Comparative Analysis of ARIMA, LSTM, and RNN Models in Time-Series Quantamental Indicator Forecasting

COMS 4995, Group 23
Project Deliverable #2

Team Members: David Dávila-García, Shrinjay Kaushik, Yuyi Tang, Ziyao Tang, Rain Wei

Columbia University in the City of New York

Outline - [Link]

Overall draft by Monday Night

Train/Val/Test Based on Time

Cleaning has been done already

Data Visualizations - include titles, axes labels, legend (if applicable)

- Target: DU05YXR_VT10
- Mean/Std for indicators - Rain
  - Boxplot, violin plot, etc
- Time Series plots - Ziyao
- Note interesting trends from plots

ML Techniques Proposed

- Prior work with economic forecasting
  - ARIMA - David
  - LSTM - Yuyi
  - Baseline OLS - Shrinjay
- Comparison between models
  - Evaluation metrics - Shrinjay

Data Cleaning & Sampling

Preprocessing

- No missing values in the provided data
- There are missing records (rows) w.r.t. dates and currencies
    - (See Supplementary Information Slide for data availability)
    - Inappropriate to generate synthetic data to fill these missing records

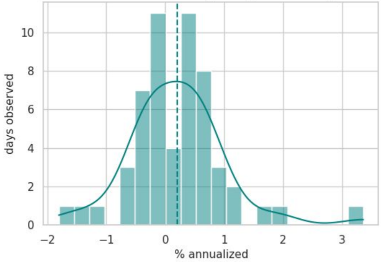
Sampling Method

- Data is split into development (90%) and test sets (10%) using TimeSeriesSplit from sci-kit learn
- K-Fold Cross Validation will be performed on the development set using TimeSeriesSplit
    - This mitigates the possibility of information leakage during model training

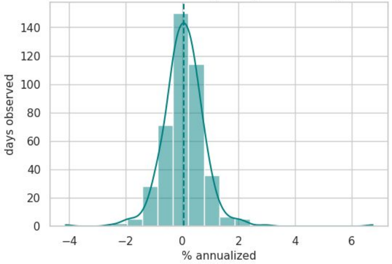
Target Variable Distribution (DU05YXR_VT10)

- DU05YXR_VT10 is the return on fixed receiver position. The plots below show the distributions of DU05YXR_VT10 for selected countries.

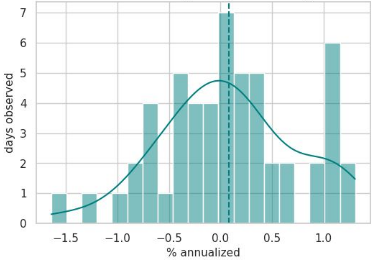
Australian duration return for 10% vol target: 5-year maturity: mean and distribution



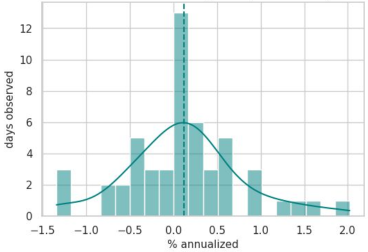
Canadian duration return for 10% vol target: 5-year maturity: mean and distribution



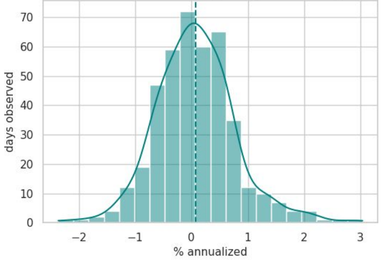
Czech duration return for 10% vol target: 5-year maturity: mean and distribution



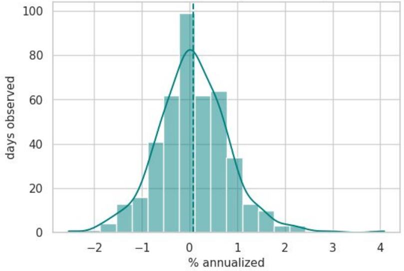
South African duration return for 10% vol target: 5-year maturity: mean and distribution



