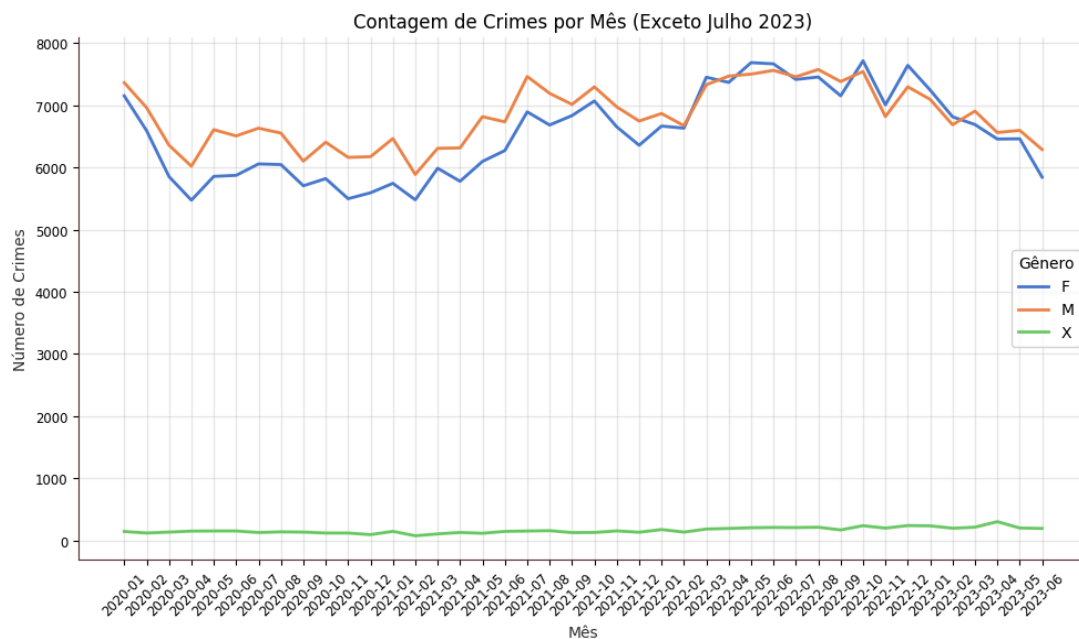


Figura 4 - Série temporal da contagem de crimes por mês (exceto julho de 2023)

## Análise de Dados

A análise dos dados visuais revela padrões importantes sobre a criminalidade. A figura 1 mostra que os crimes mais frequentes são BATTERY - SIMPLE ASSAULT e THEFT OF IDENTITY, com uma predominância de ocorrências entre homens. A distribuição dos crimes ao longo do dia, conforme a figura 2, indica que a tarde é o período com maior número de crimes, seguido pela noite. A análise mensal, representada pela figura 4, mostra variações significativas ao longo do ano, com picos em determinados meses, como agosto. Além disso, o heatmap revela duas manchas vermelhas, uma menor e outra muito maior, indicando regiões com alta concentração de crimes.

## Takeaways

1. **Principal crime:** "Battery - Simple Assault", o que aponta para a necessidade de leis mais rigorosas para coibir sua recorrência.
2. **Período do dia:** A tarde é o período com maior atividade criminosa, sugerindo a necessidade de reforço na segurança durante esse tempo. A noite também requer atenção, enquanto a manhã e a madrugada são períodos de menor risco.
3. **Tendências mensais e geográficas:** A variação mensal dos crimes, com picos, e a concentração geográfica em áreas vermelhas no heatmap, destacam a importância de estratégias sazonais e regionais de prevenção ao crime

```
# Série temporal
df2 = df.dropna(subset=['DATE OCC'])

df2['YearMonth'] = df2['DATE OCC'].dt.to_period('M').astype(str)

df_filtered = df2[df2['DATE OCC'] < '2023-07-01']

monthly_crime_counts = df_filtered.groupby(['YearMonth', 'Vict Sex']).size().reset_index(name='Count')
sns.lineplot(data=monthly_crime_counts, x='YearMonth', y='Count', hue='Vict Sex')

plt.title('Contagem de Crimes por Mês (Exceto Julho 2023)')
plt.xlabel('Mês')
plt.ylabel('Número de Crimes')

plt.xticks(rotation=45)
plt.tight_layout()

plt.legend(title='Gênero')

plt.show()

# Heatmap
df = gpd.GeoDataFrame(
    df, geometry=gpd.points_from_xy(df['LON'], df['LAT'], crs="EPSG:4326"))

df = df.to_crs(epsg=3857)

fig, ax = plt.subplots(figsize=(10, 10))

x = df.geometry.x
y = df.geometry.y

sns.kdeplot(x=x, y=y, ax=ax, fill=True, thresh=0.05, cmap="OrRd", alpha=0.6, bw_adjust=0.25)

ctx.add_basemap(ax, source=ctx.providers.OpenStreetMap.Mapnik)

ax.set_xlim(df.geometry.x.min() - buffer, df.geometry.x.max() + buffer)
ax.set_ylim(df.geometry.y.min() - buffer, df.geometry.y.max() + buffer)

plt.title("Heatmap de Densidade dos Crimes em Los Angeles")
ax.set_xlabel('Longitude')
ax.set_ylabel('Latitude')
```

Figura 5 - Trecho de código para a plotagem da figura 4 e 3 respectivamente.