

国内マーケットデータを無料で手に入れる

01_Convert_PanDiskData_to_CSV.exe

02_convert_DataSouko_to_CSV.exe

03_Divide_CSVDData_to_Eachstock.exe

04_Divide_DataSouko_to_Eachstock.exe

の使い方

目次

国内マーケットデータを無料で手に入れる.....	1
この資料について.....	2
著作権について.....	3
入手するデータについて.....	3
2000 年以降のデータ.....	3
2000 年以前のデータ.....	4
行う作業の流れ.....	4
2000 年以前のデータも使う場合.....	5
2000 年以降のデータだけ使う場合.....	5
2000 年以前のデータも使う方法.....	5
2000 年以降のデータだけを使う方法.....	20
新しいデータを追加する方法.....	26
問合せ先(ウェブサイトのお問い合わせフォーム).....	31

この資料について

日本国内の(株式・先物)マーケットのデータについて

無料データの取得先、取得方法、データの加工と保存について、紹介しています。

また、データの加工と保存を行う次の4つのプログラムが必要となります。

この資料はこれらのプログラムの使い方の説明も兼ねています。

01_Convert_PanDiskData_to_CSV.exe(9ページめ参照)

02_Convert_DataSouko_to_CSV.exe(13ページめ参照)

03_Divide_CSVData_to_Eachstock.exe(16ページめ参照)

04_Divide_DataSouko_to_Eachstock.exe(22ページめ参照)

プログラムは次のウェブサイトまたは Vector のサイトからダウンロードできます。

この資料に書かれた作業を行うと、

マーケットデータが銘柄コード別のファイルとして保存されます。

データのフォーマットは、

各項目の数値がコンマで区切られ、データごとに改行された CSV 形式となります。

1 行には、日足のデータとして

日付, コード, 始値, 高値, 安値, 終値, 出来高,

の順で値が書かれます。

このとき日付は YYYYMMDD(西暦月日)となっています。

著作権について

この資料および

01_Convert_PanDiskData_to_CSV.exe

02_Convert_DataSouko_to_CSV.exe

03_Divide_CSVData_to_Eachstock.exe

04_Divide_DataSouko_to_Eachstock.exe

については <http://t.12enjoy.net> の作者(運営管理者)が著作権をもっています。

これらが無断で販売・雑誌等への掲載、ウェブや記録メディアによる配布をすることは禁じます。

入手するデータについて

2000 年以降のデータ

株価データ倉庫

<http://www.geocities.co.jp/WallStreet-Stock/9256/data.html>

ここに掲載されているデータをダウンロードします。

(この時点ではまだダウンロードしないでください。)

このデータは、

株についてはファイルが日付ごとに分かれています。

各ファイルの中にはその当日のすべての上場株の日足データが一覧になっています。

ファイルの 1 行目は

YYYYMMDD の形で日付が書かれ

2 行目以降は

コード 会社名 始値 高値 安値 終値 出来高

と各項目が TAB で区切られて書かれています。

データ倉庫のデータに関しては、1ファイル内には各銘柄は東証と大証の両方が入っているわけではなく”代表的な方”どちらか1つが入っています。

2000 年以前のデータ

2000 年以前のデータについてはベクターのサイト

<http://www.vector.co.jp/soft/win95/business/se312808.html>

に「**パンローリング 相場データ CD-ROM 国内相場版**」というデータがあり、これを利用します。

このページに

”東証 1・2 部銘柄の 1975 年～、大証 1・2 部銘柄 1987 年 8 月～2003 年 9 月の日足データ
(四本値・出来高、東証銘柄には権利落ち乗数を含む)……を、タブ区切りテキストファイルで収録
しました。”

と書いてあります。

つまりダウンロードしたデータは銘柄コードごとのテキストファイルになっています。

(データ倉庫のデータとは異なる、ということです)

ファイルの中身は

日付 始値 高値 安値 終値 出来高

の順で書いてあり、日付は YYYYMMDD の形式になっています。

各テキストファイルの名前は

コード+市場の頭文字

となっています。

東証であればコード T、大証であればコード O です。

行う作業の流れ

2000 年以前の年月から現在までのデータがほしい場合には、上記の2つのデータ

(データ倉庫とパンローリング CD)を結合することになります。

本来ならばこの2つのデータの整合性がとれることを確認したうえで合わせるべきですが、考慮していません。

これから行う作業は次のようになります。

2000 年以前のデータも使う場合

1. 「パンローリング 相場データ CD-ROM 国内相場版」のダウンロード
2. 「データ倉庫」からデータをダウンロード
3. 1のデータの形式を変換し、1999.12.31 までのデータのみをまとめて1つの CSV ファイルにする。

この際、コードが同じで東証 T や大証 O など複数のものがあるものについては、東証以外のものは捨てる。また 2000.1.4 時点で存在しないコードのものも破棄する。

