

Exchange rate Analysis

TEAM Z

PHILIP WILSON, NUNZIO GATTI, SUSMITA GANGOPADHYAY, SUMESH RAMACHANDRA, and BANG DU

An exchange rate is the rate at which one country exchanges currency with another. In our application we have tried to capture the exchange rate of one country and found out other exchange rates which are most closely correlated with its behavior. We have considered G20 countries from a time frame of Dec 1998 to Dec 2017.

ACM Reference Format:

Philip Wilson, Nunzio Gatti, Susmita Gangopadhyay, Sumesh Ramachandra, and Bang Du. 2018. Exchange rate Analysis: TEAM Z. 1, 1 (January 2018), 6 pages. <https://doi.org/10.1145/nnnnnnn.nnnnnnn>

1 INTRODUCTION

As team we had first decided on a different topic "Will I live longer if I cycle to work?". But in the process of trying to collect data it was seen that there was a lot of scientific studies associated with the topic but we struggled to find a way of meaningfully combining different life expectancy models. Most of the data we found was already clean and we did not have much processing to do[5]. We could not find out any scientific method that directly links the factors we were considering (amount of oxygen intake, pollution, BMI etc) to the increase in life expectancy[3]. Combining all these factors into one prediction model became a challenge.

So, after spending almost 4 weeks in trying to collect data we had to shift our focus to some other topic where we could get a significantly high volume of data. Now we are trying to predict the top three indicators affecting the currency of that country. After a considerable amount of research we have taken into account the following factors: exchange rates, interest rates, employment, population, import, export, GDP and inflation[6]. The annual factors, like the employment, population, trade networks, GDP and inflation are taken in the time range of 1991-2016.

2 PROBLEM STATEMENT

2.1 High Frequency Exchange Rate Data

Our aim was to investigate how exchange rates influence each other. As we were able to gather nearly a 1000 samples we were able to perform a number of different machine learning techniques to perform this analysis.

2.2 Low Frequency Economic Indicators

From investigating economic research we were directed towards a number of different economic indicators which are understood to influence a currency's value. These indicators are only reported on an annual basis and due to a number of factors was only available

Authors' address: Philip Wilson; Nunzio Gatti; Susmita Gangopadhyay; Sumesh Ramachandra; Bang Du.

2018. This is the author's version of the work. It is posted here for your personal use. Not for redistribution. The definitive Version of Record was published in, <https://doi.org/10.1145/nnnnnnn.nnnnnnn>.

from 1999 to 2015. This means we only had 17 samples. This limited the type of analysis we were able to perform. So we decided to find the economic which was most correlated to the performance of a specific exchange rate.

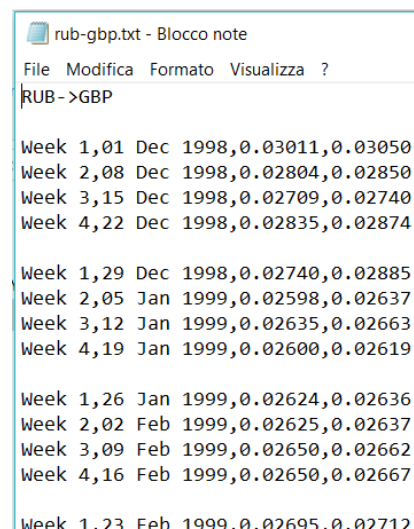
3 IMPLEMENTATION

3.1 High Frequency Exchange Rate Data

3.1.1 Cleaning with python

The data that we have collected is from oanda.com, Bank of England and World Bank websites.

First step, Download : For the currency exchange data, it was about weekly average exchange rates from December 1998 to December 2017. We considered all currency of the G20's countries and downloaded them with GBP as the base currency[1]. Every file downloaded was made in this way (see Fig1):



```
rub-gbp.txt - Blocco note
File Modifica Formato Visualizza ?
RUB -> GBP

Week 1,01 Dec 1998,0.03011,0.03050
Week 2,08 Dec 1998,0.02804,0.02850
Week 3,15 Dec 1998,0.02709,0.02740
Week 4,22 Dec 1998,0.02835,0.02874

Week 1,29 Dec 1998,0.02740,0.02885
Week 2,05 Jan 1999,0.02598,0.02637
Week 3,12 Jan 1999,0.02635,0.02663
Week 4,19 Jan 1999,0.02600,0.02619

Week 1,26 Jan 1999,0.02624,0.02636
Week 2,02 Feb 1999,0.02625,0.02637
Week 3,09 Feb 1999,0.02650,0.02662
Week 4,16 Feb 1999,0.02650,0.02667

Week 1,23 Feb 1999,0.02695,0.02712
```

