Project proposalCómo tomar notas

Goal of the Project

The goal of the Project is to develop an R script to be able to download the calendar of the Forexfactory webpage. In R you can download data of different assets but macroeconomic data is not implemented in any package, at least in the packages that I know. So this data could be good for further analysis.

As the calendar is not showing all the results in the main page, I will have to tell the algorithm that we want to jump though next page until we arrive at the end of the range of dates. I mean that we can enter a set of ranges and the algorithm should be able to download the data of the calendar for all this dates.

Data source and technology.

The data source will be the calendar page of the https://www.forexfactory.com. This calendar shows the macroeconomic events related to the currencies market, a description of the event, country that will be affected by the event/new, the impact, and a little bit more of information.

Approach/technology that you will use

To download the data, I will use R and some libraries to download, manage and clean the data like rvest to scrap the data, dplyr to manage it, tidyr to create tidy data frames and maybe stringr to clean and manage strings.

How to plan to clean the data

Making sure that the fields that are numbers are really numbers, and cleaning extrange characters and extra spaces that a sentence could have.

How you would analyse the extracted information

This information could be useful for many purposes, for instance we could use this data as a time series and then use it as an external regressor of an ARIMA model. We could also see which events affect more the returns of one currency, let's say that we want to see how affect much a new of the ECB in the EURUSD or something similar.