

Análise de Crimes em Curitiba

CSI Curitibanos

João Felipe Sarggin Machado
João Lucas Mizuguchi Ferreira
Pedro Sodré dos Santos
Thiago Ramos Bernardo

Perguntas de Pesquisa

- Como a variável “renda”, “taxa de alfabetização” e “distância em relação ao centro” influenciam nas quantidades de crimes nos bairros de Curitiba?
- Como o horário do crime (e outras variáveis do bairro) influenciam no tipo de crime?
- Como o período do ano influencia na quantidade total de crimes na cidade de Curitiba?

Como a variável “renda”, “taxa de alfabetização” e “distância em relação ao centro” influenciam nas quantidades de crimes nos bairros de Curitiba?

Modelo

Regressão múltipla

Hipótese

A renda, a taxa de analfabetização e a distância em relação ao centro influenciam na quantidade de crimes.

Dados

Serão utilizados dados dos bairros, do dataset retirado do InfoCuritiba. Além dos dados da Guarda Municipal, que foi filtrado na análise exploratória (“analise_crime_filtrado.csv”).

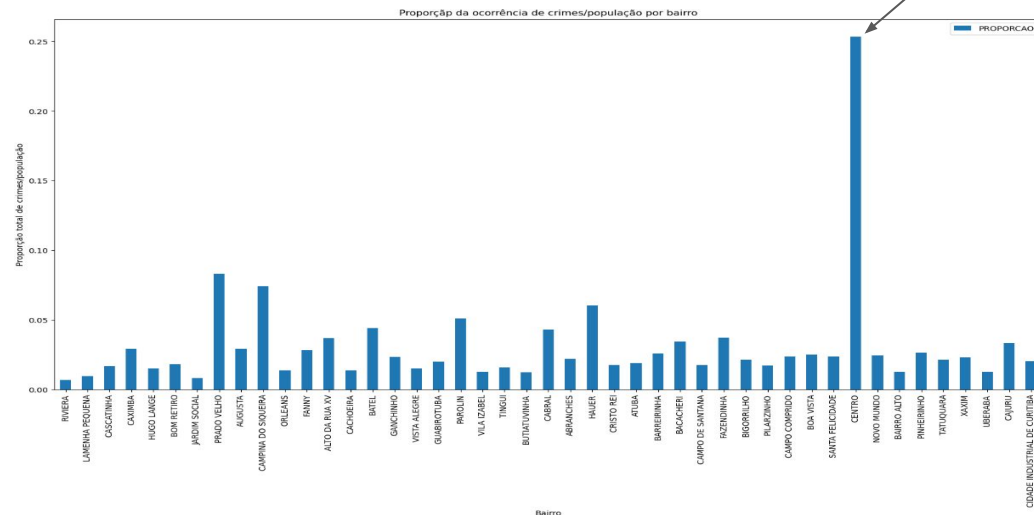
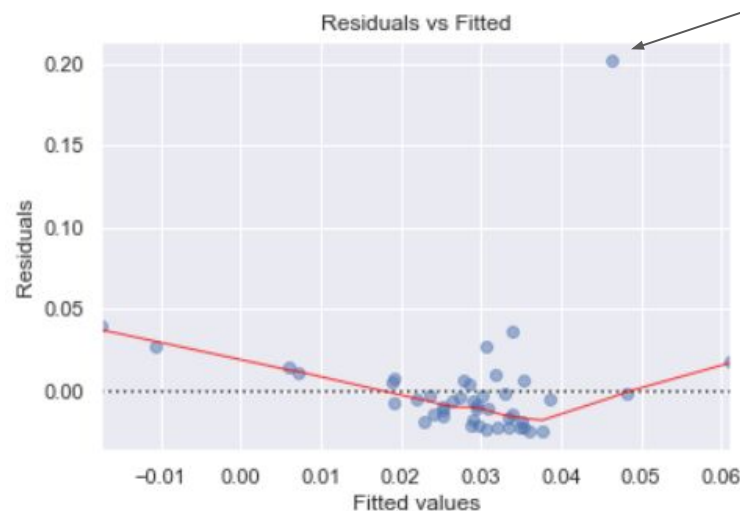
Como a variável “renda”, “taxa de alfabetização” e “distância em relação ao centro” influenciam nas quantidades de crimes nos bairros de Curitiba?

Dataframe final para aplicar no modelo

	ATENDIMENTO_BAIRRO_NOME	POPULACAO_TOTAL	RENDIMENTO_MENSAL_MEDIO	TAXA_ANALFABETISMO	TOTAL_POR_BAIRRO	CRIMINALIDADE	DISTANCIA_CENTRO
1	CIDADE INDUSTRIAL DE CURITIBA	172822	779.67	4.34	2994	0.017324	11084
2	CAJURU	96200	873.67	4.64	2969	0.030863	7848
3	PINHEIRINHO	50401	844.62	3.95	1214	0.024087	11665
4	XAXIM	57182	1040.56	3.07	1174	0.020531	9391
5	NOVO MUNDO	44063	1231.25	2.93	982	0.022286	7527
6	FAZENDINHA	28074	984.18	3.64	964	0.034338	8598
7	TATUQUARA	52780	590.06	6.11	952	0.018037	17171
8	HAUER	13315	1336.02	1.94	770	0.057830	6141
9	BACACHERI	23734	2216.40	1.27	766	0.032274	5859
10	UBERABA	72056	1101.82	3.51	713	0.009895	9209
11	BOA VISTA	31052	1600.36	1.60	659	0.021222	6102
12	SANTA FELICIDADE	31572	1652.46	2.46	645	0.020429	7731
13	CAMPO COMPRIDO	28816	1784.40	3.68	587	0.020371	7342
14	BIGORRILHO	28336	3729.56	0.74	560	0.019763	3170
15	CABRAL	13060	3519.61	0.70	544	0.041654	3835

Como a variável “renda”, “taxa de alfabetização” e “distância em relação ao centro” influenciam nas quantidades de crimes nos bairros de Curitiba?

Centro como outlier



Como a variável “renda”, “taxa de alfabetização” e “distância em relação ao centro” influenciam nas quantidades de crimes nos bairros de Curitiba?

