

Introdução ao *tidyverse*

Igo da Costa Andrade

1 O *dataset* gapminder

Visão geral dos dados:

```
gapminder

## # A tibble: 1,704 x 6
##   country      continent  year lifeExp      pop gdpPercap
##   <fct>        <fct>    <int>  <dbl>    <int>    <dbl>
## 1 Afghanistan Asia      1952   28.8  8425333    779.
## 2 Afghanistan Asia      1957   30.3  9240934    821.
## 3 Afghanistan Asia      1962   32.0 10267083    853.
## 4 Afghanistan Asia      1967   34.0 11537966    836.
## 5 Afghanistan Asia      1972   36.1 13079460    740.
## 6 Afghanistan Asia      1977   38.4 14880372    786.
## 7 Afghanistan Asia      1982   39.9 12881816    978.
## 8 Afghanistan Asia      1987   40.8 13867957    852.
## 9 Afghanistan Asia      1992   41.7 16317921    649.
## 10 Afghanistan Asia      1997   41.8 22227415    635.
## # ... with 1,694 more rows
```

Filtragem por ano:

1.1 Filtragem de dados

```
gapminder %>%
  filter(year==2007)

## # A tibble: 142 x 6
##   country      continent  year lifeExp      pop gdpPercap
##   <fct>        <fct>    <int>  <dbl>    <int>    <dbl>
## 1 Afghanistan Asia      2007   43.8  31889923    975.
## 2 Albania      Europe      2007   76.4   3600523   5937.
## 3 Algeria      Africa      2007   72.3  33333216   6223.
## 4 Angola       Africa      2007   42.7  12420476   4797.
## 5 Argentina    Americas    2007   75.3  40301927  12779.
## 6 Australia    Oceania     2007   81.2  20434176  34435.
## 7 Austria      Europe      2007   79.8   8199783  36126.
## 8 Bahrain      Asia      2007   75.6    708573  29796.
## 9 Bangladesh   Asia      2007   64.1 150448339   1391.
## 10 Belgium     Europe      2007   79.4  10392226  33693.
## # ... with 132 more rows
```

Filtragem por país:

```
gapminder %>%
  filter(country == "United States")
```

```
## # A tibble: 12 x 6
```

	country	continent	year	lifeExp	pop	gdpPercap
	<fct>	<fct>	<int>	<dbl>	<int>	<dbl>
## 1	United States	Americas	1952	68.4	157553000	13990.
## 2	United States	Americas	1957	69.5	171984000	14847.
## 3	United States	Americas	1962	70.2	186538000	16173.
## 4	United States	Americas	1967	70.8	198712000	19530.
## 5	United States	Americas	1972	71.3	209896000	21806.
## 6	United States	Americas	1977	73.4	220239000	24073.
## 7	United States	Americas	1982	74.6	232187835	25010.
## 8	United States	Americas	1987	75.0	242803533	29884.
## 9	United States	Americas	1992	76.1	256894189	32004.
## 10	United States	Americas	1997	76.8	272911760	35767.
## 11	United States	Americas	2002	77.3	287675526	39097.
## 12	United States	Americas	2007	78.2	301139947	42952.

Filtragem por ano e por país:

```
gapminder %>%
  filter(year == 2007, country == 'United States')
```

```
## # A tibble: 1 x 6
```

	country	continent	year	lifeExp	pop	gdpPercap
	<fct>	<fct>	<int>	<dbl>	<int>	<dbl>
## 1	United States	Americas	2007	78.2	301139947	42952.

