# Importing Financial Data

Mehmet F. Dicle       John Levendis

Loyola University New Orleans       Loyola University New Orleans

New Orleans, LA / USA       New Orleans, LA / USA

mfdicle@loyno.edu       jlevendi@loyno.edu

**Abstract.** This note describes two programs – *fetchyahooquotes* and *fetchyahookeystats* – which import historical financial data and key current financial statistics from Yahoo! Finance.

**Keywords:** st0001, finance, financial data, stocks, ETF, historical data

## 1 Introduction

Yahoo! Finance offers the public large amounts of financial and economic data. Often, small amounts of this data can be imported into Stata without much difficulty. However, to be able to use this data in econometric analysis, some automated download procedure may prove useful. Below, we describe two programs[1], *fetchyahooquotes* and *fetchyahookeystats*, that automate the process of importing financial data from Yahoo! Finance.

These two programs are useful for instructors who work with financial data during their lectures and need quick access to current data. They are especially useful for finance project assignments. Students do not have to spend long hours to download data into Excel spreadsheets and to merge them into portfolios. These programs are also important for researchers. Many teaching universities do not have subscriptions to costly finance databases. While Yahoo! Finance is not a substitute, it is a useful alternative. Finally, they are important for investors. *fetchyahooquotes* and *fetchyahookeystats* make financial data access fast and easy.

The *fetchyahooquotes* command is used to download time series of the more common financial statistics of multiple financial instruments. Thus, one could download the daily opening and closing prices over the past ten years of IBM and Microsoft stocks, for example. *Fetchyahookeystats* is similar to *fetchyhooquotes*, but downloads only the current day's key financial statistics for multiple financial instruments (for example, today's opening and closing prices for IBM and Microsoft).

---

1. *fetchyahooquotes* and *fetchyahookeystats* commands can be installed from http://www.loyno.edu/~mfdicle/stata

      st0001

## 2 The fetchyahooquotes and fetchyahookeystats commands

### 2.1 Syntax

fetchyahooquotes *namelist*, freq(d/w/m/v) $\left[\right.$ field(string) chg(ln/per/sper)

   save(filename) start(string) end(string) ff3$\left.\right]$

fetchyahookeystats *namelist*, field(string) $\left[\right.$ save(filename)$\left.\right]$

### 2.2 Options

namelist is a list of ticker symbols to be downloaded from Yahoo! Finance's API. Ticker symbols are separated by spaces. The ticker symbol will become part of the variable name. Regardless of whether the symbol includes special characters, ticker is typed (including the special character) exactly as it is used by Yahoo! Finance's API. Any special characters will be replaced for the variable name with underscores "_" since they are not allowed as Stata variable names.

freq(d/w/m/v) specifies the frequency of the historic price including (d)daily, (w)weekly, (m)monthly or (v)dividends only. Note: If the frequency is either daily, weekly or monthly then the variable with the symbol name is the adjusted closing price. If the frequency is to include only the dividends then the symbol name is the dividend payment.

chg(ln/per/sper) is the periodic return. Three different periodic changes can be calculated for the adjusted closing price; natural log difference, percentage change and symmetrical percentage change. The change is based on the continuous trading assumption. Thus, even though the tsset is the date (which may contain gaps), returns are calculated assuming there are no gaps in the data.

save(*filename*) is the output file name. Stata data file is created under the current working directory.

start(*string*) is the starting date for the prices (i.e. 1mar2010) in DMY format.

end(*string*) is the ending date for the prices (i.e. 23feb2011) in DMY format.

ff3 Fama/French daily factors are downloaded from "Kenneth R. French - Data Library[2]".

field(*string*), for fetchyahooquotes, is the variables to download in addition to the adjusted close and the date: These include: (o) open, (h) high, (l) low, (c) close, (v) volume.

field(*string*), for fetchyahookeystats, is the Yahoo! Finance's specific field code corre-

---

2. Available through http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html

sponding to a key statistic. These include (but are not limited to):

| | | | |
|---|---|---|---|
| a | Ask | l1 | Last Trade Price |
| b | Bid | m3 | 50-Day Moving Average |
| b4 | Book Value | m4 | 200-Day Moving Average |
| c | Percent Change | n | Name |
| c1 | Change | o | Open |
| d | Dividend per Share | p | Previous Close |
| d1 | Last Trade Date | p5 | Price/Sales |
| e | Earning per Share | p6 | Price/Book |
| f6 | Float Shares | q | Ex-Dividend Date |
| g | Day's Low | r | Price/Earnings |
| h | Day's High | s | Symbol |
| j | 52-Week Low | s7 | Short Ratio |
| j1 | Market Capitalization | v | Volume |
| j4 | EBITDA | x | Exchange |
| k | 52-Week High | y | Dividend Yield |