Regressão Linear - Distância

Dep. Variable:	CRIMINALIDADE	R-squared:	0.103
Model:	OLS	Adj. R-squared:	0.082
Method:	Least Squares	F-statistic:	4.949
Date:	Wed, 07 Dec 2022	Prob (F-statistic):	0.0314
Time:	21:26:20	Log-Likelihood:	87.269
No. Observations:	45	AIC:	-170.5
Df Residuals:	43	BIC:	-166.9
Df Model:	1		
Covariance Type:	nonrobust		

C/Centro

Regressão Linear - Taxa de Analfabetismo

OLS Regression Results			
Dep. Variable:	CRIMINALIDADE	R-squared:	0.009
Model:	OLS	Adj. R-squared:	-0.014
Method:	Least Squares	F-statistic:	0.4009
Date:	Wed, 07 Dec 2022	Prob (F-statistic):	0.530
Time:	21:26:21	Log-Likelihood:	85.027
No. Observations:	45	AIC:	-166.1
Df Residuals:	43	BIC:	-162.4
Df Model:	1		
Covariance Type:	nonrobust		

C/Centro

Regressão Linear - Renda

OLS Regression Results			
Dep. Variable:	CRIMINALIDADE	R-squared:	0.022
Model:	OLS	Adj. R-squared:	-0.001
Method:	Least Squares	F-statistic:	0.9552
Date:	Wed, 07 Dec 2022	Prob (F-statistic):	0.334
Time:	21:26:21	Log-Likelihood:	85.312
No. Observations:	45	AIC:	-166.6
Df Residuals:	43	BIC:	-163.0
Df Model:	1		
Covariance Type:	nonrobust		

C/Centro

OLS Regression Results			
Dep. Variable:	CRIMINALIDADE	R-squared:	0.067
Model:	OLS	Adj. R-squared:	0.045
Method:	Least Squares	F-statistic:	3.029
Date:	Wed, 07 Dec 2022	Prob (F-statistic):	0.0891
Time:	22:47:00	Log-Likelihood:	121.40
No. Observations:	44	AIC:	-238.8
Df Residuals:	42	BIC:	-235.2
Df Model:	1		
Covariance Type:	nonrobust		

S/Centro

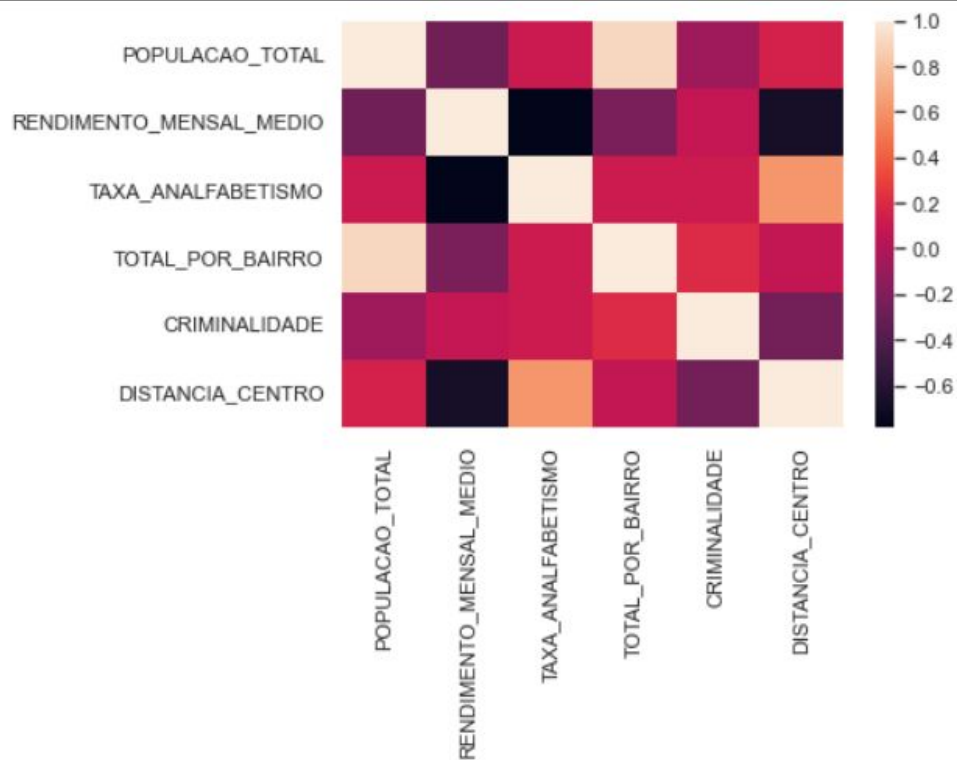
OLS Regression Results			
Dep. Variable:	CRIMINALIDADE	R-squared:	0.013
Model:	OLS	Adj. R-squared:	-0.011
Method:	Least Squares	F-statistic:	0.5431
Date:	Wed, 07 Dec 2022	Prob (F-statistic):	0.465
Time:	22:47:01	Log-Likelihood:	120.15
No. Observations:	44	AIC:	-236.3
Df Residuals:	42	BIC:	-232.7
Df Model:	1		
Covariance Type:	nonrobust		

S/Centro

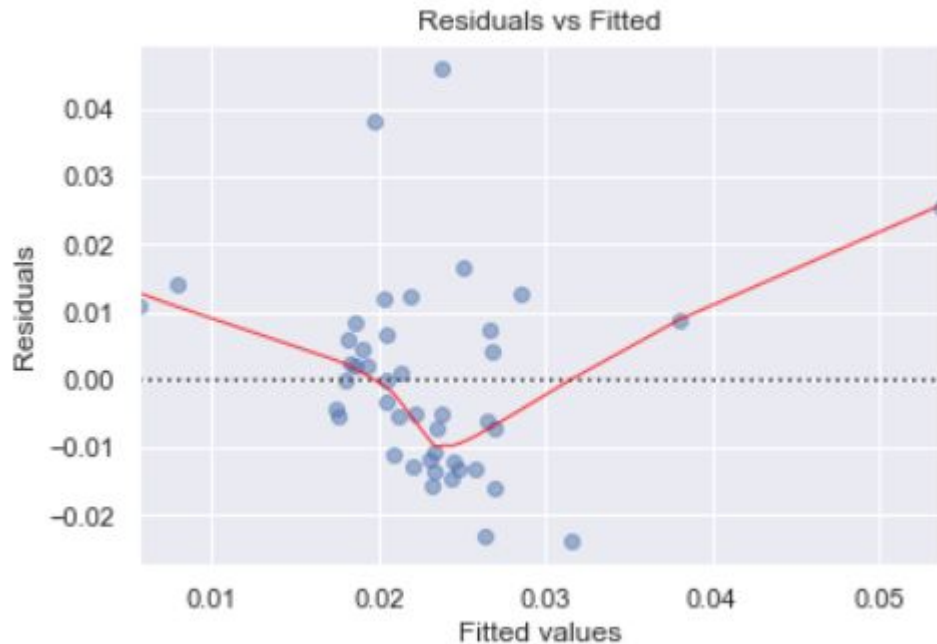
OLS Regression Results			
Dep. Variable:	CRIMINALIDADE	R-squared:	0.006
Model:	OLS	Adj. R-squared:	-0.018
Method:	Least Squares	F-statistic:	0.2443
Date:	Wed, 07 Dec 2022	Prob (F-statistic):	0.624
Time:	22:47:01	Log-Likelihood:	119.99
No. Observations:	44	AIC:	-236.0
Df Residuals:	42	BIC:	-232.4
Df Model:	1		
Covariance Type:	nonrobust		