New Zealand duration return for 10% vol target: 5-year maturity: mean and distribution



US duration return for 10% vol target: 5-year maturity: mean and distribution

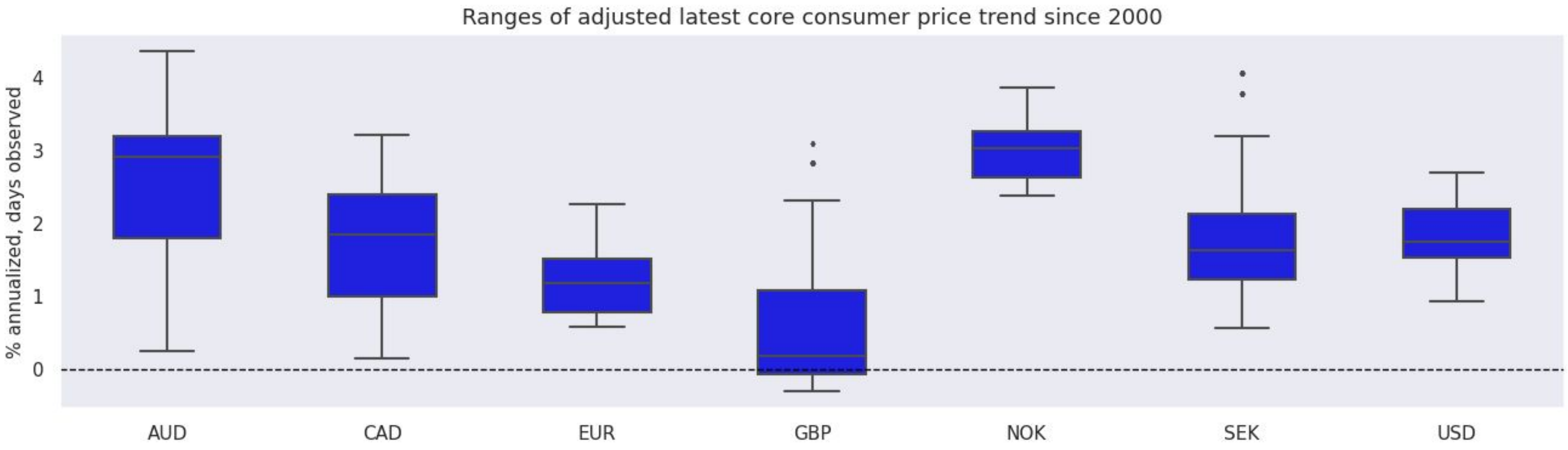


- Duration return for 10% vol target:5-year maturity(DU05YXR_VT10) versus
  FX forward return, % of notional: dominant cross(FXXR_NSA)
  FX forward return for 10% vol target: dominant cross(FXXR_VT10)

