Codebook for Attrition Dataset

Data Overview

Credentials

The data for each stock from S&P 500 is from Yahoo finance – https://finance.yahoo.com

Business goal

This data was collected to answer the following questions:

1. We wish to know on a daily basis which stocks will raise by tomorrow/various times, so that we can invest the money accordingly.

Data description

1. The data include information over 10 years of each stock from S&P 500.
2. Data Frame of 6 variables over 2800 rows for each stock (open, high, low, volume, close and adjusted).
3. We used in QUANTMOD R package in order to add some Technical indicators to improve our prediction. We added Technical indicators such as: EMA, ROC, RSI, CCI etc.

Variables description

Variable Name	Description	Type	Possible values
Date	Trading Date (format- dd/mm/YY)	Date	25/11/2015
Symbol	The name of the company	String	"AAPL"
Open	The open price at this trading date	Numeric	Positive, USD
High	The highest price reached in this trading date	Numeric	Positive, USD
Low	The lowest price reached in this trading date	Numeric	Positive, USD
Close	The close price at this trading date	Numeric	Positive, USD
Volume	The number of total shares that are traded on a given day	Numeric	Positive, USD
Change	The percent number of change in the price relative to the previous day price	Numeric	Positive/ Negative percent
Open Change	The percent number of change in the previous day close price relative to the open day price	Numeric	Positive/ Negative percent
Change within	The percent number of change in the open day price relative to the close day price	Numeric	Positive/ Negative percent
nextday_open_change	The percent number of change in the close price relative to the next open day price	Numeric	Positive/ Negative percent
nextday_close_change	The percent number of change in the close price relative to the close next day price	Numeric	Positive/ Negative percent
Market Cup	Total equity market value	Numeric	Positive, USD

RSI	The relative strength index (RSI) is a momentum indicator that compares the magnitude of recent gains and losses over a specified trading date to measure speed and change of price movements of a security	Numeric	Positive/ Negative percent
EMA	Indicator in technical analysis that helps smooth out price action by filtering out the "noise" from random price fluctuations. The two basic and commonly used MAs are the simple moving average (SMA), which is the simple average of a security over a defined number of time periods, and the exponential moving average (EMA), which gives bigger weight to more recent prices	Numeric	Positive, USD
S&P 500 close	The close price of S&P 500 index at this trading date	Numeric	Positive, USD
order	The right order in every day depending on the price of tomorrow	factor	"buy","sell","hold"
High_close	True if high price equal to close price	logical	"0", "1"

```
> str(TVIXDF)
'data.frame':    1759 obs. of  11 variables:
 $ Date  : Date, format: "2010-11-30" "2010-12-01" "2010-12-02" "2010-12-03" ...
 $ Open  : num  26567500 24995000 24577500 22090000 19942500 ...
 $ High  : num  28087500 25172500 24577500 22090000 20000000 ...
 $ Low   : num  26097500 24700000 21625000 19530000 18655000 ...
 $ Close : num  28087500 25172500 21647500 19695000 18922500 ...
 $ Volume: int  0 0 0 0 0 0 0 0 0 0 ...
 $ Change: num  NA -10.96 -15.09 -9.45 -4 ...
 $ Order : Factor w/ 3 levels "buy","hold","sell": NA 3 3 3 3 3 3 3 1 2 ...
 $ RSI   : num  NA NA NA NA NA NA NA NA NA NA ...
 $ EMA   : num  28087500 27744559 27027258 26164639 25312623 ...
 $ S.P500: num  1181 1206 1222 1225 1223 ...
> summary(TVIXDF)
      Date                 Open                High                Low                Close
 Min.   :2010-11-30   Min.   :       7   Min.   :       7   Min.   :       7   Min.   :       7
 1st Qu.:2012-08-27   1st Qu.:    1436   1st Qu.:    1501   1st Qu.:    1390   1st Qu.:    1430
 Median :2014-05-30   Median :   11575   Median :   11725   Median :   10500   Median :   11450
 Mean   :2014-05-28   Mean   : 1983509   Mean   : 2082883   Mean   : 1897832   Mean   : 1976873
 3rd Qu.:2016-02-27   3rd Qu.:  648750   3rd Qu.:  662500   3rd Qu.:  618750   3rd Qu.:  640000
 Max.   :2017-11-22   Max.   :26567500   Max.   :28087500   Max.   :26097500   Max.   :28087500

     Volume              Change            Order          RSI              EMA               S.P500
 Min.   :        0   Min.   :-35.3878   buy :541   Min.   :10.05   Min.   :       8   Min.   :1099
 1st Qu.:        0   1st Qu.: -4.6520   hold:387   1st Qu.:32.61   1st Qu.:    1561   1st Qu.:1406
 Median :     4200   Median : -1.1210   sell:830   Median :38.63   Median :   12053   Median :1882
 Mean   :  8700756   Mean   : -0.8628   NA's:  1   Mean   :40.16   Mean   : 2096633   Mean   :1813
 3rd Qu.: 11557850   3rd Qu.:  2.5022              3rd Qu.:46.38   3rd Qu.:  687362   3rd Qu.:2097
 Max.   :160781800   Max.   : 34.2533              Max.   :85.07   Max.   :28087500   Max.   :2599
                     NA's   :1                     NA's   :14
```

Who needs to review the business question?

Investment fund and stock traders