

Homicídios por 100mil habitantes em gráficos

Estado de São Paulo e Capital - (2000 - 2010)

*Raul de Sá Durló**

14 janeiro 2018

Abstract

Série de gráficos da taxa de homicídios por cem mil habitantes nas localidades: i) Estado de São Paulo (Capital, Interior e Região Metropolitana - exclusive Capital), do ano 2000 até o ano 2010 e ii) na Capital (Município de São Paulo, Distritos Policiais e Seccionais), somente nos anos 2003 e 2013.

*Mestre em Economia - Unesp/FCLAr

1 códigos: Introdução e pacotes utilizados

O objetivo deste projeto é reescrever minha dissertação de mestrado em ambiente 100% R. O relatório é gerado em `Rmarkdown` e tem como *output* um arquivo `.pdf`.

A primeira parte deste documento apresenta os códigos rodados para obtenção do texto final. O texto final está na segunda parte.

Os seguintes pacotes foram utilizados:

- `readr`: para ler extensão `.rds`;
- `tidyverse`: para manipular `data.frames`;
- `huxtable`: para montar tabelas;
- `lubridate`: para séries temporais.
- `ggpubr`: para organizar figuras
- `ggrepel`: para lidar com `labels` em obj's. `ggplot`
- `treemapify`: para criar gráficos de área

```
library(readr)
library(tidyverse)
library(huxtable)
library(lubridate)
library(ggpubr)
library(ggrepel)
library(treemapify)
```

2 códigos: Estado de São Paulo

2.1 Carregando arquivo:

2.1.1 Para taxas anuais de homicídio por 100mil habitantes

O código abaixo carrega dados anuais de homicídios e população por região no estado de São Paulo:

```
# lendo .rds
estado_sp <- read_rds("C:\\Users\\raul\\Google Drive\\meu_projeto\\dados e scripts\\tabelas_output\\ta

# subset
estado_sp <- estado_sp[,seq(4)]

# agregando
estado_sp <- estado_sp %>%
  group_by(ano) %>%
  summarise(populacao = sum(populacao),
            homicidio = sum(homicidio)) %>%
  ungroup() %>%
  mutate(local = rep("Total", 11)) %>%
  bind_rows(estado_sp)
```

Um objeto do tipo tibble:

```
glimpse(estado_sp)

## Observations: 44
## Variables: 4
## $ ano      <int> 2000, 2001, 2002, 2003, 2004, 2005, 2006, 2007, 2008...
## $ populacao <int> 36351316, 37630106, 38177742, 38709320, 39825226, 40...
## $ homicidio <dbl> 12638, 12475, 11847, 10954, 8753, 7076, 6057, 4877, ...
## $ local     <chr> "Total", "Total", "Total", "Total", "Total", "Total"...
```

Onde:

- ano é o ano de registro,
- população é a contagem da população residente.
- homicidio é o número total de registros de homicídio doloso e
- local são as localidades, com:
 - *Capital*: município de São Paulo,
 - *Grande SP*: para os municípios da região Metropolitana de São Paulo (exclusive MSP),
 - *Interior*: para os demais municípios e
 - *Total*: Todo o estado de São Paulo.

2.1.1.1 Adicionando variáveis

A taxa de homicídio é calculada com o código:

```
estado_sp$tx_homicidio <- (estado_sp$homicidio/estado_sp$populacao)*100000
```

O resultado no objeto tibble:

```
glimpse(estado_sp)
```

```
## Observations: 44
## Variables: 5
```

```
## $ ano          <int> 2000, 2001, 2002, 2003, 2004, 2005, 2006, 2007, 2...
## $ populacao    <int> 36351316, 37630106, 38177742, 38709320, 39825226,...
## $ homicidio     <dbl> 12638, 12475, 11847, 10954, 8753, 7076, 6057, 487...
## $ local         <chr> "Total", "Total", "Total", "Total", "Total", "Tot...
## $ tx_homicidio  <dbl> 34.76628, 33.15165, 31.03117, 28.29809, 21.97853,...
```

2.1.2 Para números totais de homicídios, por trimestre:

O código abaixo carrega dados absolutos das Estatísticas Trimestrais da Secretaria de Segurança Pública.

A variável `trimestre` apresenta valores correspondentes aos trimestres do ano de referência (p.e. `trimestre = "1"` → 1º Trimestre) e a periodicidade total é de 3º Trimestre de 1995 até 1º Trimestre de 2016:

```
# arq .rds
estado_sp_trim <- read_rds(
  "C:\\Users\\raul\\Google Drive\\meu_projeto\\dados e scripts\\tabelas_output\\tab_análise_1\\tab_trim"
)

# subset
estado_sp_trim <- estado_sp_trim[,seq(4)]

# agregando
estado_sp_trim <- estado_sp_trim %>%
  group_by(trimestre, ano) %>%
  summarise(homicidio = sum(homicidio)) %>%
  ungroup() %>%
  mutate(local = rep("Total", 83)) %>%
  bind_rows(estado_sp_trim) %>%
  arrange(local, ano, trimestre)
```

Aplicamos `lubridate::quarter()` e `lubridate::ymd()` para lidar com anos e trimestres:

```
trim <- quarter(
  seq(ymd("1995/7/1"), ymd("2016/1/30"), by = "quarter"),
  with_year = T,
  fiscal_start = 1
)

estado_sp_trim <- estado_sp_trim %>%
  mutate(data = rep(trim, 4)) %>%
  select(local, trimestre, ano, data, homicidio)
```

O resultado em uma tibble

```
estado_sp_trim

## # A tibble: 332 x 5
##   local  trimestre  ano  data homicidio
##   <chr>      <dbl> <int> <dbl>      <dbl>
## 1 Capital    3.00  1995  1995      1134
## 2 Capital    4.00  1995  1995      1142
## 3 Capital    1.00  1996  1996      1331
## 4 Capital    2.00  1996  1996      1109
## 5 Capital    3.00  1996  1996      1150
## 6 Capital    4.00  1996  1996      1092
## 7 Capital    1.00  1997  1997      1140
```

```
## 8 Capital      2.00 1997 1997      1051
## 9 Capital      3.00 1997 1997      1145
## 10 Capital     4.00 1997 1997      1217
## # ... with 322 more rows
```

2.2 Estatísticas descritivas

2.2.1 Agrupando os dados com dplyr

Agrupa-se a taxa de homicídio segundo as localidades:

```
estat_descr <- estado_sp %>%
  group_by(local) %>%
  # summarize() define as variáveis
  summarise(`Média` = mean(tx_homicidio),
            `Desvio padrão` = sd(tx_homicidio),
            Mediana = median(tx_homicidio),
            `IQR` = IQR(tx_homicidio),
            `Mínimo` = min(tx_homicidio),
            `Máximo` = max(tx_homicidio)) %>%
  ungroup() %>%
  rename("Localidade" = local)
```

2.2.2 Figura 1: estatísticas trimestrais

Os gráficos com dados por trimestre são gerados a partir da função `homic_trimestre()`, que aceita como argumentos as variáveis de `estado_sp_trim$local`:

```
homic_trimestre <- function(x) { # x: ("Capital", "Interior", "Grande SP", "total")

  ggplot(filter(estado_sp_trim, local == x), aes(x = data, y = homicidio, fill = local)) +
    geom_line() +
    theme(axis.text.x = element_text(angle = 90)) +
    labs(title = paste(x),
         y = "Homicídio",
         x = "Ano")
}
```

Utilizamos o `ggpubr::ggarrange` para enquadrar as localidades.

```
quadro_homic_trimestre <- ggarrange( # os gráficos
  homic_trimestre("Capital"),
  homic_trimestre("Grande SP"),
  homic_trimestre("Interior"),
  homic_trimestre("Total"),
  # colunas, linhas, alinhamento e legenda
  ncol = 2,
  nrow = 2,
  align = "hv",
  legend = "top",
  common.legend = TRUE)
```