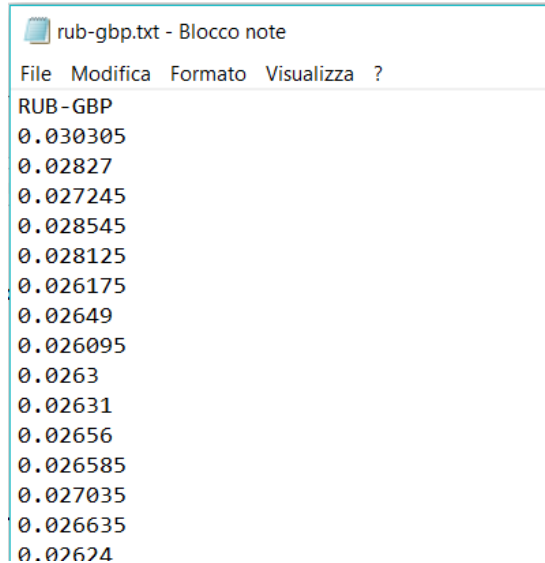
Fig. 1. Raw currency data

For currency exchange data below technique was observed

- Size of table : 1240 x 4 columns
- Col[0]= Week(1,2,3,4)
- Col[1]= Month-Year (i.e may 2002)
- Col[2]= Bid (Bid is the price a buyer is willing to pay for a security)
- Col[3]= Ask (Ask is the price a seller is willing to accept for a security)

Second step, Aggregation : We had 16 files (16 because in the G20 group France, Italy, Germany and European Union has currency

Euro). Using python we gave in input a file formed by 1240(included blank space between rows) x 4 columns and received the output below in Fig 2.

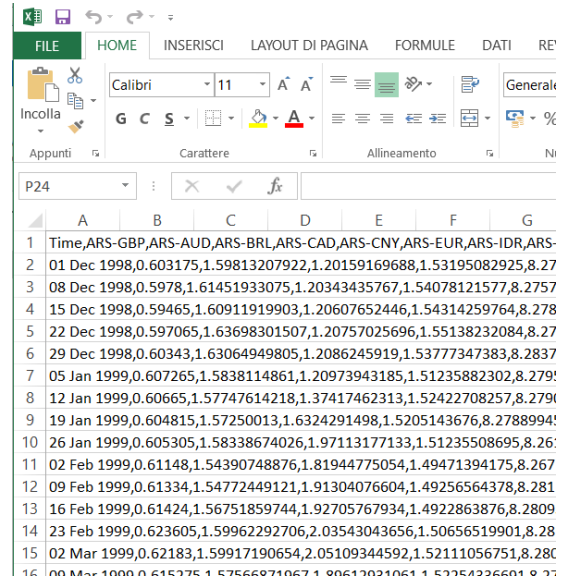


| File | Modifica | Formato | Visualizza ? |
|----------|----------|---------|--------------|
| RUB-GBP | | | |
| 0.030305 | | | |
| 0.02827 | | | |
| 0.027245 | | | |
| 0.028545 | | | |
| 0.028125 | | | |
| 0.026175 | | | |
| 0.02649 | | | |
| 0.026095 | | | |
| 0.0263 | | | |
| 0.02631 | | | |
| 0.02656 | | | |
| 0.026585 | | | |
| 0.027035 | | | |
| 0.026635 | | | |
| 0.02624 | | | |

Fig. 2. Intermediate Processed Data

All spaces were removed using python, saved in one different file, and an average between bid and ask was calculated. Third step, Creation of a complete matrix : We wrote another python script called complete Matrix.py that was used to create all possible currency pair. Remembering that the currency were 16 , we had a output matrix with 16x16 column and 994 rows (see Fig3).

3.1.2 Modeling. The target for this assignment was, given a currency exchange rate as an input, find the **three** most relevant currency exchanges rates that most influences it. To do that we considered the matrix containing all possible combinations of currency exchanges and we used Matlab for this analysis. We used linear regression technique to create a model to find if there is a clear relationship between one exchange rate and the other G20 exchange rates. We then look to find what are the top 3 currencies that influences the given exchange rate. Procedure: To find the most relevant exchange rates we created a Matlab script using CVX tool. CVX is a modeling system for constructing and solving disciplined convex programs (DCPs)[7].



