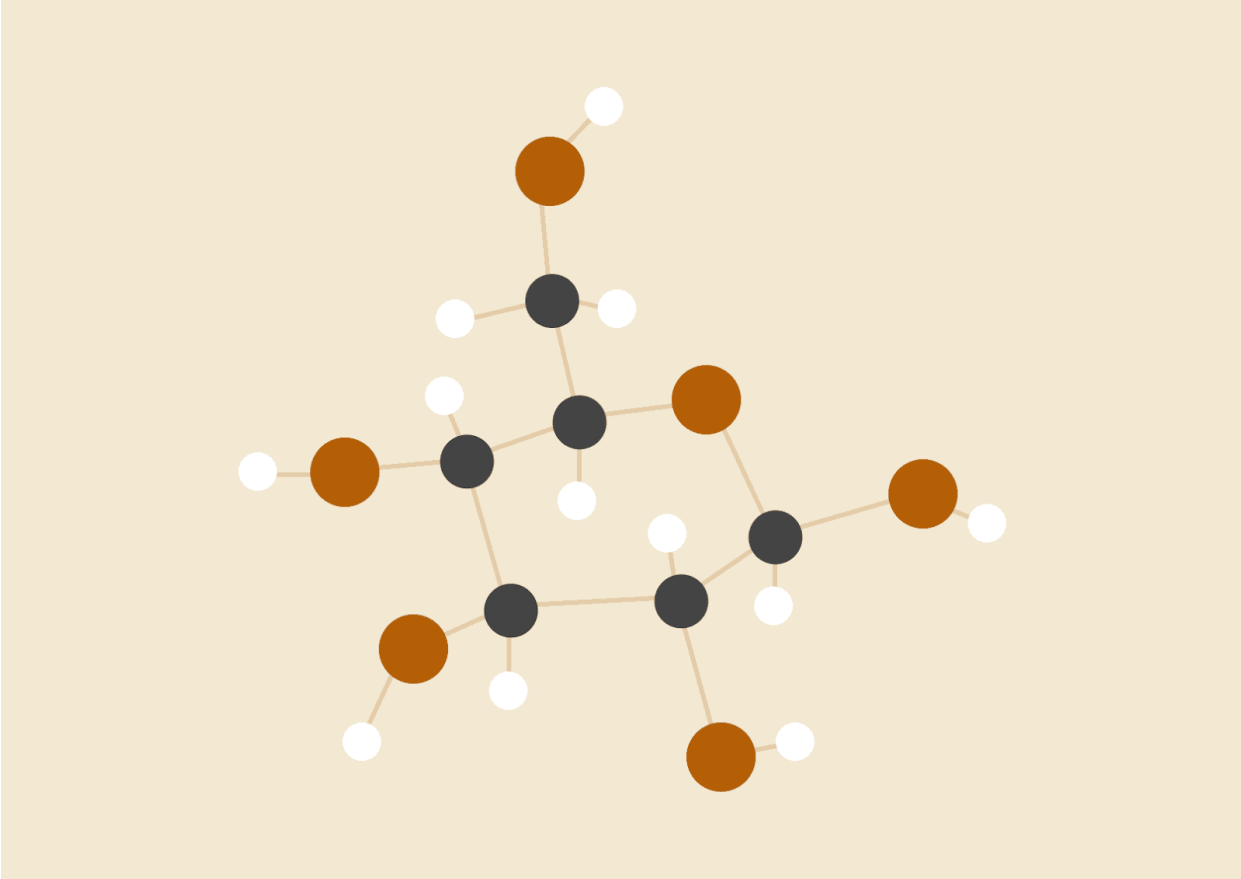
Interest Rates and ... ?

Seek correlation between interest rates and economic indicators, and Build/Evaluate models for predicting interest rates of different countries and the relationships between those economic indicators



Peter Poliakov, Hiromu Sugiyama, Raymond Smith

03/13/2023

CSE 163: Intermediate Data Programming

Summary of Questions and Results

1. Can we use economic indicators to predict interest rates?
    a. How does GDP affect interest rates?
        i. We want to explore the relationship between GDP (Gross Domestic Product) and interest rates as higher interest rates in economic theory, all else being equal, bring down GDP.
        ii. We were unable to determine a correlation.
    b. How does the unemployment rate affect interest rates?
        i. We want to explore the relationship between the unemployment rate and interest rates as higher interest rates in economic theory, all else being equal, increase the unemployment rate, meaning there should be a negative correlation.
        ii. Unemployment rates are correlated negatively with interest rates, meaning Central Banks decrease their interest rates when unemployment rates increase.
    c. How does the inflation rate affect interest rates?
        i. We want to explore the relationship between the inflation rate and interest rates as higher interest rates in economic theory, all else being equal, should decrease the inflation rate, meaning there should be a positive correlation.
        ii. Inflation rates and interest rates are positively correlated, meaning Central Banks increase their interest rates when inflation increases.
    d. How do changing interest rates correlate with stock indices?
        i. Higher interest rates should decrease stock market indices according to economic/financial theory.
        ii. We were unable to determine a correlation.
2. How did covid change these trends? How do real-life covid data compare to our model?

