Exercise 13

Your task is to test the Okun's Law (Okun, 1962), i.e., that if unemployment rates are stable, real GDP grows at 3% annually. However, for each percentage point increase in the unemployment rate, real GDP will grow by two percentage points less.

Step 1: load the following libraries:

```
library(jtools)
library(stargazer)
library(utils)
library(lmtest)
install.packages("sandwich")
```

*Note: the sandwich package is required to calculate our robust standard errors.

Step 2: Load the data and attach it to the variable okun.

```
load('/Users/maxoliverstapyltonnorris/Downloads/okun.RData')
okun<-data
attach(okun)
```

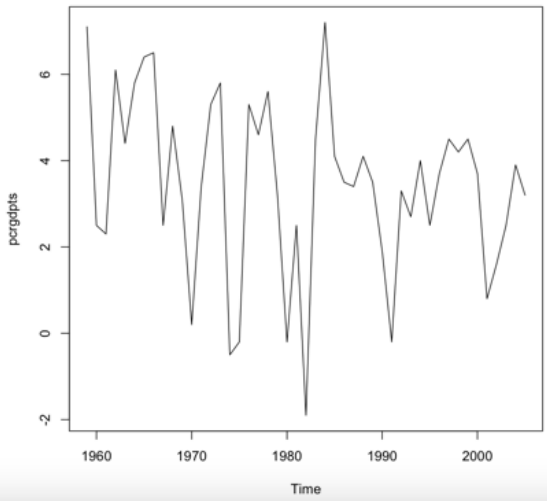
A data.frame: 47 × 4

	year	pcrgdp	unem	cunem
	<int>	<dbl>	<dbl>	<dbl>
1	1959	7.1	5.5	NA
2	1960	2.5	5.5	0.00000000
3	1961	2.3	6.7	1.19999981
4	1962	6.1	5.5	-1.19999981
5	1963	4.4	5.7	0.19999981
6	1964	5.8	5.2	-0.50000000
7	1965	6.4	4.5	-0.69999981
8	1966	6.5	3.8	-0.70000005
9	1967	2.5	3.8	0.00000000
10	1968	4.8	3.6	-0.20000005
11	1969	3.1	3.5	-0.09999990
12	1970	0.2	4.9	1.40000010
13	1971	3.4	5.9	1.00000000
14	1972	5.3	5.6	-0.30000019

Step 3: Visualise the data to check for general correlations and any outliers.

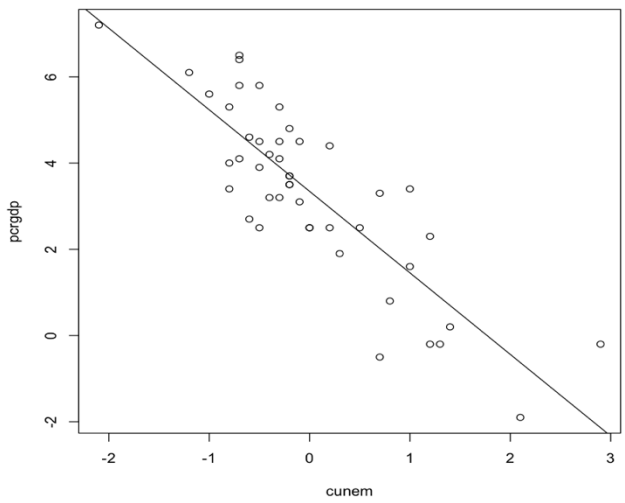
```
pcrgdpts<- ts(pcrgdp,start=c(1959),frequency=1)
plot(pcrgdpts)
```