| | A | B | C | D | E | F | G |
|----|---|---|---|---|---|---|---|
| 1 | Time,ARS-GBP,ARS-AUD,ARS-BRL,ARS-CAD,ARS-CNY,ARS-EUR,ARS-IDR,ARS- | | | | | | |
| 2 | 01 Dec 1998,0.603175,1.59813207922,1.20159169688,1.53195082925,8.27 | | | | | | |
| 3 | 08 Dec 1998,0.5978,1.61451933075,1.20343435767,1.54078121577,8.2757 | | | | | | |
| 4 | 15 Dec 1998,0.59465,1.60911919903,1.20607652446,1.54314259764,8.278 | | | | | | |
| 5 | 22 Dec 1998,0.597065,1.63698301507,1.20757025696,1.55138232084,8.27 | | | | | | |
| 6 | 29 Dec 1998,0.60343,1.63064949805,1.2086245919,1.53777347383,8.2837 | | | | | | |
| 7 | 05 Jan 1999,0.607265,1.5838114861,1.20973943185,1.51235882302,8.279 | | | | | | |
| 8 | 12 Jan 1999,0.60665,1.57747614218,1.37417462313,1.52422708257,8.279 | | | | | | |
| 9 | 19 Jan 1999,0.604815,1.57250013,1.6324291498,1.5205143676,8.2788994 | | | | | | |
| 10 | 26 Jan 1999,0.605305,1.58338674026,1.97113171133,1.51235508695,8.26 | | | | | | |
| 11 | 02 Feb 1999,0.61148,1.54390748876,1.81944775054,1.49471394175,8.267 | | | | | | |
| 12 | 09 Feb 1999,0.61334,1.54772449121,1.91304076604,1.49256564378,8.281 | | | | | | |
| 13 | 16 Feb 1999,0.61424,1.56751859744,1.92705767934,1.4922863876,8.2809 | | | | | | |
| 14 | 23 Feb 1999,0.623605,1.59962292706,2.03543043656,1.50656519901,8.28 | | | | | | |
| 15 | 02 Mar 1999,0.62183,1.59917190654,2.05109344592,1.52111056751,8.28 | | | | | | |
| 16 | 09 Mar 1999,0.615275,1.57566071067,1.90617031061,1.52754236601,8.27 | | | | | | |

Fig. 3. Extrapolated exchange rate data

```

gamma=10;
for k=1:100
    cvx_begin quiet
        variable w2( p+1 )
        minimize( norm(Y*w2-f) + gamma*norm(w2,1) );
    cvx_end

    [iNzero] = find(abs(w2) > 1e-5);
    length(iNzero')
    if length(iNzero')<=3
        if length(iNzero')==3
            %check the accuracy of the model calculating the mean
            %squared error on the test set
        end
        if rangeTooBig==1
            incr=incr-0.3;
            rangeTooBig=0;
        end
        gamma=gamma+incr;
    end
end

```

Explanation of the procedure :

- **cvx_begin** : Must be written as the first instruction of a CVX model
- **cvx_begin quiet** : Prevents the model from producing any screen output while it is being solved.
- **cvx_end** : Must be the last instruction of the CVX procedure
- **variable** : It is used to declare the variable, it includes the name of the variable, an optional dimension list, and one or more keywords that provide additional information about the content or structure of the variable.
- **minimise** : It is a command used to declare an objective function (can be also maximise) N.B. The objective function in a call to minimise must be convex; the objective function in a call to maximise must be concave. In this case $Y*w2-f$ is convex.

Goal :To find the best weight vector that minimises the error

- $+gamma(w2)$: We used gamma for the L1 regularisation like **Lasso**. This technique is normally used to solve the overfitting problem in statistical models. **In this particular case we thought that it was a good idea since we built a model using 211 different variables and so the model was complex and the risk of overfitting was high**
- $find\ abs(w2)>(1e-5)$: we found which weights are not switched off by the regulariser, that are the most relevant variables.

After obtaining the most relevant values, we split the data into training and test datasets and calculated the MSE (Means Squared Error) on the test dataset. After executing the script 15 times ,we had an average error of 20 for the test data. set[9].

Selection of gamma value:

One of the most tedious part of this process was setting an appropriate gamma value(a bias noise to reduce overfitting of data) for the regularisation. After a considerable amount of research we understood that it is almost impossible set a good gamma because it is totally dependent on both the training set and all parameters that were used[8].

- Gamma is dependent on both the training set and the other parameters you use.
- There is no "good Gamma" for any data set alone
- In mathematica terms "Gamma" is called the "Lagrangian multiplier" (complexity control).
- The higher Gamma, the higher would our data be regularised. Which means increasing Gamma results in less overfitting but also greater bias.
- Gamma values around 20 are extremely high, and should be used only when you are using high depth or if you want to directly control the features which are dominating in the data set (i.e too strong feature engineering).

To find the most relevant values, based on the dataset, gamma was changed automatically in every iteration until exactly three variables were selected (feature selection).

Create a model:

The second analysis that we did using the high frequency data values, was about creating a model that was able to derive the currency exchange trends, establishing the top three influential currency exchange. To do this, we again use the CVX tool, to regularise the data but in this case, where the gamma was fixed. We did some research and we saw many illustrations, and ran multiple tests with different gamma values to arrive at the best value, which was 8.0.