a. We will analyze our data between 2009 and 2017 and use a model to predict what should have happened in 2018-19 under normal circumstances, as well as in the following years (2020-2022) if Covid did not occur. We will then compare this to the actual data and see what the differences are if there are any.
b. We found that our model could not predict the sharp rise of interest rates in the past year, however, it may have foreseen the sudden drop of interest rates in wake of the Pandemic.

Motivations

We as a group are all somewhat interested in Economics and how to better understand how interest rates affect all sorts of other economic variables between a multitude of countries with similar basic economic structures (Monetary Policy independence, free-floating currencies). We want to see if these trends hold between countries or not. We also want to know the impact of the COVID-19 Pandemic on these trends and relationships and what the consequences were and if they are still felt in macroeconomic indicators.

Over the past few years, the world has seen a massive rise in inflation following the Pandemic. In the wake of this, there has been massive confusion on whether or not interest rates are the primary tool to help reverse this trend. By doing our project, we are hoping to find some correlation between the two variables, as well as any other variables that are correlated with interest rates to see the relationship between them all for various countries.

Since Central Banks are the institutions that manage interest rates for all of these countries, we want to see how they respond to these various economic indicators when setting their interest rates.

Dataset

All the Datasets were exported/downloaded from the various websites with a few exceptions. All the data we were searching for spanned from 1/1/2009 until the most recent available date. If data from a certain time frame was not present, we omitted that time period (happened with start and end dates).. We took quarterly data for the main economic indicators

(Interest rates, Unemployment rates, GDP per capita, and Inflation rates (CPI)) and monthly data for stock indices; for this, we would average out the index per quarter and use that in our project.

Some of the datasets were not downloadable (mostly stock market index data) so we had to copy and paste the data into excel and convert it into a CSV format. This is because although the data was there, it either would not download properly or could not be downloaded in the first place. The data is still correct however and it was very simple to access it still, even though it was annoying that this happened.

Once we had all the datasets, we edited the key column names in the files and formatted the dates such that it was easy to aggregate them. For the dates, we labeled them as "year_quarter" (2009-Q1 or 2016-Q4 for example) such that it was easy to combine them. On top of this, the key indicators in the datasets were oftentimes labeled weirdly and nonuniformly, therefore we changed the names of the columns to make our job easier when working with the data later on. We changed the column names to as follows depending on what the data was:

- Date, Close, Unemployment Rate, Interest Rate, CPI, GDP per capita

Australia:

- GDP per capita
    - https://stats.oecd.org/index.aspx?queryid=66948#
        - Filtered for quarterly and just the Australia and expanded range of dates
            - Under customize tab -> selection -> country, period & frequency
        - Exported as CSV
            - Under export tab
- Inflation/CPI
    - https://data.oecd.org/price/inflation-cpi.htm
        - Filtered for quarterly and just Australia and expanded dates
            - Made sure everything else was toggled off; no background and no baseline
        - Made sure no other countries were selected
- Unemployment
    - https://data.oecd.org/unemp/unemployment-rate.htm
        - Filtered for only australia and expanded ranges and such
            - Made sure everything else was toggled off; no background and no baseline

        - Made sure no other countries were selected
- Interest Rate
    - https://data.oecd.org/interest/short-term-interest-rates.htm#indicator-chart
        - Filter for Australia and dates
        - Monthly
            - Will take a quarterly average later
- Stock index
    - https://www.marketwatch.com/investing/index/xjo/download-data?startDate=1/1/2009&endDate=3/3/2023&countryCode=au
        - Copy pasted into excel and turned it into a CSV after filtering for dates

Canada:

- Interest Rate: https://fred.stlouisfed.org/series/IRSTCB01CAM156N#0
    - Units: Percent
    - Modified monthly data into quarterly by using the average for aggregation method
    - Selected a specific period
    - Exported as CSV file
    - Modify in VSCode
        - Changed Date format (2019-01-01 > 2019-Q1)
        - Changed the column name to “Interest Rate”
- GDP per capita: https://stats.oecd.org/index.aspx?queryid=66948#
    - Selected a specific country and period
        - 56 quarters from the most recent data (Q3-2022)
    - Exported as CSV file
    - Modify on VSCode
        - Extract only the Date and GDP per capita column
        - Added one row to uniform the size of the table (the last quarter, 2022-Q4, was missing) > null value
- Unemployment rate: https://data.oecd.org/unemp/unemployment-rate.htm
    - Selected a specific country and period (Q1 2009 - Q4 2020)
    - Exported as CSV file

