Assessing Predictability of Stock Market Returns with Macroeconomic Variables by Country Economic Profile

2021 Winter ECO421H Final Project

Emre Asci (1004412888), emre.asci@mail.utoronto.ca

Jia Yi Ong (1003856614), jiayi.ong@mail.utoronto.ca

April 12, 2021

1 Introduction

Stock markets are one of the ways to quantify the health of economies and by examining the movements in stock market indices, governments/investors can gain vital information on the empirical link between real and financial sectors. The information from the link can shed light on the utility of some policies and provide a new path for governments to evaluate their actions, whereas investors can simply profit off of the possible results by following strong indicators. Governments could create healthier and more stable business cycles by preventing the lack of funding due to downturns on stock markets. Strong correlation between indicators and stock markets could reveal the interactions within monetary transmission mechanism and help analyze changes in indicators to optimize portfolios for the macroeconomic risk.

Numerous studies have been published to identify correlations between macroeconomic variables and their relative impact on stock market indices, and lately machine learning techniques have been utilized more commonly in these researches. Chen (2008) investigates various macroeconomic variables in predicting recessions by using the Markov-Switching model, and concludes that yield curve spreads and inflation rates are the most useful predictors of recessions for SP500. She also deduces that predicting bear/bull markets is easier than predicting stock returns which may involve a highly non-linear relationship as Enke and Zhong's (2017) research would support. Enke and Zhong (2017) employ logistic regression and artificial neural networks (ANNs) to predicting the direction of the SPY for the next day. Their results show that the ANNs give significantly higher classification accuracy than logistic regression.

Instead of analyzing the stock market movements with macroeconomic variables (MV), Nti, Adekoya, Weyori, (2019) acknowledges the possible relationship between some macroeconomic variables and different sector stocks. Their study applies the random forest based feature selection of macroeconomic variables for the Ghana Stock Exchange and identifies different MVs for fluctuations on different sectors. Also, Pilinkus (2010) investigates causality between some macroeconomic indicators and stock market indices in the Baltic States and concludes the relationships can depend on the different monetary and fiscal policies. Likewise, Dovolil (2016) analyzes the correlation between economic indicators with Spearman's correlation for predicting the SP500 stock index, and finds 4 sub-economic indicators having the highest correlation

However, predicting stock market growth has not been an easy research topic as it does not follow any specific model or distribution. Stock markets are affected by many related and non-related factors, and this depends between countries. Even though we see Chen (2008) and Enke and Zhong (2017) apply techniques to analyze significant macro variables and predict somewhat return correlation, their research does not characterize the importance of distinct variables for different economic profiles. Haider and Tariq (2018) examines the performance of different macroeconomic indicators on stock market indices for different countries but their method does not employ any machine learning methods, whereas Ptak and Matuszyk (2019) uses clustering and logistic regression to select high importance macro predictors to analyze bankruptcy risk but not stock market indices. We believe machine learning algorithms can be effective to analyze specific indicators for predicting bull/bear markets and guide countries to implement policies regarding to their economic situation. We understand that if we find certain predictors are only important for countries of a certain economic profile, the different drivers of financial growth can be identified and future models for policies/investments can be shaped on significant predictors. To explore the trend, we utilize different machine learning algorithms and investigate the research question of *how well macroeconomic variables can predict the direction of the stock market index returns and how the best predictors differ by an economic profile of a country.*

In contrast to the literature discussed, which focused on predicting returns on the country-level, we seek to identify predictability on a cross-country level. That is, we attempt to find consistency in the best predictors of returns for countries of similar economic profiles. We construct economic profiles by fitting principal components on high-dimensional data (including HDI, GNI, and labor statistics) for countries around the world and clustering them. By analyzing the impact of indicators on the direction of stock market indices for different economic profiles, we expect to find consistency in the best predictors of countries in the same cluster. We think by analyzing most important predictors for different clusters, some predictability in the stock market returns can be achieved which could lead to identifying different drivers of financial growth. We believe policymakers could implement respective actions to foster the economic health of their own countries while investors can adjust their portfolio risk.