COLUMBIA UNIVERSITY
IN THE CITY OF NEW YORK

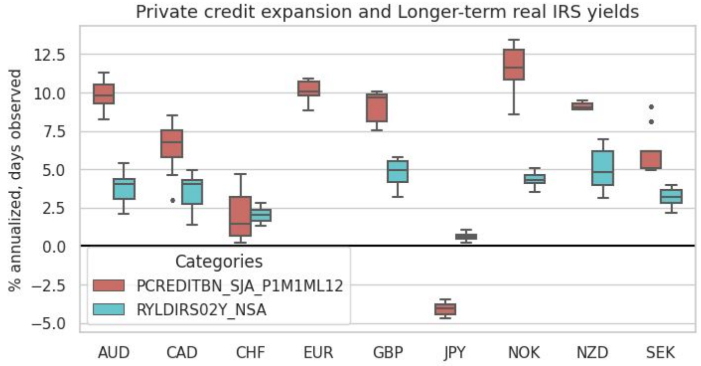
Exploratory Data Analysis of Predictors

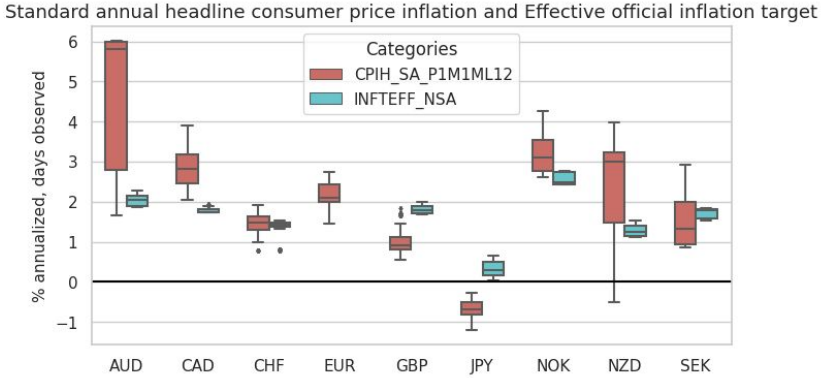
- Core CPI (Consumer Price Index)