Com `ggpubr::annotate_figure` edita-se o quadro:

```
quadro_homic_trimestre <- annotate_figure( # A figura:
  quadro_homic_trimestre,
  # Configurando o quadro:
  top = text_grob(
    "Ocorrências de homicídios - 3T-1995 até 1T-2016",
    color = "black",
    face = "bold",
    size = 12,
    family = "Times"),
  bottom = NA,
  left = NA,
  right = NA,
  fig.lab = NA,
  fig.lab.face = NA)
```

2.2.3 Figura 2: Distribuição percentual de ocorrências de homicídio e da população residente

A função `perc(y,z)` aceita argumentos do vetor de interesse (`$homicidio` ou `$populacao`) e sua posição (2=homicidios e 3=populacao):

```
perc <- function(y,z) { # y: .$homicidio; .$populacao;
  # z: 3=homicidio; 2=populacao;

  ggplot(estado_sp[-seq(11),], aes(x = ano, fill = local)) +
    geom_area(position = 'fill', aes(y = as.numeric(eval(y)))) +
    theme(legend.position = c(.8, .75),
          axis.text.x = element_text(angle = 90)) +
    labs(x = "Ano",
         fill = "Região",
         y = eval(paste("% de", colnames(estado_sp)[z])))
}
```

Gerando o quadro:

```
quadro_percs <- ggarrange(perc(estado_sp[-seq(11),]$homicidio,3),
  perc(estado_sp[-seq(11),]$populacao,2),
  ncol = 2,
  nrow = 1,
  align = "hv",
  legend = "top",
  common.legend = TRUE)
```

Editando o quadro:

```
quadro_percs <- annotate_figure(quadro_percs,
  top = text_grob(
    "% de ocorrências de homicídios e população residente por região (2000 até 2010)",
    color = "black",
    face = "bold",
    size = 12,
    family = "Times"),
  bottom = NA,
  left = NA,
  right = NA,
```

```
fig.lab = NA,
fig.lab.face = NA)
```

2.2.4 Tabela 1: estatísticas descritivas

Cria a tabela `huxtable::hux()`:

```
ht <- hux(estat_descr, add_colnames = TRUE)
```

Para Editar a tabela, basta alterar os parâmetros no código abaixo:

```
ht <- ht                                %>%
  set_bold(1, everywhere, TRUE)         %>% # negrito
  set_number_format(3)                  %>% # casas decimais
  set_top_border(1, everywhere, 1)      %>% # borda superior
  set_bottom_border(c(1,5), everywhere, 1) %>% # borda inferior
  set_align(everywhere, everywhere, 'center') %>% # alinhamento de texto na célula
  set_right_padding(3)                  %>% # para posicionar
  set_left_padding(3)                   %>% # para posicionar
  set_width(.9)                         %>% # para posicionar no pdf
  set_position('center')                %>% # para posicionar no pdf
  set_caption(
'Estatísticas descritivas - homicídios por 100.000 habitantes no Estado de São Paulo - Capital, RMSP e
  )
```

2.2.5 Figura 3: Taxa de homicídios anuais - de 2000 até 2010

O Código abaixo faz os gráficos de taxas de homicídio anuais, novamente temos uma função(`homic_tx(x)`). Ela aceita como argumento as localidades "Capital", "Interior", "Grande SP" e "Total":

```
homic_tx <- function(x){
  # x: nome da Localidade
  ggplot(filter(estado_sp, local == x)) +
    geom_line(aes(x = ano, y = tx_homicidio)) +
    theme(axis.text.x = element_text(angle = 90)) +
    labs(title = paste(x,
      y      = "Homicídio",
      x      = "Ano")
  )
}
```

Para gerar o quadro, analogamente às figuras anteriores:

```
quadro_homic_tx <- ggarrange(homic_tx("Capital"),
  homic_tx("Grande SP"),
  homic_tx("Interior"),
  homic_tx("Total"),
  ncol = 2,
  nrow = 2,
  align = "hv",
  legend = "top",
  common.legend = TRUE)
```

Editando o quadro:

```
quadro_homic_tx <- annotate_figure(quadro_homic_tx, #####  
  top = text_grob(  
    "Taxa de homicídios anuais - de 2000 até 2010",  
    color = "black",  
    face = "bold",  
    size = 12,  
    family = "Times"),  
  bottom = NA,  
  left = NA,  
  right = NA,  
  fig.lab = NA,  
  fig.lab.face = NA)
```


3 códigos: Município de São Paulo

3.1 carregando arquivo:

```
# lê os dados
msp <- read_rds("C:\\Users\\raul\\Google Drive\\meu_projeto\\dados e scripts\\tabelas_output\\tab_FINAL.rds")

# subset
msp <- msp %>%
  select(ano, distrito, dpol = distrito_num, seccional, homicidio, populacao)

# taxa de homicídio
msp$tx_homicidio <- (msp$homicidio/msp$populacao)*100000
```

3.2 Estatísticas descritivas

3.2.1 Figura 1: Histograma da taxa de homicídios por ano de referência:

```
msp$ano <- as.character(msp$ano)

histograma <- ggplot(msp, aes(x = tx_homicidio)) +
  geom_histogram(aes(y = ..density.., fill = ano, color = ano),
    alpha = .3, position = "identity", bins = 80) +
  theme(plot.title = element_text(hjust = .5, size = 10),
    axis.text = element_text(colour = "black", size = 8),
    axis.title = element_text(colour = "black", size = 8),
    axis.line = element_line(size = .5, colour = "black"),
    panel.background = element_rect(fill = "white"),
    legend.position = "top") +
  labs(title = "Taxa de homicídios por 100.000 habitantes (2003 e 2013)",
    x = "Taxa de homicídio (100.000/hab)",
    y = "Densidade") +
  geom_rug(aes(color = ano)) +
  scale_color_manual(values = c("#E69F00", "black")) +
  scale_fill_manual(values = c("#E69F00", "#999999")) +
  geom_vline(data = filter(msp, ano == "2003"),
    aes(xintercept = mean(tx_homicidio), color = ano), linetype = "dashed") +
  geom_vline(data = filter(msp, ano == "2013"),
    aes(xintercept = mean(tx_homicidio), color = ano), linetype = "dashed")
```

3.2.2 Figura 2: Histograma e boxplot da taxa de homicídios por ano de referência:

```
msp$ano <- as.character(msp$ano)

# {boxplot} #####
boxplot <- ggplot(msp, aes(x = ano, y = tx_homicidio)) +
  geom_boxplot() +
  geom_text(check_overlap = T, aes(label = distrito, color = ano), size = 2.5) +
  scale_color_manual(values = c("#E69F00", "#999999")) +
  theme(plot.title = element_text(hjust = .5, size = 8),
```

```

plot.subtitle = element_text(hjust = .5, size = 6),
axis.text     = element_text(colour = "black", size = 8),
axis.title    = element_text(colour = "black", size = 8),
axis.line     = element_line(size = .5, colour = "black"),
panel.background = element_rect(fill = "white"),
legend.position = "none" +
labs(title     = "Taxa de homicídios por 100.000 habitantes",
      subtitle = "(2003 e 2013)",
      x       = "Ano",
      y       = "Taxa de homicídio (100.000/hab)")

```

3.2.3 Figura 3: Taxas de homicídio por seccional (2000 e 2010):

A função abaixo é criada para ordenar as barras dos gráficos da maior para a menor, isso permite uma visualização mais adequada, ressaltando os maiores e menores valores da taxa de homicídios por Seccional:

```

limits <- function(x) { # x: nome da seccional

  msp %>%
  filter(ano == 2003 & seccional == x) %>%
  select(distrito, tx_homicidio) %>%
  arrange(desc(tx_homicidio)) %>%
  select(distrito)
  # Ordena barras em geom_bar()+scale_x_discrete()

} # first function!