The data was divided in training and test datasets with a ratio of 3:1 respectively before running the script using the CVX tool.

3.2 Consideration about the model

Depending on the input, the program creates a model to forecast future values. The validity of the model is calculated based on two types of errors :

- MAPE (Mean Absolute Percent Error) measures the error in percentage terms.

```
%model
ii=randperm(N);
%splitting in training and test set
training=Y(ii(1:(N*0.75)),:);
test=Y(ii((N*0.75)+1:N),:);
fTraining=f(ii(1:(N*0.75)),:);
fTest=f(ii((N*0.75)+1:N),:);

gamma=8.0;

cvx_begin quiet
    variable w2( p+1 )
    minimize( norm(training*w2-fTraining) + gamma*norm(w2,1) );
cvx_end

%storing all errors of all currencies in a vector to check
%the highest and the lowest
e(z)=sum(((test*w2)-fTest).^2)/(N/2))*100;
```

Fig. 4. Error Calculation

$$\left(\frac{1}{N} \sum \left(\frac{|Actual - Forecast|}{|Actual|} \right) \right) * 100$$

- MAP (Mean Absolute Deviation) measures the size of the error in units)

$$\frac{1}{N} \sum (|Actual - Forecast|$$

Both these values, when calculated resulted in producing very low errors (lesser than 1). Which means that our model performed perfectly (with minimal error) when trying to predict the unknown data (test data) using the other currency exchange rate values. The reasons of that may be numerous:

- The first reason could be that a lot of currencies were indexed linked with USD and GBP(Eg: the saudi arabia currency)
- Influence of major economic trends and activities (i.e. financial crises, economic growth)

Justification of the model:

In the beginning there were some doubts about using regression technique or an ANN(Artificial Neural network)[7]. After considering both merits and de-merits, in the end we choose to use linear regression for two reasons :

- The main reason was that the ANN is a black box method and it would be very difficult to find any relationship between variables, on the contrary these relationships can be easily shown using regression models.
- The method of least squared regression converge much faster than a neural network, and this means saving resources and time

3.3 Low Frequency data

The low frequency data included population, employment rate, trade network, GDP, inflation and interest rates from the G-20 countries. Every files downloaded was made in this way (see Fig 5):

The size of each factor of G-20 dataset table is 265 rows time 61 columns and each row is a country's factor data from 1960 to 2016. Firstly, we used python code to select the data of G-20 and

