Homework 7 - Drew Kearny - Due December 1

Name: Drew Kearny netID: 231005935  
Collaborated with:

Your homework **must be submitted in Word or PDF format, created by calling “Knit Word” or “Knit PDF” from RStudio on your R Markdown document. Submission in other formats may receive a grade of 0**. Your responses must be supported by both textual explanations and the code you generate to produce your result. Note that all R code used to produce your results must be shown in your knitted file.

Gross domestic product (GDP) is a measure of the total market value of all goods and services produced in a given country in a given year. The percentage growth rate of GDP in year is

An important claim in economics is that the rate of GDP growth is closely related to the level of government debt, specifically with the ratio of the government’s debt to the GDP. The file debt.csv contains measurements of GDP growth and of the debt-to-GDP ratio for twenty countries around the world, from the 1940s to 2010. Note that not every country has data for the same years, and some years in the middle of the period are missing data for some countries but not others.

This data is also used and discussed in Lab. In this homework, we will use the package dplyr for some of the data manipulation and ggplot2 for visualization.

1. Load the data into a **tibble** named debt and make a scatter-plot of the GDP growth rate (vertical axis) against the debt ratio (horizontal axis) with ggplot2 and appropriate axis titles.

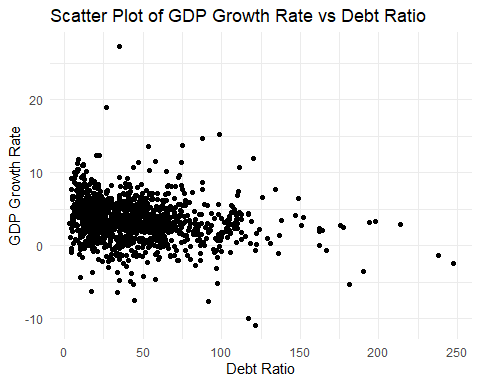
library(dplyr)

##   
## Attaching package: 'dplyr'

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

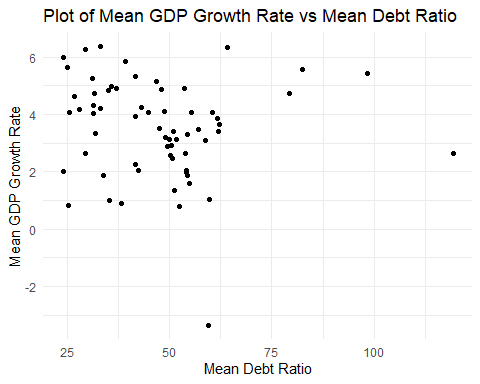
## The following objects are masked from 'package:base':  
##   
## intersect, setdiff, setequal, union

library(ggplot2)  
  
debt <- read.csv("C:\\Users\\kearn\\R Projects\\School Projects\\debt.csv") %>% as\_tibble()  
  
ggplot(debt, aes(x = ratio, y = growth)) +  
 geom\_point() +  
 labs(x = "Debt Ratio", y = "GDP Growth Rate", title = "Scatter Plot of GDP Growth Rate vs Debt Ratio") +  
 theme\_minimal()



1. Use group\_by() and summarise() to compute the mean growth rate and mean debt ratio for each **year** in the data set. Plot the results as a scatter-plot of the mean GDP growth rate (vertical axis) against the mean debt ratio (horizontal axis) with ggplot2 and appropriate axis titles. You only have to submit the code and the plot.

# Get the averages  
yearly\_averages <- debt %>%  
 group\_by(Year) %>%  
 summarise(average\_growth = mean(growth, na.rm = TRUE),  
 average\_debt\_ratio = mean(ratio, na.rm = TRUE))  
  
# Plot  
ggplot(yearly\_averages, aes(x = average\_debt\_ratio, y = average\_growth)) + geom\_point() +  
 labs(x = "Mean Debt Ratio", y = "Mean GDP Growth Rate", title = "Plot of Mean GDP Growth Rate vs Mean Debt Ratio") +  
 theme\_minimal()



1. Fit a linear model of growth on the debt ratio, using lm(). Report the intercept and slope. Add a line to your scatterplot from Q1 showing the fitted regression line. (You may have to redraw the plot in Q1 here in order to add a line.)

The Intercept is 4.27929 and the slope is -0.01835518.

# Fit the linear model  
linear\_model <- lm(growth ~ ratio, data = debt)  
  
summary <- summary(linear\_model)  
intercept <- summary$coefficients[1, 1]  
slope <- summary$coefficients[2, 1]  
  
intercept

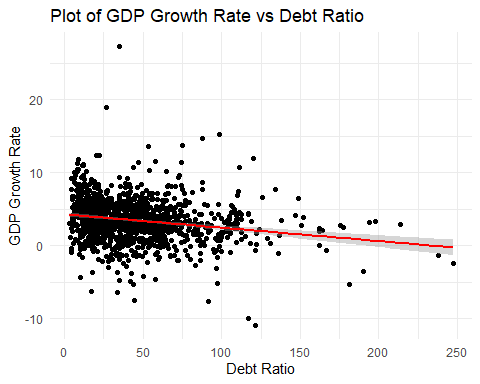
## [1] 4.27929

slope

## [1] -0.01835518

# Plot with regression line  
ggplot(debt, aes(x = ratio, y = growth)) +  
 geom\_point() +  
 geom\_smooth(method = "lm", color = "red") +  
 labs(x = "Debt Ratio", y = "GDP Growth Rate", title = "Plot of GDP Growth Rate vs Debt Ratio") +  
 theme\_minimal()

## `geom\_smooth()` using formula = 'y ~ x'



* 1. Using filter(), create a new tibble named France.debt which just contains the rows of debt for France. It should have 54 rows and 4 columns.

