Exercise 2: R Markdown for Gapminder Exploration

Carleena Ortega

September 25 2019

Table of Contents

## -- Attaching packages ---------------------------------------------------------------- tidyverse 1.2.1 --

## v ggplot2 3.2.1 v purrr 0.3.2  
## v tidyr 1.0.0 v dplyr 0.8.3  
## v readr 1.3.1 v stringr 1.4.0  
## v ggplot2 3.2.1 v forcats 0.4.0

## -- Conflicts ------------------------------------------------------------------- tidyverse\_conflicts() --  
## x dplyr::filter() masks stats::filter()  
## x dplyr::lag() masks stats::lag()

#Data on Gapminder ## Population of Countries Here we show that gapminder has information on life expectancy, population, and GDP per capita of 142 countries from 1952

str(gapminder)

## Classes 'tbl\_df', 'tbl' and 'data.frame': 1704 obs. of 6 variables:  
## $ country : Factor w/ 142 levels "Afghanistan",..: 1 1 1 1 1 1 1 1 1 1 ...  
## $ continent: Factor w/ 5 levels "Africa","Americas",..: 3 3 3 3 3 3 3 3 3 3 ...  
## $ year : int 1952 1957 1962 1967 1972 1977 1982 1987 1992 1997 ...  
## $ lifeExp : num 28.8 30.3 32 34 36.1 ...  
## $ pop : int 8425333 9240934 10267083 11537966 13079460 14880372 12881816 13867957 16317921 22227415 ...  
## $ gdpPercap: num 779 821 853 836 740 ...

## Summary of Gapminder Data

We summarize the gapminder data and show that life expectancy, populations and GDP per capita were collected between 1952 and 2007 in 5 continents.

summary(gapminder)

## country continent year lifeExp   
## Afghanistan: 12 Africa :624 Min. :1952 Min. :23.60   
## Albania : 12 Americas:300 1st Qu.:1966 1st Qu.:48.20   
## Algeria : 12 Asia :396 Median :1980 Median :60.71   
## Angola : 12 Europe :360 Mean :1980 Mean :59.47   
## Argentina : 12 Oceania : 24 3rd Qu.:1993 3rd Qu.:70.85   
## Australia : 12 Max. :2007 Max. :82.60   
## (Other) :1632   
## pop gdpPercap   
## Min. :6.001e+04 Min. : 241.2   
## 1st Qu.:2.794e+06 1st Qu.: 1202.1   
## Median :7.024e+06 Median : 3531.8   
## Mean :2.960e+07 Mean : 7215.3   
## 3rd Qu.:1.959e+07 3rd Qu.: 9325.5   
## Max. :1.319e+09 Max. :113523.1   
##

## Number of data collected per continent

#{r} #<<<<<<< HEAD #group\_by(gapminder$continent) @ n\_distinct(gapminder$country) ======= gapminder %>% group\_by(gapminder$continent) counts<-n\_distinct(gapminder$country) plot(gapminder$continent,counts) #ggplots today >>>>>>> 8f88bcf05c1073c579a702624062a6846bf3deb8

#try to use gg plot #use a unique fnc instead of making a vector for datapt (no penalty but good practice)

#use n\_distinct to give the # of countries