4. データ倉庫のデータの形式を変換し、2で作成した CSV ファイルへデータを加える。
5. 4のファイルからコードごとの CSV ファイルを作成する。

2000 年以降のデータだけ使う場合

- 1 「データ倉庫」からデータをダウンロード
- 2 データの形式を変換し、さらにコードごとの CSV ファイルを作成する

それでは実際の作業手順を説明します。

2000 年以前のデータも使う方法

1. 「パンローリング 相場データ CD-ROM 国内相場版」のダウンロード

<http://www.vector.co.jp/soft/win95/business/se312808.html>

上記のページから自己解凍方式の圧縮ファイル pandata1.exe をダウンロードし、適当なフォルダに解凍する。

解凍後2つのフォルダと2つのファイルができる。

先物については Futures、株については STOCK のフォルダの中にデータがある。

2. 「データ倉庫」からのダウンロード

<http://www.geocities.co.jp/WallStreet-Stock/9256/data.html>

このページから圧縮されたデータをダウンロードするが、データ数が多いため、一括してダウンロードする Fuerza というソフトを使うのがお勧め。

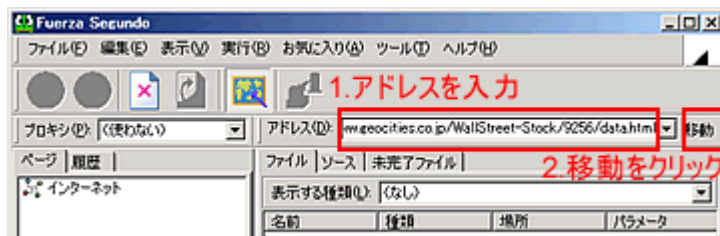
<http://www1.bbiq.jp/jylki/>

上記の作者のページから Furza をダウンロードし、適当なフォルダに解凍する。

Isabel.exe のアイコンをダブルクリックして Fuerza を起動する。

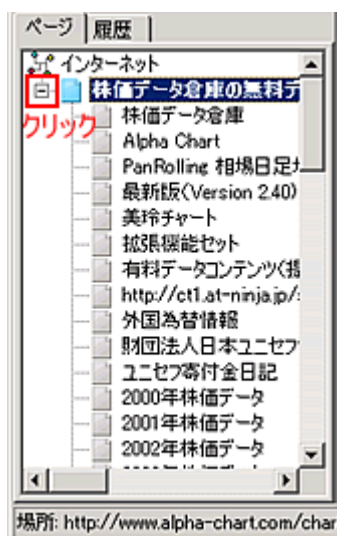


アドレスの入力フォームに上記のデータ倉庫のアドレスを入力し、フォーム横の移動ボタンをクリックする。



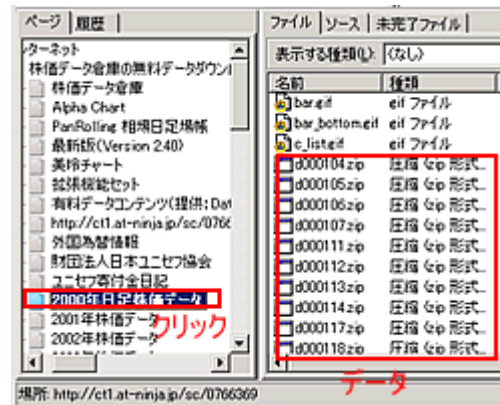
”インターネット”の下に「株価データ倉庫の無料データ」のファイルが現れるので、その左横の＋をクリック。