        - Extracted only the Date and Unemployment Rate column
        - Sliced only the rows of Canada (eliminated OECD ones)
- Stock Indices (S&P/TSX):

  https://ca.finance.yahoo.com/quote/%5EGSPTSE/history?p=%5EGSPTSE
    - Searched specific stock (KOSPI) and selected specific period (Jan 1, 2009 - Dec 31, 2022) and frequency (month)
    - Exported as CSV file
    - Modify on VSCode
        - Extracted Open, Close, Adj Close and Volume columns (only Date, High, and High)
        - Changed Column name > can understand when the table merge
        - The calculated average of every three months
- CPI: Total All Items for Canada:

  https://fred.stlouisfed.org/series/CPALTT01CAQ657N#0
    - Units: Growth Rate Previous Period
    - Default data collected quarterly
    - Selected a specific period
    - Exported as CSV file
    - Modify on VSCode
        - Changed Date format (2019-01-01 > 2019_Q1)

Euro Zone (19):

- GDP per capita
    - https://stats.oecd.org/index.aspx?queryid=66948#
        - Filtered for quarterly and just the euro area (19) and an expanded range of dates
        - Exported as CSV
- Inflation/CPI
    - https://data.oecd.org/price/inflation-cpi.htm
        - Filtered to be only EU 19
        - Expanded dates to 2009 Q1
- Unemployment
    - https://data.oecd.org/unemp/unemployment-rate.htm

        - Filtered for only EU 19 and expanded ranges and such
        - Made sure no other countries were selected
- Interest Rate
    - https://data.oecd.org/interest/short-term-interest-rates.htm#indicator-chart
        - Make sure its only EU 19
        - Monthly data
- Stock index
    - Omitted; Euro Zone does not have one stock index that tracks all public companies in the economic block therefore it is not part of our analysis

Japan:

- Interest Rate: https://www.stat-search.boj.or.jp/ssi/mtshtml/ir01_m_1.html
    - Exported as CSV file
    - Modify in VSCode
        - Extract the first few rows with unknown text
        - Extract the only period necessary (1980~ > 2009~)
        - Converted Monthly data into Quarterly data
            - Using groupby and mean aggregation
- GDP per capita: https://stats.oecd.org/index.aspx?queryid=66948#
    - Selected a specific country and period
        - 56 quarters from the most recent data (Q3-2022)
    - Exported as CSV file
- Unemployment: https://data.oecd.org/unemp/unemployment-rate.htm
    - Selected a specific country and period
    - Exported as CSV file
- Stock indices (Nikkei 225): https://finance.yahoo.com/quote/%5EN225/history?p=%5EN225
    - Searched specific stock (Nikkei 225) and selected specific period (Jan 1, 2009 - Dec 31, 2022) and frequency (month)
    - Exported as CSV file
- CPI: Total All Items for Japan: https://fred.stlouisfed.org/series/JPNCPIALLQINMEI#0
    - Units: Percent Change

    - Searched by country and selected a specific period
    - Exported as CSV file

Republic of Korea (South Korea):

- Interest Rate: https://fred.stlouisfed.org/series/INTDSRKRM193N#
    - Modified monthly data into quarterly by using the average for aggregation method
    - Selected a specific period
    - Exported as CSV file
- GDP per capita: https://stats.oecd.org/index.aspx?queryid=66948#
    - Selected a specific country and period
        - 56 quarters from the most recent data (Q3-2022)
    - Exported as CSV file
- Unemployment rate: https://data.oecd.org/unemp/unemployment-rate.htm
    - Selected a specific country and period (Q1 2009 - Q4 2020)
    - Exported as csv file
- Stock Index (KOSPI): https://finance.yahoo.com/quote/%5EKS11/history/
    - Searched specific stock (KOSPI) and selected specific period (Jan 1, 2009 - Dec 31, 2022) and frequency (month)
    - Exported as CSV file
- CPI: Total All Items for the Republic of Korea: https://fred.stlouisfed.org/series/CPALTT01KRQ657N
    - Units: Growth Rate Previous Period
    - Default data collected quarterly
    - Selected a specific period
    - Exported as CSV file

South Africa:

- GDP per capita
    - https://stats.oecd.org/index.aspx?queryid=66948#

        - Filtered for quarterly and just the euro area (19) and an expanded range of dates
        - Exported as CSV
- Inflation
    - https://data.oecd.org/price/inflation-cpi.htm
        - Filtered for quarterly and just South Africa and expanded dates
        - Made sure no other countries were selected
- Unemployment
    - https://fred.stlouisfed.org/series/LRUN64TTZAQ156S
        - Filter for dates and download
- Interest Rate
    - https://data.oecd.org/interest/short-term-interest-rates.htm#indicator-chart
        - Filter for South Africa and the years
        - monthly
- Stock Index
    - https://finance.yahoo.com/quote/%5EJ580.JO/history?period1=1356998400&period2=1678665600&interval=1mo&filter=history&frequency=1mo&includeAdjustedClose=true
        - Copy pasted into excel and saved it as a CSV after filtering it

