Analysis of gapminder data

David Mawdsley

In this notebook we will explore the gapminder data.

First we'll read in the data:

library(readr)  
library(ggplot2)  
library(dplyr)

##   
## Attaching package: 'dplyr'

## The following objects are masked from 'package:stats':  
##   
## filter, lag

## The following objects are masked from 'package:base':  
##   
## intersect, setdiff, setequal, union

setwd("\_episodes\_rmd/")  
gapminder <- read\_csv("./data/gapminder-FiveYearData.csv",   
 col\_types = cols(  
 country = col\_character(),  
 year = col\_integer(),  
 pop = col\_double(),  
 continent = col\_character(),  
 lifeExp = col\_double(),  
 gdpPercap = col\_double()  
)  
 )

Let's investigate how life expectancy varies per continent:

gapminder %>%   
 filter(year == 1997) %>%   
 group\_by(continent, year) %>%   
 summarise(medianLifeExp=median(lifeExp))

## # A tibble: 5 x 3  
## # Groups: continent [?]  
## continent year medianLifeExp  
## <chr> <int> <dbl>  
## 1 Africa 1997 52.759  
## 2 Americas 1997 72.146  
## 3 Asia 1997 70.265  
## 4 Europe 1997 76.116  
## 5 Oceania 1997 78.190

And let's plot a graph:

gapminder %>% filter(continent %in% c("Europe", "Americas")) %>%   
 ggplot(aes(x=year, y=gdpPercap, colour=country)) +  
 geom\_line() + facet\_wrap( ~ continent) + guides(colour="none")