| Country Code | ARI | AUS | BRA | CAN | CHN | DEU | EUU | FRA | GBR | IDN | IND | ITA |
|--------------|-------------|-------------|-------------|----------|-------------|----------|----------|----------|----------|-------------|--------------|----------|
| 1991 | 15166000000 | 52678895811 | 26251000000 | 1.52E+11 | 54297000000 | 4.32E+11 | 1.88E+12 | 2.83E+11 | 2.51E+11 | 31180000000 | 27031902767 | 2.13E+11 |
| 1992 | 19319500000 | 55168803181 | 27964000000 | 1.57E+11 | 73819000000 | 4.85E+11 | 2.11E+12 | 2.98E+11 | 2.67E+11 | 34874000000 | 29656593966 | 2.32E+11 |
| 1993 | 21975400000 | 56603959409 | 34856000000 | 1.66E+11 | 98349000000 | 4.20E+11 | 1.89E+12 | 2.61E+11 | 2.55E+11 | 38222000000 | 30054048002 | 1.85E+11 |
| 1994 | 27273600000 | 66750323102 | 43495000000 | 1.83E+11 | 1.12E+11 | 4.64E+11 | 2.06E+12 | 2.81E+11 | 2.84E+11 | 43738000000 | 37872294417 | 2.05E+11 |
| 1995 | 26025117000 | 75423012738 | 63283000000 | 1.99E+11 | 1.35E+11 | 5.39E+11 | 2.51E+12 | 3.34E+11 | 3.27E+11 | 54401000000 | 48225107437 | 2.45E+11 |
| 1996 | 30212204066 | 80378140345 | 66018000000 | 2.06E+11 | 1.54E+11 | 5.53E+11 | 2.59E+12 | 3.34E+11 | 3.55E+11 | 59379000000 | 54999897568 | 2.59E+11 |
| 1997 | 37511586709 | 82747229907 | 75139000000 | 2.37E+11 | 97790000000 | 5.37E+11 | 2.59E+12 | 3.23E+11 | 3.80E+11 | 62830000000 | 58172791088 | 2.54E+11 |
| 1998 | 36787937537 | 79601560689 | 74415000000 | 2.41E+11 | 97527000000 | 5.85E+11 | 2.74E+12 | 3.46E+11 | 3.96E+11 | 44030357373 | 59367895754 | 2.68E+11 |
| 1999 | 32903076781 | 84902583708 | 62807000000 | 2.55E+11 | 1.16E+11 | 5.79E+11 | 2.84E+12 | 3.65E+11 | 4.19E+11 | 42974513486 | 62827495276 | 2.69E+11 |
| 2000 | 33968803787 | 86280804343 | 71576900702 | 2.87E+11 | 1.61E+11 | 5.95E+11 | 2.89E+12 | 3.77E+11 | 4.40E+11 | 56003631930 | 73979192329 | 2.84E+11 |
| 2001 | 27594942433 | 79477801918 | 71620648627 | 2.68E+11 | 1.80E+11 | 5.87E+11 | 2.97E+12 | 3.74E+11 | 4.39E+11 | 50548822609 | 71311160598 | 2.85E+11 |
| 2002 | 13372220000 | 89218900918 | 60778494949 | 2.71E+11 | 2.10E+11 | 5.89E+11 | 3.11E+12 | 3.91E+11 | 4.71E+11 | 52869733633 | 757414901268 | 3.02E+11 |
| 2003 | 18724510000 | 1.09E+11 | 62707203451 | 2.95E+11 | 4.10E+11 | 7.28E+11 | 3.70E+12 | 4.61E+11 | 5.29E+11 | 56846585710 | 92959121914 | 3.61E+11 |
| 2004 | 27823640000 | 1.34E+11 | 78895240218 | 3.37E+11 | 5.54E+11 | 8.58E+11 | 4.50E+12 | 5.42E+11 | 6.25E+11 | 67472192811 | 1.31E+11 | 4.23E+11 |
| 2005 | 34786990000 | 1.52E+11 | 96610925191 | 3.85E+11 | 6.49E+11 | 9.34E+11 | 4.95E+12 | 5.93E+11 | 6.86E+11 | 86268317494 | 1.82E+11 | 4.60E+11 |
| 2006 | 43281872520 | 1.70E+11 | 1.19E+11 | 4.30E+11 | 7.83E+11 | 1.08E+12 | 5.68E+12 | 6.58E+11 | 7.84E+11 | 87614055150 | 2.25E+11 | 5.38E+11 |
| 2007 | 53551778520 | 2.06E+11 | 1.58E+11 | 4.71E+11 | 9.49E+11 | 1.25E+12 | 6.46E+12 | 7.60E+11 | 8.41E+11 | 1.01E+11 | 2.76E+11 | 6.14E+11 |
| 2008 | 68242884908 | 2.48E+11 | 2.20E+11 | 5.08E+11 | 1.15E+12 | 1.41E+12 | 7.49E+12 | 8.69E+11 | 8.67E+11 | 1.36E+11 | 3.79E+11 | 6.89E+11 |

Fig. 5. Raw economic indicators data

used append function to put the G-20 data into a new array. Then, in order to clean out the data suitable for using and analysis, we used the zip function to transpose the array. Next, Due to the low frequency data were analysed from 1991 to 2016, we filtered the data and put it into the new csv file. Furthermore, we used replace function to remove the space, which may cause the Index error. Using this python script all low frequency data were filtered and we put the all cleaned data into one file (see Fig 6).