```

Uma função para gerar as 8 seccionais:

```

homic_seccional <- function(x, y, z){

  ggplot(data = filter(msp, seccional == x),
        aes(x = distrito, y = tx_homicidio, fill = ano)) +
  scale_x_discrete(limits = as_vector(limits(x))) +
  geom_bar(stat = "identity",
        position = "dodge",
        show.legend = y,
        alpha = .5,
        color = "black") +
  theme(plot.title = element_text(hjust = .5),
        plot.subtitle = element_text(hjust = .5, margin = margin(b = -10)),
        axis.text = element_text(colour = "black"),
        axis.text.x = element_text(size = 10, angle = 90, hjust = 1, vjust = .3),
        axis.text.y = element_text(size = 10),
        axis.ticks = element_line(),
        axis.line = element_line(size = 1, colour = "black"),
        axis.title.x = element_blank(),
        axis.title.y = element_text(size = 10),
        panel.background = element_rect(fill = "white")) +
  labs(color = "Ano",
        subtitle = paste("Seccional: ", x),
        y = z) +
  scale_fill_manual(values = c("#E69F00", "#999999")) +

```

```

scale_color_manual(msp$ano, values = c("#E69F00", "black")) +
scale_y_continuous(limits = c(0,255),
                    breaks = c(0,25,50,75,100,125,150,175,200,225,250))

}

```

E a figura

```

quadro_msp_secc <- ggarrange(homic_seccional("1 CENTRO", T, "Taxa de Homicídio"),
                             homic_seccional("2 SUL", F, ""),
                             homic_seccional("3 OESTE", F, ""),
                             homic_seccional("4 OESTE", F, ""),
                             homic_seccional("5 LESTE", F, "Taxa de Homicídio"),
                             homic_seccional("6 SANTO AMARO", F, ""),
                             homic_seccional("7 ITAQUERA", F, ""),
                             homic_seccional("8 SÃO MATEUS", F, ""),
                             ncol = 4,
                             nrow = 2,
                             align = "hv",
                             common.legend = TRUE,
                             legend= "top")

```

Editando a figura

```

quadro_msp_secc <- annotate_figure(quadro_msp_secc,
                                   top = text_grob(
"Taxa de Homicídios (100000 habitantes) - Seccionais do Município de São Paulo (2003 e 2013)",
                                   color = "black",
                                   face = "bold",
                                   size = 16,
                                   family = "Times"),
                                   bottom = text_grob("Fonte: SSP/SP",
                                                       color = "black",
                                                       face = "italic",
                                                       size = 10),
                                   left = NA,
                                   right = NA,
                                   fig.lab = NA,
                                   fig.lab.face = NA
                                   )

```

3.2.4 Figura 4: Município de São Paulo - Seccionais

O gráfico de área é gerado com:

E a figura

```

quadro_msp_area <- ggarrange(graf_area("2003"),
                             graf_area("2013"),
                             ncol = 1,
                             nrow = 2,
                             align = "hv",
                             common.legend = TRUE,
                             legend= "top")

```

Editando a figura

```
quadro_msp_area <- annotate_figure(quadro_msp_area,  
  top = text_grob(  
    "Número de homicídios - Seccionais do Município de São Paulo (2003 e 2013)",  
    color = "black",  
    face = "bold",  
    size = 22,  
    family = "Times"),  
  bottom = text_grob("Fonte: SSP/SP",  
    color = "black",  
    face = "italic",  
    size = 12),  
  left = NA,  
  right = NA,  
  fig.lab = NA,  
  fig.lab.face = NA  
)
```

4 Estado de São Paulo

4.1 Dados

Os dados referem-se ao número de ocorrências de homicídio registradas entre os anos de 2000 e 2010. Como a interpretação de ocorrências criminais é sensível à mudanças demográficas, os dados foram normalizados em relação à população residente, sendo calculado, portanto, uma taxa de homicídios por 100.000 habitantes:

$$txhomicidio_{ij} = \left(\frac{homicidio_{ij}}{populacao_{ij}} \right) 100000$$

Na equação acima, a taxa de homicídio no período i é calculada para a localidade j por 100.000 habitantes.

- Os dados de ocorrências criminais são provenientes das Estatísticas Trimestrais¹ da Secretaria Estadual de Segurança Pública do Estado de São Paulo. para esta análise os dados trimestrais foram agrupados em anos.
- Já os dados da população residente foram extraídos das estimativas utilizadas pelo Tribunal de Contas da União para determinação das cotas do Fundo de Participação dos Municípios².

4.2 Estatísticas descritivas

Principais pontos:

- As estatísticas trimestrais mostram queda significativa dos registros de homicídio.
- As localidades no interior do estado apresentam um aumento na proporção de homicídios registrados.
- A distribuição da população nas localidades permaneceram relativamente estáveis.
- A queda da taxa de homicídios é persistente em todas as localidades. há uma pequena resistência no decréscimo da taxa de homicídio no interior, com aumento a partir do ano de 2008.

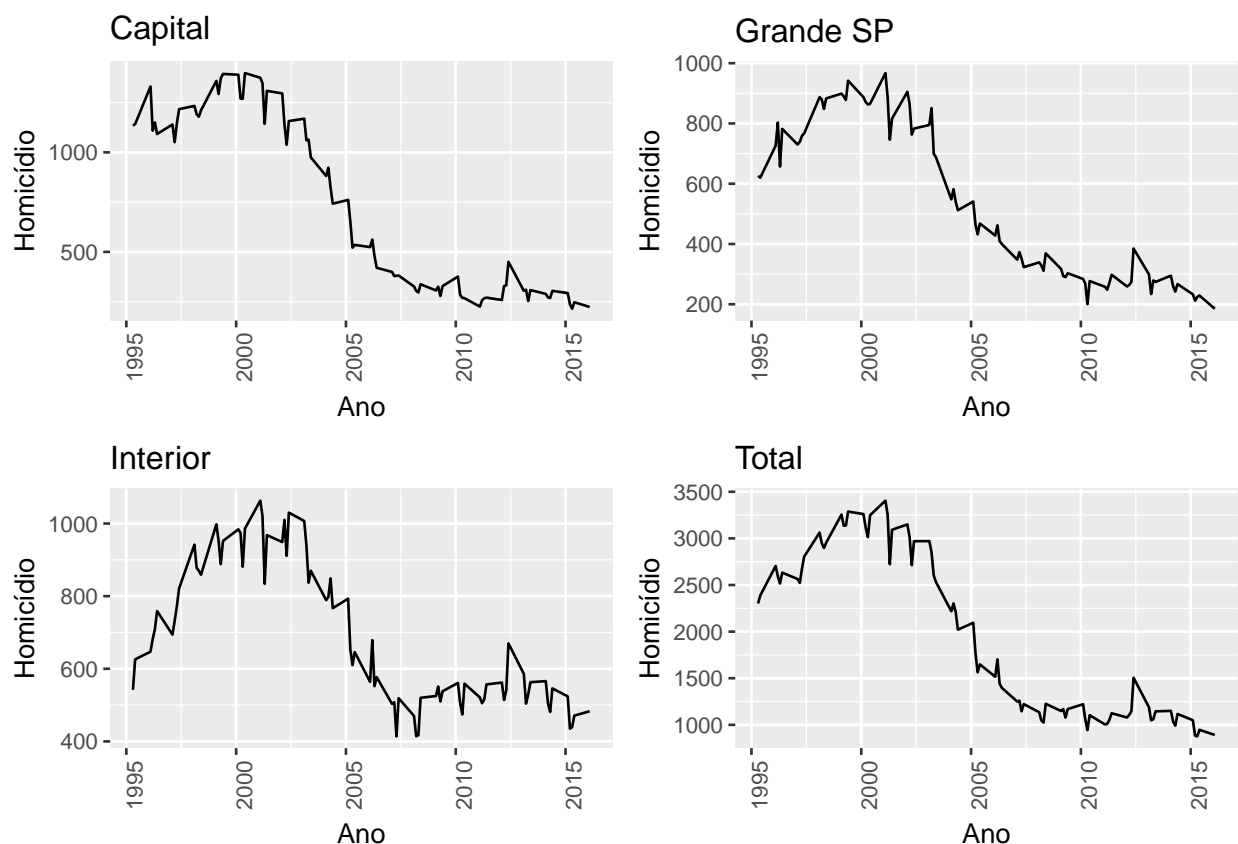
4.2.1 Estatísticas Trimestrais da Secretaria de Segurança Pública.