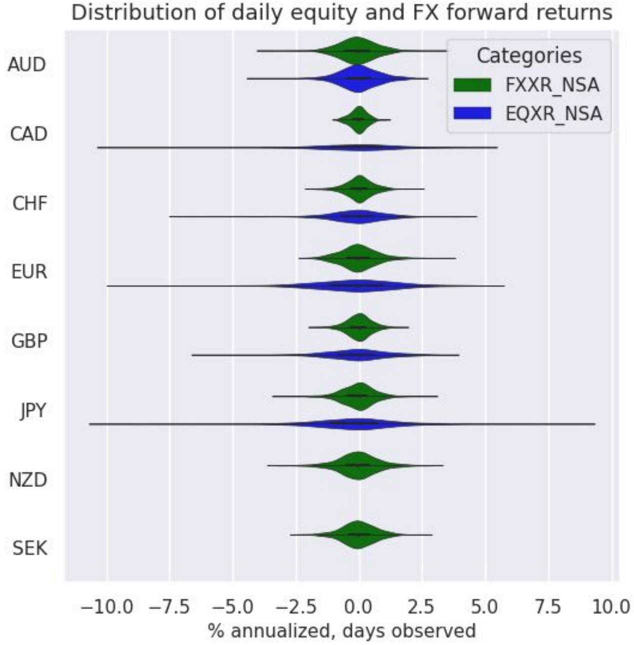


COLUMBIA UNIVERSITY
IN THE CITY OF NEW YORK

Exploratory Analysis of Predictors

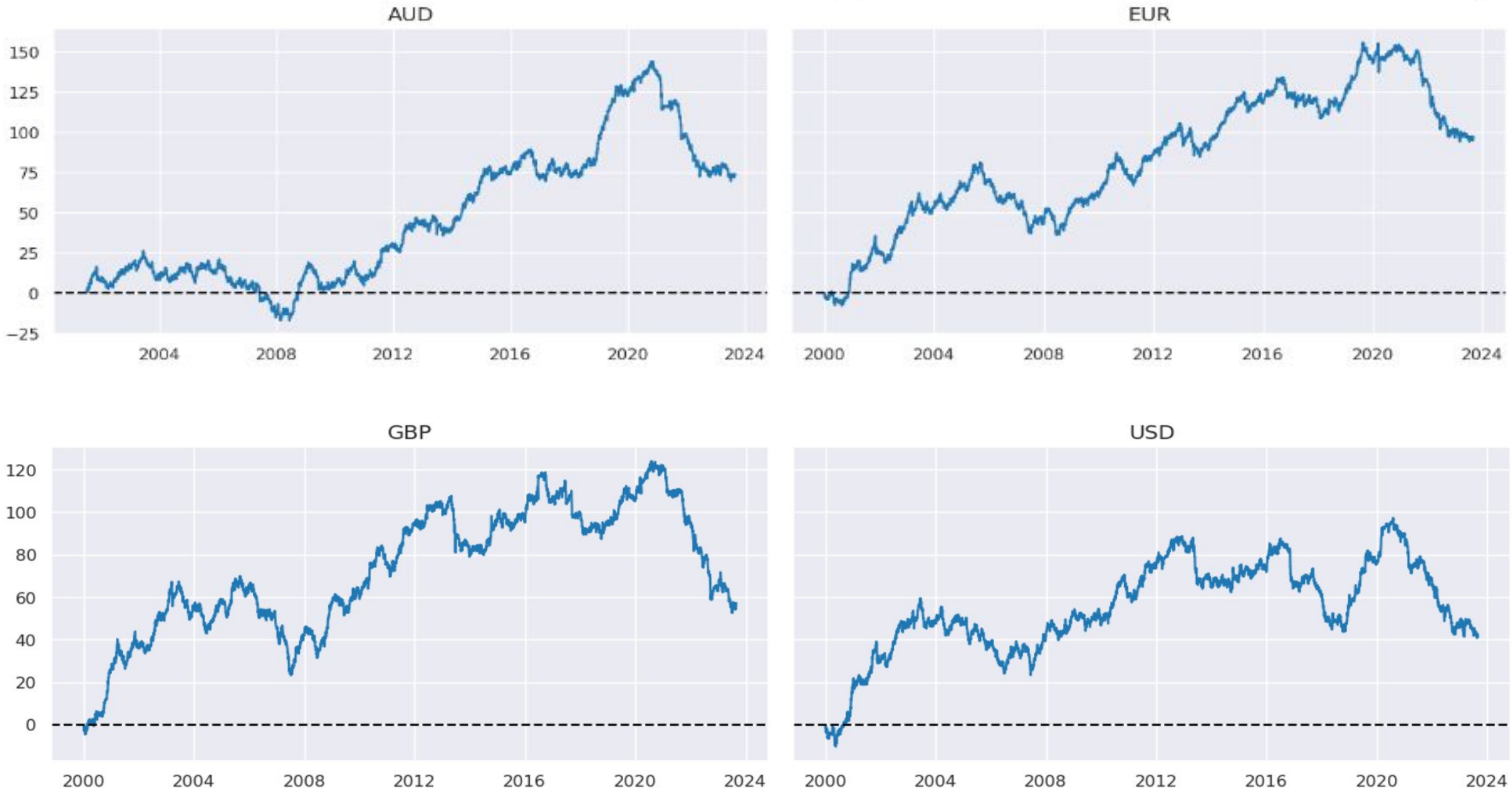






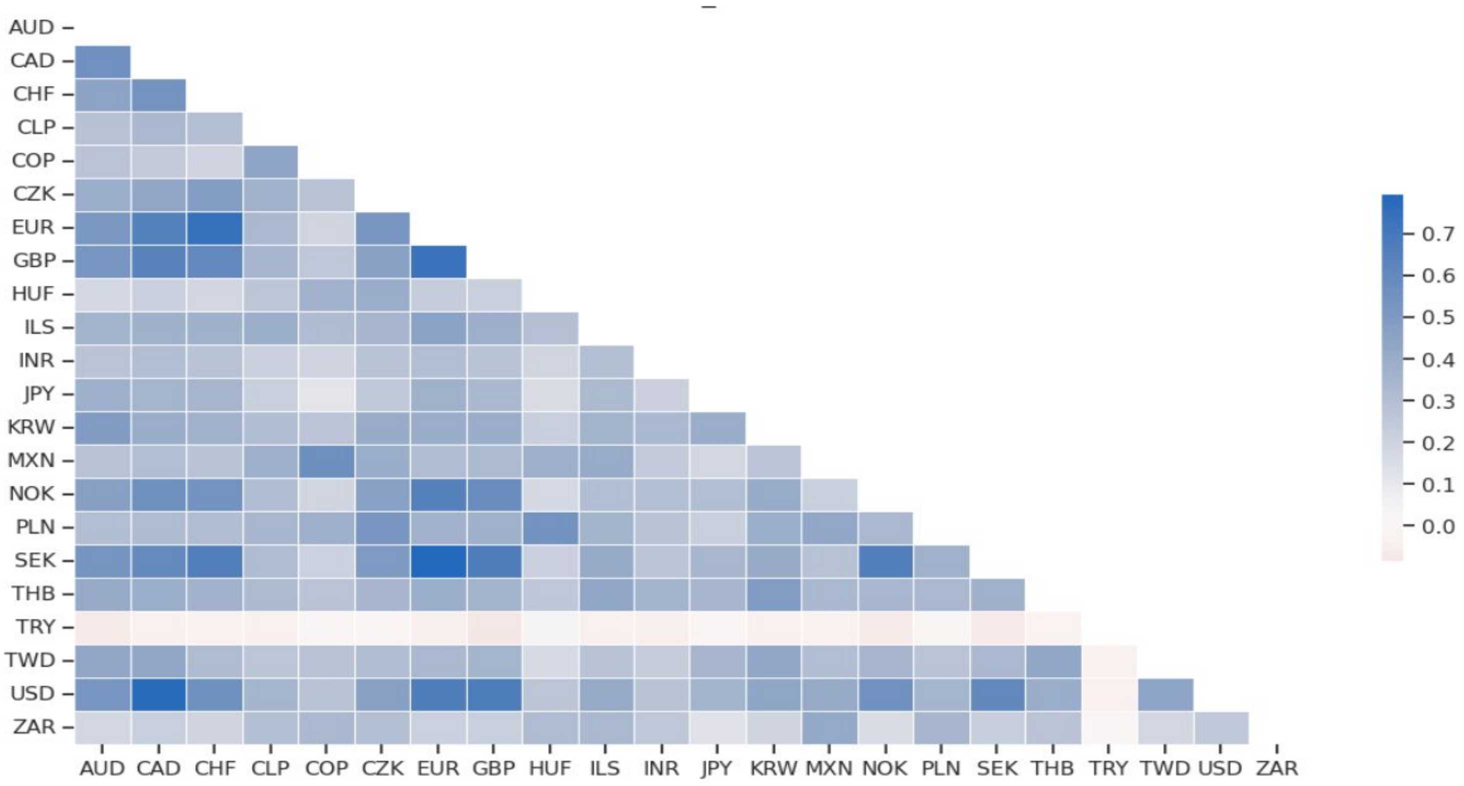
Columbia University in the City of New York

- The trend of DU05YRX_VT10 for selected markets



COLUMBIA UNIVERSITY
IN THE CITY OF NEW YORK

- Correlation of DU05YRX_VT10 across markets



Proposed ML Methods

1. Baseline - Ordinary Least Squares (OLS)
2. Autoregressive Integrated Moving Average (ARIMA)
3. Long Short-Term Memory (LSTM)

Baseline OLS

- The baseline OLS model is a basic regression model that ignores the relationships between the target variable and the independent variables with time. Thereby, treating the data points as if there is a linear relationship between the independent variables and the dependent variable.

$$\hat{\beta} = \left(\mathbf{X}^{\top}\mathbf{X}\right)^{-1}\mathbf{X}^{\top}\mathbf{y}$$

$\hat{\beta}$ = ordinary least squares estimator

X = matrix regressor variable X

$\top$ = matrix transpose

y = vector of the value of the response variable

- "A Contribution to the Empirics of Economic Growth" by N. Gregory Mankiw, et al, 1992, is a seminal work in the field of empirical economic growth analysis. This paper presents a comprehensive analysis of the determinants of economic growth by examining a large dataset of cross-country and time-series data while using OLS as the model for implementation.
- "Government Expenditure and Economic Growth in the European Union: An Empirical Investigation" by Paulo Reis Mourão, et al, 2008 is another great work in the field of econometrics which implements OLS to predict the relationship between the GDP of a country and the expenditure of a government.

- In this approach we treat the data features as if they have a linear relationship with the independent feature.
- With respect to this model we also do away with the assumption that there is any dependence between the dependent variables
- We also make the assumption that the residuals follow a normal distribution.
- Before we train our OLS Model we need to prepare our data to ensure that the model yields the best results. To do so we first pre-process the data and deal with the scaling issue of econometric datasets. We need the data to be on a similar scale to ensure that the features are given unbiased weights.
- Post the preprocessing we can proceed to train our model. Which can then need to be interpreted to determine the strength and direction of the relationships between the quantamental indicators and the outcome variable.
- Once we have interpreted the results we can move onto evaluating the results of the trained model to decide whether or not our hypothesis holds true. There are multiple evaluation metrics that can be utilized to obtain a reasonable conclusion.
- Once we have validated our model with the help of the evaluation metric(s), we can make the predictions for our target variable.
- Finally, we can consider some risk management strategies to further optimize the quantamental analysis.

