```r
library(knitr)
```

```r
data <- read.csv("uval.csv")
```

Question 1

```r
lm.fit.1 <- lm(growth ~ underval + log(gdp), data = data)

kable(coef(summary(lm.fit.1)),
      digits = 5, caption = "Coefficients & Std. Error of linear model")
```

Table 1: Coefficients & Std. Error of linear model

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	-0.03525	0.00665	-5.30037	0.00000
underval	0.00476	0.00218	2.18614	0.02898
log(gdp)	0.00630	0.00079	7.96591	0.00000

Since the coeficient of $log(gdp)$ is positive, this model does not seem to support the idea of "catching-up" as countries with higher GDP have a higher economic growth rate. However, it does support the idea that under-valuing a currency boosts economic growth as the coefficient of *underval* is positive, indicating a positive *underval* index, which represents undervaluing, leads to higher economic growth.

Question 2

Q2 a)

```r
lm.fit.2 <- lm(growth ~ underval + log(gdp) + factor(country) + factor(year),
               data = data)

kable(coef(summary(lm.fit.2))[2:3, ],
      digits = 5, caption = "Coefficients & Std. Error of linear model")
```

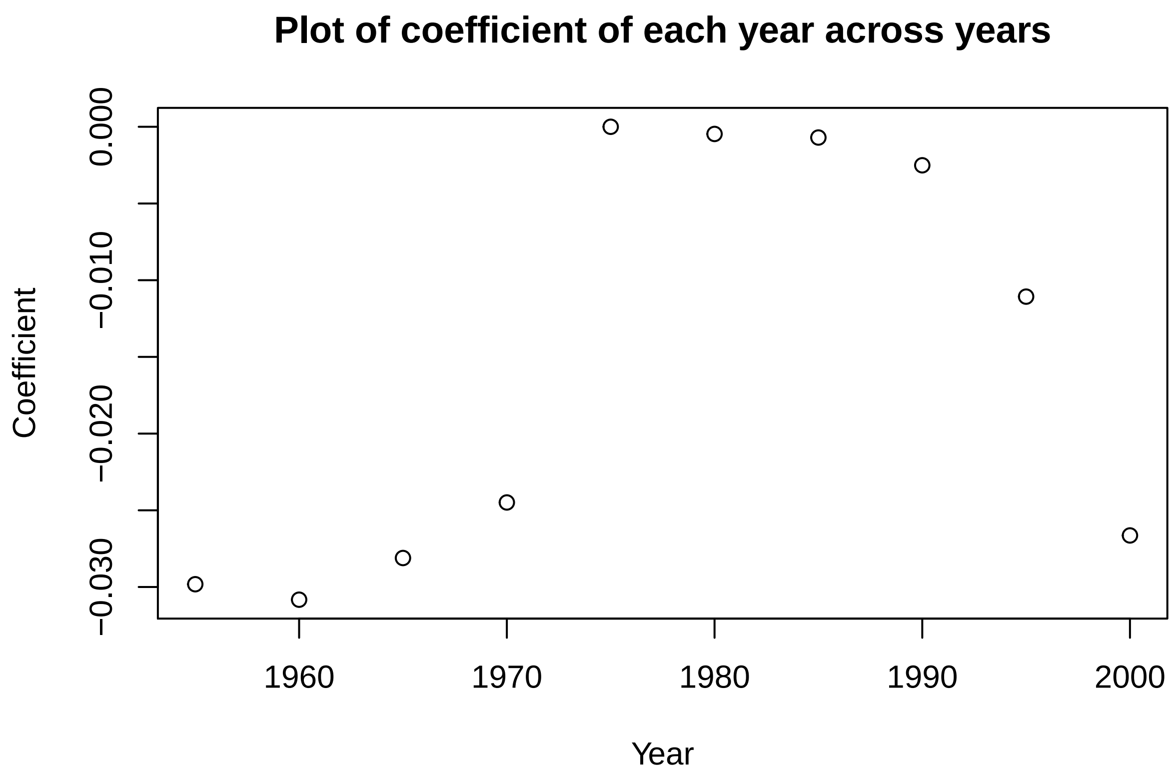
Table 2: Coefficients & Std. Error of linear model

	Estimate	Std. Error	t value	Pr(>|t|)
underval	0.01361	0.00290	4.69667	0
log(gdp)	0.02892	0.00317	9.13254	0

Q2 b) It is more appropriate to use `factor(year)` as there are only unique years that are 5 years apart. As such, modelling this way we will have a slope for each 5-year interval rather than a single slope for each increment of year.

Q2 c)

```r
plot(unique(data$year), c(0, coef(lm.fit.2)[182:190]),
     xlab = "Year", ylab = "Coefficient",
     main = "Plot of coefficient of each year across years")
```



Q2 d)

Since the coeficient of $log(gdp)$ is positive, this model does not seem to support the idea of "catching-up" as countries with higher GDP have a higher economic growth rate. However, it does support the idea that under-valuing a currency boosts economic growth as the coefficient of *underval* is positive, indicating a positive *underval* index, which represents undervaluing, leads to higher economic growth.

Question 3

Q3 a)

```r
lm.r.squareds <- matrix(nrow = 2, ncol = 2)
rownames(lm.r.squareds) <- c("$R^2$", "Adj. $R^2$")
colnames(lm.r.squareds) <- c("Model 1", "Model 2")
lm.r.squareds[, 1] <- with(summary(lm.fit.1), c(r.squared, adj.r.squared))
lm.r.squareds[, 2] <- with(summary(lm.fit.2), c(r.squared, adj.r.squared))

kable(lm.r.squareds, digits = 5,
      caption = "$R^2$ values for each linear model")
```

Table 3: R^2 values for each linear model

	Model 1	Model 2
R^2	0.04855	0.42924
Adj. R^2	0.04709	0.33214

Q3 b)

```
# Taken from textbook chapter 3 page 77
cv.lm <- function(data, formulae, nfolds = 5) {
    data <- na.omit(data)
    formulae <- sapply(formulae, as.formula)
    n <- nrow(data)
    fold.labels <- sample(rep(1:nfolds, length.out = n))
    mses <- matrix(NA, nrow = nfolds, ncol = length(formulae))
    colnames <- as.character(formulae)
    for (fold in 1:nfolds) {
        test.rows <- which(fold.labels == fold)
        train <- data[-test.rows, ]
        test <- data[test.rows, ]
        for (form in 1:length(formulae)) {
            current.model <- lm(formula = formulae[[form]], data = train)
            predictions <- predict(current.model, newdata = test)
            test.responses <- eval(formulae[[form]][[2]], envir = test)
            test.errors <- test.responses - predictions
            mses[fold, form] <- mean(test.errors^2)
        }
    }
    return(colMeans(mses))
}
```

```
loocv.mse <- cv.lm(data, c("growth ~ underval + log(gdp)", "growth ~ underval + log(gdp) + factor(count
```

```
names(loocv.mse) <- c("Model 1", "Model 2")
kable(loocv.mse, caption = "$\\hat{MSE}$ of linear models by LOOCV")
```

Table 4: $\hat{MSE}$ of linear models by LOOCV

	x
Model 1	0.001030348892
Model 2	0.000952766927