すると入力したアドレスのページにリンクしたファイル類が表示される。

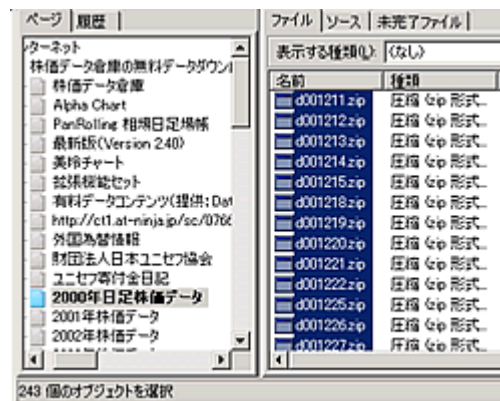


ここに 2000 年株価データ～2008 年株価データがあるのがわかる。

このうち任意の年をクリックすると、右側に日にちごとの圧縮データが一覧で表示される。



Zip 形式の圧縮ファイルをすべて選ぶ。



ダウンロード用のボタンをクリックし、適当なフォルダを保存先として指定する。



ダウンロード中は次のウィンドウが表示される。

ダウンロードが完了したら閉じるのボタンをクリックする。



ほしい年すべてのファイルを1つのフォルダにダウンロードする。

ダウンロードが終わったら Fuerza を終了する。

ダウンロードされたファイルは圧縮ファイルなので解凍する必要がある。

今回は 7-zip という圧縮解凍ソフトを使い解凍するが、別に他のソフトでも構いません。

圧縮ファイルをすべて選択し、右クリックして現れたメニューから「7-zip」を選び「現在のフォルダに解凍」を選択すればよい。



解凍時に上書きしていいか？というメッセージが現れたら「はい」を選び、～は解凍できなかったというメッセージが現れても気にする必要はない。

解凍終了直後、圧縮ファイルが選択された状態にあるので、そのまま削除(Del キー)を行うと圧縮ファイルは削除される。

これでデータのダウンロードは終了。

3. パンローリング 相場データを変換する。

1のデータ「パンローリング 相場データ CD-ROM 国内相場版」略して PanDiskData を自作のプログラムで形式を変換し、1999.12.31 までのデータをまとめて1つの CSV ファイルにする。

この際、2の「データ倉庫」の 2000.1.4 のファイル d000104.txt を参照にして、1の PanDiskData 中にコードが同じで東証 T や大証 O など複数のものがあるものについては、東証以外のものは自動で捨てる。

また 2000.1.4 時点で存在しないコードのものも自動で破棄する。

これは後で2の「データ倉庫」のデータと結合するために必要。

そして PanDiskData ではデータに欠損もしくは不明なところがある場合、始値，高値，安値，終値，出来高のいずれかが0と表示されている。

これも不要であるため自動で捨てる。

自作したプログラムは 01_Convert_PanDiskData_to_CSV.exe と名づけた。

01_Convert_PanDiskData_to_CSV.zip

zip 形式で圧縮されているので 7zip などのソフトウェアで解凍する。

お決まりの文句ですが、このプログラムは私個人用に作ったものであるため、使用する人は自己責任でお使いください。

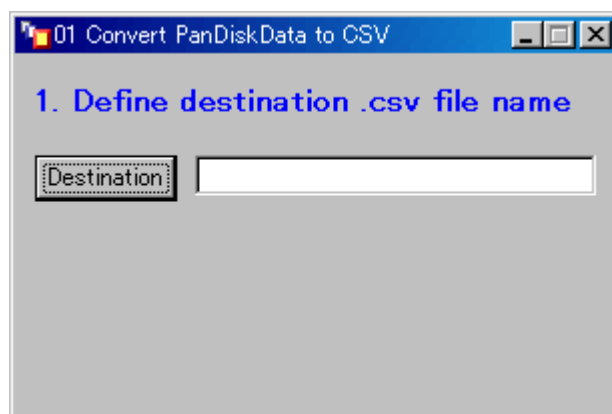
不具合がある可能性もあり、動作を保障するものではありません。

以下に使用方法を書く。

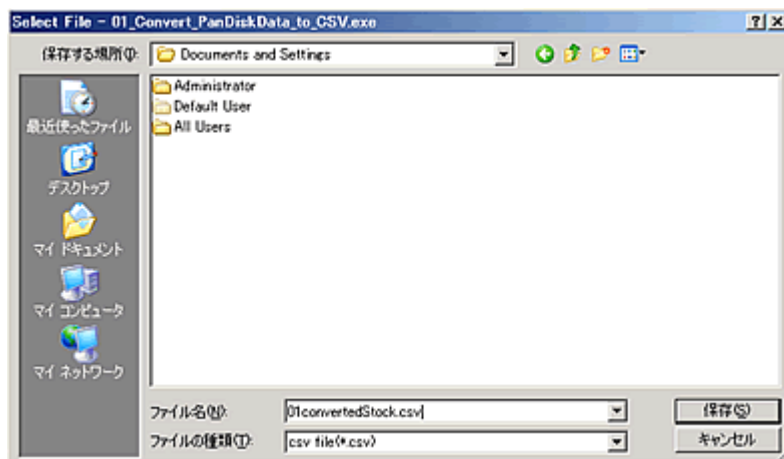
01. 01_Convert_PanDiskData_to_CSV.exe のアイコンをダブルクリックして起動する。



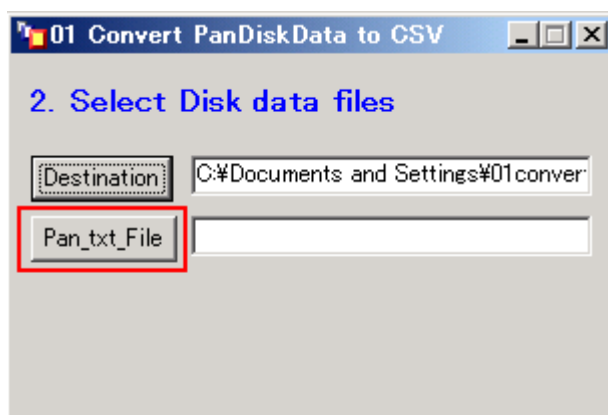
02. 次のウィンドウが現れるので Destination のボタンをクリックする。