S/Centro

Como a variável “renda”, “taxa de alfabetização” e “distância em relação ao centro” influenciam nas quantidades de crimes nos bairros de Curitiba?



Como a variável “renda”, “taxa de alfabetização” e “distância em relação ao centro” influenciam nas quantidades de crimes nos bairros de Curitiba?



Como a variável “renda”, “taxa de alfabetização” e “distância em relação ao centro” influenciam nas quantidades de crimes nos bairros de Curitiba?

- Em pequena escala - cidade - as variáveis socioeconômicas podem ser não tão relevantes
- Falta de variáveis relevantes e difíceis de obter (policimento, fluxo de pessoas, iluminação)
- Falta de outros índices socioeconômicos (IDH, grau de escolaridade, etc.)

Como o horário do crime influencia no tipo de crime?

Modelo

Clusterização Hierárquica

Hipótese

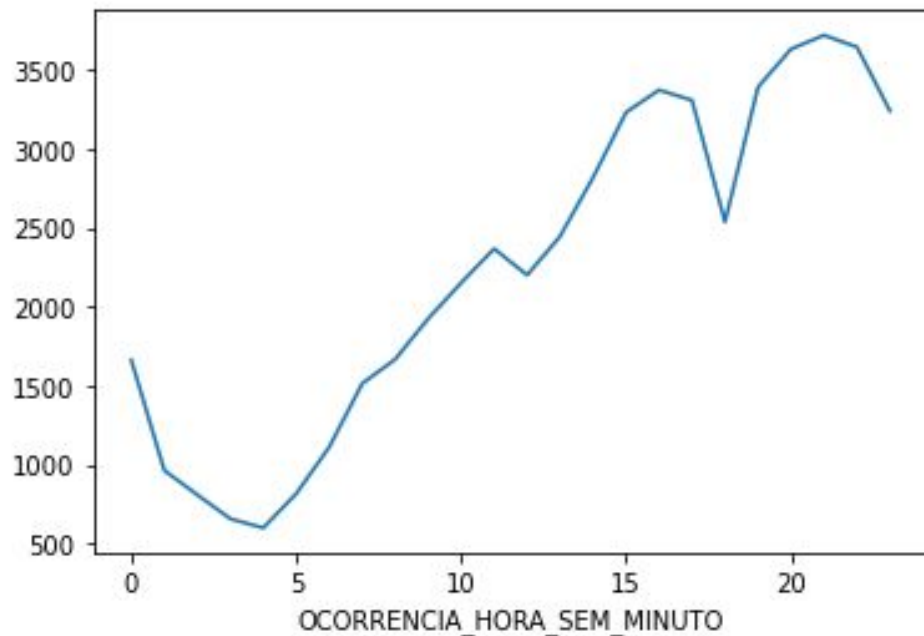
Certos tipos de crime acontecem com mais frequência após o horário de expediente (depois das 17h).

Dados

Serão utilizados dados dos bairros, do dataset retirado do InfoCuritiba. Além dos dados da Guarda Municipal, que foi filtrado na análise exploratória (“analise_crime_filtrado.csv”).

Como o horário do crime influencia no tipo de crime?

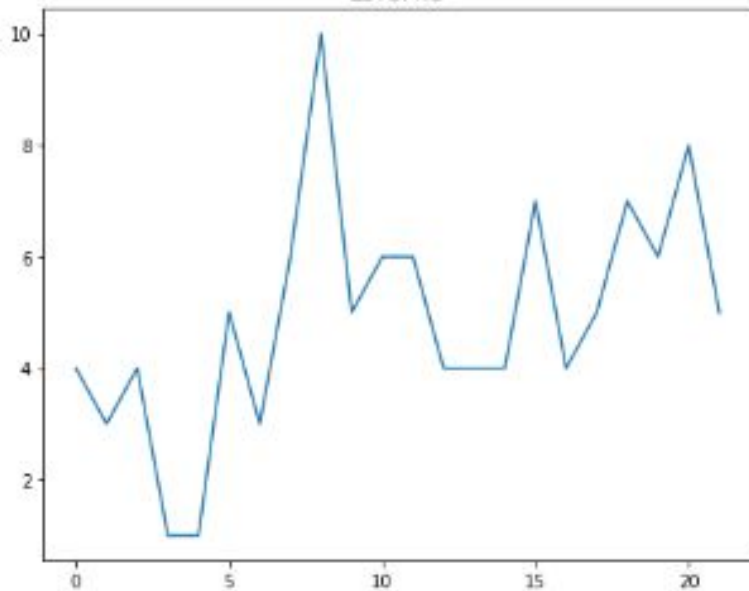
Série temporal de todos os crimes:



