Population Health Dashboard

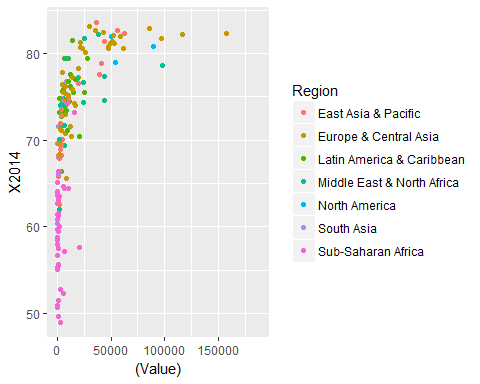
## Exploring the World Development Indicators

#set working directory  
setwd('~/UN Project/national global health risks')  
  
wdi <- read.csv("API\_8\_DS2\_en\_csv\_v2.csv", header=TRUE,sep="\t",stringsAsFactors = TRUE)  
per.cap.gdp <- read.csv('~/UN Project/per capita GDP from UNdata.csv')  
  
#subset 2014 values  
gdp2014 <- subset(per.cap.gdp, per.cap.gdp$Year==2014)  
  
#load population data  
pop <- read.csv("population.csv",skip=4)  
#load region data  
region <- read.csv("Metadata\_Country\_API\_8\_DS2\_en\_csv\_v2.csv")  
colnames(region) <- c('Country.Code','Region','IncomeGroup','SpecialNotes','TableName')  
  
#lets look at life expectancy  
lifexp <- "Life expectancy at birth, total (years)"  
  
library(ggplot2)

## Warning: package 'ggplot2' was built under R version 3.3.2

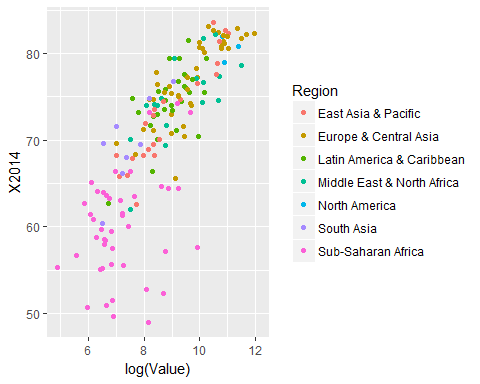
#Create subset of total life expectancy for each country  
life.x <- subset(wdi, wdi$Indicator.Name==lifexp)  
  
#merge with GDP & region data  
mergedx <- merge(life.x, gdp2014, by.x="Country.Name",by.y="Country.or.Area")  
merged1x <- merge(region, mergedx, by.x="Country.Code",by.y="Country.Code")  
  
#Plot 2014 life expectancy versis GDP.   
ggplot(merged1x, aes(x=(Value),y=X2014)) + geom\_point(aes(color=Region))

## Warning: Removed 13 rows containing missing values (geom\_point).



#plot 2014 life expectancy versus GDP. Converted GDP to log scale because most points were  
#bunched at the lower end of GDP scale.  
ggplot(merged1x, aes(x=log(Value),y=X2014)) + geom\_point(aes(color=Region))

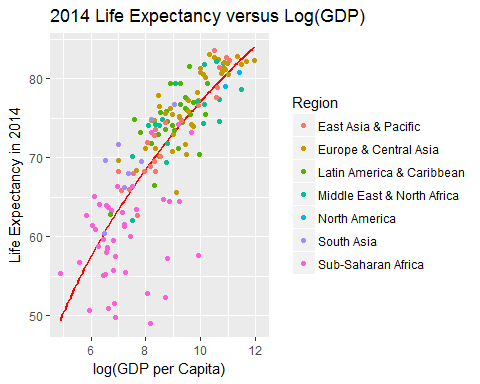
## Warning: Removed 13 rows containing missing values (geom\_point).