United Kingdom: Same as above for oecd links

- Inflation
    - https://stats.oecd.org/index.aspx?queryid=66948#
- Unemployment
    - https://data.oecd.org/unemp/unemployment-rate.htm
- Interest Rates
    - https://fred.stlouisfed.org/series/IRSTCI01GBQ156N
        - Filter for dates and download
- Stock index
    - https://www.marketwatch.com/investing/index/ukx/download-data?startDate=1/1/2009&endDate=03/10/2023&countryCode=uk
        - Copy pasted manually after filtering for dates and frequency into excel and then saved it as a CSV as download did not work
- GDP per capita

- : https://stats.oecd.org/index.aspx?queryid=66948#

United States: Same as above for oecd links

- Inflation
  - https://stats.oecd.org/index.aspx?queryid=66948#
- GDP
  - https://stats.oecd.org/index.aspx?queryid=66948#
- Unemployment
  - https://data.oecd.org/unemp/unemployment-rate.htm
- Interest rates
  - https://fred.stlouisfed.org/series/IRSTFR01USQ156N
    - Filter for dates and download
- Stock index
  - https://finance.yahoo.com/quote/%5EGSPC/history?period1=1230768000&period2=1678060800&interval=1mo&filter=history&frequency=1mo&includeAdjustedClose=true
    - Copy pasted into excel and turned into CSV

Methods

1. **Research Question 1:** Can we use economic indicators to find correlations with interest rates?
   a. Combine and clean datasets - challenge goal #1
      i. Combine per-country economic indicators and stock indices to per-country interest rates per the year 2009-2022
   b. Create ML regression model - challenge goals #2 & 3
      i. Multivariable regression
      ii. Used heatmap to show results
2. How did Covid change trends?
   a. Create ML prediction model (Ridge Regression) - challenge goals #2 & 3
      i. Features = economic indicators

ii. Value = interest rate

iii. Trained from 2009-2017

b. Compare to real data from 2018-2022

RESULTS

1. Can we use economic indicators to predict interest rates?
    a. How does GDP affect interest rates?
        i. Through our results, we were not able to determine a clear correlation between interest rates and GDP. Some nations, Australia, Europe, and South Korea, had negative correlation coefficients, meaning as interest rates decreased, GDP increased. On the other hand, Canada, the United Kingdom, and the United States had positive correlation coefficients, meaning as interest rates increased, GDP also increased. South Africa had a near 0 correlation coefficient, meaning changes in interest rates had no correlation with GDP.
        ii. This result is interesting as it could showcase how Central Banks do not just only look at GDP when making decisions on interest rates. The variety of correlations shows that each country has a wide array of circumstances that could influence this decision-making process. Although we should see a negative correlation according to theory, in practice it is hard to find this trend due to how many variables influence one another.
    b. How does the unemployment rate affect interest rates?
        i. Our results indicate a negative correlation between the interest rate and unemployment. This trend holds for every country, except for Europe, although it changes in magnitude. This means that if the unemployment rate increases, interest rates decrease, and vice-versa.
        ii. This result holds with economic theory as Central Banks decrease their interest rate when unemployment increases as the decrease in the interest

rate allows businesses to borrow money at cheaper rates, allowing them to hire more employees and pursue more projects. These institutions usually have a mandate of "full employment," which in simpler terms means to have a low unemployment rate without an increase in inflation. As unemployment decreases, the risk for inflation starts to increase, causing Central Banks to increase their interest rates again to balance the two (unemployment and inflation).

c. How does the inflation rate affect interest rates?
    i. Our results indicate a slight positive correlation between inflation and interest rates. This trend holds for every nation except for Europe, although again there are magnitude changes across all the countries. This correlation shows that as inflation rates increase, interest rates follow suit upwards and vice-versa.
    ii. This result holds with economic theory as the Central Bank's primary response to rising inflation rates is to increase interest rates. This is such that firms will borrow less money, causing them to employ fewer people and pursue fewer projects, all in order to bring inflation down. The reason they do this is the fact that low and stable inflation helps with long-term economic/financial planning and high inflation imposes sizable, real costs onto economies, which wastes people's time and resources.
        1. In various unorthodox schools of economic thought, there is a theory that high-interest rates are correlated with/cause high inflation rates. This is false due to the fact that Central Banks respond to higher inflation rates with higher interest rates in order to ease inflation, meaning the interest rate is a response variable. These schools of economics sometimes fall for the causation = correlation fallacy.