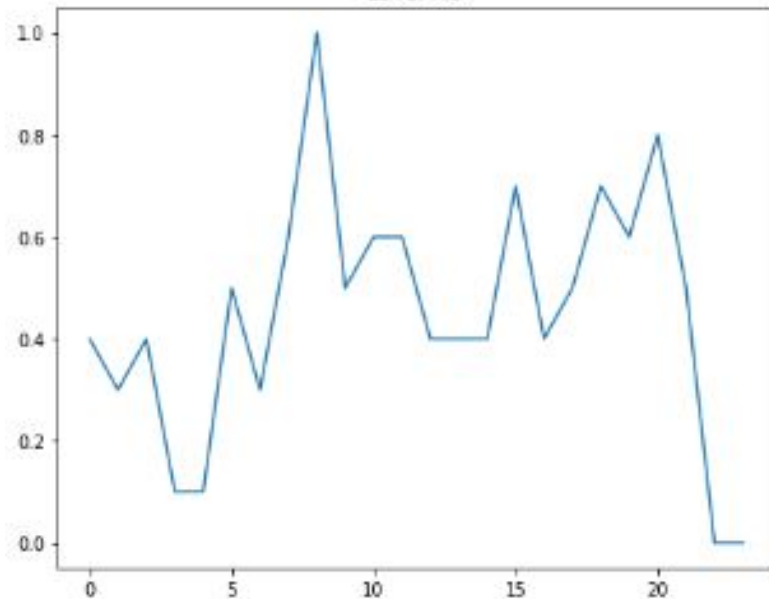
Como o horário do crime influencia no tipo de crime?

Normalização:

ESTUPRO

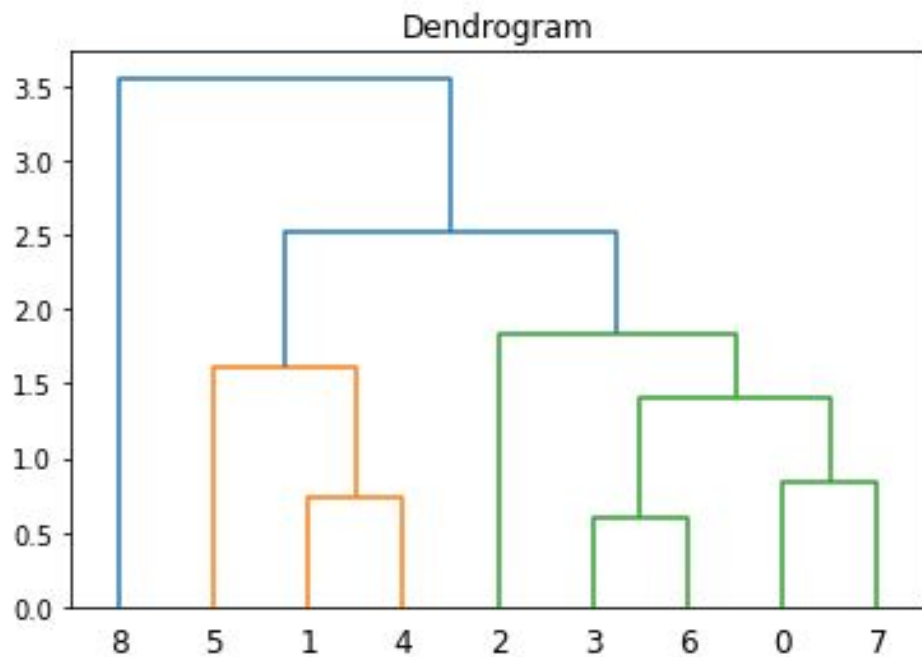


ESTUPRO



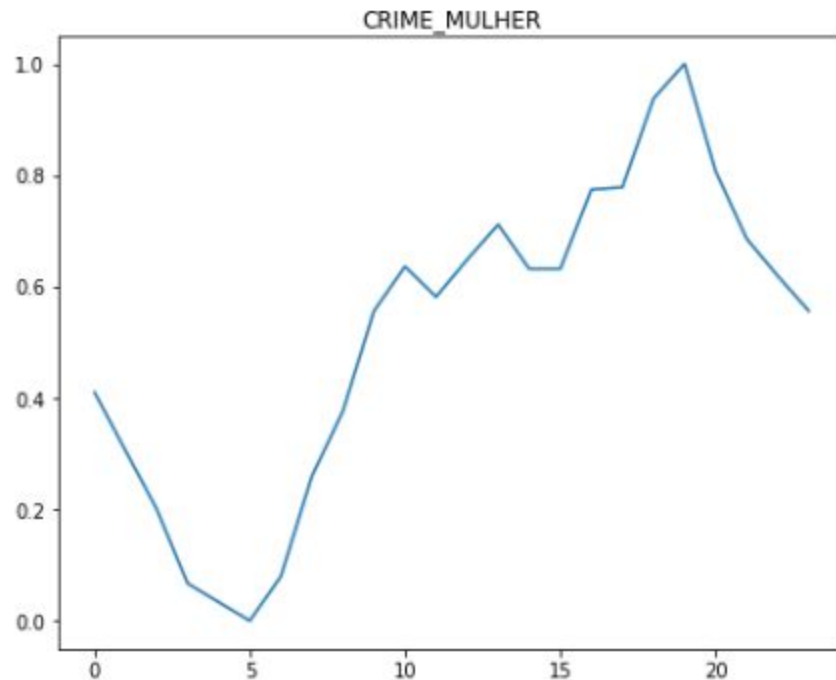
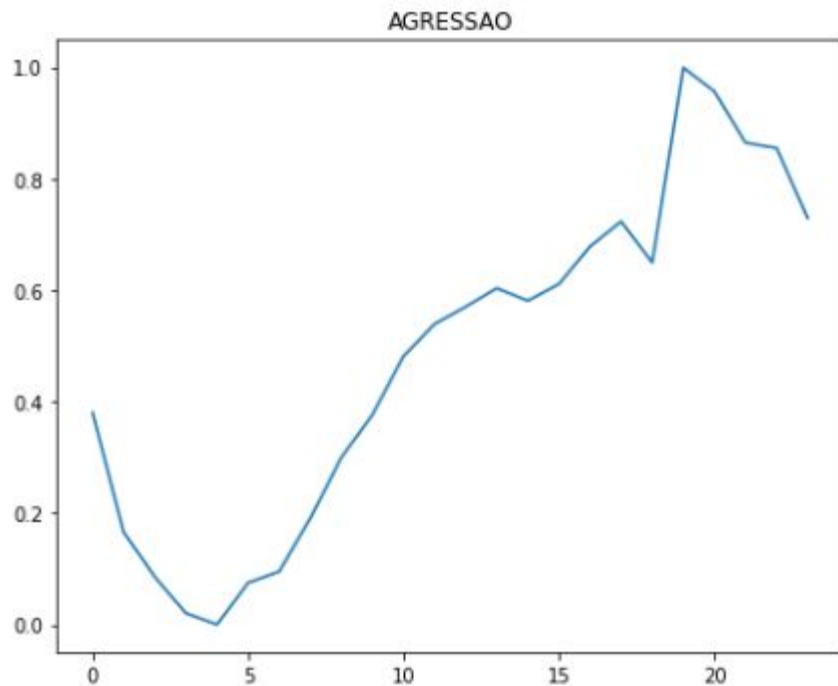
Como o horário do crime influencia no tipo de crime?

