Lab 8, Solutions

```r
library(tidyverse)
library(stringr)
```

Graphing youth unemployment data

1. Read the youth unemployment data in the file `API_ILO_country_YU.csv`, in the Lab 8 folder on Canvas, into a tibble called `youthUI`.

```r
youthUI <- read_csv("API_ILO_country_YU.csv")
```

2. Use an appropriate pivot function to reshape `youthUI` into a longer table with columns corresponding to `Country Name`, `Country Code`, `year`, and `Unemployment Rate`. When you pivot, automatically convert the newly created `year` column to be an integer column vector (hint: look at the help files for the appropriate pivot function, specifically the `names_transform, values_transform` argument). After pivoting, arrange the new tibble such the rows are ordered by (increasing) `year`, followed by `Country Name` within each `year`.

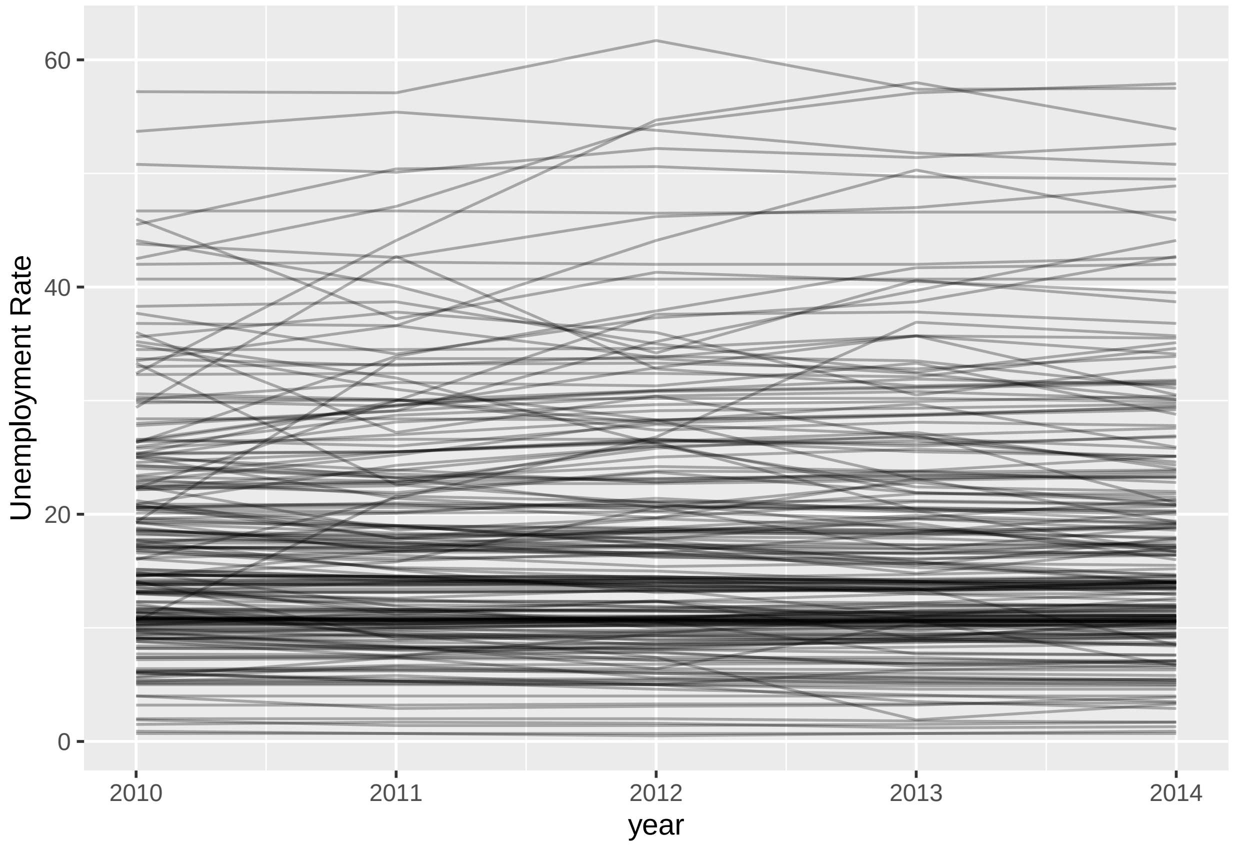
```r
youthUI <- pivot_longer(youthUI,c(`2010`:`2014`),
                        names_to="year",
                        values_to="Unemployment Rate",
                        names_transform = list(year = as.integer)) %>%
  arrange(year,`Country Name`)
youthUI
```

```
## # A tibble: 1,095 x 4
##    `Country Name` `Country Code`  year `Unemployment Rate`
##    <chr>          <chr>          <int>               <dbl>
##  1 Afghanistan    AFG             2010               20.6
##  2 Albania        ALB             2010               25.8
##  3 Algeria        DZA             2010               22.2
##  4 Angola         AGO             2010               10.8
##  5 Arab World     ARB             2010               25.0
##  6 Argentina      ARG             2010               19.5
##  7 Armenia        ARM             2010               38.3
##  8 Australia      AUS             2010               11.4
##  9 Austria        AUT             2010                8.80
## 10 Azerbaijan     AZE             2010               14.6
## # ... with 1,085 more rows
```

```r
#summarize(youthUI,sd(year))
```