| Country Code | ARI | AUS | BRA | CAN | CHN | DEU | EUU | FRA | GBR | IDN | IND | ITA |
|--------------|-------------|-------------|-------------|----------|-------------|----------|----------|----------|----------|-------------|--------------|----------|
| 1991 | 15166000000 | 52678895811 | 26251000000 | 1.52E+11 | 54297000000 | 4.32E+11 | 1.88E+12 | 2.83E+11 | 2.51E+11 | 31180000000 | 27031902767 | 2.13E+11 |
| 1992 | 19319500000 | 55168803181 | 27964000000 | 1.57E+11 | 73819000000 | 4.85E+11 | 2.11E+12 | 2.98E+11 | 2.67E+11 | 34874000000 | 29656593966 | 2.32E+11 |
| 1993 | 21975400000 | 56603959409 | 34856000000 | 1.66E+11 | 98349000000 | 4.20E+11 | 1.89E+12 | 2.61E+11 | 2.55E+11 | 38222000000 | 30054048002 | 1.85E+11 |
| 1994 | 27273600000 | 66750323102 | 43495000000 | 1.83E+11 | 1.12E+11 | 4.64E+11 | 2.06E+12 | 2.81E+11 | 2.84E+11 | 43738000000 | 37872294417 | 2.05E+11 |
| 1995 | 26025117000 | 75423012738 | 63283000000 | 1.99E+11 | 1.35E+11 | 5.39E+11 | 2.51E+12 | 3.34E+11 | 3.27E+11 | 54401000000 | 48225107437 | 2.45E+11 |
| 1996 | 30212204066 | 80378140345 | 66018000000 | 2.06E+11 | 1.54E+11 | 5.53E+11 | 2.59E+12 | 3.34E+11 | 3.55E+11 | 59379000000 | 54999897568 | 2.59E+11 |
| 1997 | 37511586709 | 82747229907 | 75139000000 | 2.37E+11 | 97790000000 | 5.37E+11 | 2.59E+12 | 3.23E+11 | 3.80E+11 | 62830000000 | 58172791088 | 2.54E+11 |
| 1998 | 36787937537 | 79601560689 | 74415000000 | 2.41E+11 | 97527000000 | 5.85E+11 | 2.74E+12 | 3.46E+11 | 3.96E+11 | 44030357373 | 59367895754 | 2.68E+11 |
| 1999 | 32903076781 | 84902583708 | 62807000000 | 2.55E+11 | 1.16E+11 | 5.79E+11 | 2.84E+12 | 3.65E+11 | 4.19E+11 | 42974513486 | 62827495276 | 2.69E+11 |
| 2000 | 33968803787 | 86280804343 | 71576900702 | 2.87E+11 | 1.61E+11 | 5.95E+11 | 2.89E+12 | 3.77E+11 | 4.40E+11 | 56003631930 | 73979192329 | 2.84E+11 |
| 2001 | 27594942433 | 79477801918 | 71620648627 | 2.68E+11 | 1.80E+11 | 5.87E+11 | 2.97E+12 | 3.74E+11 | 4.39E+11 | 50548822609 | 71311160598 | 2.85E+11 |
| 2002 | 13372220000 | 89218900918 | 60778494949 | 2.71E+11 | 2.10E+11 | 5.89E+11 | 3.11E+12 | 3.91E+11 | 4.71E+11 | 52869733633 | 757414901268 | 3.02E+11 |
| 2003 | 18724510000 | 1.09E+11 | 62707203451 | 2.95E+11 | 4.10E+11 | 7.28E+11 | 3.70E+12 | 4.61E+11 | 5.29E+11 | 56846585710 | 92959121914 | 3.61E+11 |
| 2004 | 27823640000 | 1.34E+11 | 78895240218 | 3.37E+11 | 5.54E+11 | 8.58E+11 | 4.50E+12 | 5.42E+11 | 6.25E+11 | 67472192811 | 1.31E+11 | 4.23E+11 |
| 2005 | 34786990000 | 1.52E+11 | 96610925191 | 3.85E+11 | 6.49E+11 | 9.34E+11 | 4.95E+12 | 5.93E+11 | 6.86E+11 | 86268317494 | 1.82E+11 | 4.60E+11 |
| 2006 | 43281872520 | 1.70E+11 | 1.19E+11 | 4.30E+11 | 7.83E+11 | 1.08E+12 | 5.68E+12 | 6.58E+11 | 7.84E+11 | 87614055150 | 2.25E+11 | 5.38E+11 |
| 2007 | 53551778520 | 2.06E+11 | 1.58E+11 | 4.71E+11 | 9.49E+11 | 1.25E+12 | 6.46E+12 | 7.60E+11 | 8.41E+11 | 1.01E+11 | 2.76E+11 | 6.14E+11 |
| 2008 | 68242884908 | 2.48E+11 | 2.20E+11 | 5.08E+11 | 1.15E+12 | 1.41E+12 | 7.49E+12 | 8.69E+11 | 8.67E+11 | 1.36E+11 | 3.79E+11 | 6.89E+11 |

Fig. 6. Cleaned Economic data

3.4 Web Application Design

We created a simple web application using the python flask framework.

The final analysis and displaying of results was implemented in JavaScript and D3

When viewing the low frequency economic indicator data we also loaded the high frequency (weekly samples) exchange rate data, this is then aggregated down to yearly samples for analysis with the economic indicators.

The web app has the following two functions:

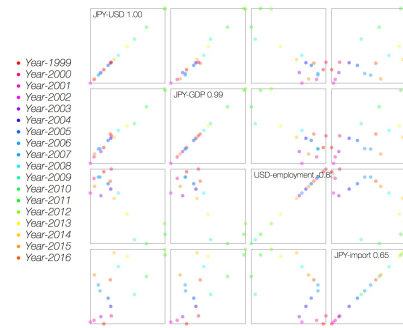


Fig. 7. Covariance Scatter plot of Economic Indicators with respect to a specific exchange rate