Dendrograma da Clusterização Hierárquica:



Como o horário do crime influencia no tipo de crime?

Agressão X Crime Mulher :

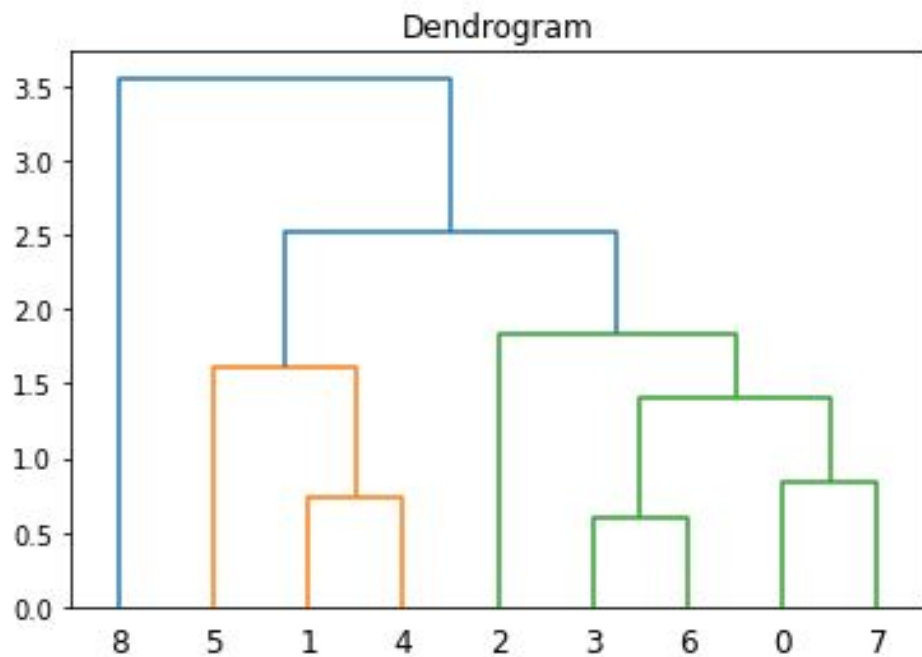


Como o horário do crime influencia no tipo de crime?

- Semelhança poderia ser explicada se os crimes contra a mulher também envolvessem denúncias de violência doméstica ou agressões;
- *“Entre as agressões relatadas nos últimos doze meses, a maioria ocorreu em casa (42%), enquanto 29% na rua e o restante em ambientes como o trabalho, um bar/ balada ou a internet – na lista dos locais mais comuns.”*
- Lei Maria da Penha.
- *“Das que buscaram ajuda, 23,8% disseram ter procurado algum órgão especializado (em instituições como delegacias da mulher ou delegacias comuns, além de ligações para o 190). Já 15% compartilharam a situação com alguém da família ou pessoas próximas.”*
- Trecho retirado do site: [Veja Abril](#)

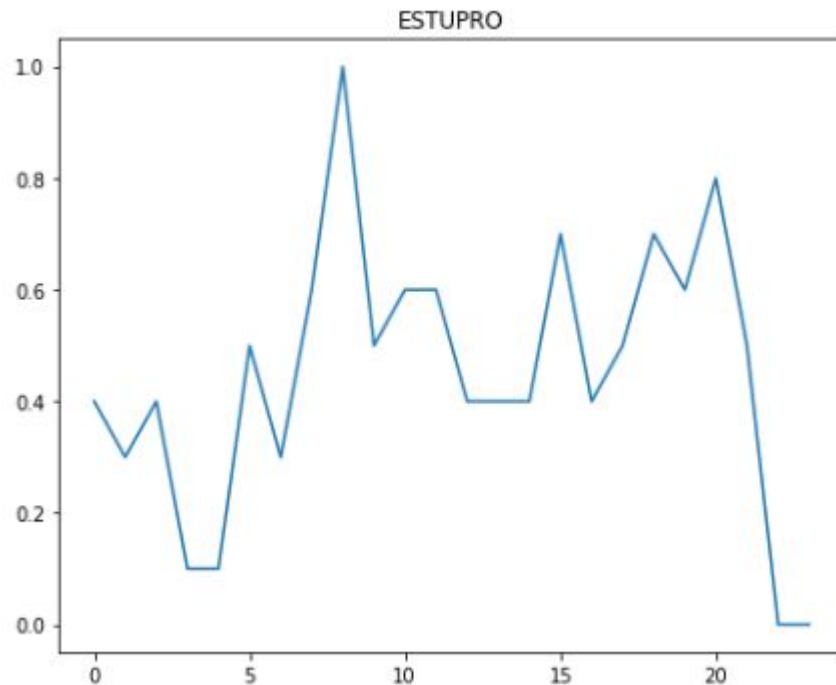
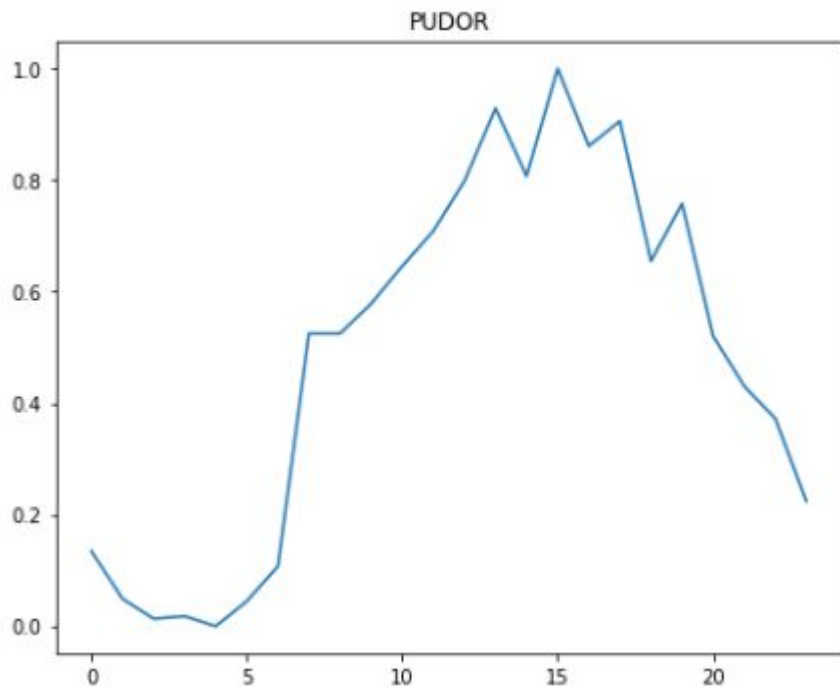
Como o horário do crime influencia no tipo de crime?