2 Data and Methodology

We adopted a three-stage approach to answer our research question. In the first stage, we clustered 78 countries around the world by applying the K-Means algorithm on principal components constructed from 19 selected indexes, obtained from the 2020 Human Development Data of the United Nations Development Programme[1]. The 19 selected indexes, which cover human development, equality, inflation, tax, labour and employment, capital flow, and technology (as of 2019), are intended to characterize the economic profiles of the countries. We employed PC clustering on the standardized data set to achieve better profiling based on a summary of the indexes, rather than giving each index an equal weight. Thus, we can treat countries belonging to the same PC cluster as having similar economic profiles in our analysis. The data set contained

[1] for full list of countries and indexes used in this stage, view appendix Section 1

189 countries, but only 143 countries had data for all 19 variables. We then used the top 55% of the 143 countries (resulting in 78 countries) based on their Gross National Income per capita to ensure sufficient financial data can be found. Once the optimal clustering has been chosen and estimated, we proceeded to analyze the countries closest to each cluster centroid, treating them as representative of every country in their respective clusters. The top 6 countries closest to each centroid were considered, and amongst them 2-3 countries were chosen as representatives based on data availability. The analyses made on the representative countries will be generalized to all countries in the respective clusters.

The second stage involves determining the most important predictors of stock market returns for the countries closest to their respective centroids (the representative countries of each economic profile). We drew data from multiple data sets of the International Financial Statistics database (International Monetary Fund, 2021) for candidates of the most important predictors of returns. In addition to a month control[2], the candidate variables are monthly indicators belonging to 7 different categories: 1) exchange rates; 2) interest rates; 3) price indexes; 4) trade; 5) liquidity; 6) GDP components; 7) balance of payments. The date ranges from 2014 January to 2020 December, inclusive, with a total of $7 \times 12 = 84$ data points for each variable. The countries in the database do not have the same variables available under each category. Thus, to enable cross-country comparison of the most important predictor of returns, we ensured that we collected at least one series in each category for each representative country, supplementing missing data from FRED Economic Data when needed[3]. In other words, we used the categories as the basis of comparison. In choosing the variables, we also ensured no obvious collinearity in the predictors, e.g. excluding Total expenditure when using Consumption and Government expenditure. The target variable is derived from daily major stock indexes of each representative country[4]. The target is a monthly binary indicator for the direction of change from the starting price to the ending price of a month. To predict one-month ahead, we shifted the target variable forward, with the idea that information of predictors collected at end of month t will predict the direction of change in starting price to ending price in month $t+1$. Finally, We employed the logistic LASSO classifier for variable selection.

For the third and final stage, we evaluate the predictability of stock return direction using machine learning algorithms. Using the same data set in stage 2, for each representative country, we fit five types of classification models using only the selected variables to predict the country's stock return direction. The types of classification models are: 1) Logistic, 2) Random Forest, 3) K-Nearest Neighbours, 4) Naive Bayes, 5) Multilayer Perceptron. For types 3 and 5, we fit multiple models of each type while varying key parameters such as K neighbours and number of

[2] an integer from 1-12 indicating the month of the data point

[3] for complete list of variables and their labels used for each country, view appendix Section 3