# Filter for France  
France.debt <- debt %>%  
 filter(Country == "France")  
dim(France.debt)

## [1] 54 4

1. Create a new column named next.growth and this column gives next year’s growth *if* the next year is in the data frame, or NA if the next year is missing. (next.growth for 1971 should be (rounded) , but for 1972 it should be NA.) Use mutate(). lead() may be useful here, check it’s help page. Replace the tibble France.debt by the new tibble with both the original columns and the new column next.growth. Print the rows corresponding to Years 1971, 1972, 1990 and 2009 using filter().

# Add the next.growth column  
France.debt <- France.debt %>%  
 mutate(next.year = Year + 1,   
 next.growth = ifelse(next.year %in% Year, round(lead(growth), 3), NA)) %>%  
 select(-next.year)  
  
years <- France.debt %>%  
 filter(Year %in% c(1971, 1972, 1990, 2009))  
  
print(years)

## # A tibble: 4 × 5  
## Country Year growth ratio next.growth  
## <chr> <int> <dbl> <dbl> <dbl>  
## 1 France 1971 5.37 10.8 5.89  
## 2 France 1972 5.89 8.76 NA   
## 3 France 1990 1.93 49.4 3.10  
## 4 France 2009 -1.91 60.0 NA

1. Add a next.growth column, as in Q4, to the **whole** debt tibble. (The next.growth for France in 2009 should be NA, not .) Note that the tibble debt should be replaced by the new tibble with both the original columns and the new column next.growth. Print two rows using filter(): (1) Year 2006 of Austria and (2) Year 1990 of UK.

debt <- debt %>%  
 group\_by(Country) %>%  
 mutate(next.year = Year + 1, # Calculate the next year  
 next.growth = ifelse(next.year %in% Year, round(lead(growth), 3), NA)) %>%  
 select(-next.year) %>%  
 ungroup()  
  
# Print 2006 Austria & 1990 UK  
rows <- debt %>%  
 filter((Country == "Austria" & Year == 2006) | (Country == "UK" & Year == 1990))  
  
print(rows)

## # A tibble: 2 × 5  
## Country Year growth ratio next.growth  
## <chr> <int> <dbl> <dbl> <dbl>  
## 1 Austria 2006 3.25 56.7 NA   
## 2 UK 1990 0.779 33.8 -1.39

1. Linearly regress next year’s growth rate on the current year’s debt ratio. Using ggplot2, make a scatter-plot of next year’s GDP growth against this year’s debt ratio, and add the fitted regression line to the plot. Report the intercept and slope. How do the slope compare to the slope from the regression of the current year’s growth on the current year’s debt ratio?

The new intercept is 3.924729 and new slope is -0.01160844 which is a lower (closer to 0) intercept than the regression of the current year and a slightly less steep negative slops than the regression line of the current year.

# Fit the linear model  
new\_linear <- lm(next.growth ~ ratio, data = debt)  
  
model\_sum <- summary(new\_linear)  
new\_intercept <- model\_sum$coefficients[1, 1]  
new\_slope <- model\_sum$coefficients[2, 1]  
  
new\_intercept

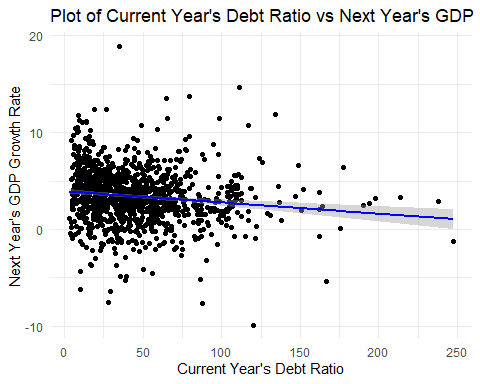
## [1] 3.924729

new\_slope

## [1] -0.01160844

# Plot  
ggplot(debt, aes(x = ratio, y = next.growth)) +  
 geom\_point() +  
 geom\_smooth(method = "lm", color = "blue") +  
 labs(x = "Current Year's Debt Ratio", y = "Next Year's GDP Growth Rate", title = "Plot of Current Year's Debt Ratio vs Next Year's GDP Growth Rate") +  
 theme\_minimal()

## `geom\_smooth()` using formula = 'y ~ x'



1. Add a column, delta.growth, to the debt tibble, giving the difference between next year’s GDP growth rate and this year’s GDP growth rate (i.e., next year’s growth - this year’s growth). Note that the tibble debt should be replaced by the new tibble with both the original columns (including next.growth due to Q5) and the new column delta.growth. Then regress the change in GDP growth on the current GDP growth and the current debt ratio. Report the coefficients.

The intercept is 2.232 so when growth and ratio are 0 the estimated delta.growth value is 2.232. The growth coefficient is -.610 so for each one unit increase in the current year’s growth rate, the change in delta.growth is expected to decrease by about 0.610 units if we hold the debt ratio constant. The coefficient for ratio is -.00484 so for each one unit increase in the current year’s debt ratio, the change in delta.growth is expected to decrease by about 0.00485 units if we hold the current year’s growth rate constant. All coefficients have very low p-values so they are statistically significant.

debt <- debt %>%  
 mutate(delta.growth = next.growth - growth)  
  
# Fit the model  
delta\_model <- lm(delta.growth ~ growth + ratio, data = debt)  
  
# Report coefficients  
summary\_delta <- summary(delta\_model)  
coefficients\_delta <- summary\_delta$coefficients  
  
print(coefficients\_delta)

## Estimate Std. Error t value Pr(>|t|)  
## (Intercept) 2.232346353 0.176442865 12.651950 1.961892e-34  
## growth -0.609954286 0.026910940 -22.665663 3.045325e-94  
## ratio -0.004845935 0.002395063 -2.023302 4.327421e-02