Dendrograma da Clusterização Hierárquica:



Como o horário do crime influencia no tipo de crime?

Pudor X Estupro:

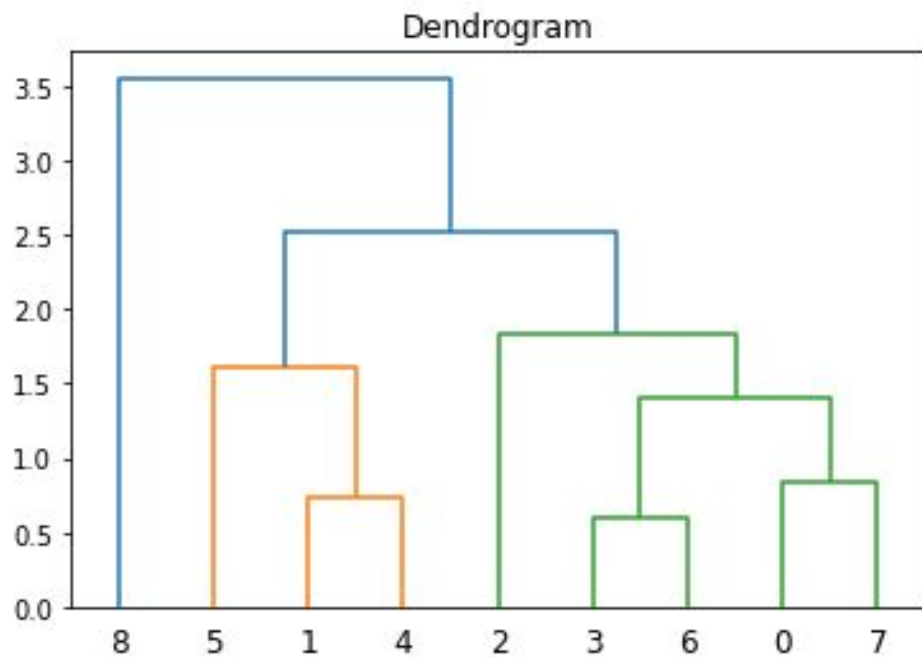


Como o horário do crime influencia no tipo de crime?

- Dentro da categoria de estupro, também está incluso os atentados violentos ao pudor e por isso possui alguma semelhança o grupo de crimes do pudor (menos violentos), porém tendo um número menor de casos;
- *“O crime de atentado violento ao pudor foi inserido dentro do tipo penal do crime de ESTUPRO a partir do ano de 2009 e, dessa forma, a conduta do autor que atuara sob o revogado artigo 214 do Código Penal passou a ser classificada como crime de ESTUPRO (art. 213) e não de estupro de vulnerável (artigo 217-A).”*
- Importunação Sexual vs Estupro
- Trecho retirado do site: [ConJur](#)

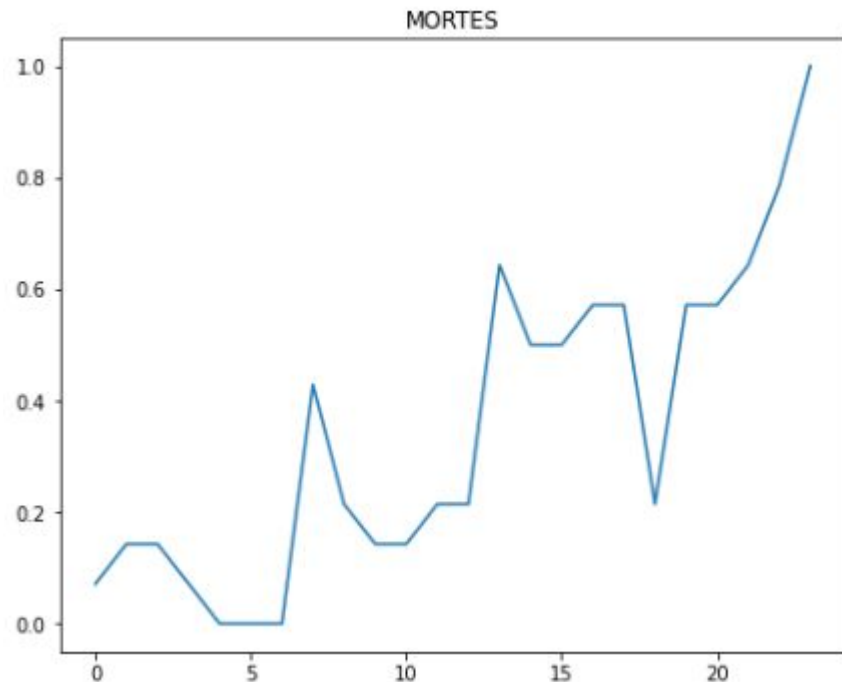
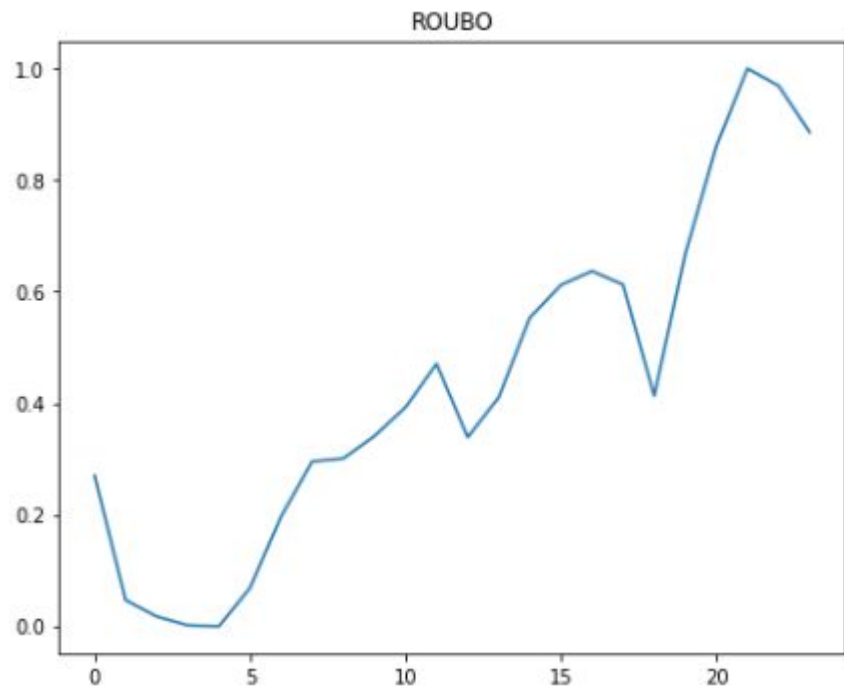
Como o horário do crime influencia no tipo de crime?

