

DATA 606 Data Project Proposal

Sam Reeves

Data Preparation

```
library(zoo)
library(tidyverse)
rates <- read.csv(
  "https://raw.githubusercontent.com/TheWerefriend/exchange-rate-prediction/master/rates.csv") %>%
  na.locf()

colnames(rates)[colnames(rates) == "X"] <- "date"
rates <- mutate(rates, date=as.Date(date[[1]]))
str(rates)
```

Research question

You should phrase your research question in a way that matches up with the scope of inference your dataset allows for.

Is it reasonable to predict the USD/MMK exchange rate using the historical data for all ***/MMK exchange rates? Does the prediction get better when you remove some currencies as independent variables?

Cases

What are the cases, and how many are there? Each day (this has halted since the coup), the Central Bank of Myanmar declares an exchange rate for each of 38 currencies which can be used to by Kyats. We have nearly 10 years of data, each day is an observation, and there are 38 variables in each observation.

Data collection

Describe the method of data collection.

This data was scraped from the Central Bank of Myanmar's main website.

Type of study

What type of study is this (observational/experiment)?

This is an observational study.

Data Source

If you collected the data, state self-collected. If not, provide a citation/link.

I collected this from a website which is currently inactive.

Dependent Variable

What is the response variable? Is it quantitative or qualitative?

The dependent variable is the USD/MMK exchange rate, which is quantitative.

Independent Variable

You should have two independent variables, one quantitative and one qualitative.

The independent variables are the 38 exchange rates given in previous observations.

Relevant summary statistics

Provide summary statistics for each the variables. Also include appropriate visualizations related to your research question (e.g. scatter plot, boxplots, etc). This step requires the use of R, hence a code chunk is provided below. Insert more code chunks as needed.

```
xAndY <- function(days = 1, input = rates) {
  x <- input[1:(nrow(input)-days), 2:39]
  return(mutate(x, y = input[(days+1):nrow(input), "USD"]))
}

dataSplit <- function(data, ratio = 0.7) {
  n <- nrow(data) * ratio
  train <- data[1:n,]
  test <- data[(n+1):nrow(data),]
  return(list(train, test))
}

data <- xAndY(5) %>% dataSplit()
model <- lm(formula = y ~ ., data = data[[1]])
summary(model)
```

```
##
## Call:
## lm(formula = y ~ ., data = data[[1]])
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -971.20   -6.27    0.14    7.40   254.67
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) 731.333974  20.645845  35.423  < 2e-16 ***
## AUD           0.151422   0.050567   2.994  0.00278 **
## BDT           9.558132   7.026031   1.360  0.17385
## BND           0.019637   0.045605   0.431  0.66681
## BRL           0.116824   0.066048   1.769  0.07707 .
## CAD          -0.032364   0.047783  -0.677  0.49828
## CHF          -0.055661   0.034690  -1.605  0.10874
## CNY           0.751970   0.542076   1.387  0.16552
## CZK           4.882774   1.159796   4.210 2.66e-05 ***
## DKK          -0.515311   0.658887  -0.782  0.43425
## EGP           0.129335   0.084175   1.536  0.12456
## EUR           0.115569   0.087683   1.318  0.18763
## GBP           0.028130   0.019845   1.417  0.15649
## HKD           1.621013   4.192098   0.387  0.69903
## IDR           4.887404   4.319901   1.131  0.25803
## ILS           0.945622   0.161575   5.853 5.59e-09 ***
```

```

## INR          3.736317   6.761463   0.553   0.58060
## JPY          -0.085245   0.027454  -3.105   0.00193 **
## KES         -27.797521   5.643598  -4.925  9.06e-07 ***
## KHR          -8.970387   2.949512  -3.041   0.00238 **
## KRW          -0.002126   0.465226  -0.005   0.99636
## KWD          -0.261693   0.065146  -4.017  6.10e-05 ***
## LAK          45.754378  10.858411   4.214  2.62e-05 ***
## LKR         -11.895599   8.834494  -1.346   0.17829
## MYR          -0.228767   0.114190  -2.003   0.04526 *
## NOK          -1.326378   0.288748  -4.594  4.61e-06 ***
## NPR          3.385088  10.729146   0.316   0.75241
## NZD          -0.038123   0.043031  -0.886   0.37575
## PHP         -25.360148   2.719297  -9.326 < 2e-16 ***
## PKR         -11.224259   4.658354  -2.409   0.01606 *
## RSD          12.178645   7.053832   1.727   0.08440 .
## RUB          1.273497   0.706271   1.803   0.07151 .
## SAR          0.811575   2.376808   0.341   0.73279
## SEK          -1.910379   0.336862  -5.671  1.61e-08 ***
## SGD          -0.496608   0.161564  -3.074   0.00214 **
## THB         -12.554782   2.078591  -6.040  1.81e-09 ***
## USD          1.706139   0.820625   2.079   0.03773 *
## VND         -20.844373  18.593320  -1.121   0.26238
## ZAR          0.464775   0.214310   2.169   0.03022 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 24.41 on 2136 degrees of freedom
## Multiple R-squared:  0.9828, Adjusted R-squared:  0.9825
## F-statistic: 3215 on 38 and 2136 DF, p-value: < 2.2e-16

```