Project 2

Jackson Dial

10/7/2021

library(gapminder)  
library(tidyverse)  
library(wesanderson)  
library(RColorBrewer)

## Background

This project uses the gapminder dataset, which contains data on life expectancy, populations, and GDP for various countries.

glimpse(gapminder)

## Rows: 1,704  
## Columns: 6  
## $ country <fct> "Afghanistan", "Afghanistan", "Afghanistan", "Afghanistan", ~  
## $ continent <fct> Asia, Asia, Asia, Asia, Asia, Asia, Asia, Asia, Asia, Asia, ~  
## $ year <int> 1952, 1957, 1962, 1967, 1972, 1977, 1982, 1987, 1992, 1997, ~  
## $ lifeExp <dbl> 28.801, 30.332, 31.997, 34.020, 36.088, 38.438, 39.854, 40.8~  
## $ pop <int> 8425333, 9240934, 10267083, 11537966, 13079460, 14880372, 12~  
## $ gdpPercap <dbl> 779.4453, 820.8530, 853.1007, 836.1971, 739.9811, 786.1134, ~

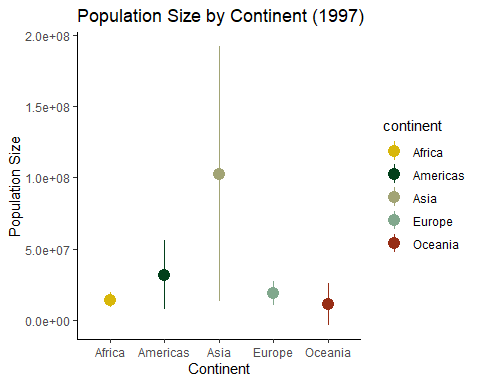
## Data Exploration/Cleaning

sum(is.na(gapminder)) #there are no NA values in the dataframe, lucky me

## [1] 0

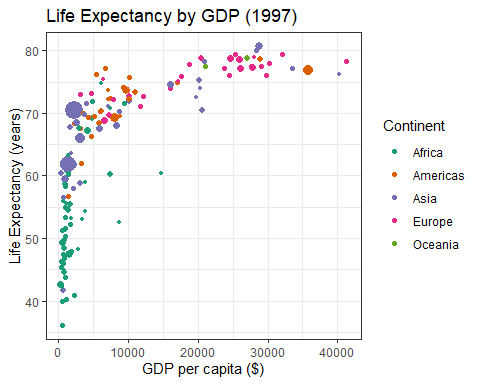
## First Plot

df1 <- gapminder %>%   
 filter(year == 1997) %>%   
 group\_by(continent) %>%   
 summarise(AvgPop = mean(pop),  
 SD = sd(pop),  
 upper = mean(pop) + 1.96 \* (sd(pop) / sqrt(n())),  
 lower = mean(pop) - 1.96 \* (sd(pop) / sqrt(n())))   
  
  
  
ggplot(df1, aes(continent, AvgPop, col = continent))+  
 geom\_linerange(aes(ymin = lower, ymax = upper))+  
 geom\_point(size = 4)+  
 scale\_color\_manual(values = wes\_palette("Cavalcanti1", n = 5, type = "discrete"))+  
 labs(x = "Continent",  
 y = "Population Size",  
 title = "Population Size by Continent (1997)")+  
 theme(legend.position = "none")+  
 theme\_classic()



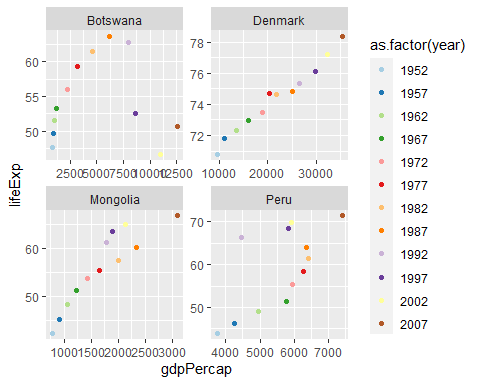
## Second Plot

df2 <- gapminder %>%   
 filter(year == 1997) %>%   
 mutate(dummyVar = 1)  
   
ggplot(df2)+  
 geom\_point(aes(x = gdpPercap, y = lifeExp, col = continent, size = pop))+  
 scale\_color\_brewer(palette = "Dark2")+  
 guides(size = "none")+  
 labs(x = "GDP per capita ($)",  
 y = "Life Expectancy (years)",  
 col = "Continent",  
 title = "Life Expectancy by GDP (1997)")+  
 theme\_bw()