Dendrograma da Clusterização Hierárquica:



Como o horário do crime influencia no tipo de crime?

Roubo x Mortes:

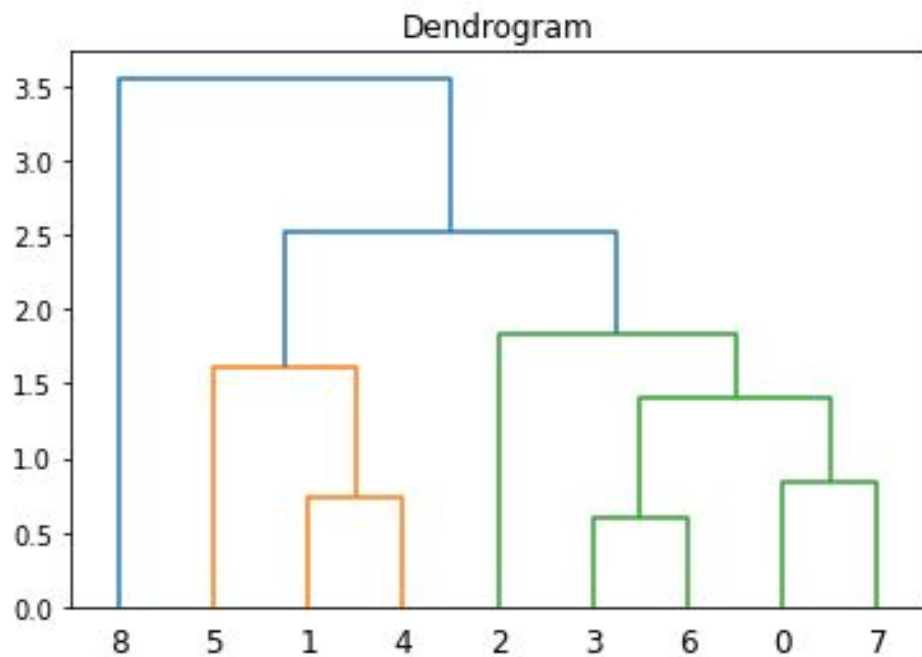


Como o horário do crime influencia no tipo de crime?

- Comportamento é esperado pois esses crimes ocorrem no fim da noite e uma parte dos roubos são assaltos a mão armada, aumentando as chances de morte da vítima;
- *“Somente na capital, foram 79 casos de roubos ou tentativa de roubos com tiros, que resultaram em 28 mortos e 52 feridos — 80 baleados no total.*
- *Em toda a Região Metropolitana, somente no mês de junho houve 20 casos de tiroteios durante tentativas de assalto, com oito mortos e 12 feridos.” RJ*
- Trecho retirado do site: [G1](#)

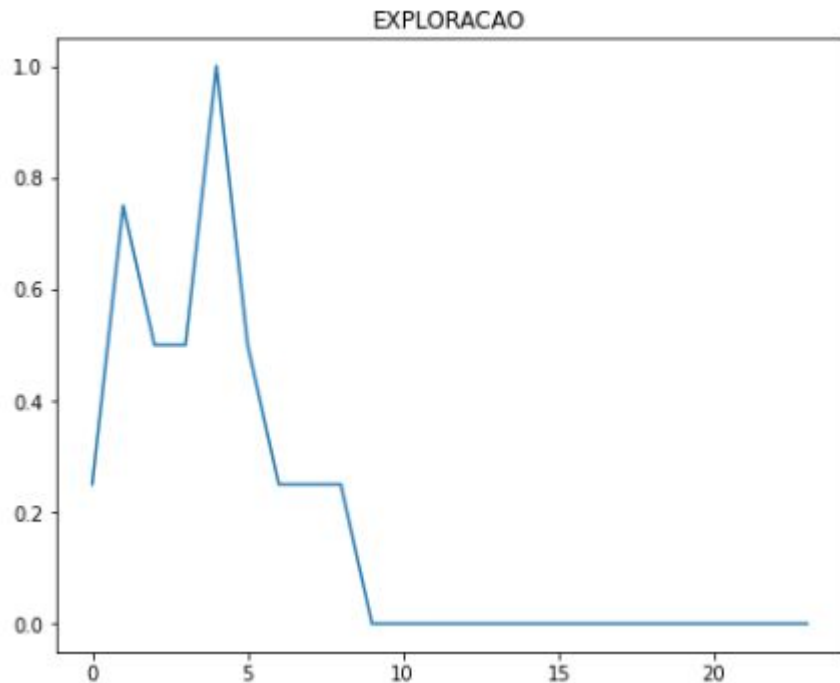
Como o horário do crime influencia no tipo de crime?

Dendrograma da Clusterização Hierárquica:



Como o horário do crime influencia no tipo de crime?

Outlier



- Poucos dados registrados (17 ocorrências) e por isso temos um resultado zerado por grande parte do dia, dificultando verificar a semelhança deste com os demais crimes;

Como o período do ano influencia na quantidade total de crimes na cidade de Curitiba?

Modelo

Regressão por dados de painel

Hipótese

Períodos festivos, épocas do ano (estações do ano) e período de férias (julho, agosto, dezembro e janeiro) tem impacto sobre o número semanal de casos.

