

GEOPROCESSAMENTO - BICT Mar

Aula 2: Manipulação de atributos não espaciais em tabelas

Fabio Cop

23-08-2023

Pacotes utilizados

```
library(tidyverse)
```

Lendo a base de dados.

Fonte: <https://www.ibge.gov.br/estatisticas/sociais/populacao/22827-censo-demografico-2022.html>

```
munibge <- read_csv('dados/municipios_ibge2022.csv', col_types = cols(uf_cod = 'c'))
munibge
```

```
## # A tibble: 5,570 x 9
##   uf    uf_cod municipio_cod municipio      capital area_km2 populacao_2010
##   <chr> <chr>   <chr>         <chr>      <chr>      <dbl>         <dbl>
## 1 RO    11     00015      Alta Floresta D'O~ não         7067         24392
## 2 RO    11     00023      Ariquemes        não         4427         90353
## 3 RO    11     00031      Cabixi            não         1314          6313
## 4 RO    11     00049      Cacoal            não         3793        78574
## 5 RO    11     00056      Cerejeiras        não         2783        17029
## 6 RO    11     00064      Colorado do Oeste não         1451        18591
## 7 RO    11     00072      Corumbiara        não         3060          8783
## 8 RO    11     00080      Costa Marques     não         4987        13678
## 9 RO    11     00098      Espigão D'Oeste   não         4518        28729
## 10 RO   11     00106      Guajará-Mirim     não        24857        41656
## # i 5,560 more rows
## # i 2 more variables: populacao_2022 <dbl>, domicilios_2022 <dbl>
```

Descrevendo o data frame.

```
dim(munibge) # Dimensões da tabela
```

```
## [1] 5570    9
```

```
glimpse(munibge) # Descrição dos atributos (colunas)
```

```
## Rows: 5,570
## Columns: 9
## $ uf          <chr> "RO", "RO", "RO", "RO", "RO", "RO", "RO", "RO", "RO", ~
## $ uf_cod      <chr> "11", "11", "11", "11", "11", "11", "11", "11", "11", ~
## $ municipio_cod <chr> "00015", "00023", "00031", "00049", "00056", "00064", ~
## $ municipio    <chr> "Alta Floresta D'Oeste", "Ariquemes", "Cabixi", "Cocoa~
## $ capital      <chr> "não", "não", "não", "não", "não", "não", "não", "não"~
## $ area_km2     <dbl> 7067, 4427, 1314, 3793, 2783, 1451, 3060, 4987, 4518, ~
## $ populacao_2010 <dbl> 24392, 90353, 6313, 78574, 17029, 18591, 8783, 13678, ~
## $ populacao_2022 <dbl> 21495, 96833, 5363, 86895, 15890, 15663, 7519, 12627, ~
## $ domicilios_2022 <dbl> 7695, 34768, 1967, 31919, 5873, 5991, 2840, 4161, 1046~
```

Contando número de estados.

```
munibge %>%
  distinct(uf) %>%
  count()
```

```
## # A tibble: 1 x 1
##       n
##   <int>
## 1     27
```

Contando número de municípios pelo nome.

```
munibge %>%
  distinct(municipio) %>%
  count()
```

```
## # A tibble: 1 x 1
##       n
##   <int>
## 1  5297
```

Contando número de municípios pelo código.

```
munibge %>%
  select(municipio_cod) %>%
  count()
```

```
## # A tibble: 1 x 1
##       n
##   <int>
## 1  5570
```

Verificando nomes repetidos

```
munibge %>%  
  count(municipio) %>%  
  arrange(desc(n))
```

```
## # A tibble: 5,297 x 2  
##   municipio      n  
##   <chr>      <int>  
## 1 Bom Jesus      5  
## 2 São Domingos    5  
## 3 Bonito          4  
## 4 Planalto        4  
## 5 Santa Helena    4  
## 6 Santa Inês       4  
## 7 Santa Luzia      4  
## 8 Santa Terezinha  4  
## 9 São Francisco    4  
## 10 Vera Cruz       4  
## # i 5,287 more rows
```