- (1) Finding Strongest Correlation : When this button is pressed the web app iterates through all the countries currency exchange rates between each other and calculates the correlation coefficient between the exchange rates and their countries respective economic indicator. This value is then sorted in order of strength (closest to either 1.0 or -1.0) before being displayed. This allows us to find what are the strongest and weakest correlations. The app then also aggregates the absolute value of each correlation coefficient (to prevent positive and negative correlations cancelling each other out and allowing us to find the average strength of the correlation) with respect to the type of economic indicator. Which allows us to find the type of economic indicator which is tightly correlated with the exchange rates.
- (2) Go! : This button takes a selected exchange rate and then generates a scatter plot for each of the economic indicators that are most correlated with those exchange rate. Which is then present in a 4 x 4 grid (as seen in Fig 7), the selected exchange rate is in the top right corner, and each economic indicator is plotted on the diagonal, labelled with their respective correlation coefficient related to the exchange rate. The user is then able to see a covariance scatter plot for each variable by tracing the horizontal and vertical intersection.

4 RESULTS

4.1 Observation

As the data we have used is not very huge, we chose to find the strongest correlation between the exchange rates and the indicators which led to this correlation. We have used Pearson's correlation coefficient as a mathematical approach, since the data was not suitable for any sort of mathematical regression or any other kind of predictive analysis. The result that we have found can be summarised below showing the ranking of the indicators in terms of those that are most strongly correlate to the exchange rate data which are annually averaged over all the exchange rates.

- Population = 0.62
- GDP = 0.60
- Imports = 0.59

- Exports = 0.58
- Foreign Trade = 0.50
- Inflation GDP Deflator = 0.40
- Employment = 0.40
- Interest Rate = 0.40
- Inflation Consumer Prices = 0.38

However in some particular cases we have seen a very strong correlation which are worth mentioning

Table 1. Strongest correlation and related indicator

| Currency Code | Indicator | Value |
|---------------|----------------|-------|
| JPY-USD | JPY GDP | 0.99 |
| JPY-SAR | JPY GDP | 0.99 |
| SAR-JPY | JPY GDP | -0.99 |
| MXN-CNY | MXN Population | -0.98 |
| AUD-TRY | TRY Population | 0.98 |
| CAD-TRY | TRY Population | 0.98 |
| SAR-CNY | SAR imports | -0.98 |
| AUD-IDR | IDR Population | 0.98 |
| AUD-TRY | AUD Population | 0.98 |
| MXN-CNY | CNY Population | -0.98 |
| RUB-CAD | CAD Population | -0.98 |
| CAD-TRY | CAD Population | 0.98 |
| CNY-SAR | SAR Imports | 0.97 |
| TRY-MXN | MXN Inflation | 0.97 |
| RUB-AUD | AUD Population | -0.97 |
| AUD-IDR | AUD Population | 0.97 |
| AUD-ZAR | AUD Population | 0.97 |
| AUD-ZAR | ZAR Population | 0.97 |

From the table above, a few things can be concluded

- GDP, Population, Imports and Inflation have been the strongest indicators in the past[4]. These factors have played the most important role in determining the exchange rates for the two currencies
- A high correlation value indicates that the corresponding indicator is the strongest factor in determining the exchange rates between those two countries[2]. For example, We find Japanese GDP has a very strong correlation value (0.99) between Japanese Yen and US Dollar and also between Japanese Yen and Saudi Riyal. This does not necessarily mean there is a strong correlation, to analysis this we must then look at the covariance scatter plot to understand the relationship and in this case we observe a strong correlation.

4.2 Analysis

4.2.1 Currency Exchange Trends. If we look at the currency exchange rates between Argentina (ARS) and USA (USD) we observe that until 2001 the currency exchange value is 1 and then we see a sudden dip in the value. What could be the reason for this dip? Well, that was because each peso was index-linked to USD at 1ARS=1USD. However, after the financial crisis of 2001, the fixed exchange rate system was abandoned. Since 2002, the exchange rate started to

fluctuate, keeping the exchange rate at between 2.90 and 3.10 pesos per US dollar at that time. This is the same case in terms of Saudi Riyal, where even today it is index linked to the USD @ 1 USD = 3.75 SAR

4.2.2 Correlation Coefficient. If we look at the currency exchange rates JPY-USD, we observe that it has a .99 correlation value against the GDP of Japan. Yes, without any arguments we can agree that the GDP of a country has a direct impact on its currency value. Where we saw that the overall influence (correlation coefficient value) of GDP to its countries currency is calculated to be .6. But in this case, we see a value of .99.