O quadro abaixo apresenta a evolução da taxa de homicídio nas localidades:

¹<http://www.ssp.sp.gov.br/estatistica/trimestrais.aspx>

²<http://tabnet.datasus.gov.br/cgi/deftohtm.exe?ibge/cnv/poptsp.def>

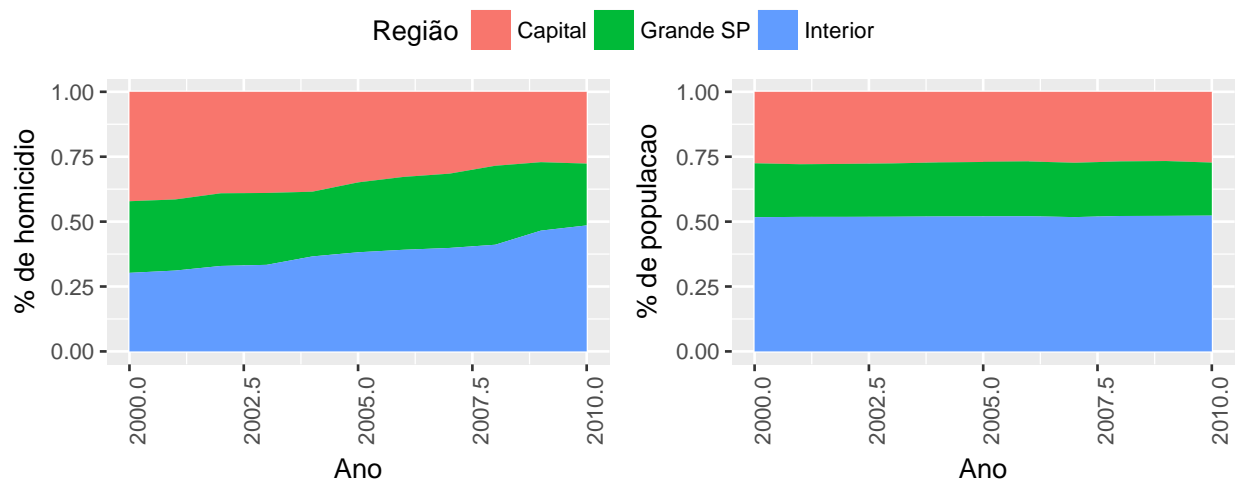
Ocorrências de homicídios – 3T-1995 até 1T-2016



4.2.2 Taxas de homicídio, por localidade, no período 2000 a 2010:

A evolução da população e das ocorrências de homicídios, em termos proporcionais, por localidade no período 2000-2010.

% de ocorrências de homicídios e população residente por região (2000 até 2010)



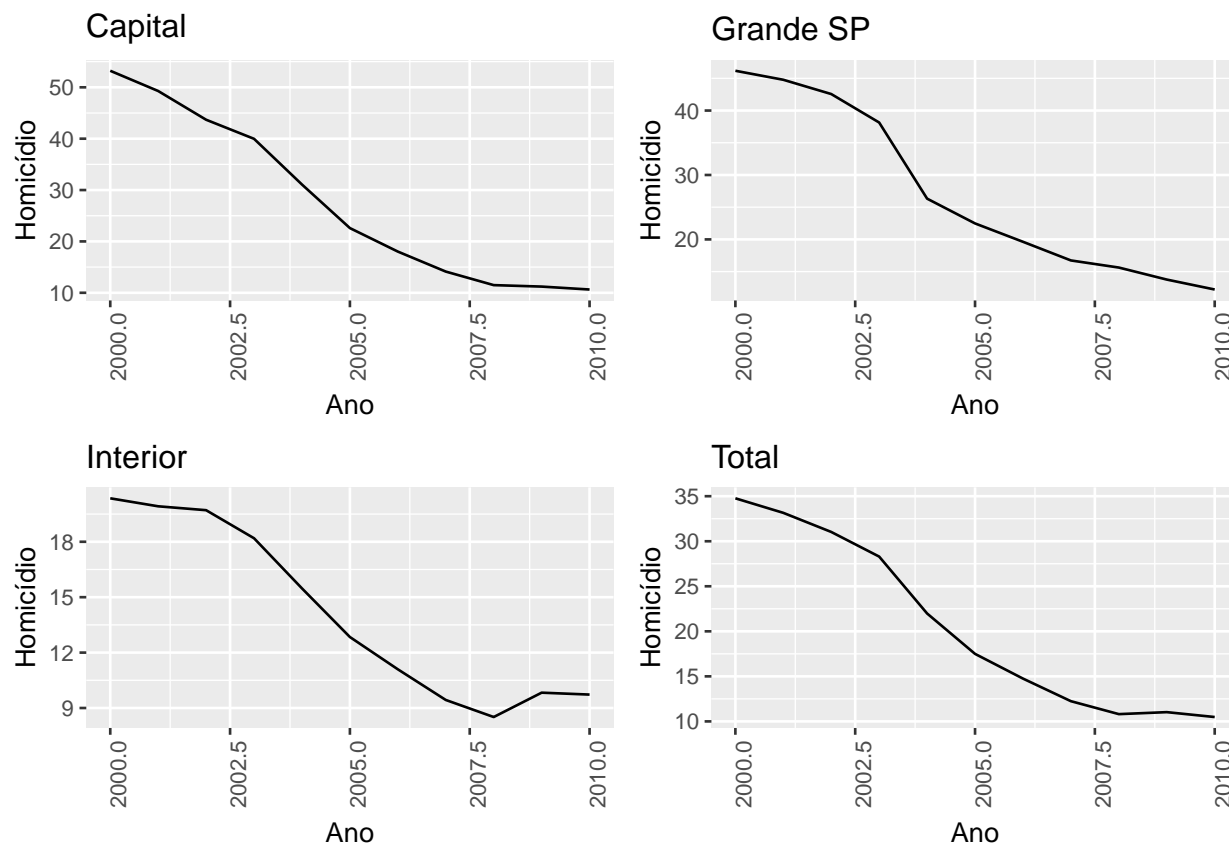
Foram calculadas as taxas de homicídio por localidade. Os resultados estão na tabela 1:

O quadro abaixo mostra a evolução da taxa de homicídios no período 2000-2010 por localidade:

Table 1: Estatísticas descritivas - homicídios por 100.000 habitantes no Estado de São Paulo - Capital, RMSP e Interior - no período de 2000 a 2010

Localidade	Média	Desvio padrão	Mediana	IQR	Mínimo	Máximo
Capital	27.756	16.323	22.593	29.021	10.636	53.221
Grande SP	27.130	13.242	22.477	24.174	12.221	46.173
Interior	14.099	4.734	12.843	9.172	8.511	20.354
Total	20.549	9.652	17.496	18.028	10.484	34.766