Filtrando linhas

```
acre <- munibge %>%  
  filter(uf == 'AC')
```

Ordenando linhas

```
acre %>% arrange(populacao_2022)
```

```
## # A tibble: 22 x 9  
##   uf    uf_cod municipio_cod municipio      capital area_km2 populacao_2010  
##   <chr> <chr>   <chr>      <chr>      <chr>      <dbl>      <dbl>  
## 1 AC    12     00435      Santa Rosa do Pur~ não        6156        4691  
## 2 AC    12     00054      Assis Brasil      não        4979        6072  
## 3 AC    12     00328      Jordão            não        5357        6577  
## 4 AC    12     00179      Capixaba          não        1706        8798  
## 5 AC    12     00393      Porto Walter      não        6446        9176  
## 6 AC    12     00344      Manoel Urbano     não       10631       7981  
## 7 AC    12     00138      Bujari            não        3035        8471  
## 8 AC    12     00013      Acrelândia        não        1812       12538  
## 9 AC    12     00427      Rodrigues Alves   não        3076       14389  
## 10 AC   12     00385      Plácido de Castro não        1953       17209  
## # i 12 more rows  
## # i 2 more variables: populacao_2022 <dbl>, domicilios_2022 <dbl>
```

```
acre %>% arrange(desc(populacao_2022))
```

```
## # A tibble: 22 x 9
##   uf    uf_cod municipio_cod municipio      capital area_km2 populacao_2010
##   <chr> <chr>   <chr>         <chr>      <chr>    <dbl>      <dbl>
## 1 AC     12     00401      Rio Branco    sim        8835      336038
## 2 AC     12     00203      Cruzeiro do Sul não         8783      78507
## 3 AC     12     00609      Tarauacá      não        20169      35590
## 4 AC     12     00500      Sena Madureira não        23760      38029
## 5 AC     12     00302      Feijó         não        27977      32412
## 6 AC     12     00104      Brasiléia     não         3928      21398
## 7 AC     12     00450      Senador Guimard não         2320      20179
## 8 AC     12     00336      Mâncio Lima   não         5452      15206
## 9 AC     12     00252      Epitaciolândia não         1653      15100
## 10 AC    12     00708      Xapuri        não         5351      16091
## # i 12 more rows
## # i 2 more variables: populacao_2022 <dbl>, domicilios_2022 <dbl>
```

Sumarizando um data frame

```
acre %>%
  summarise(populacao_total = sum(populacao_2022))
```

```
## # A tibble: 1 x 1
##   populacao_total
##   <dbl>
## 1      830026
```

Mais de uma estatística descritiva para a coluna

```
acre %>%
  summarise(populacao_total = sum(populacao_2022),
            populacao_media = mean(populacao_2022))
```

```
## # A tibble: 1 x 2
##   populacao_total populacao_media
##   <dbl>          <dbl>
## 1      830026      37728.
```

Mais de uma coluna

```
acre %>%
  summarise(across(c(area_km2, populacao_2010, populacao_2022, domicilios_2022), ~sum(.)))
```

```
## # A tibble: 1 x 4
##   area_km2 populacao_2010 populacao_2022 domicilios_2022
##   <dbl>      <dbl>      <dbl>      <dbl>
## 1  164174    733559    830026    260997
```

Todas as colunas numericas

```
acre %>%
  summarise(across(where(is.numeric), ~sum(.)))

## # A tibble: 1 x 4
##   area_km2 populacao_2010 populacao_2022 domicilios_2022
##   <dbl>         <dbl>         <dbl>         <dbl>
## 1    164174         733559         830026         260997
```

Todas as colunas com tamanhos populacionais

```
acre %>%
  summarise(across(starts_with('populacao'), ~sum(.)))

## # A tibble: 1 x 2
##   populacao_2010 populacao_2022
##   <dbl>         <dbl>
## 1         733559         830026
```

Sumarisando um data frame por grupos

```
estados <- munibge %>%
  group_by(uf) %>%
  summarise(across(where(is.numeric), ~sum(.)))
```

Criando colunas

```
capitais <- munibge %>%
  filter(capital == 'sim')
```

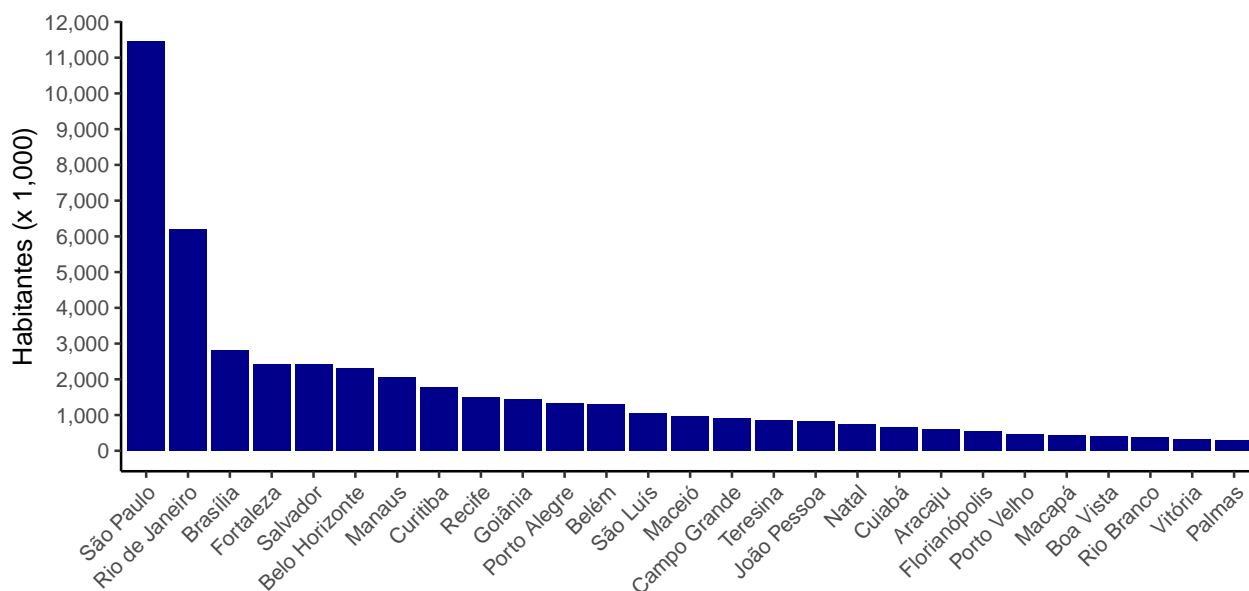
Crescimento entre 2010 e 2022

```
capitais <- capitais %>%
  mutate(crescimento = populacao_2022 - populacao_2010) %>%
  mutate(cresc_percentual = round((populacao_2022 / populacao_2010 - 1) * 100, 2)) %>%
  mutate(densidade_2022 = round(populacao_2022 / area_km2, 2)) %>%
  arrange(desc(densidade_2022))
```