#add trend line  
ggplot(merged1x, aes(x=log(Value),y=X2014)) + geom\_smooth(aes(group=1),   
 method='lm',  
 se=FALSE,  
 formula= y~ log(x),  
 color='red') +  
 geom\_point(aes(color=Region)) +  
 ggtitle("2014 Life Expectancy versus Log(GDP)")+  
 xlab("log(GDP per Capita)")+  
 ylab("Life Expectancy in 2014")

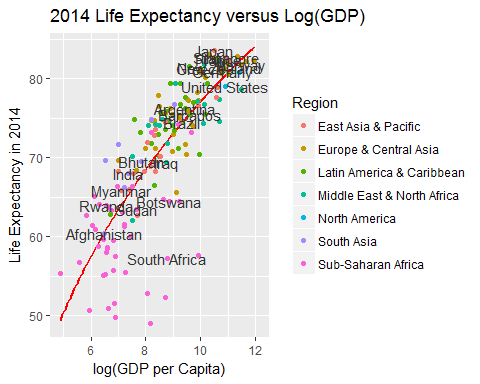
## Warning: Removed 13 rows containing non-finite values (stat\_smooth).

## Warning: Removed 13 rows containing missing values (geom\_point).



# Label some countries  
pointsToLabel <- c("Russia", "Venezuela", "Iraq", "Myanmar", "Sudan",  
 "Afghanistan", "Congo", "Greece", "Argentina", "Brazil",  
 "India", "Italy", "China", "South Africa", "Spane",  
 "Botswana", "Cape Verde", "Bhutan", "Rwanda", "France",  
 "United States", "Germany", "Britain", "Barbados", "Norway", "Japan",  
 "New Zealand", "Singapore")   
#Label some points  
ggplot(merged1x, aes(x=log(Value),y=X2014)) + geom\_smooth(aes(group=1),   
 method='lm',  
 se=FALSE,  
 formula= y~ log(x),  
 color='red') +  
 geom\_point(aes(color=Region)) +  
 ggtitle("2014 Life Expectancy versus Log(GDP)")+  
 xlab("log(GDP per Capita)")+  
 ylab("Life Expectancy in 2014") +  
 geom\_text(aes(label = Country.Name),  
 color = "gray20",  
 data = subset(merged1x, Country.Name %in% pointsToLabel))

## Warning: Removed 13 rows containing non-finite values (stat\_smooth).  
  
## Warning: Removed 13 rows containing missing values (geom\_point).

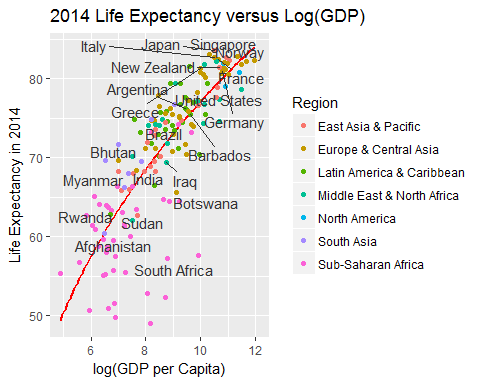


#Labels are overlapping, need to alter  
library("ggrepel")

## Warning: package 'ggrepel' was built under R version 3.3.2

ggplot(merged1x, aes(x=log(Value),y=X2014)) + geom\_smooth(aes(group=1),   
 method='lm',  
 se=FALSE,  
 formula= y~ log(x),  
 color='red') +  
 geom\_point(aes(color=Region)) +  
 ggtitle("2014 Life Expectancy versus Log(GDP)")+  
 xlab("log(GDP per Capita)")+  
 ylab("Life Expectancy in 2014") +  
 geom\_text\_repel(aes(label = Country.Name),  
 color = "gray20",  
 data = subset(merged1x, Country.Name %in% pointsToLabel),  
 force=10)

## Warning: Removed 13 rows containing non-finite values (stat\_smooth).  
  
## Warning: Removed 13 rows containing missing values (geom\_point).



#What is the highest life expectancy?   
merged1x[which.max(merged1x$X2014),'X2014']

## [1] 83.5878

#Where is it?  
merged1x[80,"Country.Name"] #It's Japan

## [1] Japan  
## 264 Levels: Afghanistan Albania Algeria American Samoa Andorra ... Zimbabwe

#What is the lowest?  
merged1x[which.min(merged1x$X2014),'X2014']

## [1] 48.93473

#Where is it?  
merged1x[153,'Country.Name'] #It's Swaziland

## [1] Swaziland  
## 264 Levels: Afghanistan Albania Algeria American Samoa Andorra ... Zimbabwe

knitr::opts\_chunk$set(echo = TRUE)