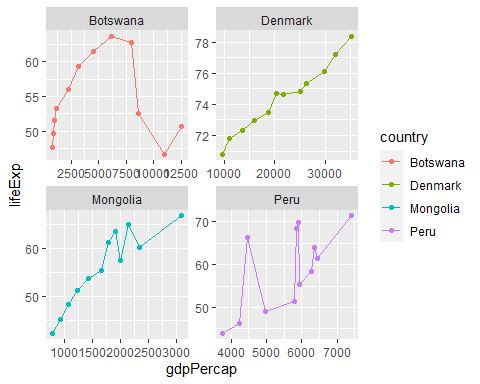


## Third Plot

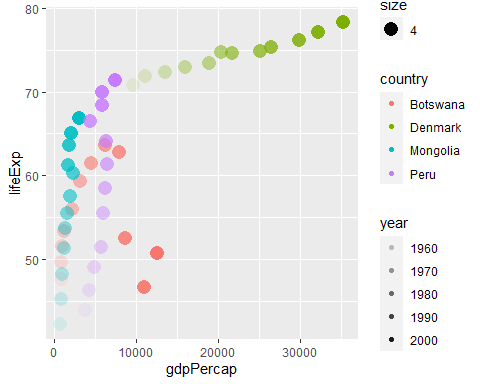
df3 <- gapminder %>%   
 filter(country %in% c("Denmark", "Peru", "Botswana", "Mongolia"))  
  
ggplot(df3, aes(x = gdpPercap, y = lifeExp))+  
 geom\_point(aes(col = as.factor(year)))+  
 facet\_wrap(~country, scales = "free")+  
 scale\_color\_brewer(palette = "Paired")



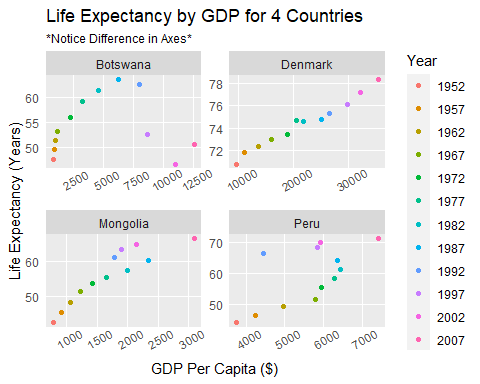
ggplot(df3, aes(x = gdpPercap, y = lifeExp, group = country))+  
 geom\_point(aes(col = country))+  
 geom\_line(aes(col = country))+  
 facet\_wrap(~country, scales = "free")



ggplot(df3, aes(x = gdpPercap, y = lifeExp))+  
 geom\_point(aes(alpha = year, col = country, size = 4), shape = 16)



ggplot(df3, aes(x = gdpPercap, y = lifeExp))+  
 geom\_point(aes(col = as.factor(year)))+  
 facet\_wrap(~country, scales = "free")+  
 scale\_fill\_gradientn(colors = wes\_palette("Zissou1", 12, type = "continuous"))+  
 labs(x = "GDP Per Capita ($)",  
 y = "Life Expectancy (Years)",  
 title = "Life Expectancy by GDP for 4 Countries",  
 subtitle = "\*Notice Difference in Axes\*",  
 col = "Year")+  
 theme(axis.text.x = element\_text(angle = 30),  
 panel.grid.minor = element\_blank(),  
 axis.ticks = element\_blank(),  
 plot.subtitle = element\_text(size = 9))



ggplot(df3, aes(x = gdpPercap, y = lifeExp))+  
 geom\_point()+  
 facet\_wrap(~year)

