

Data Description

Overview

You are provided with historical data ranging from 2016/01/01 to 2018/04/22. The challenge is to predict the customer interest in some bond transactions (which can be either Buy or Sell orders) on 2018/04/23.

The Challenge table

There 6 columns in this table:

- PredictionIdx is a hash computed from the CustomerIdx, IsinIdx and BuySell columns. It serves as a unique identifier for the predictions that will be used for scoring (see `sample_submission.csv` for a submission example).
- DateKey date in format YYYYMMDD at which the interaction happens. In this table, it is always equal to 20180423
- CustomerIdx is the identifier of the Customer
- IsinIdx is the identifier of the bond
- BuySell is the direction of the interaction: Buy (Sell) means the Customer would like to buys (sells) the Bond.
- CustomerInterest is the variable you should predict. Customer is thought to has an interest when he has an interaction of type Done, NotTraded, Unknown or IOI (but not of type Holding). See Trade table documentation for details about interaction types.

The Trade table

There are 8 columns in this table:

- TradeDateKey: date in format YYYYMMDD at which the interaction happened
- CustomerIdx: identifier of the Customer
- IsinIdx: identifier of the Bond
- BuySell: direction of the interaction: Buy (Sell) means the Customer would like to buy (sell) the Bond.
- NotionalEUR: notional in EUR of the transaction
- Price: proposed price of the transaction expressed in percent. This may not be always be available or set to a stupid value, depending on Market condition and on Trading coverage.
- TradeStatus: type of the interaction. There are 6 different possible interaction:
 - *Done*: the transaction was traded between the Customer and BNP Paribas
 - *NotTraded*: the Customer looked at the Price but decided not to Trade
 - *Unknown*: the Customer sent an electronic request for a Quote to BNP Paribas and some competitors, but because BNP Paribas did not reply and we do not know what the Customer decided to do

- *IOI*: (Indication Of Interest): Customer expressed his interest to buy or sell a given Bond in the future
- *Holding*: Some of our Customer have a legal obligation to report at a regular frequency (monthly, quarterly) what they have in their portfolio. This information is represented in our database as a serie of transaction happening on those reporting dates: if Customer suddenly has a holding of 10M on 31st of Jan in a given Bond, we will record a Buy transaction for 10M on 31st of Jan. If on the 31st of March the holding is now 5M, we will record a Sell transaction for 5M on the 31st of March. Those are not real transaction but just a way to represent Customer's holdings.
- CustomerInterest is equal to 1 when the Customer is thought to has an interest, i.e. when he has an interaction of type Done, NotTraded, Unknown or IOI (but not of type Holding). It is equal to 0 otherwise.

The Isin Table

This table gives all kind of information on each Bond:

- IsinIdx: the Index of the Bond
- TickerIdx: the index of the Issuer of the Bond
- IssueDateKey: the date in format YYYYMMDD at which Bond was issued
- ActualMaturityDateKey: the date in format YYYYMMDD at which Bond will mature.
- Currency: the currency of the Bond
- Seniority: the seniority of the Bond (GOV means issued by a Government, SEN for Senior Debt, SUB for subordinated Debt, SEC for secured Debt, MOR for Mortgage...)
- IssuedAmount: how much of this Bond was issued in the Bond currency
- CouponType: type of coupon being paid: can be fixed (FIXED), indexed on a floating rate (FLOATING), can be a deterministic coupon increasing with time (STEP CPN), or can have no coupon at all (ZERO COUPON)
- CompositeRating: the composite Rating of the Bond based on the main rating agencies (Moody's, S&P, Fitch): best Rating is AAA, the worse is D. Some bonds may not be rated (NR).
- MarketIssue: market standard classification of the Bond
- IndustrySector: industrial sector of the Issuer
- IndustrySubgroup: more detailed activity of the Issuer
- ActivityGroup / Region / Activity / RiskCaptain / Owner: internal BNPP Paribas Bond hierarchy: we have 3 main Activity Group: Corporate Bond activity in G10 countries (FLOW G10), in local markets (FLOW LOCAL MARKET). Business in Bonds with SupraNationals or on Covered Bonds is a separate activity called SAS & COVERED BONDS. Then those Business can be split by Region (EUROPE, AMERICAS...)...

The Customer Table

This table gives a few information of each Customer

- CustomerIdx: the index of the Customer

- Sector: the sector of activity of the Customer
- Subsector: the sub-sector of activity of the Customer
- Country: the country of incorporation of the Customer

The Market Table

This table gives historical pricing information for the Bond:

- IsinIdx: the index of the Bond
- DateKey: the Date in format YYYYMMDD at which pricing is done: pricing is the one contributed by BNP Paribas Trading floor at the end of the day.
- Price: price of the Bond expressed in percent: the cost in EUR of entering in a given Bond Transaction is approximately equal to $\text{NotionalEUR} * \text{Price} / 100$
- Yield: yield of the Bond expressed in percent: the yield represents the yearly expected return on investment when entering a Bond Transaction
- ZSpread: difference between the Yield of the Bond and the corresponding Risk-Free rate expressed in Percent. It represents the extra return asked by investor to compensate for the Risk of holding this Bond (as the Issuer may fail to pay the coupon and notional of the Bond)

The Macro Market Database

This database contains pricing information on different Equity, Foreign Exchange and Yield Curve instruments.

The instruments are:

- Equity: main stock indices as quoted on the corresponding exchange: for example DOWJONES_IND
- Foreign Exchange: the instrument FX_USD.CCY is the value of CCY expressed in USD: for example if FX_USD.EUR is worth 1.2 it means that 1 EUR is equal to 1.2 USD.
- Yield Curve: the interbank collateralized interest rate level (~risk free rate) in percent: for example a Swap_JPY5Y at 0.2 means that the 5Y Rate in JPY is 0.2%, and a MoneyMarket_CHF3M at -0.7 means that the rate at 3 months in CHF is worth -0.7%.

Sample submission file

This file contains two columns, the PredictionIdx corresponding to the ones found in the table `Challenge_20180423.csv` and CustomerInterest corresponding to the associated predicted probability.