[4] for list of which stock indexes are used, view appendix Section 2

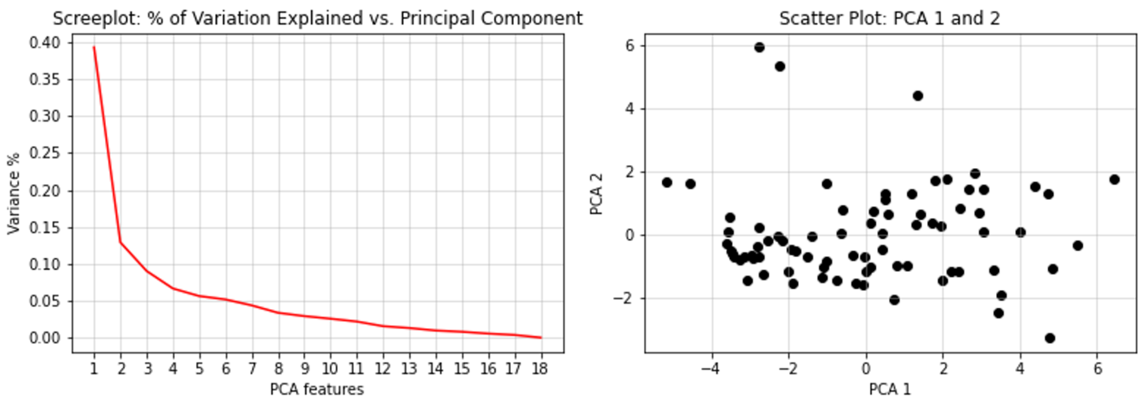
layers[5]. To evaluate predictability, we compute training and validation misclassification rates and Area-Under-Curve scores for each Country-Model combination (a total of 180 combinations). The best model (of 20) for each country is then chosen by balancing (a) minimum validation misclassification rate and (b) maximum AUC score. The exact criteria is: if a model has a higher (a) by an absolute value of 0.05, then it is strictly preferred, and if (a) is greater but by less than 0.05, then the model with higher absolute $(b) - 0.5$ is preferred.

We coded a walk-forward cross-validation function to compute the training and validation misclassification rates for each Country-Model combination. In following the 70-30 split rule, we chose the initial training sample to be the first $0.7 \times 84 \approx 59$ months in our data set. The trained model is validated on the next-month data point, and any misclassification is recorded. This process is iterated until the end of the 84-point series is reached. The total number of misclassification is summed and divided by the number of steps forward to obtain the validation error rate. The training error is simply the average of training errors across every step. The AUC score for each Country-Model is computed once using a 70-30 split of the whole sample.

3 Results

3.1 Stage 1: PCA / K-Means

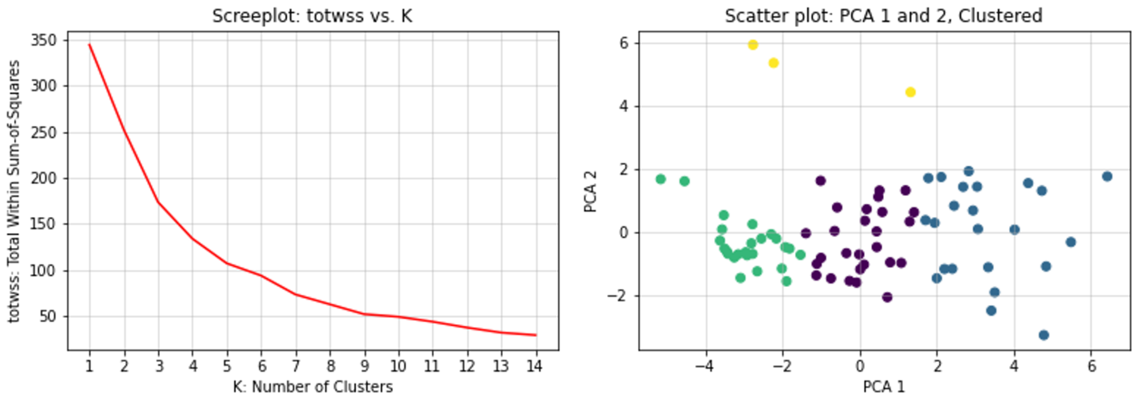
Figure 1: Choosing the number of Principal Components



Looking at the left panel of Figure 1 above, an obvious 'elbow' in the curve of the screeplot occurs at the second principal component (PC). Thus, we have chosen to use two principal components to characterize economic profiles. These two PCs account for approximately $(0.4 + 0.13) \times 100\% = 53\%$ of the variation in the data with the 19 indexes. Although the variation accounted is not very high, we have decided to use 2 PCs as a parsimonious approach and for ease of visual-

[5]for detailed model specifications, view appendix Section 4

Figure 2: Choosing the number of K-Means clusters



ization. The right panel of Figure 1 shows the scatter plot of the resulting ‘economic profiles‘ of each country.