03. 作成する CSV ファイルの保存先と名前を設定する。名前の後には.csv を必ずつける。「保存(S)」ボタンを押す。



04. Pan_txt_File のボタンをクリックする。

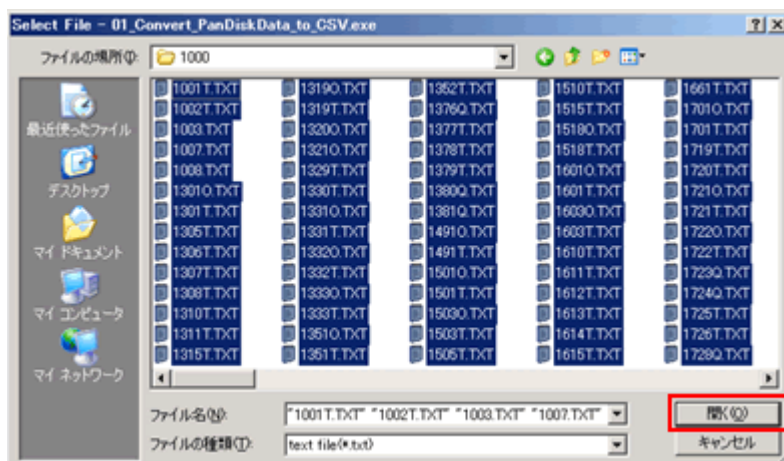


05. PanDiskData 中のデータを選択する。

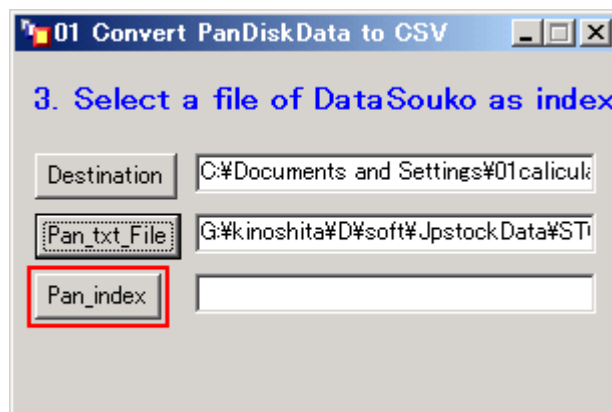
PanDiskData のファイルは Stock フォルダの中に 1000,2000...,8000,9000 という名前の計9つのフォルダがあり、その中に入っている。

まずは 1000 のフォルダの中のファイルすべてを選択する。

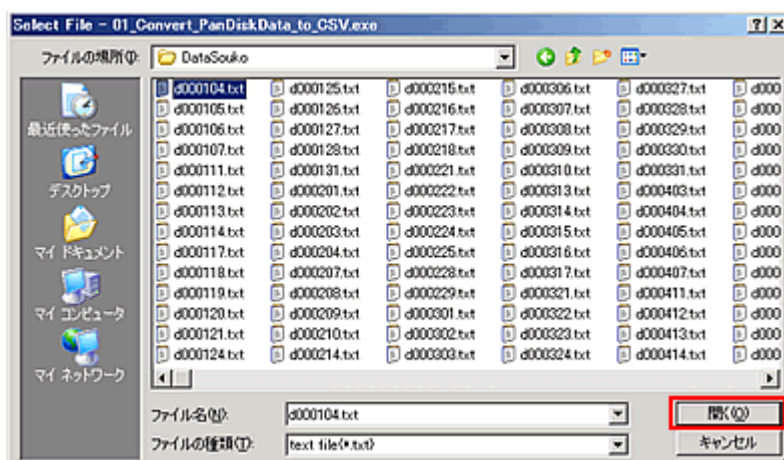
そして「開く(O)」ボタンをクリックする。



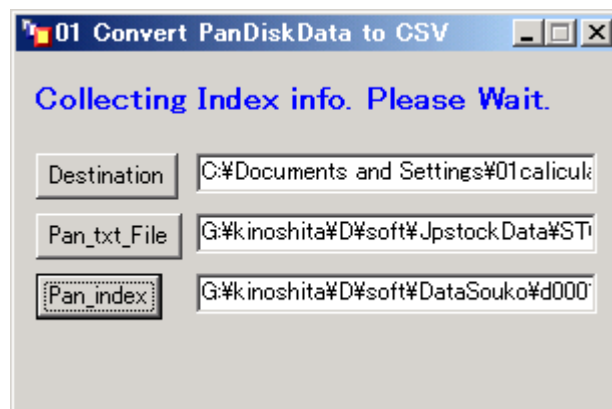
06. Pan_index のボタンをクリックする。



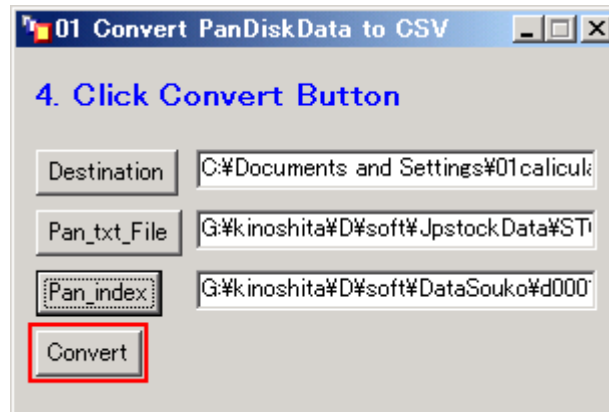
07. データ倉庫の d000104.txt を選択し、「開く(O)」ボタンをクリックする。