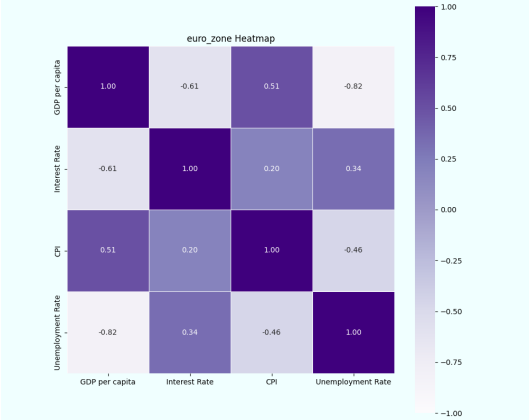
d. How do changing interest rates correlate with stock indices?

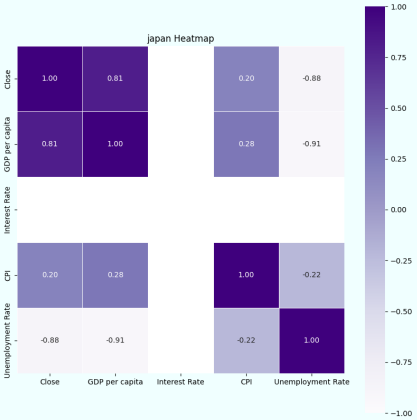
i. Through our results, we were not able to determine a correlation between stock indices and interest rates. Some countries, more specifically Australia and South Korea had negative correlations, while Canada, South Africa, the United Kingdom, and the United States had positive correlations. This means that the level of all the stock prices of companies located in those countries had different correlations with interest rates.

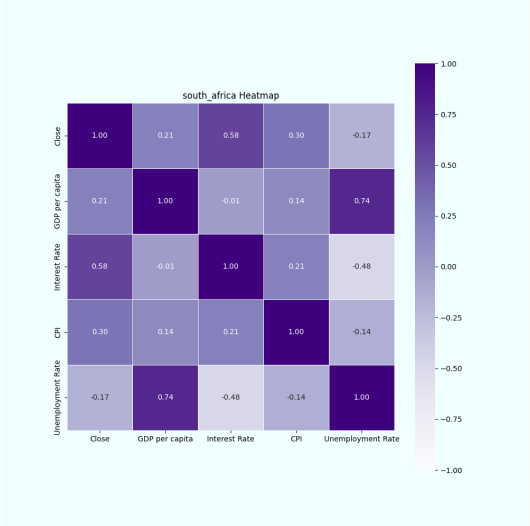
ii. This result diverges a bit from economic/financial theory as lower interest rates should decrease the prices of stocks on the stock market. However, there could be other variables at play that we don't see such firm profits or how the underlying economy is doing. A Central Bank usually does not change interest rates in response to the stock market, but it tries to limit bubbles forming (stock prices increasing rapidly without good reason) or prevent financial crises (stock prices falling too rapidly).

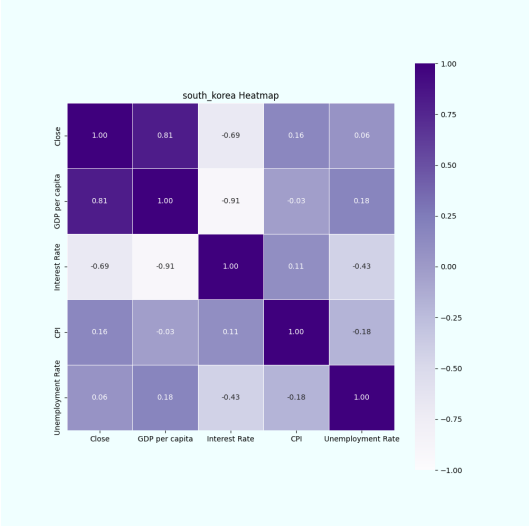

australia Heatmap
Close
GDP per capita
Interest Rate
CPI
Unemployment Rate
1.00 0.92 -0.83 0.22 -0.20
0.92 1.00 -0.81 0.39 -0.36
-0.83 -0.81 1.00 0.13 -0.13
0.22 0.39 0.13 1.00 -0.79
-0.20 -0.36 -0.13 -0.79 1.00
1.00
0.75
0.50
0.25
0.00
−0.25
−0.50
−0.75
−1.00

