

## cm007 Exercises: Practice with dplyr

```
suppressPackageStartupMessages(library(tidyverse))
suppressPackageStartupMessages(library(gapminder))
suppressPackageStartupMessages(library(tsibble))
```

This worksheet contains exercises aimed for practice with dplyr.

1. (a) What's the minimum life expectancy for each continent and each year? (b) Add the corresponding country to the tibble, too. (c) Arrange by min life expectancy.

```
gapminder %>%
  group_by(continent, year) %>%
  summarize(min_life = min(lifeExp),    #compress into a single value

            country = country[lifeExp == min_life]) %>%
  arrange(min_life)
```

```
## # A tibble: 60 x 4
## # Groups:   continent [5]
##   continent year min_life country
##   <fct>      <int>   <dbl> <fct>
## 1 Africa    1992    23.6 Rwanda
## 2 Asia      1952    28.8 Afghanistan
## 3 Africa    1952    30   Gambia
## 4 Asia      1957    30.3 Afghanistan
## 5 Asia      1977    31.2 Cambodia
## 6 Africa    1957    31.6 Sierra Leone
## 7 Asia      1962    32.0 Afghanistan
## 8 Africa    1962    32.8 Sierra Leone
## 9 Asia      1967    34.0 Afghanistan
## 10 Africa   1967    34.1 Sierra Leone
## # ... with 50 more rows
```

2. Calculate the growth in population since the first year on record *for each country* by rearranging the following lines, and filling in the FILL\_THIS\_IN. Here's another convenience function for you: `dplyr::first()`.

```
gapminder %>%
  group_by(country) %>%
  arrange(year) %>%
  mutate(rel_growth = pop - first(pop))
```

```
## # A tibble: 1,704 x 7
## # Groups:   country [142]
##   country      continent year lifeExp      pop gdpPercap rel_growth
##   <fct>        <fct>    <int>   <dbl>   <int>   <dbl>    <int>
## 1 Afghanistan Asia      1952    28.8  8425333    779.        0
## 2 Albania     Europe    1952    55.2  1282697   1601.        0
## 3 Algeria     Africa    1952    43.1  9279525   2449.        0
## 4 Angola      Africa    1952    30.0  4232095   3521.        0
## 5 Argentina   Americas  1952    62.5 17876956   5911.        0
## 6 Australia   Oceania   1952    69.1  8691212  10040.        0
## 7 Austria     Europe    1952    66.8  6927772   6137.        0
## 8 Bahrain     Asia      1952    50.9  120447    9867.        0
```

```
## 9 Bangladesh Asia 1952 37.5 46886859 684. 0
## 10 Belgium Europe 1952 68 8730405 8343. 0
## # ... with 1,694 more rows
```

```
?first
```

```
## starting httpd help server ... done
```

3. Determine the country that experienced the sharpest 5-year drop in life expectancy, in each continent, sorted by the drop, by rearranging the following lines of code. Ensure there are no NA's. Instead of using `lag()`, use the convenience function provided by the `tsibble` package, `tsibble::difference()`:

```
gapminder %>%
  group_by(country) %>%
  arrange(year) %>%
  mutate(inc_life_exp = difference(lifeExp)) %>%
  drop_na() %>%
  ungroup() %>%
  group_by(continent) %>%
  filter(inc_life_exp == min(inc_life_exp)) %>%
  arrange(inc_life_exp) %>%
  knitr::kable()
```

country	continent	year	lifeExp	pop	gdpPercap	inc_life_exp
Rwanda	Africa	1992	23.599	7290203	737.0686	-20.421
Cambodia	Asia	1977	31.220	6978607	524.9722	-9.097
El Salvador	Americas	1977	56.696	4282586	5138.9224	-1.511
Montenegro	Europe	2002	73.981	720230	6557.1943	-1.464
Australia	Oceania	1967	71.100	11872264	14526.1246	0.170