Exercise 2: R Markdown for Gapminder Exploration

Carleena Ortega

September 27 2019

```
## -- Attaching packages ---------------------------------

## 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 ------------------------------------------
## 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. o
##  $ country  : Factor w/ 142 levels "Afghanistan",..: 1 1
##  $ continent: Factor w/ 5 levels "Africa","Americas",..
##  $ year     : int  1952 1957 1962 1967 1972 1977 1982 19
```

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          li
##  Afghanistan:  12   Africa  :624   Min.   :1952   Min.
##  Albania    :  12   Americas:300   1st Qu.:1966   1st Qu
##  Algeria    :  12   Asia    :396   Median :1980   Median
##  Angola     :  12   Europe  :360   Mean   :1980   Mean
##  Argentina  :  12   Oceania : 24   3rd Qu.:1993   3rd Qu
##  Australia  :  12                  Max.   :2007   Max.
##  (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
```

Number of data collected per continent

```
#{r} <<<<<<< HEAD ======= #<<<<<<< HEAD
#group_by(gapminder$continent) @
n_distinct(gapminder$country) ======= >>>>>>>
8d0d2153fc09f15e9d6f7f76d387924fbe9a2b5b 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