3. Plot unemployment rates by year for each "country" in `youthUI`. Represent each time series by a line. Use an appropriate alpha level to manage overplotting.

```r
ggplot(youthUI,aes(x=year,y=`Unemployment Rate`,group=`Country Name`)) + geom_line(alpha=0.3)
```



```
# you may have used a slightly different alpha level
```

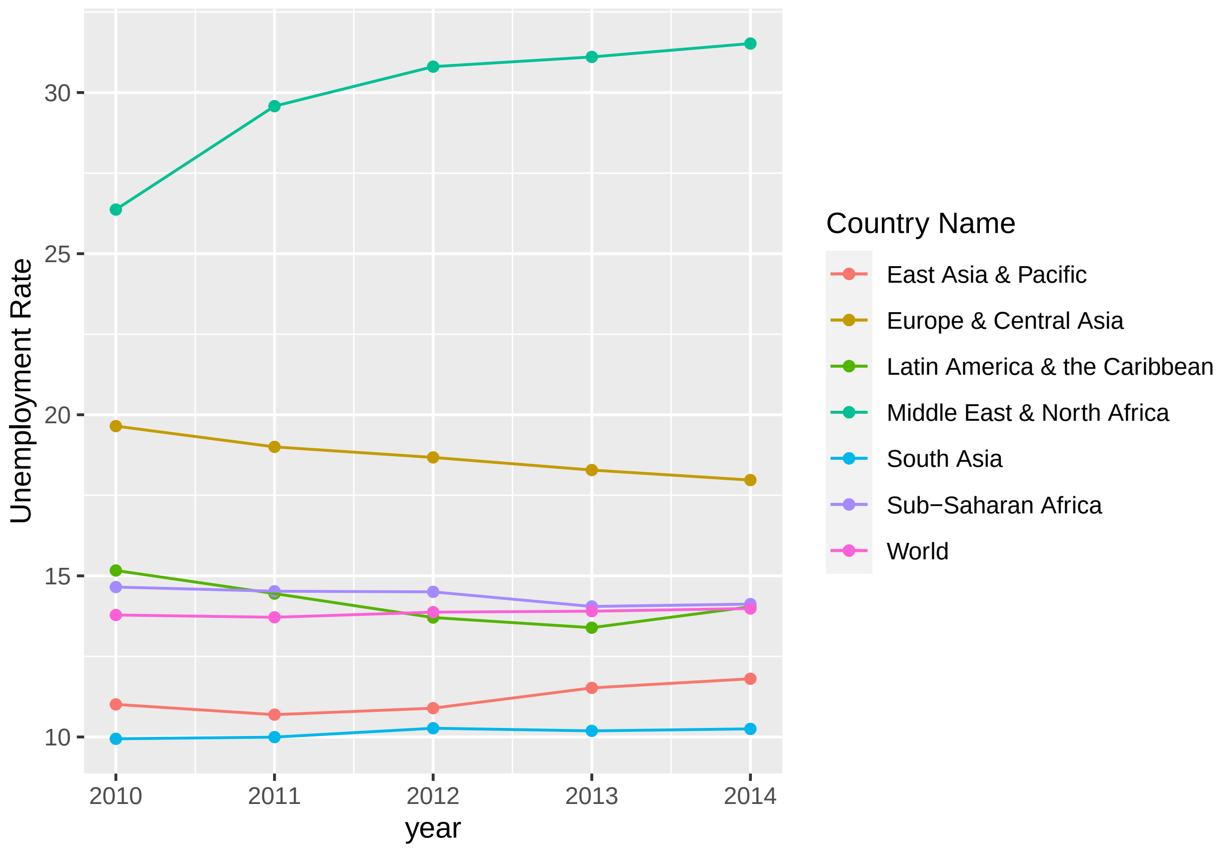
4. Using a **regular expression**, extract the subset of "Countries" whose `Country Name` contains the string "(IDA & IBRD countries)" or "(IDA & IBRD)", and save in a data frame `youthDevel`. (No cheating by using `fixed()`. Hint: `(` is a special character string, so a character string representation of a regexp involving `(` would include "'\(".) Then, using a **regular expression**, remove the "(IDA & IBRD countries)" or "(IDA & IRBD)" from the country names. Notes: IDA stands for International Development Association. Countries that qualify for IDA loans are considered among the poorest developing countries in the world. IBRD stands for International Bank for Reconstruction and Developent. IBRD countries are considered middle-income developing countries.

```
# There are probably several possible regexps you could use. One
# thing to watch for is that there are multiple country names
# that include parentheses, so it is not enough to use "\\(.*\\)".

my_pattern <- " \\(IDA.*\\)"
youthDevel <- filter(youthUI,str_detect(`Country Name`,my_pattern)) %>%
  mutate(`Country Name` = str_replace(`Country Name`,my_pattern,"")) %>%
  select(-`Country Code`)
```

5. Plot unemployment rates by year for each region in `youthDevel` with different colors for each region. Your plot should include both points and lines for each region. Then add a layer that plots the world-wide unemployment data from `youthUI` (with `Country.Name==World`).

```
wd <- filter(youthUI,`Country Name`=="World")
ggplot(youthDevel,aes(x=year,y=`Unemployment Rate`, color=`Country Name`))  + geom_point() +
  geom_line() +  geom_line(data=wd) + geom_point(data=wd)
```


Unemployment Rate
30
25
20
15
10
2010
2011
2012
2013
2014
year
Country Name
East Asia & Pacific
Europe & Central Asia
Latin America & the Caribbean
Middle East & North Africa
South Asia
Sub−Saharan Africa
World