The left panel of Figure 2 is the screeplot for clustering the ‘economic profiles’. This screeplot has no obvious ‘elbow’, but we chose four clusters to balance complexity and data requirements (2-3 new data sets of nearest countries are required for each new centroid).

Table 1: Top 6 countries closest to each cluster centroid

Cluster	Countries (Euclidean distance)
1	Canada (0.13), Germany (0.26), U.S.A. (0.35), U.K. (0.36), Ireland: (0.39), Israel (0.42)
2	Slovakia (0.45), Barbados (0.46), Belarus (0.49), Portugal (0.56), Russia (0.63), Argentina (0.75)
3	Tunisia (0.24), Mongolia (0.71), China (0.73), South Africa (1.16), Armenia (1.17), Mexico (1.38)
4	Kuwait (1.01), Qatar (1.69), Maldives (2.67)

Table 2: Countries selected as representative of each cluster

Cluster	Countries
1	Canada (CAN), Germany (GER), U.S.A. (US)
2	Portugal (POR), Russia (RUS), Argentina (ARG)
3	China (CHN), South Africa (SAFR)
4	Qatar (QAT)

Of the top 6 countries seen in Table 1, the selected representative countries are listed in Table 2. The criterion is closeness to centroid and data availability. There are only three countries in cluster 4, thus, only one country is made representative.

3.2 Stage 2: Variable Selection

Figure 3: Selected variables by Logistic LASSO per representative country

	Shrinkage	Top 10
CAN	19 -> 17	C2, mon, H2, F3, F2, E1, E3, G4, G3, H5
GER	29 -> 18	A6, C2, C1, F2, E3, E1, F4, F3, H2, F1
US	25 -> 18	C1, mon, A6, F2, H2, F4, E1, E3, F3, F1
ARG	26 -> 24	B5, B7, mon, B9, H2, B10, C2, B3, A2, B1
POR	28 -> 21	B19, B17, mon, A6, C2, C1, H2, F2, F3, E1
RUS	23 -> 20	mon, A6, C2, B3, C1, H2, F3, F2, E3, H4
CHN	18 -> 15	A6, C2, mon, H2, F2, F3, E2, E1, H4, H3
SAFR	26 -> 18	B11, C2, A6, mon, F2, F3, E1, E3, H1, H4
QAT	20 -> 19	B9, B5, B7, B1, B3, mon, F2, F3, H2, E1

The variables under 'Top 10' are encoded. The letters are the categories described in the Methodology section[6]. The 'Shrinkage' column of Figure 3 above shows the change in number of variables, from the number of candidate variables to the number of variables after shrinkage.

Looking at the pattern of categories present in the top 10 predictors for cluster 2 (ARG, POR, RUS), we see that the top 4 predictors consistently contain variables in the B category, the category of interest rates related variables[7]. This suggests that the major stock indexes of countries of the economic profile of cluster 2 are highly predictable by interest rates related variables. Similarly, since Qatar is representative of cluster 4, this may be evidence that the stock price of the other two countries with similar economic profile as Qatar may be highly predictable with interest rate related variables. Other than these observations, there are no other clear patterns of a certain cluster having unique important predictors. These results are not strong enough to robustly generalize to *every* country in each clusters, but may be sufficient as additional evidence to inform governments or stakeholders of the key variables to watch for when analyzing returns for those particular countries.

3.3 Stage 3: Evaluation of Predictability

In Table 3 below, Err_T, Err_V and AUC are the training and validation misclassification rates and AUC score, respectively. 'NN(d)' represents a Multilayer Perceptron with depth (number of layers) equals to d and 'KNN(k)' represents a K-Nearest Neighbour classifier with K equals to k.