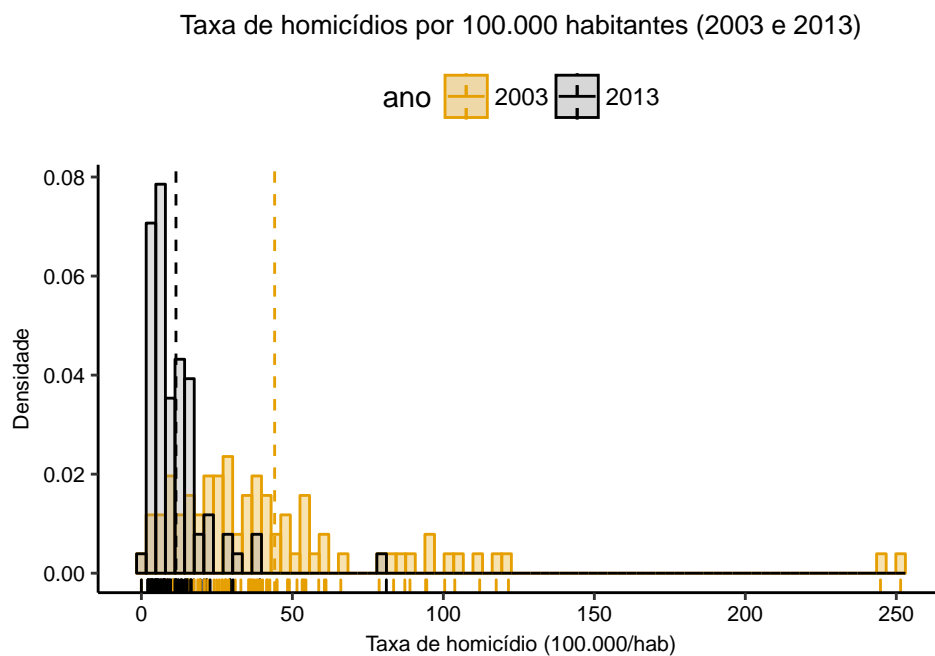
Taxa de homicídios anuais – de 2000 até 2010



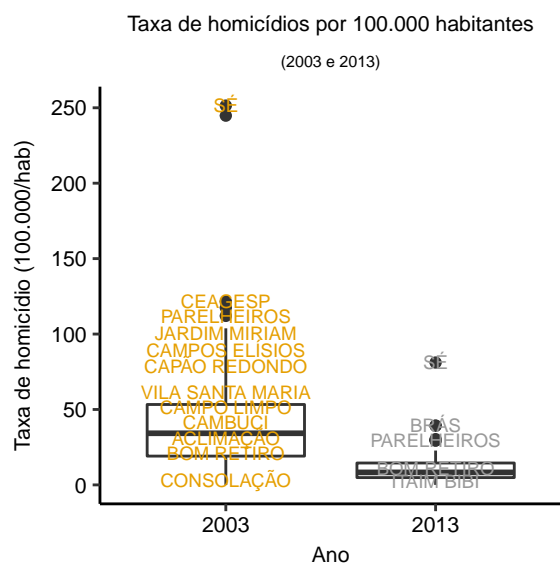
5 Município de São paulo

5.1 Análise Exploratória

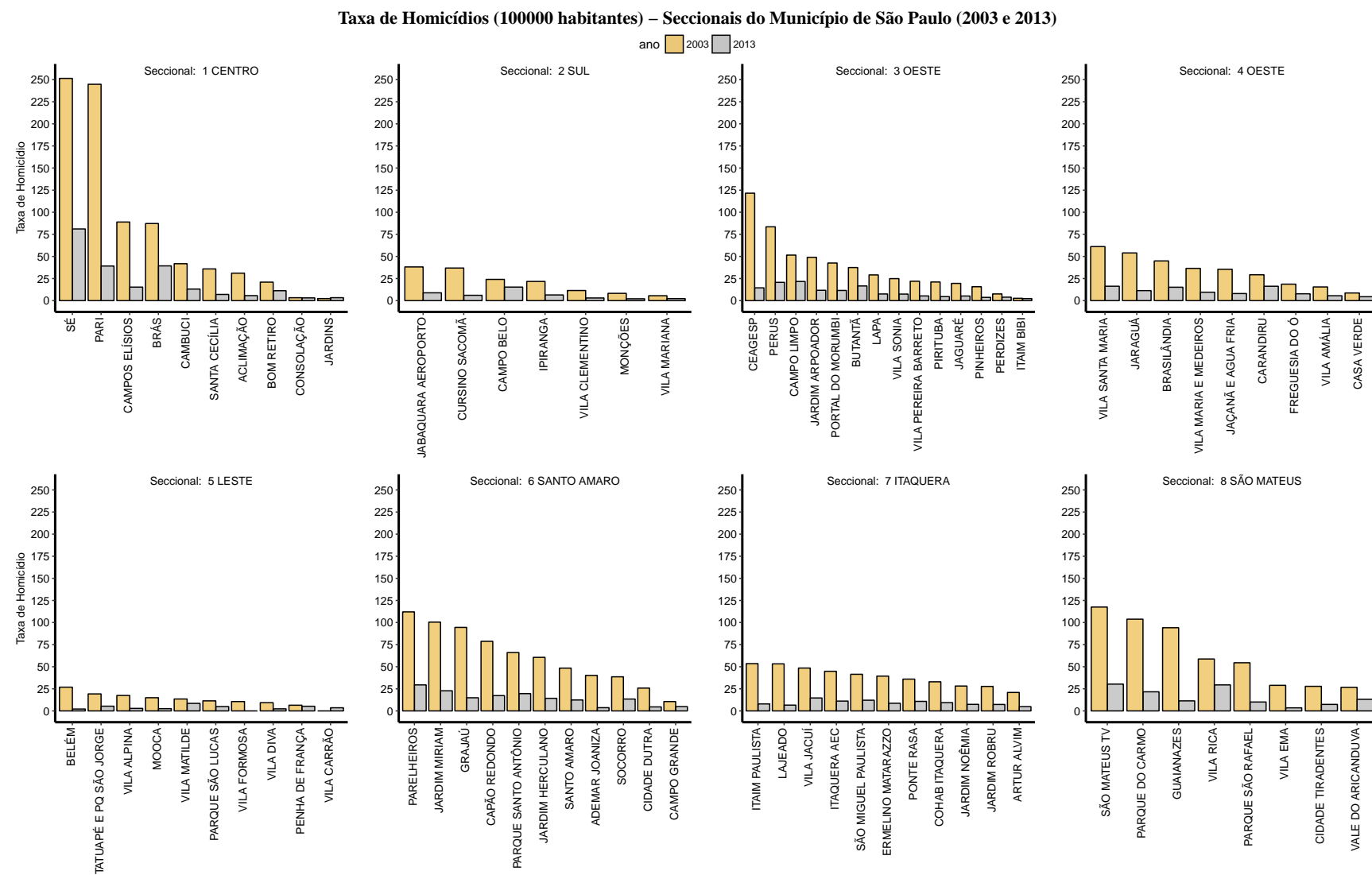
5.1.1 Figura 1: Distribuição da taxa de homicídio no Município de São Paulo (2003 - 2013)