# 3 Using fetchyahooquotes to import historical prices

▷ **Example**

*Single company historic share prices*: This example uses *fetchyahooquotes* to import the adjusted daily closing price of Microsoft shares from the beginning of 2010 to the end of 2010. Fama/French daily factors are downloaded as well as the Microsoft share prices. The program also calculates the log difference change of the daily closing share prices.

```
. fetchyahooquotes  MSFT, freq(d) chg(ln) start("01jan2010") end("31dec2010") ff3
Fama/French daily factors are downloaded from `Kenneth R. French - Data Library´.
MSFT is downloaded.

. sum
```

| Variable | Obs | Mean | Std. Dev. | Min | Max |
|---|---|---|---|---|---|
| date | 252 | 18446.69 | 104.8328 | 18266 | 18627 |
| adjclose_M~T | 252 | 26.26262 | 2.06914 | 22.33 | 30.32 |
| ln_MSFT | 251 | −.0003306 | .013845 | −.0419884 | .0515297 |
| ff3_Mkt_RF | 252 | .0007242 | .0117564 | −.0408 | .0453 |
| ff3_SMB | 252 | .000479 | .00558 | −.0202 | .0165 |
| ff3_HML | 252 | −.0000512 | .0051936 | −.0162 | .0154 |
| ff3_RF | 252 | 7.94e-07 | 2.71e-06 | 0 | .00001 |

By specifying the *chg(ln)* option, we requested to calculate the log difference changes in these prices. With the *ff3* option, Fama/French factors are also downloaded. Fama/Frenh original variable names are used with prefix of "ff3_" [3].

▷ **Example**

---

3. Explanations of these variables are available through
   http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/Data_Library/f-f_factors.html

*Multiple companies' historic share prices*: This example downloads monthly historical data on the closing share prices of Microsoft and IBM, and computes the log difference changes. The monthly data is actually the data for the first day of trading in the month.

```
. fetchyahooquotes MSFT IBM XYZ, freq(m) chg(ln)
MSFT is downloaded.
IBM is downloaded.
XYZ does not have sufficient number of observations.

. sum
```

| Variable | Obs | Mean | Std. Dev. | Min | Max |
|---|---|---|---|---|---|
| date | 590 | 9679.19 | 5175.287 | 732 | 18630 |
| adjclose_M~T | 299 | 14.43371 | 12.00902 | .08 | 46.74 |
| ln_MSFT | 298 | .0196599 | .1049801 | −.4211015 | .430783 |
| adjclose_IBM | 589 | 32.60188 | 37.39051 | 1.55 | 155.5 |
| ln_IBM | 588 | .0070381 | .0708084 | −.303776 | .3030213 |

```
. list in 1/3
```

| | date | adjclo~T | ln_MSFT | adjclo~M | ln_IBM |
|---|---|---|---|---|---|
| 1 | 02jan1962 | . | . | 2.48 | . |
| 2 | 01feb1962 | . | . | 2.46 | −.0080972 |
| 3 | 01mar1962 | . | . | 2.44 | −.0081633 |

Since Microsoft did not exist in 1962, missing values are reported. Since symbol XYZ does not exist, no data could be downloaded for XYZ.

▷ **Example**

*Multiple companies' indices and a foreign symbol's historic values*: This example downloads the adjusted closing prices of IBM and Google, as well as the S&P-500 (denoted as ^GSPC) and shares of BMW.DE (BMW shares that are listed on the XETRA in Germany). We requested to download, not just the closing prices of these, but also the daily high, low, and opening prices.

```
. fetchyahooquotes IBM GOOG ^GSPC BMW.DE, freq(d) field(h l o)
IBM is downloaded.
GOOG is downloaded.
^GSPC is downloaded.
BMW.DE is downloaded.

. sum
```