[6] For full list of encodings and descriptions, view label dictionary in appendix Section 5
[7] For full list of encodings and descriptions, view label dictionary in appendix Section 5

Table 3: The best classifier for each country

Countries	Model	Err_T	Err_V	AUC
CAN	NN(d=2)	0.4733	0.44	0.6309
GER	NN(d=6)	0.4838	0.36	0.5000
USA	Logistic	0.3194	0.28	0.5242
POR	NN(d=3)	0.3411	0.36	0.4583
RUS	KNN(k=8)	0.3305	0.48	0.3954
ARG	NN(d=5)	0.4927	0.28	0.5000
CHN	Logistic	0.3824	0.36	0.5500
SAFR	KNN(k=6)	0.32897	0.36	0.5000
QAT	KNN(k=7)	0.3317	0.24	0.5303

The first observation is that of the five trialed classifiers, Naive Bayes and Random Forest are not the best classifier for any of the countries. Secondly, it is clear that predictive error as reflected by validation misclassification rate is high (moderate at best) for many countries. AUC scores are also unsatisfactory, with many near 0.5 (weak to no discriminatory capacity).

4 Conclusion

We found no clear consistency in the top predictors of major stock index returns for countries in the same clusters. This provides some evidence that the importance of predictors is country-specific and not dependent on economic profiles that are characterized by a summary of general indexes. This also means governments/investors belonging to one country may only obtain limited insights into the process of their home country's index by extrapolating from another country's index movement, even when both countries have similar economic profiles. Verification of this conclusion warrants additional research that uses a more comprehensive set of indexes to construct economic profiles, or an entirely different method for grouping countries. A simple extension may also replicate our variable selection segment with *all* countries in each cluster instead of generalizing form a few, so that the true distribution of best predictors can be found in each cluster. The high misclassification rates of the models across countries may be indicative of several factors, including: stock processes are inherently high-frequency, so low predictability of returns on a monthly basis; strong lag in effect of monthly variables on returns beyond the one-month ahead return predicted in our analysis; more than 85 months are needed to robustly train the models. Extensions of this project may consider predicting returns computed longer into the future to capture any lagged effects, or to use higher-frequency data sets. Another conclusion is that the macroeconomic variables used in our analysis do not predict stock returns well. This implies that stakeholders should use a different set of predictors or add more robust predictors.

5 Appendix

5.1 Section 1

78 countries and 19 variables used in K-Means clustering:

Countries: Qatar, Luxembourg, Switzerland, Ireland, Norway, United States, Hong Kong, China (SAR), Denmark, Kuwait, Netherlands, Austria, Germany, Iceland, Sweden, Belgium, Canada, Finland, Australia, Saudi Arabia, France, United Kingdom, Korea (Republic of), Japan, Italy, Spain, New Zealand, Israel, Malta, Cyprus, Czechia, Slovenia, Estonia, Lithuania, Portugal, Slovakia, Poland, Hungary, Latvia, Greece, Panama, Romania, Croatia, Turkey, Malaysia, Russian Federation, Mauritius, Bulgaria, Chile, Kazakhstan, Montenegro, Argentina, Uruguay, Mexico, Belarus, Costa Rica, Thailand, Maldives, Serbia, Botswana, China, Barbados, Bosnia and Herzegovina, Brazil, Colombia, Albania, Armenia, Moldova (Republic of), Ukraine, Sri Lanka, Peru, Paraguay, South Africa, Egypt, Indonesia, Mongolia, Iraq, Bhutan, Tunisia

Variables (2019): Human Development Index, Life Expectancy at Birth, Expected Years of Schooling, Mean Years of Schooling, Gross National Income per capita, Female Labour Participation Rate, Male Labour Participation Rate, Urban Life Percentage, Employment in Agriculture Percentage, Employment in Service Percentage, Exports and imports (% of GDP), Foreign Direct Investment Net Inflows (% of GDP), Remittances-inflows (% of GDP) , Net Migration Rate (per 1k people), Immigrant Percentage, International Inbound Tourists (thousands) , Internet Users Percentage, Mobile Phone Subscriptions