Dados

Dados de ocorrência da Guarda Municipal de Curitiba e calendário de feriados municipais/nacionais.

Como o período do ano influencia na quantidade total de crimes na cidade de Curitiba?

Dataframe final para aplicar no modelo

	DATA	ANO	MES	DIA_SEMANA	CASOS	FERIADO	TEMPERATURA_MEDIA	CHUVA
0	2009-01-01	2009	1	4	2	1	20.4	233
1	2009-01-02	2009	1	5	4	0	20.4	233
2	2009-01-03	2009	1	6	3	0	20.4	233
3	2009-01-04	2009	1	7	9	0	20.4	233
4	2009-01-05	2009	1	1	7	0	20.4	233

Como o período do ano influencia na quantidade total de crimes na cidade de Curitiba?

OLS Regression Results

Dep. Variable:	CASOS	R-squared:	0.002
Model:	OLS	Adj. R-squared:	0.001
Method:	Least Squares	F-statistic:	2.611
Date:	Wed, 07 Dec 2022	Prob (F-statistic):	0.0337
Time:	22:38:57	Log-Likelihood:	-15027.
No. Observations:	4793	AIC:	3.006e+04
Df Residuals:	4788	BIC:	3.010e+04
Df Model:	4		
Covariance Type:	nonrobust		

	coef	std err	t	P> t	[0.025	0.975]
Intercept	9.9224	0.681	14.574	0.000	8.588	11.257
FERIADO	-0.2351	0.453	-0.519	0.604	-1.122	0.652
TEMPERATURA_MEDIA	0.1396	0.049	2.846	0.004	0.043	0.236
CHUVA	-0.0068	0.003	-2.458	0.014	-0.012	-0.001

Como o período do ano influencia na quantidade total de crimes na cidade de Curitiba?

Dep. Variable:	CASOS	R-squared:	0.477
Model:	OLS	Adj. R-squared:	0.475
Method:	Least Squares	F-statistic:	272.3
Date:	Wed, 07 Dec 2022	Prob (F-statistic):	0.00
Time:	22:42:20	Log-Likelihood:	-13478.
No. Observations:	4793	AIC:	2.699e+04
Df Residuals:	4776	BIC:	2.710e+04
Df Model:	16		
Covariance Type:	nonrobust		

	coef	std err	t	P> t	[0.025	0.975]
Intercept	3.6319	0.490	7.419	0.000	2.672	4.592
C(ANO)[T.2010]	0.2166	0.301	0.721	0.471	-0.373	0.806
C(ANO)[T.2011]	1.1771	0.300	3.928	0.000	0.590	1.765
C(ANO)[T.2012]	2.2564	0.299	7.550	0.000	1.670	2.842
C(ANO)[T.2013]	3.7487	0.299	12.517	0.000	3.162	4.336
C(ANO)[T.2014]	8.0435	0.299	26.895	0.000	7.457	8.630
C(ANO)[T.2015]	10.9558	0.299	36.633	0.000	10.369	11.542

C(ANO)[T.2016]	9.4913	0.299	31.758	0.000	8.905	10.077
C(ANO)[T.2017]	8.4818	0.299	28.361	0.000	7.896	9.068
C(ANO)[T.2018]	7.9257	0.299	26.501	0.000	7.339	8.512
C(ANO)[T.2019]	8.2900	0.299	27.720	0.000	7.704	8.876
C(ANO)[T.2020]	10.4995	0.299	35.132	0.000	9.914	11.085
C(ANO)[T.2021]	8.1613	0.299	27.289	0.000	7.575	8.748
C(ANO)[T.2022]	9.6087	0.577	16.667	0.000	8.478	10.739
FERIADO	-0.2528	0.328	-0.771	0.441	-0.896	0.390
TEMPERATURA_MEDIA	0.1499	0.035	4.243	0.000	0.081	0.219
CHUVA	-0.0084	0.002	-4.194	0.000	-0.012	-0.004

Como o período do ano influencia na quantidade total de crimes na cidade de Curitiba?

```
=====
                        PanelOLS Estimation Summary
=====
Dep. Variable:          CASOS      R-squared:                  0.0001
Estimator:              PanelOLS   R-squared (Between):       -0.1256
No. Observations:       26860      R-squared (Within):        0.0001
Date:                   Wed, Dec 07 2022  R-squared (Overall):      -2.182e-05
Time:                   23:37:39      Log-likelihood             -4.464e+04
Cov. Estimator:         Unadjusted

                        F-statistic:          1.3642
                        P-value               0.2556
Entities:               11              Distribution:              F(2,26847)
Avg Obs:                2441.8
Min Obs:                22.000
Max Obs:                4466.0
                        F-statistic (robust):  1.3642
                        P-value               0.2556
Time periods:           4793           Distribution:              F(2,26847)
Avg Obs:                5.6040
Min Obs:                1.0000
Max Obs:                11.000
```

```
=====
                        Parameter Estimates
=====
                        Parameter  Std. Err.   T-stat   P-value   Lower CI   Upper CI
-----
const          1.9225    0.0595   32.299   0.0000    1.8059    2.0392
T_MEDIA        0.0076    0.0047    1.6118   0.1070   -0.0016    0.0168
CHUVA         -0.0004    0.0003   -1.4501   0.1470   -0.0009    0.0001
=====
```

```
F-test for Poolability: 1139.4
P-value: 0.0000
Distribution: F(10,26847)
```

Como o período do ano influencia na quantidade total de crimes na cidade de Curitiba?

```
: model = smf.ols("CASOS ~ T_MEDIA + CHUVA + C(REGIONAL_FATO_NOME) + C(OCORRENCIA_ANO)", data=df3)
response = model.fit()
response.summary()
```

: OLS Regression Results

Dep. Variable:	CASOS	R-squared:	0.830
Model:	OLS	Adj. R-squared:	0.827
Method:	Least Squares	F-statistic:	291.0
Date:	Wed, 07 Dec 2022	Prob (F-statistic):	0.00
Time:	23:22:54	Log-Likelihood:	-6021.6
No. Observations:	1515	AIC:	1.210e+04
Df Residuals:	1489	BIC:	1.223e+04
Df Model:	25		
Covariance Type:	nonrobust		

	T_MEDIA	0.3885	0.203	1.917	0.055	-0.009	0.786
	CHUVA	-0.0276	0.012	-2.394	0.017	-0.050	-0.005