If we look at the history for the currency of Japan (YEN) we see that, following World War II the Yen lost much of its value. To stabilise the Japanese economy the exchange rate of the yen was fixed at 360Yen per 1USD as part of the Bretton Woods system. When that system was abandoned in 1971, the Yen became undervalued and was allowed to float. The Yen had appreciated to a peak of 271Yen per 1 USD in 1973, then underwent periods of depreciation and appreciation due to the 1973 oil crisis, arriving at a value of 227Yen per 1 USD by 1980. Since 1973, the Japanese government has maintained a policy of currency intervention, and the yen is therefore under a "dirty float" regime. This intervention continues until today and that is the reason we see such a tight correlation between the currency exchange of JPY-USD against the GDP of Japan.

But the question is does it stand good for other cases? Well, yes it does? For instance, if we look at the currency of China(CNY) and USA(USD) we see that both these two countries have a relatively huge GDP values, which does have an impact on their respective currencies. But when we look at the exchange rate which is a copula of both these countries currency. There is a possibility that the influence of the GDP values on the exchange rate tend to be slightly lower (Correlation coefficient of China's GDP and US GDP against the exchange rate of CNY-USD is .92 and .87 respectively) but still have a significant impact the exchange rates.

5 CONCLUSION

5.1 Economic Indicator

We have built a tool which allows for the analysis of 254 different exchange rates and there relationship to 9 different economic indicators. The tool enables you to quickly see for a given exchange rate what are the most correlated economic indicators and the visually validate the nature of that correlation. It also allows you observe the relationship between the economic indicators them selves. This can then help guide further investigation into what underlying world events helped drive fluctuations in the exchange rate.

5.2 Exchange Rate Interrelationships

6 LIMITATIONS AND FUTURE WORK

- Limitation : Our data pipeline is not as automated as we would have liked, there is still a significant amount of manual effort required to update our web application with fresh data.
- Improvement : The cleaning we performed was very systematic and with more time could be fully automated in python such that the data is downloaded, clipped and

cleaned and exported for use in the web application. This would allow the exchange data being used to be refreshed on a daily basis.

- Limitation : The world exchange rates market is very complex and its structure has fluctuated over time. Some of the currencies were index linked to USD prior to 1999, for which we had a limited availability of currency data. We need to be clear as to the purpose of our tool, our primary aim was to help people understand the relationship between exchange rates and economic indicators and so this is why we deliberately limited our time range so that the user is able to focus on these relationships alone.
- Limitation : The dashboard we have created is very limited only focusing on covariance of economic indicators and relative relationships between exchange rates. This is useful as a tool to help direct further investigation however there is a lot more we could show to support this investigation.
 - Improvement : we could expand our dashboard to display all the respective trend data for the correlations that were found.
 - Improvement : In the exchange rate analysis we could look to automatically find points of inflection in the exchange rate trend data that correlation with the other related exchange rates that were found. This will identify points in time where something significant might have happened. Using these time point we could then look to automatically search for financial news related to the countries at this point in time. This would significantly reduce the effort required to explore the economic history of a specific exchange rate.

REFERENCES

- [1] Angela Abbate and Massimiliano Marcellino. 2018. Point, interval and density forecasts of exchange rates with time varying parameter models. *Journal of the Royal Statistical Society: Series A (Statistics in Society)* 181, 1 (2018), 155–179.
- [2] Ariel Burstein, Martin Eichenbaum, and Sergio Rebelo. 2005. Large devaluations and the real exchange rate. *Journal of political Economy* 113, 4 (2005), 742–784.
- [3] Ryan D Edwards and Carl N Mason. 2014. Spinning the wheels and rolling the dice: Life-cycle risks and benefits of bicycle commuting in the US. *Preventive medicine* 64 (2014), 8–13.
- [4] Sebastian Edwards. 2006. *The relationship between exchange rates and inflation targeting revisited*. Technical Report. National Bureau of Economic Research.
- [5] Michael Iacono, Kevin Krizek, and Ahmed M El-Geneidy. 2008. Access to destinations: how close is close enough? Estimating accurate distance decay functions for multiple modes and different purposes. (2008).
- [6] Nuwan Kuruwitaarachchi, MKM Peiris, CN Madawala, KMAR Perera, and VUN Perera. 2018. Design and Development of an Algorithm to Predict Fluctuations of Currency Rates. *World Academy of Science, Engineering and Technology, International Journal of Computer and Information Engineering* 5, 2 (2018).
- [7] Mark T Leung, An-Sing Chen, and Hazem Daouk. 2000. Forecasting exchange rates using general regression neural networks. *Computers & Operations Research* 27, 11 (2000), 1093–1110.
- [8] Adewole Adetunji Philip, Akinwale Adio Taofiki, and Akintomide Ayo Bidemi. 2011. Artificial neural network model for forecasting foreign exchange rate. *World of Computer Science and Information Technology Journal* 1, 3 (2011), 110–118.
- [9] Barbara Rossi. 2013. Exchange rate predictability. *Journal of economic literature* 51, 4 (2013), 1063–1119.