5.2 Section 2

Table 4: Countries and their respective stock indexes used in prediction

Countries	Stock Indexes
CAN	S&P/TSX 60 Index (SPTSE)
GER	DAX Performance Index (DAXI)
USA	S&P 500 Index (SPX)
POR	Portuguese Stock Index (PSI20)
RUS	Russia Trading System Index (RTSI)
ARG	Argentina Stock Market (MERVAL)
CHN	Shangai Stock Exchange Composite Index (SHCOMP)
SAFR	FTSE/JSE All Share Index (JALSH)
QAT	Qatar Exchange Index (QE)

5.3 Section 3

The table below shows which variables are the initial candidate variables (before LASSO) for each country. For description of encoded variables, view appendix Section 5.

Table 5: Countries and their variables used in prediction

Countries	Variables
CAN	A1 A2 A3 A4 A5 A6 B1 C2 D1 E1 E3 F1 F2 F3 F4
	G1 G2 G3 G4 G5 H1 H2 H3 H4 H5
GER	A5 A6 B12 B13 B14 B15 B16 B17 B18 B19 B20 B21 C1 C2 E1
	E3 F1 F2 F3 F4 G1 G2 G3 G4 G5 H1 H2 H3 H4 H5
USA	A1 A2 A5 A6 B1 B2 B3 B4 B9 B11 C1 C2 E1 E3 F1
	F2 F3 F4 G1 G2 G3 G4 G5 H1 H2 H3 H4 H5
POR	A5 A6 B12 B13 B14 B15 B16 B17 B18 B19 B20 C1 C2 E1 E3
	F1 F2 F3 F4 G1 G2 G3 G4 G5 H1 H2 H3 H4 H5
RUS	A1 A2 A3 A4 A5 A6 B1 B3 B7 B9 C1 C2 E1 E3 F1
	F2 F3 F4 G1 G2 G3 G4 G5 H1 H2 H3 H4 H5
ARG	A1 A2 A3 A4 B1 B3 B5 B6 B7 B8 B9 B10 C2 E1 E3
	F1F2 F3 F4 G1 G2 G3 G4 G5 H1 H2 H3 H4 H5
CHN	A1 A2 A3 A4 A5 A6 B2 B7 B9 C2 E1 E2 F1 F2 F3
	F4 G1 H1 H2 H3 H4 H5
SAFR	A1 A2 A3 A4 A5 A6 B1 B3 B4 B5 B7 B9 B11 C1
	C2 E1 E3 F1 F2 F3 F4 G1 G2 G3 G4 G5 H1 H2 H3 H4 H5
QAT	A1 A2 A3 A4 A5 B1 B3 B5 B7 B9 C2 E1 E3 F1 F2
	F3 F4 G1 G3 H1 H2 H3 H4 H5

5.4 Section 4

Detailed specifications of classifiers used in Stage 3

Logistic: L2-norm penalty, with intercept

Random Forest: max depth=10, min samples in leaf=2

KNN(k): uniform weights, Euclidean distance, k neighbours

Naive Bayes: -

NN(d): Multilayer Perceptron, activation function = ReLU, size of each layer = 100 (constant), number of layers = d