1.2 Ordenação

```
gapminder %>%
  arrange(gdpPercap)
```

```
## # A tibble: 1,704 x 6
```

	country	continent	year	lifeExp	pop	gdpPercap
	<fct>	<fct>	<int>	<dbl>	<int>	<dbl>
## 1	Congo, Dem. Rep.	Africa	2002	45.0	55379852	241.
## 2	Congo, Dem. Rep.	Africa	2007	46.5	64606759	278.
## 3	Lesotho	Africa	1952	42.1	748747	299.
## 4	Guinea-Bissau	Africa	1952	32.5	580653	300.
## 5	Congo, Dem. Rep.	Africa	1997	42.6	47798986	312.
## 6	Eritrea	Africa	1952	35.9	1438760	329.
## 7	Myanmar	Asia	1952	36.3	20092996	331.
## 8	Lesotho	Africa	1957	45.0	813338	336.
## 9	Burundi	Africa	1952	39.0	2445618	339.
## 10	Eritrea	Africa	1957	38.0	1542611	344.

```
## # ... with 1,694 more rows
```

```
gapminder %>%
  arrange(desc(gdpPercap))
```

```
## # A tibble: 1,704 x 6
##   country    continent  year lifeExp      pop gdpPercap
##   <fct>      <fct>      <int>  <dbl>    <int>    <dbl>
## 1 Kuwait     Asia         1957    58.0   212846   113523.
## 2 Kuwait     Asia         1972    67.7   841934   109348.
## 3 Kuwait     Asia         1952    55.6   160000   108382.
## 4 Kuwait     Asia         1962    60.5   358266    95458.
## 5 Kuwait     Asia         1967    64.6   575003    80895.
## 6 Kuwait     Asia         1977    69.3  1140357    59265.
## 7 Norway     Europe        2007    80.2  4627926   49357.
## 8 Kuwait     Asia         2007    77.6  2505559   47307.
## 9 Singapore  Asia         2007    80.0  4553009   47143.
## 10 Norway    Europe        2002    79.0  4535591   44684.
## # ... with 1,694 more rows
```

1.3 Combinando filtragem com ordenação

```
x <- gapminder %>%
  filter(year==2007) %>%
  arrange(desc(gdpPercap)) %>%
  head(10)
```

Table 1: tabela

country	continent	year	lifeExp	pop	gdpPercap
Norway	Europe	2007	80.196	4627926	49357.19
Kuwait	Asia	2007	77.588	2505559	47306.99
Singapore	Asia	2007	79.972	4553009	47143.18
United States	Americas	2007	78.242	301139947	42951.65
Ireland	Europe	2007	78.885	4109086	40676.00
Hong Kong, China	Asia	2007	82.208	6980412	39724.98
Switzerland	Europe	2007	81.701	7554661	37506.42
Netherlands	Europe	2007	79.762	16570613	36797.93
Canada	Americas	2007	80.653	33390141	36319.24
Iceland	Europe	2007	81.757	301931	36180.79

```
mtcars %>% head(10) %>%
  kbl(format = "latex") %>%
  kable_styling(latex_options = c('striped'))
```

	mpg	cyl	disp	hp	drat	wt	qsec	
gray!6Mazda RX4	gray!621.0	gray!66	gray!6160.0	gray!6110	gray!63.90	gray!62.620	gray!616.46	g
Mazda RX4 Wag	21.0	6	160.0	110	3.90	2.875	17.02	
gray!6Datsun 710	gray!622.8	gray!64	gray!6108.0	gray!693	gray!63.85	gray!62.320	gray!618.61	g
Hornet 4 Drive	21.4	6	258.0	110	3.08	3.215	19.44	
gray!6Hornet Sportabout	gray!618.7	gray!68	gray!6360.0	gray!6175	gray!63.15	gray!63.440	gray!617.02	g
Valiant	18.1	6	225.0	105	2.76	3.460	20.22	
gray!6Duster 360	gray!614.3	gray!68	gray!6360.0	gray!6245	gray!63.21	gray!63.570	gray!615.84	g
Merc 240D	24.4	4	146.7	62	3.69	3.190	20.00	
gray!6Merc 230	gray!622.8	gray!64	gray!6140.8	gray!695	gray!63.92	gray!63.150	gray!622.90	g
Merc 280	19.2	6	167.6	123	3.92	3.440	18.30	