5.1.2 figura 2; Taxas de homicídio no município de São Paulo



5.1.3 Figura 3: Taxas de homicídio por seccional (2003 e 2013):



Fonte: SSP/SP

5.1.4 Figura 4: Homicídios no Município de São Paulo - Seccionais



5.1.5 Figura 5: Diagrama de dispersão de Moran e Estatística I de Moran

Carrega os dados:

```
# os dados
rm(list=ls())

# pacotes (news: spdep, wesanderson - warning=FALSE, message=FALSE)

library(spdep)
library(tidyverse)
library(ggrepel)
library(ggpubr)
library(wesanderson)
library(ggrepel)

# lê os dados
msp <- read_rds("C:\\Users\\raul\\Google Drive\\meu_projeto\\dados e scripts\\tabelas_output\\tab_FINAL.rds")

# subset
msp <- msp %>%
  select(ano, distrito, dpol = distrito_num, seccional, homicidio, populacao)

# taxa de homicídio
msp$tx_homicidio <- (msp$homicidio / msp$populacao) * 100000

# variáveis com características espaciais
msp <- msp %>%
  mutate(tx_homicidio_z      = rep(NA, nrow(msp)),
         lag_tx_homicidio    = rep(NA, nrow(msp)),
         lag_tx_homicidio_z  = rep(NA, nrow(msp)),
         i_moran              = rep(NA, nrow(msp)))

glimpse(msp)

## Observations: 160
## Variables: 11
## $ ano          <int> 2003, 2013, 2013, 2003, 2013, 2003, 2013, 2...
## $ distrito     <chr> "SÉ", "SÉ", "BOM RETIRO", "BOM RETIRO", "CA...
## $ dpol         <int> 1, 1, 2, 2, 3, 3, 4, 4, 5, 5, 6, 6, 7, 7, 8...
## $ seccional    <chr> "1 CENTRO", "1 CENTRO", "1 CENTRO", "1 CENT...
## $ homicidio    <dbl> 53, 20, 4, 6, 9, 45, 4, 4, 4, 20, 5, 13, 5,...
## $ populacao    <dbl> 21079, 24654, 35567, 28656, 58784, 50595, 1...
## $ tx_homicidio <dbl> 251.435078, 81.122739, 11.246380, 20.938023...
## $ tx_homicidio_z <lgl> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA,...
## $ lag_tx_homicidio <lgl> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA,...
## $ lag_tx_homicidio_z <lgl> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA,...
## $ i_moran      <lgl> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA,...
```

Carrega a matriz:

```
# a matriz queen

# diretório
queen <- read.table("C:\\Users\\raul\\Google Drive\\meu_projeto\\dados e scripts\\tabelas_output\\planilha_queen.csv", as.is=TRUE)
dim(queen) # 80x80
```

```
## [1] 80 80

# Vamos criar uma matriz de contiguidade através de uma matriz esparsa
queen <- as.matrix(queen) # matriz feita à mão
is.matrix(queen)

## [1] TRUE

# Nomeando colunas e linhas com o num. do distrito_pol correspondente
distrito_pol <- c("1","2","3","4","5","6","7","8","9","10","11","12","13","14","15",
  "16","17","21","22","23","24","25","27","28","29","31","33","34",
  "36","37","38","40","41","42","44","46","47","48","50","51","53",
  "54","55","56","58","59","62","63","65","66","67","68","70","74",
  "75","77","78","81","87","89","91","92","93","96","98","99","100",
  "102","103","1857","2073","3052","3264","3597","4380","4572","4969",
  "85101","193990","268395")

colnames(queen) <- distrito_pol
rownames(queen) <- distrito_pol

isSymmetric(queen) # pergunta se queen é simetrica

## [1] TRUE

# objeto 'listw', style="W" (padronizada na linha)
w <- mat2listw(queen, row.names = NULL, style="M")
listw <- nb2listw(w$neighbours, glist=NULL, style="W", zero.policy=NULL)
summary(listw)

## Characteristics of weights list object:
## Neighbour list object:
## Number of regions: 80
## Number of nonzero links: 414
## Percentage nonzero weights: 6.46875
## Average number of links: 5.175
## Link number distribution:
##
## 1 2 3 4 5 6 7 8 9 10
## 1 4 8 11 23 18 10 2 2 1
## 1 least connected region:
## 25 with 1 link
## 1 most connected region:
## 3264 with 10 links
##
## Weights style: W
## Weights constants summary:
##      n  nn S0      S1      S2
## W 80 6400 80 33.77879 329.0915

insere variáveis defasadas

defasando <- function(x) { # "2003" e "2013"
  msp                                     %>%
  filter(ano == x)                       %>%
  arrange(dpol)                           %>%
  mutate(tx_homicidio_z = (tx_homicidio - mean(tx_homicidio)) / sd(tx_homicidio),
```

```

lag_tx_homicidio = lag.listw(listw, tx_homicidio),
lag_tx_homicidio_z = lag.listw(listw, tx_homicidio_z),
i_moran = rep( moran(tx_homicidio,
                      listw=listw, 80, Szero(listw))[[1]],80)) %>%
bind_cols(., as_data_frame(localmoran(msp%>%
                      filter(ano==x)%>%
                      select(tx_homicidio)%>%
                      as_vector(),
                      listw = listw)) ) %>%
                      %>%
select(-E.Ii, -Var.Ii, -Z.Ii)
rename(local_moran = Ii,
       local_moran_pvalor = `Pr(z > 0)`))
}
msp <- bind_rows(defasando("2003"), defasando("2013"))

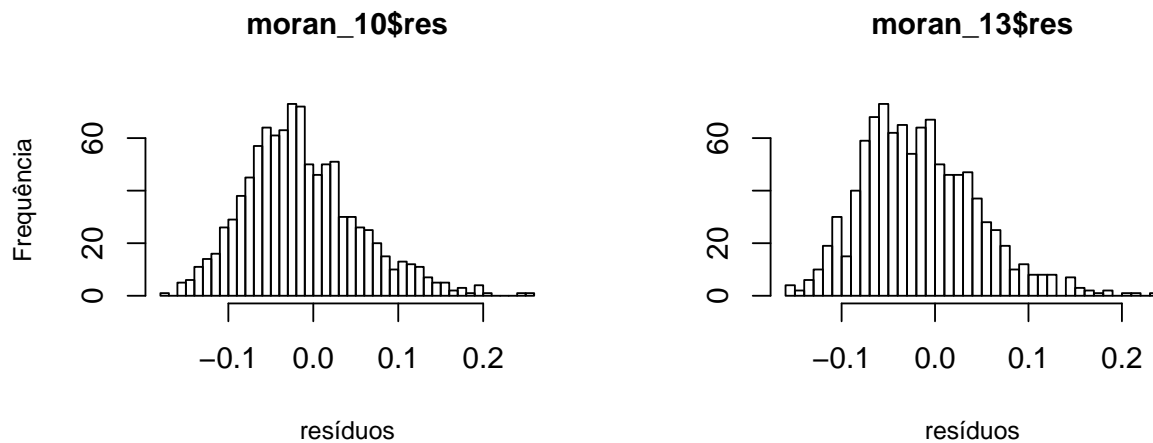
```

Veja os resíduos de Moran:

```

# simulação MC, veja a distr. dos resíduos
par(mfrow = c(1, 2))
hist(moran.mc(msp$tx_homicidio[msp$ano == "2003"], listw=listw, nsim=999)[[7]],
     breaks = 50,
     main = list("moran_10$res", cex = 1),
     xlab = list("resíduos", cex = .8),
     ylab = list("Frequência", cex=.8))
hist(moran.mc(msp$tx_homicidio[msp$ano == "2013"], listw=listw, nsim=999)[[7]],
     breaks = 50,
     main = list("moran_13$res",cex = 1),
     xlab = list("resíduos", cex = .8),
     ylab = list("", cex = .8))