5.5 Section 5

Figure 4: Descriptions of variable encodings

Key	Value
mon	Integer 1-12 for month
A2	National Currency per SDR, Period Average
A6	Real Effective Exchange Rate, based on Consumer Price Index
B1	Central Bank Policy Rate
B2	Discount Rate
B3	Money Market Rate
B4	Treasury Bill Rate
B5	Savings Rate
B6	Savings Rate, Foreign Currency
B7	Deposit Rate
B8	Deposit Rate, Foreign Currency
B9	Lending Rate
B10	Lending Rate, Foreign Currency
B11	Government Bonds
B12	Harmonized Euro Area Rates, Outstanding Amounts, Deposits, Households, Agreed Maturity, Up to 2 Years
B13	Harmonized Euro Area Rates, New Business, Deposits, Households, Agreed Maturity, Up to 1 Year
B14	Harmonized Euro Area Rates, Outstanding Amounts, Deposits, Non-Financial Corporations, Agreed Maturity, Up to 2 Years
B15	Harmonized Euro Area Rates, New Business, Deposits, Non-financial Corporations, Agreed Maturity, Up to 1 Year
B16	Harmonized Euro Area Rates, Loans, Households, Consumer Credit and Other, Up to 1 Year
B17	Harmonized Euro Area Rates, New Business, Loans, Households, Consumption, Floating Rate and up to 1 Year
B18	Harmonized Euro Area Rates, Loans, Households, House Purchase, Over 5 Years
B19	Harmonized Euro Area Rates, New Business, Loans, Households, House Purchase, Over 5 Years
B20	Harmonized Euro Area Rates, Loans, Non-Financial Corporations, Up to 1 Year
B21	Harmonized Euro Area Rates, New Business, Loans, Non-financial Corporations, Other Than Bank Overdrafts, Over EUR 1 Million, Over 3 Months and up to 1 Year
C1	Prices, Producer Price Index, All Commodities, Index
C2	Prices, Consumer Price Index, All items, Index
E1	Goods, Value of Exports, US Dollars
E2	Goods, Value of Exports, National Currency
E3	Goods, Value of Imports, CIF, US Dollars
F1	International Liquidity, Total Reserves excluding Gold, US Dollars
F2	International Reserves, Official Reserve Assets, SDRs, US Dollars
F3	International Reserves, Official Reserve Assets, IMF Reserve Position, US Dollars
F4	International Liquidity, Total Reserves excluding Gold, Foreign Exchange, US Dollars
G1	Gross Domestic Product, Nominal, Domestic Currency
G2	Household Consumption Expenditure, incl. NPISHs, Nominal, Domestic Currency
G3	Government Consumption Expenditure, Nominal, Domestic Currency
G4	Gross Fixed Capital Formation, Nominal, Domestic Currency
G5	Change in Inventories, Nominal, Domestic Currency
G6	Exports of Goods and Services, Nominal, Domestic Currency
G7	Imports of Goods and Services, Nominal, Domestic Currency
H1	Balance of Payments, Current Account, Goods and Services, Net, US Dollars
H2	Balance of Payments, Capital Account, Total, Debit, US Dollars
H3	Balance of Payments, Supplementary Items, Financial Account, Net (excluding exceptional financing), US Dollars
H4	Balance of Payments, Financial account, Net lending (+) / net borrowing (-) (balance from financial account), Direct investment, Net acquisition of financial assets, US Dollars
H5	Balance of Payments, Supplementary Items, Current Acct + Capital Acct + Financial Acct, US Dollars

6 References

Altinbas, H., Biskin, O. T. (2015). Selecting macroeconomic influencers on stock markets by using feature selection algorithms. Procedia Economics and Finance, 30, 22-29. doi:10.1016/s2212-5671(15)01251-4

Chen, S. (2009). Predicting the bear stock market: Macroeconomic variables as leading indicators. Journal of Banking Finance, 33(2), 211-223. doi:10.1016/j.jbankfin.2008.07.013

Dovolil, J. (2016). The use of economic indicators as a tool for predicting SP 500 stock index. ACC Journal, 22(2), 7-14. doi:10.15240/tul/004/2016-2-001

Fred Economic Data (2021). Consumer Price Index: All items: Total: Total for Argentina Retrieved from https://fred.stlouisfed.org/series/ARGCPALTT01IXNBM

Fred Economic Data (2021). Current Price Gross Domestic Product in China Retrieved from https://fred.stlouisfed.org/series/CHNGDPNQDSMEI

Haider, Z., Tariq, R. (june 2018). Macroeconomic Indicators and their Impact on Stock Market Performance: (Comparative Study between Pakistan and India). International Journal of Advance Study and Research Work, 1(3). doi:10.1787/503744818245