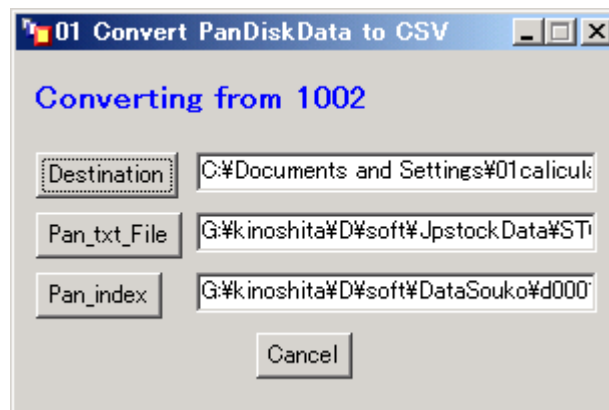
08. Collecting Index info. Please Wait.. のメッセージが出るのでしばらく待つ。



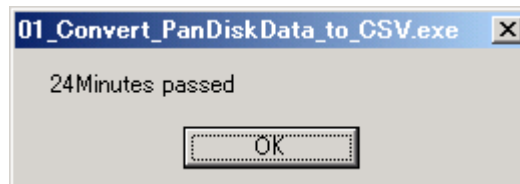
09. 4.Click Convert Button のメッセージが出たら Convert ボタンをクリックする。



10. Converting from ”ファイル名” のメッセージが出てファイルの変換・作成が実行される。

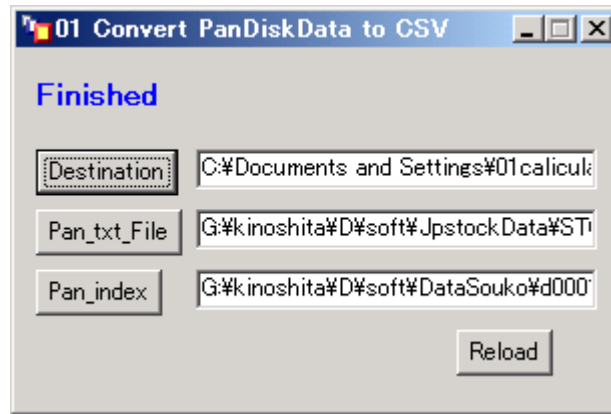


11. 変換・作成が終了すると、作業にかかった時間が表示される。Ok ボタンをクリックする。



12. Finished のメッセージが表示される。終了する場合はウィンドウ右上のクローズボタンをクリックする。

続けて作業する(例えば 2000 のフォルダの中のファイルについて行う)場合は Reload ボタンをクリックする。



13. Reload ボタンをクリックした場合には、02. からの手順を繰り返すことになる。

今回、PanDiskData 中の 1000,2000...,8000,9000 という計9つのフォルダ内のデータについて、02から13までの手順を行ってください。

このときの注意として、03. で保存する CSV ファイルの名前は毎回同じにします。

05. では1回目は 1000 フォルダ内のデータ、2回目は 2000 フォルダ内のデータ、...9回目は 9000 フォルダ内のデータを指定しますが、他はすべて同じ作業です。