ARIMA Model - traditional statistical time-series model

$$ARIMA(p,d,q) : (1 - \sum_{i=1}^{p} \phi_i L^i)(1-L)^d X_t = \delta t + (1 + \sum_{j=1}^{q} \theta_j L^j)\varepsilon_t \quad (1)$$

X_t: prediction target at time t
$\phi_1, ..., \phi_p$: autoregressive coefs
$\theta_1, ..., \theta_q$: moving average coefs
L: Lag Operator, such that $L^k X_t = X_{t-k}$
ε_t: error term at time t
d: Order of differencing
δ: drift coef (mean change in series per unit time)

Assumptions

1. **Stationarity:** prediction target is "stationary" after differencing
   a. (i.e. const. mean, variance, autocorrelation w.r.t. time)
2. **Linearity:** prediction target can be represented as a linear function of past values and errors
3. **Homoscedasticity:** error term has constant variance
4. **Normality of Errors:** error term follows the normal distribution
5. **Independence of Errors:** error terms are uncorrelated with one another
   a. No autocorrelation in residuals

Prior Work

- **Siami-Namini et al. (2018):** empirically compared ARIMA and LSTM model performance in forecasting economic and financial time series data using the evaluation metric of minimizing prediction error rates. The authors found that the average reduction in error for LSTM was 84-87% as compared to ARIMA.
  - The authors noted that economic forecasting has historically been challenging due to market volatility and incomplete information

- **Adebiyi et al. (2014):** showed that an ARIMA model trained on NY Stock Exchange and Nigeria Stock Exchange data yielded promising results in predicting short-term stock price.
  - Notable that the model's prediction ability on the short-term did not generalize to long-term prediction

References:

[1] Siami-Namini, Sima, and Akbar Siami Namin. "Forecasting economics and financial time series: ARIMA vs. LSTM." arXiv preprint arXiv:1803.06386 (2018).
[2] Adebiyi A.A., Adewumi A.O., Ayo C.K., 2014. Stock Price Prediction Using the ARIMA Model. UKSim-AMSS 16th International Conference on Computer Modelling and Simulation.

LSTM Model- recurrent neural network model with memory cells

Prior Works

- LSTM model is a type of neural network model that by nature captures the long distance dependency between different events. Prior researches have used LSTM to model sequential macroeconomic events and make predictions.

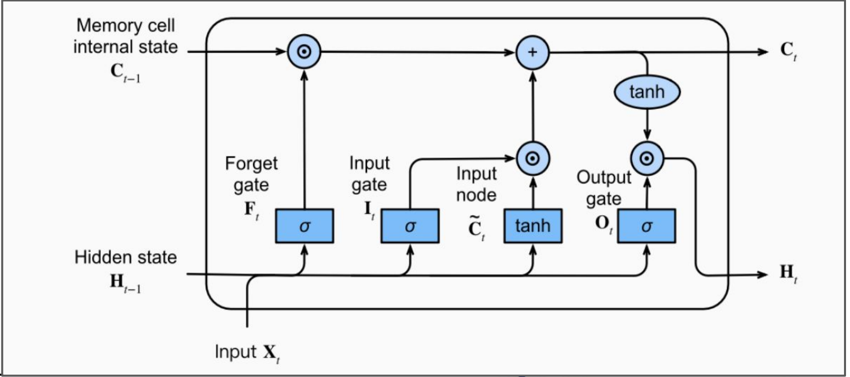
- Park's paper utilized LSTM to predict macroeconomic trends such as GDP growth and inflation based on predictors like unemployment, stock, private debt and interest rate. Park leveraged the SHAP values to explain the feature importances. From his research, LSTM is able to outperform other machine learning models. (Park 2022)

- Qiao performed a similar research on predicting stock returns using LSTM. Qiao measured the performance of the LSTM model using the NMSE metrics on the target stock returns and predicted values, and he has concluded that LSTM performs more accurately than classic loss functions. (Qiao 2022)

- Extending Park and Qiao's research to ours, we can also leverage LSTM to predict macroeconomic trends using quantitative indicators and measure the feature importance using SHAP values.