International Monetary Fund. (2021). International Financial Statistics (IFS): Exchange Rates incl. Effective Ex. Rates. Retrieved from https://data.imf.org/regular.aspx?key=61545850

International Monetary Fund. (2021). International Financial Statistics (IFS): External Sector. Retrieved from https://data.imf.org/regular.aspx?key=61545851

International Monetary Fund. (2021). International Financial Statistics (IFS): GDP and Components Retrieved from https://data.imf.org/regular.aspx?key=61545852

International Monetary Fund. (2021). International Financial Statistics (IFS): Interest Rates Retrieved from https://data.imf.org/regular.aspx?key=61545855

International Monetary Fund. (2021). International Financial Statistics (IFS): Prices, Production, and Labor. Retrieved from https://data.imf.org/regular.aspx?key=61545849

International Monetary Fund. (2021). International Financial Statistics (IFS): Trade of Goods, Retrieved from https://data.imf.org/regular.aspx?key=61545859

Investing.com. (2021). IRTS historical rates. (n.d.). Retrieved April, 2021, from https://ca.investing.com/indices/rtsi-historical-data

Investing.com (2021).PSI 20 historical rates. (n.d.). Retrieved April, 2021, from https://ca.investing.com/indices/psi-20-historical-data

Investing.com (2021) . QE total RETURN historical Rates (QETR). (n.d.). Retrieved April, 2021, from https://www.investing.com/indices/qe-total-return-historical-data

Investing.com (2021). SP 500 historical rates. (n.d.).Retrieved April, 2021, from https://ca.investing.com/indices/us-spx-500-historical-data

Nti, K. O., Adekoya, A., Weyori, B. (2019). Random forest based feature selection of macroeconomic variables for stock market prediction. SSRN Electronic Journal. doi:10.2139/ssrn.3446053.

Pilinkus, D. (2010). MACROECONOMIC indicators and their impact on stock market performance in the short and long run: The case of the Baltic states. Technological and Economic Development of Economy, 16(2), 291-304. doi:10.3846/tede.2010.19

Ptak-Chmielewska, A., Matuszyk, A. (2019). Macroeconomic factors in modelling the SMEs BANKRUPTCY Risk. the case of the Polish market. Econometrics, 23(3), 40-49. doi:10.15611/eada.2019.3.04

Topforeignstocks.com. (n.d.). Retrieved April, 2021, from https://topforeignstocks.com/indices/

United Nations Development Programme, Human Development Reports (2021). 2020 Human Development Data: All Tables and Dashboards. Retrieved from http://hdr.undp.org/en/content/download-data?fbclid=IwAR1wY7rH-Iy0YPwEIxi59hIw5PdtVDAIWWFlw1aEBEueeUiGWjOuFFLIZM

Wall Street Journal(2021). ALSH — FTSE/JSE all share Index historical prices - WSJ. Retrieved April, 2021, from https://www.wsj.com/market-data/quotes/index/ZA/ALSH/historical-prices

Yahoo Finance.(2021) DAX PERFORMANCE-INDEX (gdaxi) historical data. (2021, April 11). Retrieved April, 2021, from https://finance.yahoo.com/quote/5EGDAXI/history/

Yahoo Finance (2021). SP/TSX composite index (gsptse) historical data. (2021, April 11). Retrieved April, 2021, from https://ca.finance.yahoo.com/quote/5EGSPTSE/history/

Yahoo Finance (2021). SP merval historical rates. (n.d.). Retrieved April, 2021, from https://ca.investing.com/indices/merv-historical-data

Yahoo Finance (2021). SSE composite Index (000001.ss) historical data. (2021, April). Retrieved April , 2021, from https://finance.yahoo.com/quote/000001.ss/history/

Zhong, X., Enke, D. (2017). A comprehensive cluster and CLASSIFICATION MINING procedure for daily stock market Return forecasting. Neurocomputing, 267, 152-168. doi:10.1016/j.neucom.2017.06.010