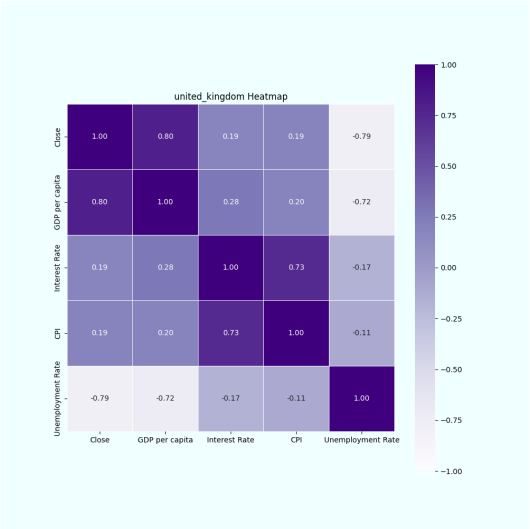

canada Heatmap
Close
GDP per capita
Interest Rate
CPI
Unemployment Rate
1.00 0.73 0.12 0.52 -0.37
0.73 1.00 0.47 0.29 -0.80
0.12 0.47 1.00 -0.02 -0.61
0.52 0.29 -0.02 1.00 -0.28
-0.37 -0.80 -0.61 -0.28 1.00
1.00
0.75
0.50
0.25
0.00
−0.25
−0.50
−0.75
−1.00

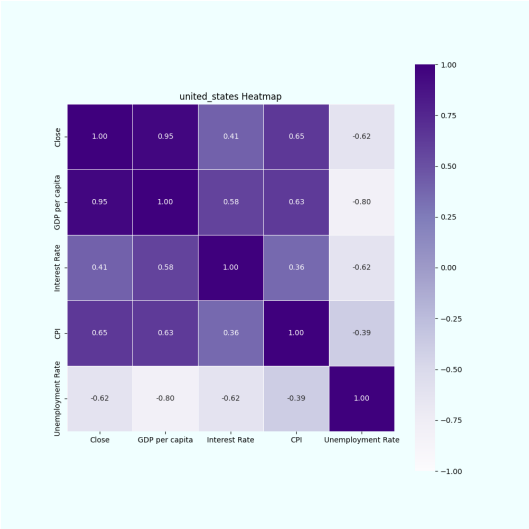

euro_zone Heatmap
GDP per capita
Interest Rate
CPI
Unemployment Rate
1.00 -0.61 0.51 -0.82
-0.61 1.00 0.20 0.34
0.51 0.20 1.00 -0.46
-0.82 0.34 -0.46 1.00
1.00
0.75
0.50
0.25
0.00
−0.25
−0.50
−0.75
−1.00


japan Heatmap
Close
GDP per capita
Interest Rate
CPI
Unemployment Rate
1.00 0.81 0.20 -0.88
0.81 1.00 0.28 -0.91
0.20 0.28 1.00 -0.22
-0.88 -0.91 -0.22 1.00
1.00
0.75
0.50
0.25
0.00
−0.25
−0.50
−0.75
−1.00





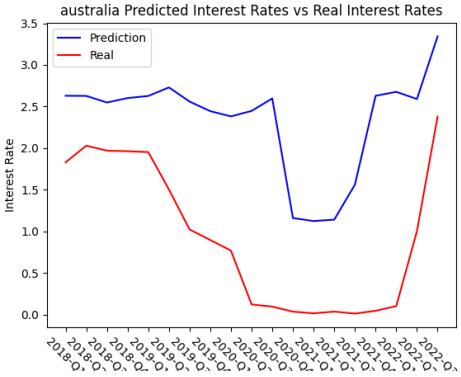


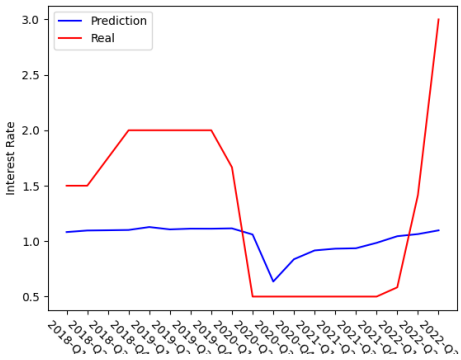


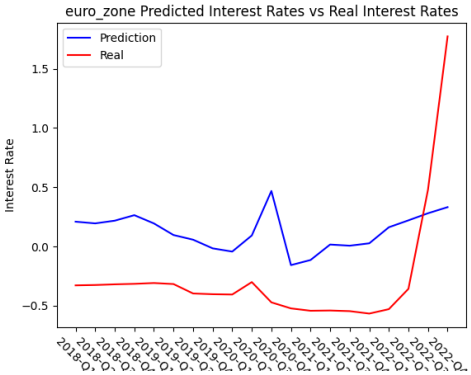
2. How did covid change these trends? How do real-life covid data compare to our model?
    a. Covid had shifted interest rates to the lowest levels in recent history, with almost every country having 0, or near-0 interest rates (with the exception of Europe where interest rates went negative). Afterward, rates started to normalize only for them to keep rising rapidly to high levels in relation to recent history at a record pace.
    b. Our model was not able to predict the exact values of the interest rates were going to be. However, there were two key aspects of our model that we found interesting:

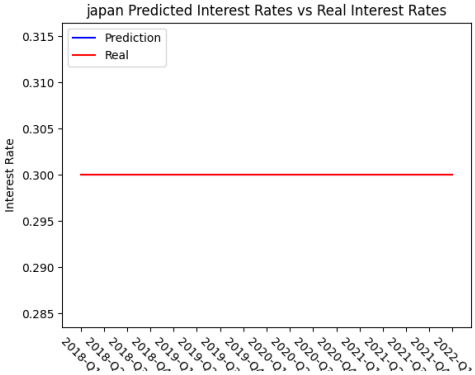
The first aspect is the fact that none of our models for the individual countries were able to predict the rapid increases in interest rates in the past year. This goes to show how unprecedented the inflation everyone has been seeing since 2021 has been; Central Banks may have been caught off guard by this inflation and may have thought it was not as bad as it actually was when it first arrived, but now they have corrected course and have increased interest rates to fight it. In total, none of our models could have predicted this, and rightfully so. The only country that is an exception to this rule is South Africa, but this is only because we did not have good data on interest rates within South Africa for 2022.

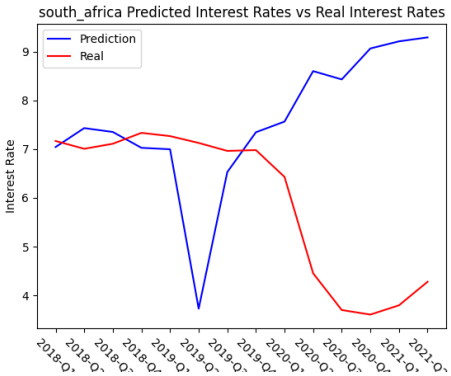
The second aspect is the fact that most of our models for the countries foresaw that interest rates were to decrease sometime between 2019 and 2021. Central Banks respond to recessions by decreasing interest rates to help spur economic growth and fight deflation (negative inflation). The models for our countries, not including Europe or the United Kingdom, we're predicting that Central Banks were to decrease interest rates during this time as the economic indicators had shown that the economic climate had turned sour. Although they did not get the exact timing of the decreases correct, it is still notable that they even predicted this in the first place.

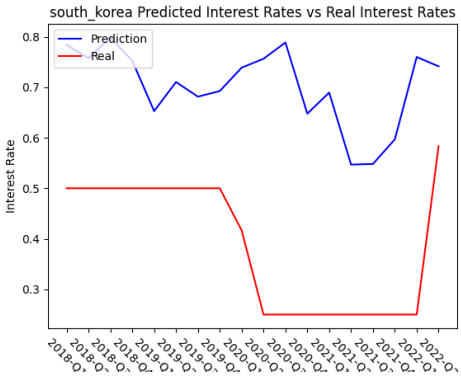

australia Predicted Interest Rates vs Real Interest Rates
Prediction
Real
Interest Rate

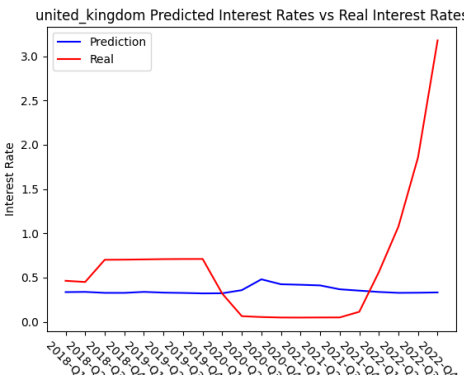

canada Predicted Interest Rates vs Real Interest Rates
Prediction
Real
Interest Rate

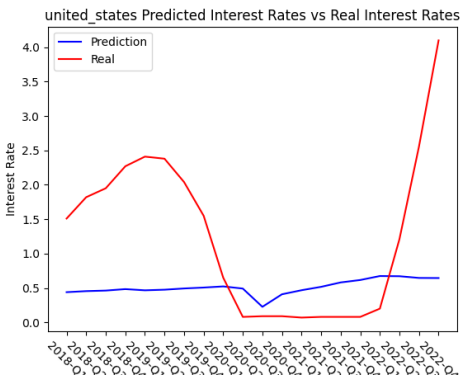

euro_zone Predicted Interest Rates vs Real Interest Rates
Prediction
Real
Interest Rate


japan Predicted Interest Rates vs Real Interest Rates
Prediction
Real
Interest Rate


south_africa Predicted Interest Rates vs Real Interest Rates
Prediction
Real
Interest Rate


south_korea Predicted Interest Rates vs Real Interest Rates
Prediction
Real
Interest Rate


united_kingdom Predicted Interest Rates vs Real Interest Rates
Prediction
Real
Interest Rate


united_states Predicted Interest Rates vs Real Interest Rates
Prediction
Real
Interest Rate