14. すべてのデータを1つの CSV ファイルにまとめることができたら作業は終了となる。

4. データ倉庫のデータを変換し1つにまとめる。

データ倉庫のサイトからダウンロードしたデータを自作のプログラムで形式を変換しまとめて1つの CSV ファイルにする。

「3. パンローリング 相場データを変換する」で作成した CSV ファイルへデータ倉庫のデータを結合するのが目的。

データに欠損もしくは不明なところがある場合、始値，高値，安値，終値，出来高のいずれかが 0と表示されている。

これは不要であるため自動で捨てる。

自作したプログラムは 02_Convert_DataSouko_to_CSV.exe と名づけた。

02_Convert_DataSouko_to_CSV.zip

zip 形式で圧縮されているので解凍する。

お決まりの文句ですが、このプログラムは私個人用に作ったものであるため、使用する人は自己責任でお使いください。

不具合がある可能性もあり、動作を保障するものではありません。

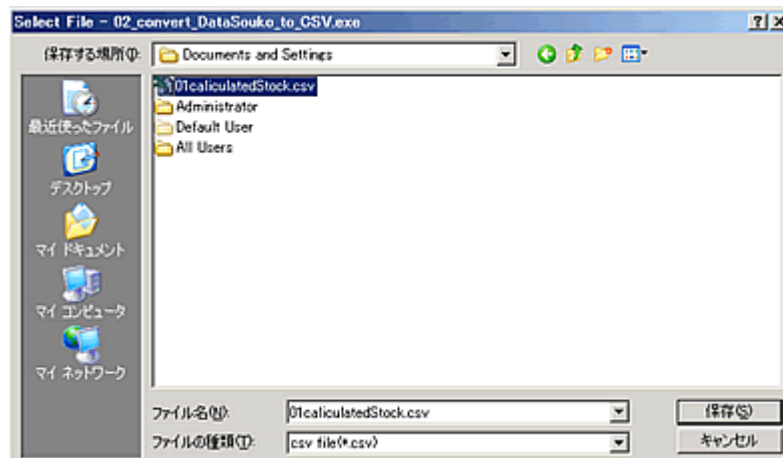
以下に使用方法を書く。

01. 02_Convert_DataSouko_to_CSV.exe のアイコンをダブルクリックして起動する。

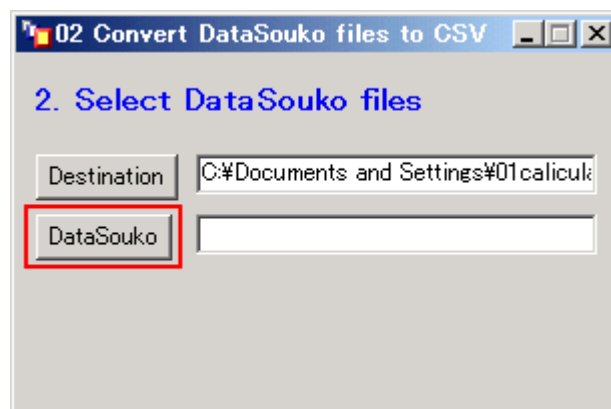


02. 次のウィンドウが現れるので Destination のボタンをクリックする。

03. 「3. パンローリング 相場データを変換する」で作成した PanDisData の CSV ファイルを指定する。「保存(S)」ボタンを押す。



04. DataSouko のボタンをクリックする。

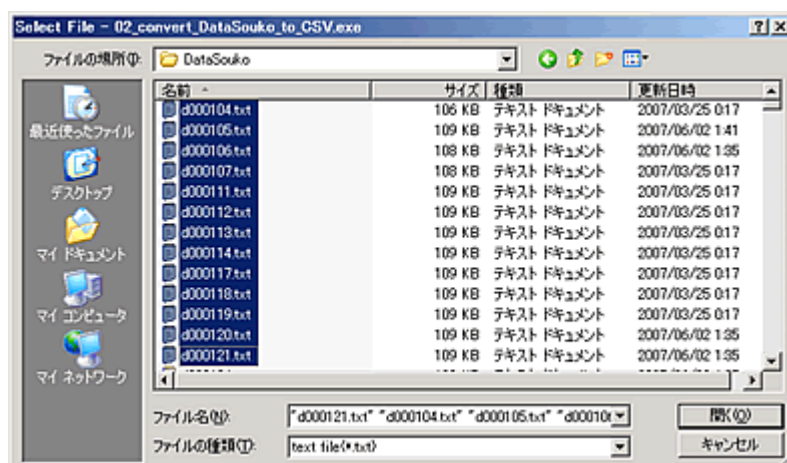


05. ダウンロードと解凍を済ませたデータ倉庫のデータを選択する。

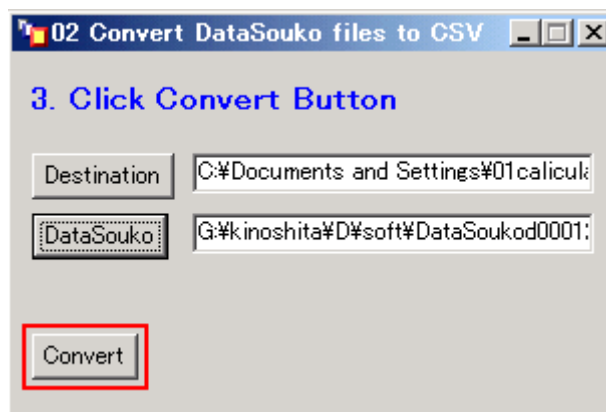
そして「開く(O)」ボタンをクリックする。

このときメモリーもしくは字数制限が原因と思われる理由で、データをすべて選ぶとエラーが出る場合があるので、データを2つか3つに分けて選択し、2～9の作業を繰り返し行う必要があります。

