POLYFINANCES

DATATHON 2022

The impact of central banks

CONTEXT

The expectations of the central banks' monetary policies have an important impact on the markets and on the evolution of currencies. Knowing this, a fund manager wants to develop an investment strategy by investing his money in CAD or USD over a given period.

Important resources

- Interesting terms to explore :
  - Fed jumping date
  - Sentometrics
  - Textual analysis
  - Huggingface
- Look at the literature on the subject!

Submission

- **Mon, Nov 7, noon:** NoteBook(s) used to train and test your models with your results.
- **Thursday, Nov. 10, 11pm:** Your PowerPoint presentation.

CHALLENGE

Your role as a data scientist is to design an algorithm to **predict fluctuations or values** of the CAD/USD exchange rate **over a two week period** based on the following data :

- US central bank press releases, speeches, announcements or transcripts.
- Canadian monetary policy reports.
- Policy rates.
- Other data on the economic state of these two countries.

After completing such a task, you must present your solution to the fund manager and the fund manager will choose the most promising solution. This will be done in the form of a jury presentation.

PRESENTATION TO JURY

1. **To show the impact that central bank texts have on the USD/CAD exchange rate**

   Ideas: Data visualization, observed statistics, macroeconomic theories

2. **Explain your prediction model based on the Fed's texts as well as other data.**

   Ideas: Statistical models to predict the sentiment of a text, deep learning models to predict the impact

3. **Show the approach you used to carry out such a project.**

   Ideas: Share how you achieved such a task and what you learned!

PRESENTED BY

NATIONAL BANK

FINANCE MONTRÉAL

CFA MONTRÉAL

SR.AI

DATATHON
Some details

AVAILABLE DATA

1 Bank of Canada data

The Bank of Canada's responsibilities focus on the following objectives: low, stable and predictable inflation; a safe and secure currency; and a stable and efficient financial system in Canada and internationally. In practice, the Bank has a narrower and more precise internal definition of this mandate: to keep inflation (as measured by the consumer price index) between 1 and 3 per cent.

The most powerful tool the Bank of Canada has to achieve this objective is its ability to set the interest rate on borrowed money. Because of the importance of trade between Canada and the United States, specific interest rate adjustments are often influenced by those of the United States at that time

→ Data set from this source:

- USD - CAD exchange rate [from 2017 to present] ← Value to predict!
- Interest rate that commercial banks charge their most creditworthy customers. [2012 to present]
- Texts extracted from the Bank of Canada's monetary policy reports [2016 to present]

2 Data from the U.S. Federal Open Market Committee (FOMC)

In the United States, the Federal Open Market Committee (FOMC), a committee of the Federal Reserve System (the Fed), is charged by U.S. law with overseeing the operations of the nation's markets. This committee of the Federal Reserve makes key decisions regarding interest rates and the growth of the U.S. money supply.

The federal funds target range is determined by a meeting of the Federal Open Market Committee (FOMC) members, which is normally held eight times a year at about seven-week intervals. The committee may also hold additional meetings and implement changes to the target rate outside its normal schedule.

→ Dataset

- All FOMC texts [from 1980 to present]
- Interest rate at which commercial banks charge their most creditworthy customers [2012 to present]

3 Other data

→ Dataset

- Loughran-McDonald master dictionary with sentiment word lists from a finance perspective (Feel free to use other dictionaries and even create your own!)

For this mandate, you have the option of supplementing these data sets with any data you deem relevant to the task. If you add data, it will be important that you demonstrate to the jury that the sources of your data are reliable and that you convince them that these data can be used in the future when they need to make a decision.

→ Data sets: Link on Discord

Other information

YOUR MODEL

As you may have noticed, this year's Datathon is complex. In order to give the teams the best possible chance to learn, we will separate this one into three distinct steps.

1. Using the **US FED texts**, extract a **market sentiment.** (40% of the points)
    - It is up to you to choose the best way to extract this sentiment.
2. Do a **similar analysis**, but on the **Bank of Canada's texts** this time. (20% of the points)
3. **Use these sentiments** (and other metrics of your choice) to **make a prediction about the conversion rate** of the two countries' currencies **over a two-week interval**. (40% of the points)
    - Predict the exchange rate & if possible an uncertainty interval (X + 14 days)
      → 2022-05-02 (Predict: 2022-05-16), 2022-05-16 (Predict: 2022-05-30), ...
      → Output the MSE (mean square error) error metric and the bias of your predictions.

EVALUATION METRICS

For the presentation

20% Quality of presentation and expression
60% Content of the presentation

- Your approach to the problem
    - Summary of the problem
    - Strategy used
    - External elements you added
- Presentation of the models
    - Explanation of the models
    - Validation performed
- Presentation of the results
    - Prediction
    - Model limitations
    - What you would have liked to explore
- What you learned

For the results

20% In order to be able to compare the different approaches between teams, you will have to present the sentiment and forecasting results "out of sample" for the last 6 months.

This means:

- you can use all data before May 2022 to design / train / validate your models.
- the results must come only from the last 6 months, May 2022 to October 2022

PRIZES

1ST PLACE
→ 2000$

2ND PLACE
→ 1000$

3RD PLACE
→ 500$

PRIZES OF PARTICIPATION
→ Best visualization
→ Best outreach
→ Mini-games

SCHEDULE

SATURDAY, NOVEMBER 5 - ATRIUM OF THE J.-A. BOMBARDIER P.

9:00 am - Welcome of the participants with pastries and coffee
9:30 am - Presentation of the competition by PolyFinances
10:00 am - NLP course by SR.AI
10:30 am - Break
10:45 am - Finance course by National Bank
11:30 am - Meeting with sponsors
~~~ The rest of the day will be virtual
Noon - Beginning of the competition + release of the data
6:00 pm - Mini game #1
10:00 pm - Mini game #2

SUNDAY, NOVEMBER 6 - VIRTUAL

10:00 am - Question period with National Bank
2:00 pm - Voting for the best data visualization + Prize!

MONDAY, NOVEMBER 7 - VIRTUAL

Noon - End of the competition + submission of projects

FRIDAY, NOVEMBER 11 - POLYTECHNIQUE MONTREAL

1:30 pm - Jury presentations begin at the Rolland Gallery
→ Detailed schedule to come
→ Submission of PowerPoints on Thursday evening
4:30 pm - End of the presentations
~~~
5:30 pm - Networking evening at the Atrium of the Lassonde Pavilion
→ Prizes ceremony
→ 10+ companies will be present
→ Drinks will be provided
7:30 pm - End of the networking evening