| Variable | Obs | Mean | Std. Dev. | Min | Max |
|---|---|---|---|---|---|
| date | 15430 | 7554.218 | 6451.202 | -3650 | 18648 |
| open_IBM | 12350 | 191.5857 | 139.1102 | 41 | 649 |
| high_IBM | 12350 | 193.2161 | 139.9349 | 41.75 | 649.88 |
| low_IBM | 12350 | 190.052 | 138.4056 | 40.63 | 645.5 |
| adjclose_IBM | 12350 | 32.50723 | 37.26821 | 1.4 | 155.8 |
| open_GOOG | 1619 | 429.4347 | 133.5289 | 99.19 | 741.13 |
| high_GOOG | 1619 | 434.461 | 134.1036 | 101.74 | 747.24 |
| low_GOOG | 1619 | 423.8753 | 132.3732 | 95.96 | 725 |
| adjclose_G~G | 1619 | 429.1093 | 133.2038 | 100.01 | 741.79 |
| open__GSPC | 15363 | 388.8959 | 453.9477 | 16.66 | 1564.98 |
| high__GSPC | 15363 | 391.5098 | 457.0745 | 16.66 | 1576.09 |
| low__GSPC | 15363 | 386.15 | 450.5819 | 16.66 | 1555.46 |
| adjclose__~C | 15363 | 388.9835 | 454.0107 | 16.66 | 1565.15 |
| open_BMW_DE | 2085 | 36.26493 | 7.734583 | 17.28 | 64.98 |
| high_BMW_DE | 2085 | 36.72145 | 7.740518 | 17.82 | 65.49 |
| low_BMW_DE | 2085 | 35.7866 | 7.736921 | 16 | 64.3 |
| adjclose_B~E | 2085 | 34.90521 | 7.879823 | 16.71 | 64.8 |

Since variable names in Stata cannot have special characters such as "`.`", and since our convention is to include the ticker name in the variable names, this presents a problem. That is, we would like to report the daily open price of BMW.DE as a variable entitled "*open_BMW.DE*", however Stata does not allow periods in variable names. Thus, *fetchyahooquotes* replaces special characters with an underscore: *open_BMW.DE* becomes *open_BMW_DE*. Likewise, *open_^GSPC* becomes *open__GSPC*.

▷ **Example**

*Multiple companies' historic dividends*: This example downloads the dividend payments of IBM, BMW and Ford.

```
. fetchyahooquotes IBM VWS.CO BMW.DE F, freq(v)
IBM is downloaded.
VWS.CO does not have sufficient number of observations.
BMW.DE does not have sufficient number of observations.
F is downloaded.

. sum
```

| Variable | Obs | Mean | Std. Dev. | Min | Max |
|---|---|---|---|---|---|
| date | 312 | 10583.16 | 4671.158 | 767 | 18847 |
| dividends_IBM | 196 | .1837389 | .1535859 | .001 | .75 |
| dividends_F | 116 | .2428889 | 1.113203 | .01907 | 12.06693 |

# 4   Using fetchyahookeystats to import historical prices

▷ **Example**

*Using fetchyahookeystats to import key statistics for multiple companies*: This example uses *fetchyahookeystats* to download today's (Jan. 24$^{th}$, 2011) key statistics for IBM, Google, BMW and the S&P-500.

```
. fetchyahookeystats IBM GOOG ^GSPC BMW.DE, field(n s l1 a b d1 g h k e n s k e L)
save(my_portfolio)
```

```
    (17 vars, 4 obs)
```

Under the *field* option, we specified that we wanted the full name of the item (n), the ticker symbol (s), the last traded price (l1), the day's low (g), the day's high (h), the 52-week high (k), and the earnings per share (e). Note that, n, s, k and e are included twice into the field option. Also, L is included into the field option which is not listed as a known Yahoo! Finance key statistics field.

```
    . describe

                        storage   display   value
    variable name       type      format    label    variable label
    Symbol              str6      %9s
    Name                str17     %17s
    Last_Trade_Pr~e     float     %9.0g
    Ask                 double    %10.0g
    Bid                 double    %10.0g
    Days_Low            float     %9.0g
    Days_High           float     %9.0g
    _52_Week_High       float     %9.0g
    Earnings_per_~e     double    %10.0g
    Unknown_Field_1     str6      %9s
    Last_Trade_Date     double    %td
    Sorted by:
```

Notice that the fields options that are included twice (n, s, k and e) are ignored for their second time. The option L is downloaded and named Unknown_Field_1.

```
    . sum

        Variable │     Obs        Mean    Std. Dev.       Min         Max
        Symbol   │       0
          Name   │       0
    Last_Trad~ce │       4    528.4316     561.952      55.23     1290.27
           Ask   │       2     332.285    391.7301      55.29      609.28
           Bid   │       2     332.205    391.6735      55.25      609.16
      Days_Low   │       4    523.3575     559.071       54.4     1282.47
     Days_High   │       4    529.6325    562.3061      55.83     1291.14
    _52_Week_H~h │       4    540.3225    563.9871      65.49     1296.06
   Earnings_p~e  │       3       12.61    13.18882          0       26.31
  Unknown_Fi~1   │       0
   Last_Trad~te  │       4       18651           0      18651       18651
```

**About the authors**

Mehmet F. Dicle is an assistant professor of finance at Loyola University New Orleans. John Levendis is an assistant professor of economics at Loyola University New Orleans.