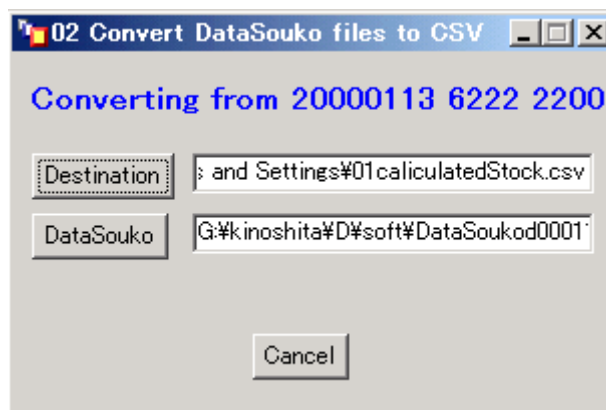
また、データを選ぶ順序は時系列順にしてください。



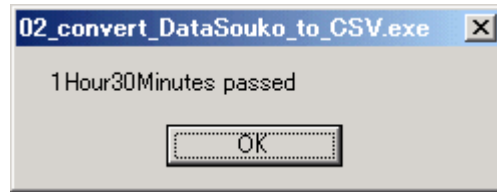
06. 3.Click Convett Button のメッセージが出たら Convert ボタンをクリックする。



07. Converting from データ内容 のメッセージが出てファイルの変換・作成が実行される。



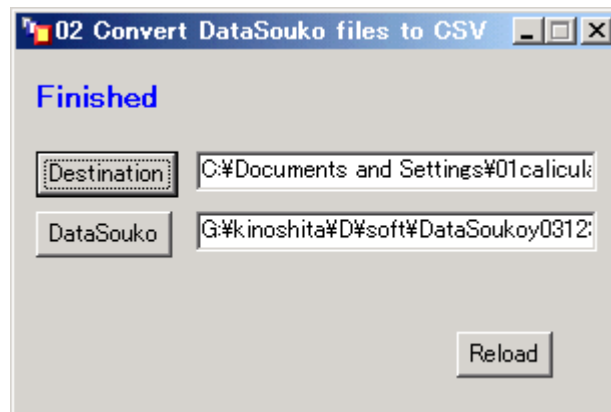
08. 変換・作成が終了すると、作業にかかった時間が表示される。OK ボタンをクリックする。



09. Finished のメッセージが表示される。

終了する場合はウィンドウ右上のクローズボタンをクリックする。

続けて作業する場合は Reload ボタンをクリックする。



10. Reload ボタンをクリックした場合には、02. からの手順を繰り返すことになる。

11. すべてのデータを1つの CSV ファイルにまとめることができたなら作業は終了となる。

5. コード毎の CSV ファイルに分割する

いよいよデータ準備の最終段階。

「4. データ倉庫のデータを変換し1つにまとめる。」で作成した CSV ファイルには PanDiskData とデータ倉庫のデータがまとめて入っている。

この CSV ファイルのデータを分割し、各銘柄コードごとに時系列データの CSV ファイルとして保存する。

そのために新たに自作したプログラムを使用する。

自作したプログラムは 03_Divide_CSVDData_to_Eachstock.exe と名づけた。

03_Divide_CSVDData_to_Eachstock.zip

ここからダウンロードできるが、zip 形式で圧縮されているので解凍する。

お決まりの文句ですが、このプログラムは私個人用に作ったものであるため、使用する人は自己責任でお使いください。

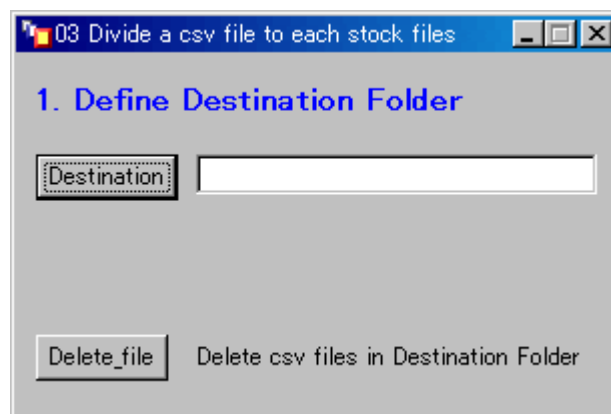
不具合がある可能性もあり、動作を保障するものではありません。

以下に使用方法を書く。

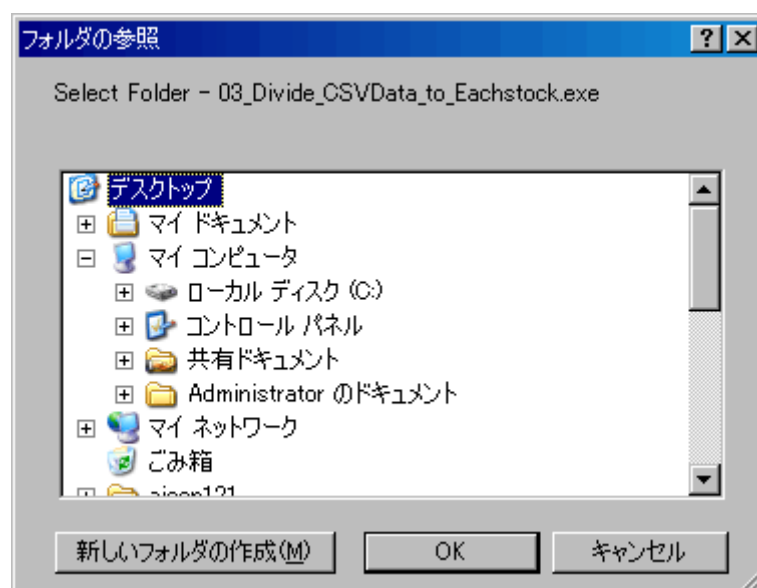
01. 03_Divide_CSVData_to_Eachstock.exe のアイコンをダブルクリックして起動する。