Model Details

- After we've cleaned and normalized the data, we can feed the inputs into the LSTM model. At each step of the LSTM, we choose to forget some information and learn some new inputs. As we feed in the old time series data, we're able to model the long distance dependencies and the non-linear correlations between the variables. Since we've divided the data using structured split by time, there is no data leakage and we can predict the target returns based on our modeled relationships.
- The input/output/forget gates and the memory cells allow for capability to model long distance relationships, which makes it very applicable to model time series data.
- The details of LSTM memory cell is illustrated below.



COLUMBIA UNIVERSITY
IN THE CITY OF NEW YORK

Evaluation Metrics

Evaluation metrics are an important part of the quantamental analysis. They can help us confidently come to a conclusion about the performance of a model on the dataset. We need to choose the best evaluation metrics for our project to ensure an optimized outcome.

1. Mean Squared Error (**MSE**): MSE calculates the average of the squared differences between predicted and actual values. Squaring the errors gives more weight to larger errors, which may be useful when outliers are of particular concern. Hence, this evaluation metric could end up providing us with biased results due to the presence of outliers, unless, they are of importance.
2. Root Mean Squared Error (**RMSE**): RMSE is the square root of the MSE and provides a metric on the same scale as that of the dependent variable and hence it is preferable because it penalizes large errors more than Mean Average Error (**MAE**).
3. R-squared (**R^2**): R-squared gives the amount of variability in the dependent variable that can be attributed to the independent variables. R^2 can be utilized to evaluate the effectiveness of our model in accounting for fluctuations in quantamental metrics.
4. **Sharpe Ratio**: The sharpe ratio, although commonly used for assessment of the risk-adjusted return of an investment or portfolio, can be considered for assessing the volatility of national currencies over a time period for cross-sectional study.
5. **Information Ratio**: The information ratio provides us the exact same evaluation metric data as the sharpe ratio with the only difference being that the sharpe ratio accounts for total risk, information ratio provides us with relative performance. Hence, it can prove to be quite beneficial to our quantamental analysis study.

Supplementary Information

Start years of quantamental indicators.