Graficos

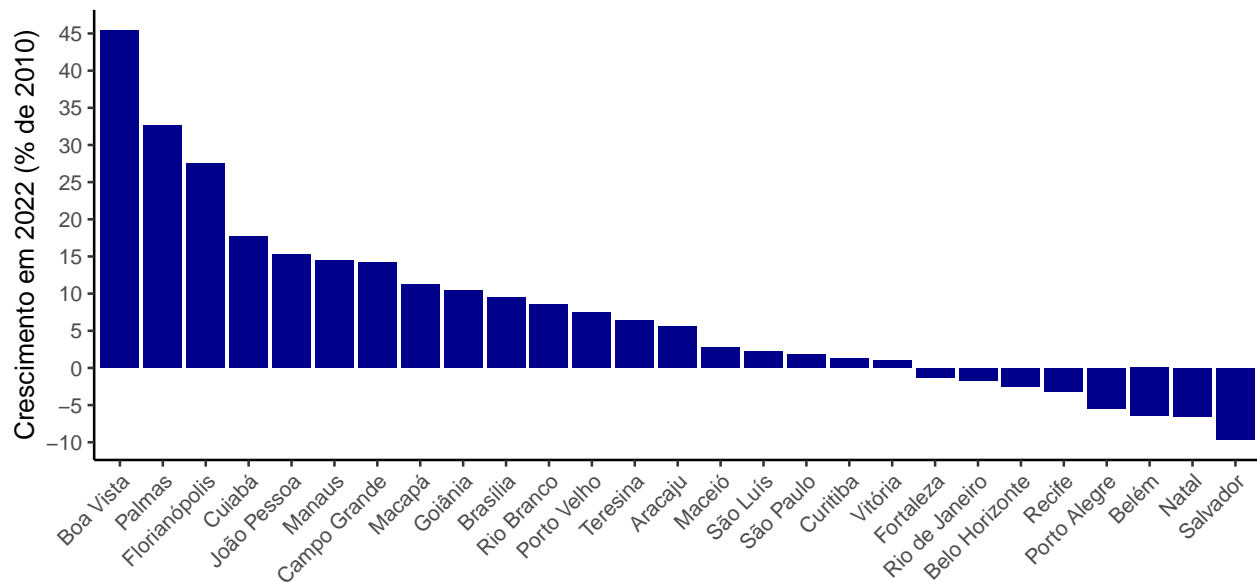
Tamanho populacional em 2022

```
ggplot(capitais, aes(y = populacao_2022/1000, x = reorder(municipio, -populacao_2022))) +
  geom_col(fill = 'darkblue') +
  labs(y = "Habitantes (x 1,000)", x = '') +
  scale_y_continuous(breaks = seq(0, 12000, by = 1000),
    labels = scales::comma) +
  theme_classic(base_size = 15) +
  theme(axis.text.x = element_text(angle = 45, hjust = 1))
```



Crescimento comparado a 2010

```
ggplot(capitais, aes(y = cresc_percentual, x = reorder(municipio, -cresc_percentual))) +
  geom_col(fill = 'darkblue') +
  labs(y = "Crescimento em 2022 (% de 2010)", x = '') +
  scale_y_continuous(breaks = seq(-10, 50, by = 5),
    labels = scales::comma) +
  theme_classic(base_size = 15) +
  theme(axis.text.x = element_text(angle = 45, hjust = 1))
```



Densidade populacional

```
ggplot(capitais, aes(y = densidade_2022, x = reorder(municipio, -densidade_2022))) +
  geom_col(fill = 'darkblue') +
  labs(y = expression("Densidade populacional (Hab/km\"^2*\"))", x = '') +
  scale_y_continuous(breaks = seq(0, 10000, by = 500),
    labels = scales::comma) +
  theme_classic(base_size = 15) +
  theme(axis.text.x = element_text(angle = 45, hjust = 1))
```