Impact and Limitations

- Economics is an extremely complicated field and economists have spent decades trying to understand the relationships between all these economic variables/indicators. In our analysis, we did not prove any causation, only forms of correlation between our variables. If we wanted to prove causation, we would have never taken this class and instead would have taken an econometrics course (we are thankful to be in this course and have learned a lot). Here, we are just analyzing the trends between variables and various correlations between them, not if they actually influence one another necessarily.
- Japan has had near 0 interest rates for almost the entire time. In fact, it is true that Japan's interest rate has not changed for the entirety of the time period that we evaluated it. It is hard to draw any conclusions from their data because of its fact and any correlations between variables are non-existent for the country, therefore it is not present in our results.
- South Africa is still a developing country while the rest of the countries are all developed. That is to say that the country is an outlier in our results as it has only recently undergone industrialization and it is a greatly unequal and poor country that is still feeling the effects of colonialism and apartheid. Therefore, it could be unfair to apply the same trends from other developed nations to South Africa in our analysis.
- Europe (we labeled it as the Euro Zone 19) is not a sovereign political entity but instead is a collection of nations that all use the Euro for currency. We look at trends across the entire block and not necessarily within countries. Results may vary for individual countries such as Germany, Finland, and the Netherlands when compared to Italy, Greece, and Spain as Europe has a north/south divide between their economies.

- Economic theory says that there are clear relationships between indicators/variables and we could see them if we hold all other variables constant. However, in the real world, it is impossible to hold all other variables constant as everything is changing everywhere all at once. Therefore our analysis is very limited due to this fact.
- The Zero Lower Bound is an economic phenomenon that restricts interest rates to a floor, usually zero, due to various reasons. The years which we created the model over were defined by an extended period at this lower bound. This may have caused our predictions for the interest rate to be lower than they should have been in our models for some countries.

Challenge Goals

Challenge Goal #1: Using Multiple Datasets

We used 39 different datasets to accomplish our research goal. Combining and cleaning them was a serious challenge. We grouped them by country.

- The method for uniforming the data varies from resource because the data collecting approach differs from one another
    - One data may have quarterly as default, whereas the other, we had to modify
    - If we had monthly datasets and converted them to quarterly data, we can either take data from a month from the quarter or take an average of the quarter.
        - Extreme example: Let's say Jan:1, Feb:8, Mar:11 -> Quarterly Data may differ depending on how you calculate
        - We will still be lenient about this caveat in our case from an assumption that the economic indices will not exponentially change by month (like what we see in the example).
- Due to the number of datasets, we had to eliminate three countries from our analysis, them being Brazil, Turkey, and India. These countries had poor/unreliable data which would have made this project more complicated and longer than it needed to be.

Challenge Goal #2: Building Machine Learning models

- We used the Ridge Regression model to predict interest rates. We had never used Ridge before, so it was completely new to us. First, we created heatmaps to exemplify the correlation between indicators. Then, we trained our model on 2009-17 data and tested it on 2018-2022 data.
- Originally we attempted to use 3 other ML models. However, we found that they didn't give realistic results.

Challenge Goal #3: Result Validity

- There are unorthodox schools in economics that say interest rates have little causation with certain other economic indicators (inflation and asset prices just name two of them.) We want to see if our results for the correlation between these variables happen by chance or if is there some causation at play.
    - Evaluated using criteria: R^2 and MAE
        - R^2 is shown in the heatmap for the linear regression
        - MAE is printed into the console for the Ridge Regression

Work Plan Evaluation

Our work plan underestimated how much time we would spend on this project. Firstly, the time spent finding good, accurate data took an enormous amount of time. We found that oftentimes datasets would be inadequate or even would not exist in the framework of what we were looking for. For example, if we were looking for quarterly data and it did not exist, then we would use monthly and standardize it to make it quarterly by taking an average of the three months. Cleaning up the datasets was another huge problem as it took a bit of time to rename the columns and making sure it worked within our existing framework.

Another issue we ran into was some of the regression models we had originally used had results that we could not interpret. This made us spend a bit of time finding good models that could give us results that we could understand and present.

One positive aspect was the creation of the data visualizations, which were surprisingly easy. However, in total, the work plan underestimated the amount of time we spent on the project.

Testing

Because we were predicting interest rates, the best way to verify the validity of our results was to compare our test data to real data. It was not possible to use an assert statement because there is no objectively “right” answer - we were just making predictions. Our results can be trusted because of our strong initial data and fairly large sample size.

Collaboration

We consulted our mentor and various sources on the internet such as stack overflow. This website suggested Ridge: https://qiita.com/DS27/items/aa3f6d0f03a8053e5810 (in Japanese). Qiita is a Japanese forum used to share code and different ideas. It is similar to stack overflow but geared toward an Asian audience. We used Ridge because the documentation was neatly organized and there were detailed explanations of how each regression model calculates its result.