Quantamental Indicators	AUD	CAD	CHF	CLP	COP	CZK	EUR	GBP	HUF	IDR	ILS	INR	JPY	KRW	MXN	NOK	NZD	PLN	SEK	THB	TRY	TWD	USD	ZAR
CPIC_SA_P1M1ML12	2000	2000	2003	2002	2001	2005	2000	2000	2000	2011	2000	2014	2005	2000	2000	2000	2002	2004	2000	2000	2006	2000	2000	2000
CPIC_SJA_P3M3ML3AR	2000	2000	2003	2002	2001	2006	2000	2000	2000	2011	2000	2014	2005	2000	2000	2000	2002	2004	2000	2000	2006	2000	2000	2000
CPIC_SJA_P6M6ML6AR	2000	2000	2003	2002	2001	2006	2000	2000	2000	2011	2000	2014	2005	2000	2000	2000	2002	2004	2000	2000	2006	2000	2000	2000
CPIH_SA_P1M1ML12	2000	2000	2000	2000	2000	2000	2000	2000	2003	2000	2000	2000	2000	2000	2000	2000	2000	2000	2000	2000	2000	2000	2000	2000
CPIH_SJA_P3M3ML3AR	2000	2000	2000	2000	2000	2000	2000	2000	2003	2000	2000	2000	2000	2000	2000	2000	2000	2000	2000	2000	2000	2000	2000	2000
CPIH_SJA_P6M6ML6AR	2000	2000	2000	2000	2000	2000	2000	2000	2003	2000	2000	2000	2000	2000	2000	2000	2000	2000	2000	2000	2000	2000	2000	2000
DU02YXR_NSA	2001	2000	2000	2007	2011	2001	2000	2000	2001	2008	2006	2006	2000	2006	2006	2000	2000	2001	2000	2006	2006	2006	2000	2001
DU02YXR_VT10	2001	2000	2000	2007	2011	2001	2000	2000	2001	2008	2006	2006	2000	2006	2006	2000	2000	2001	2000	2006	2006	2006	2000	2001
DU05YXR_NSA	2001	2000	2000	2007	2011	2001	2000	2000	2001	2008	2006	2006	2000	2006	2006	2000	2000	2001	2000	2006	2006	2006	2000	2001
DU05YXR_VT10	2001	2000	2000	2007	2011	2001	2000	2000	2001	2008	2006	2006	2000	2006	2006	2000	2000	2001	2000	2006	2006	2006	2000	2001
EQXR_NSA	2000	2000	2000				2000	2000				2002	2000	2000	2000			2013	2005	2006	2005	2000	2000	2000
EQXR_VT10	2000	2000	2000				2000	2000				2002	2000	2000	2000			2013	2005	2006	2005	2000	2000	2000
FXCRR_NSA	2000	2000	2002	2000	2000	2002	2002	2002	2005	2002	2000	2000	2000	2000	2000	2002	2000	2003	2002	2000	2002	2000		2000
FXTARGETED_NSA	2000	2000	2000	2000	2000	2000	2000	2000	2000	2000	2000	2000	2000	2000	2000	2000	2000	2000	2000	2000	2000	2000		2000
FXUNTRADABLE_NSA	2000	2000	2000	2000	2000	2000	2000	2000	2000	2000	2000	2000	2000	2000	2000	2000	2000	2000	2000	2000	2000	2000		2000
FXXR_NSA	2000	2000	2000	2000	2000	2000	2000	2000	2000	2002	2000	2000	2000	2000	2000	2000	2000	2000	2000	2000	2000	2000		2000
FXXR_VT10	2000	2000	2000	2000	2000	2000	2000	2000	2000	2002	2000	2000	2000	2000	2000	2000	2000	2000	2000	2000	2000	2000		2000
INFTEFF_NSA	2000	2000	2000	2000	2000	2001	2002	2000	2005	2000	2000	2000	2000	2000	2000	2000	2000	2003	2000	2000	2000	2000	2000	2000
INTRGDP_NSA_P1M1ML12_3MMA	2000	2000	2000	2000	2005	2000	2000	2000	2000	2002	2000	2002	2000	2000	2000	2000	2000	2000	2000	2000	2003	2000	2000	2000
INTRGDPv5Y_NSA_P1M1ML12_3MMA	2000	2000	2000	2000	2006	2000	2001	2000	2001	2003	2001	2003	2000	2000	2000	2000	2000	2001	2000	2000	2004	2000	2000	2000
PCREDITBN_SJA_P1M1ML12	2000	2000	2000	2001	2003	2005	2000	2001	2001	2003	2000	2002	2000	2003	2000	2000	2001	2000	2000	2000	2000	2000	2000	2000
PCREDITGDP_SJA_D1M1ML12	2000	2000	2000	2001	2003	2005	2000	2001	2001	2003	2000	2002	2000	2003	2000	2000	2001	2000	2000	2000	2000	2000	2000	2000
RGDP_SA_P1Q1QL4_20QMA	2000	2000	2000	2000	2006	2000	2001	2000	2001	2003	2001	2003	2000	2000	2000	2000	2000	2001	2000	2000	2004	2000	2000	2000
RYLDIRS02Y_NSA	2000	2000	2000	2007	2011	2001	2002	2000	2005	2006	2006	2006	2000	2006	2006	2000	2000	2003	2000	2006	2005	2006	2000	2001
RYLDIRS05Y_NSA	2000	2000	2000	2007	2011	2001	2002	2000	2005	2006	2006	2006	2000	2006	2006	2000	2000	2003	2000	2006	2005	2006	2000	2001

Currency