executed in 477ms, finished 10:56:15 2022-02-06



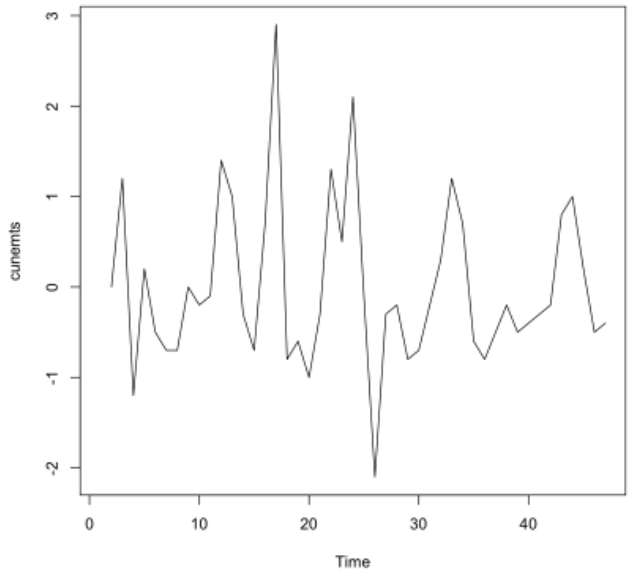

```
plot(cunem,pcrgdp)
abline(model1)
```

executed in 49ms, finished 13:54:24 2022-02-01



```
cunemts<-ts(cunem,start(1959),frequency=1)
plot(cunemts)
```

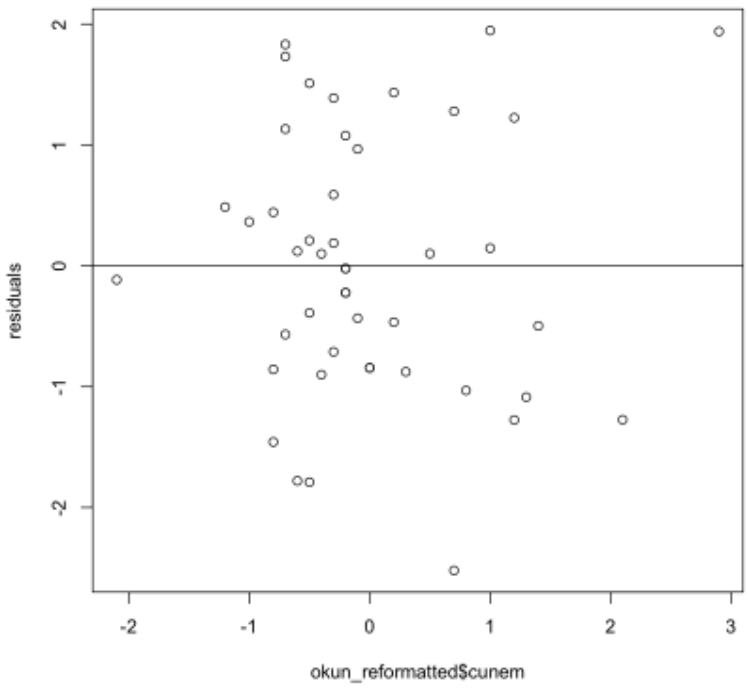
executed in 59ms, finished 10:57:23 2022-02-06



Residual plot:

```
resid(model1)
plot(okun_reformatted$cunem, resid(model1),ylab='residuals',x_lab='cunemployment')
abline(0,0)
```

executed in 58ms, finished 11:27:19 2022-02-06



$$Residual = y - \hat{y}$$

Step 4: Specify the model and run the regression with/out robust t statistics:

```
#model 1
model1<-dynlm(pcrgdp~cunem)
summ(model1, digits = 5)

coeftest(model1)
coeftest(model1, vcov=hccm (model1, type="hc0"))
```

executed in 72ms, finished 14:00:27 2022-02-01

```
MODEL INFO:
Observations: 46 (1 missing obs. deleted)
Dependent Variable: pcrgdp
Type: OLS linear regression

MODEL FIT:
F(1,44) = 107.91642, p = 0.00000
R² = 0.71037
Adj. R² = 0.70378

Standard errors: OLS
------------------------------------------------------------
                       Est.      S.E.      t val.         p
----------------- ---------- --------- ----------- ---------
(Intercept)          3.34443   0.16267    20.55903   0.00000
cunem               -1.89092   0.18202   -10.38828   0.00000
------------------------------------------------------------


t test of coefficients:

            Estimate Std. Error t value  Pr(>|t|)
(Intercept)  3.34443    0.16267  20.559 < 2.2e-16 ***
cunem       -1.89092    0.18202 -10.388 2.037e-13 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1


t test of coefficients:

            Estimate Std. Error t value  Pr(>|t|)
(Intercept)  3.34443    0.15964 20.9500 < 2.2e-16 ***
cunem       -1.89092    0.21779 -8.6824 4.319e-11 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

Step 5: Check for Heteroskedasticity using the Breusch-Pagan test.