```



```

# Moran.plot modo convencional .....
# função spdep::moran.plot

par(mfrow = c(1, 2))

moran.plot(as.vector(msp$tx_homicidio_z[msp$ano == "2003"]),

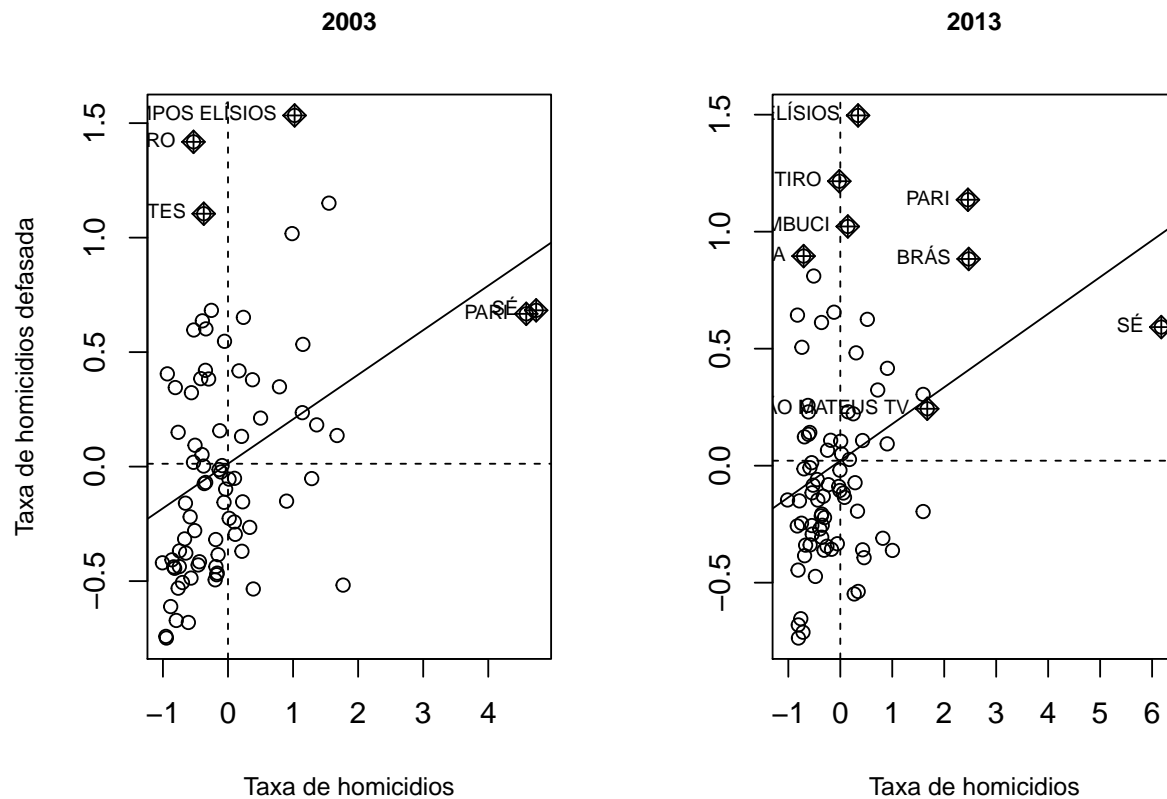
```

```

listw,
zero.policy = T,
spChk = NULL,
xlab = list("Taxa de homicídios", cex = .8),
ylab = list("Taxa de homicídios defasada", cex = .8),
labels = as.character(msp$distrito[msp$ano == "2003"]),
quiet = NULL)
title(main = list("2003", cex = .8))

moran.plot(as.vector(msp$tx_homicidio_z[msp$ano == "2013"]),
listw,
zero.policy = T,
spChk = NULL,
xlab = list("Taxa de homicídios", cex = .8),
ylab = list(""),
labels = paste(msp$distrito[msp$ano == "2013"]),
quiet = NULL)
title(main=list("2013", cex = .8))

```



```

# identify the Local Moran plot quadrant for each observation this is some
# serious slicing and illustrate the power of the bracket
msp$quad_sig <- NA
msp[(msp$tx_homicidio_z >= 0 & msp$lag_tx_homicidio_z >= 0) & (msp$local_moran_pvalor <= 0.05), "quad_s"]
msp[(msp$tx_homicidio_z <= 0 & msp$lag_tx_homicidio_z <= 0) & (msp$local_moran_pvalor <= 0.05), "quad_s"]

```

```

msp[(msp$tx_homicidio_z >= 0 & msp$lag_tx_homicidio_z <= 0) & (msp$local_moran_pvalor <= 0.05), "quad_s"]
msp[(msp$tx_homicidio_z <= 0 & msp$lag_tx_homicidio_z >= 0) & (msp$local_moran_pvalor <= 0.05), "quad_s"]
msp[(msp$local_moran_pvalor > 0.05), "quad_sig"] <- "Não sig."
glimpse(msp)

## Observations: 160
## Variables: 14
## $ ano                <int> 2003, 2003, 2003, 2003, 2003, 2003, 2003, 2...
## $ distrito           <chr> "SÉ", "BOM RETIRO", "CAMPOS ELÍSIOS", "CONS...
## $ dpol               <int> 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, ...
## $ seccional          <chr> "1 CENTRO", "1 CENTRO", "1 CENTRO", "1 CENT...
## $ homicidio          <dbl> 53, 6, 45, 4, 20, 13, 18, 23, 15, 17, 31, 3...
## $ populacao          <dbl> 21079, 28656, 50595, 120821, 64310, 31114, ...
## $ tx_homicidio       <dbl> 251.435078, 20.938023, 88.941595, 3.310683,...
## $ tx_homicidio_z     <dbl> 4.73573143, -0.52990521, 1.02361616, -0.932...
## $ lag_tx_homicidio   <dbl> 74.01164, 106.21720, 111.26871, 61.86524, 6...
## $ lag_tx_homicidio_z <dbl> 0.68254580, 1.41827228, 1.53367253, 0.40506...
## $ i_moran            <dbl> 0.194438, 0.194438, 0.194438, 0.194438, 0.1...
## $ local_moran        <dbl> 3.27326949, -0.76106316, 1.58976404, -0.382...
## $ local_moran_pvalor <dbl> 1.080103e-22, 9.793767e-01, 3.635768e-05, 8...
## $ quad_sig           <chr> "Alto-alto", "Não sig.", "Alto-alto", "Não ...

library(huxtable)
huxreg(
  lm(msp$lag_tx_homicidio_z[msp$ano=="2003"]~msp$tx_homicidio_z[msp$ano=="2003"]),
  lm(msp$lag_tx_homicidio_z[msp$ano=="2013"]~msp$tx_homicidio_z[msp$ano=="2013"])
) %>%
  set_bold(1, everywhere, TRUE)           %>% # negrito
  set_top_border(1, everywhere, 1)        %>% # borda superior
  set_right_padding(3)                    %>% # para posicionar
  set_left_padding(3)                     %>% # para posicionar
  set_width(1)                            %>% # para posicionar no pdf
  set_position('center')                  %>% # para posicionar no pdf
  set_caption(
    'Indicador I de Moran'
  )

```

Table 2: Indicador I de Moran

	(1)	(2)
(Intercept)	0.012 (0.052)	0.021 (0.049)
msp\$tx_homicidio_z[msp\$ano == "2e+03"]	0.194 *** (0.053)	
msp\$tx_homicidio_z[msp\$ano == "2.01e+03"]		0.157 ** (0.049)
N	80	80
R 2	0.148	0.117
logLik	-51.895	-45.850
AIC	109.789	97.700