```
bptest(model1)
```

executed in 15ms, finished 12:43:52 2022-02-03

```
        studentized Breusch-Pagan test

data:  model1
BP = 4.0698, df = 1, p-value = 0.04366
```

Step 6: Compute and compare robust standard errors

```
summ(model1)
```

executed in 32ms, finished 12:39:03 2022-02-03

```
MODEL INFO:
Observations: 46 (1 missing obs. deleted)
Dependent Variable: pcrgdp
Type: OLS linear regression

MODEL FIT:
F(1,44) = 107.92, p = 0.00
R² = 0.71
Adj. R² = 0.70

Standard errors: OLS
-------------------------------------------------
                      Est.   S.E.   t val.      p
----------------- -------- ------ -------- ------
(Intercept)           3.34   0.16    20.56   0.00
cunem                -1.89   0.18   -10.39   0.00
-------------------------------------------------
```

```
: summ(model1,robust=TRUE)
```

executed in 32ms, finished 12:38:37 2022-02-03

```
MODEL INFO:
Observations: 46 (1 missing obs. deleted)
Dependent Variable: pcrgdp
Type: OLS linear regression

MODEL FIT:
F(1,44) = 107.92, p = 0.00
R² = 0.71
Adj. R² = 0.70

Standard errors: Robust, type = HC3
-------------------------------------------------
                      Est.   S.E.   t val.      p
----------------- -------- ------ -------- ------
(Intercept)           3.34   0.17    19.64   0.00
cunem                -1.89   0.26    -7.16   0.00
-------------------------------------------------
```

Results:

Steps 1&2 illustrate the necessary libraries and dataset required for our analysis. We decide to load:

- sandwich to calculate robust standard errors.
- Stargazer for our export tables
- Distributed lag and non linear model to handle our time series data.
- Liner model test for Breusch-Pagan and

Before we fit any models to our dataset it is good practice to pre-process and visualise our data. Viewing the table in R shows our first entry in 1959 has a cunem value NA, while most models in r have arguments to deal with non-float values, it is worth keeping in mind that these occasionally cause output errors. Step 3 shows our initial visualisation plots which help identify any potential outliers to remove and the presence of heteroskedasticity in the residual plot. The residual plot already shows that our data contains heteroskedastic qualities which should be investigated with further diagnostic tests.

Step 4 runs the regression of cumem on gdp from our Okun dataset, we can observe from our visualisation that cumen and gdp exhibit a negative correlation, such that an increase in cunem reduces gdp. An examination of the coefficients produced from the following regression:

$$pcrgdp_t = \beta_0 + \beta_1 cunem_t + u_t$$

Show that $\beta_{1=-1.891}$ this suggests that for each percentage point increase in the unemployment rate, real GDP will grow by roughly two percentage points less.

Step 5 tests for the heteroskedasticity we suspected from our initial residuals plot. Using the BP test we are setting our null hypothesis to H0: Var(u,cunem)=constant. The p-value from our BP test shows that at the 5% level we can reject the null hypothesis.

Since we don’t necessarily know the form of heteroskedasticity we decide to use the heteroskedastic robust standard errors in our final regression model as shown by step 6. Comparing the robust statistics with the original regression, we can see that they vary only slightly and thus do not affect the respective p-values from both regressions. Furthermore, our R^2 value of 0.71 shows that 70% of the variation in our data is explained by the regression. Given our residual plot illustrating slight heteroskedasticity and the output from the robust regression supporting this, we can conclude that despite the small presence of heteroskedasticity our analysis broadly supports the initial claim a percentage point increase in unemployment decreases GDP growth by roughly 2 percentage points.