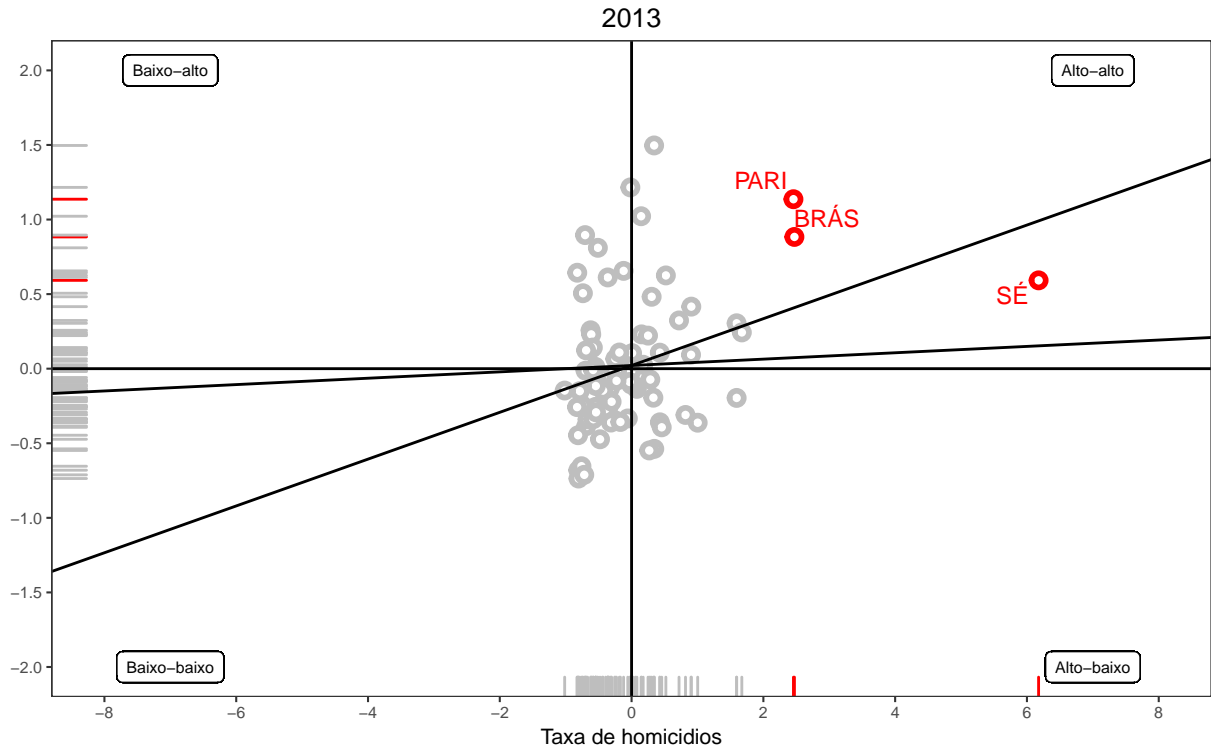
*** p < 0.001; ** p < 0.01; * p < 0.05.

```

# Moran ggplot2 ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
#2003 #filter(msp, ano==2003)
moran_ggplot <- function(x){
ggplot(filter(msp, ano==x), aes(x=tx_homicidio_z, y=lag_tx_homicidio_z,color=as.factor(quad_sig))) +
  geom_point(shape= 21,
             fill = "white",
             size = 1.5,
             stroke = 1.5) +
  geom_rug(aes(color=as.factor(quad_sig))) +
  theme_bw(base_size = 8) +
  theme(plot.title = element_text(hjust = .5),
        legend.position = "bottom",
        legend.text = element_text(size = 8),
        panel.background = element_blank(),
        panel.grid.major = element_blank(),
        panel.grid.minor = element_blank()) +
  labs(title = paste(x),
        x = "Taxa de homicidios",
        y="",
        color = "I de Moran Local (p-valor<0,05)") +
  scale_y_continuous(limits = c(-2,2), breaks=seq(-2,2, by=.5)) +
  scale_x_continuous(limits = c(-8,8), breaks=seq(-8,8, by=2)) +
  scale_color_manual(values=c('Alto-alto'='red', 'Baixo-baixo'='blue', 'Não sig.'='gray')) +
  geom_vline(xintercept = 0) +
  geom_hline(yintercept = 0) +
  geom_abline(slope=coef(lm(msp$lag_tx_homicidio_z[msp$ano==x]~msp$tx_homicidio_z[msp$ano==x])),
             intercept=coef(lm(msp$lag_tx_homicidio_z[msp$ano==x]~msp$tx_homicidio_z[msp$ano==x]))[[1]]),
  geom_text_repel(data=subset(msp, ano == x & quad_sig == "Alto-alto" | ano == x & quad_sig == "Baixo-baixo"),
                 aes(label = distrito),
                 size = 3)+
  geom_label( label = "Alto-alto", x = 7, y = 2, size = 2, colour = "black") +
  geom_label( label = "Alto-baixo", x = 7, y = -2, size = 2, colour = "black") +
  geom_label( label = "Baixo-baixo", x = -7, y = -2, size = 2, colour = "black") +
  geom_label( label = "Baixo-alto", x = -7, y = 2, size = 2, colour = "black")
}

moran_ggplot("2013")

```

I de Moran Local (p -valor<0,05) —●— Alto-alto —●— Não sig.

```
# enquadrando gráficos
fig <- ggarrange(moran_ggplot("2003") + ylab("Lag - taxa de homicídios"), moran_ggplot("2013"),
  ncol=2, nrow=1, # align="hv",
  common.legend = TRUE, legend = "top")

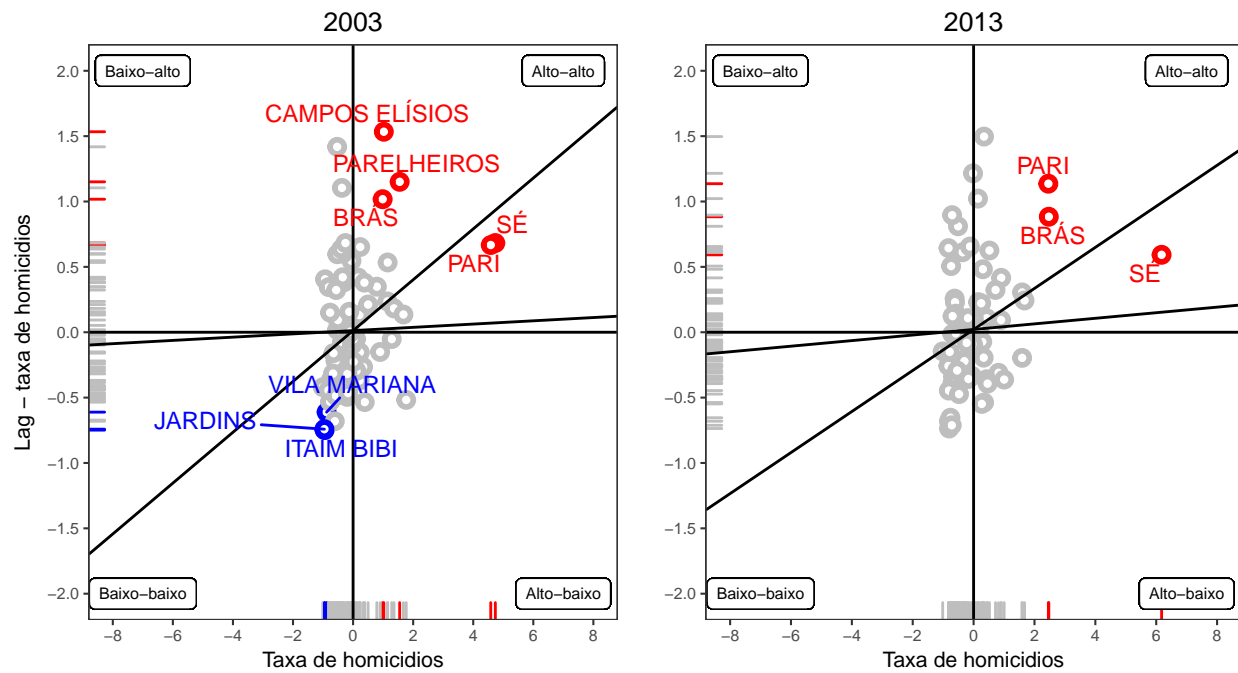
fig_completa <- annotate_figure(fig, #####
  top = text_grob("Figura: Diagramas de dispersão de Moran (2003/2013) \n Índice global",
    color = "black",
    face = "bold",
    size = 10),
  bottom = text_grob("Fonte: Elaboração própria a partir de dados da SSP/SP",
    color = "black",
    face = "italic",
    size = 8),
  left = NA,
  right = NA,
  fig.lab = NA, fig.lab.face = NA
)

fig_completa
```

```
## Warning: Removed 1 rows containing missing values (geom_text).
```

Figura: Diagramas de dispersão de Moran (2003/2013)
Índice global e local

I de Moran Local (p-valor<0,05) — Alto-alto — Baixo-baixo — Não sig.



Fonte: Elaboração própria a partir de dados da SSP/SP