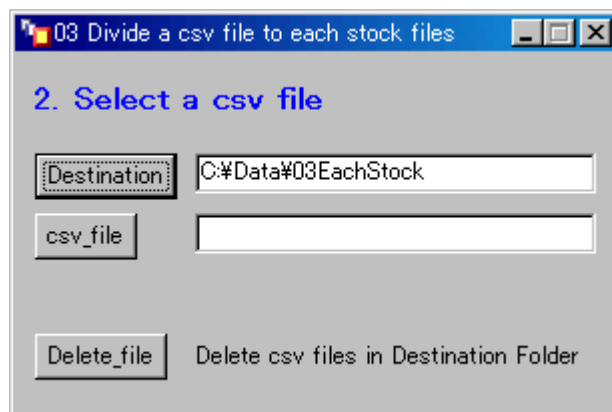
02. 次のウィンドウが現れるので Destination のボタンをクリックする。



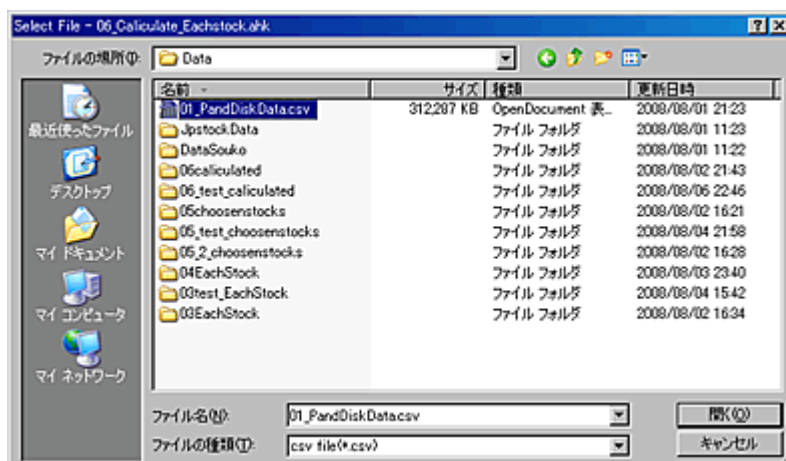
03. 作成される CSV ファイルを保存するフォルダ(ディレクトリ)を指定する。OK ボタンをクリックする。



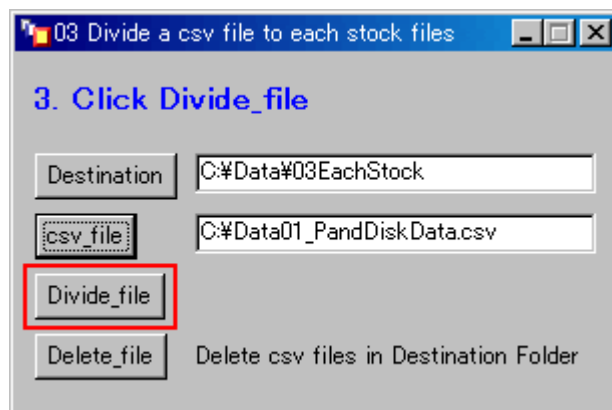
04. csv_file のボタンをクリックする。



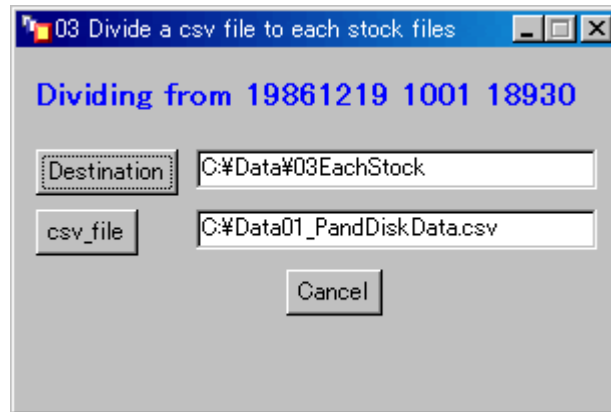
05. PanDiskData とデータ倉庫のデータから作った CSV ファイルを指定する。そして「開く(O)」ボタンをクリックする。



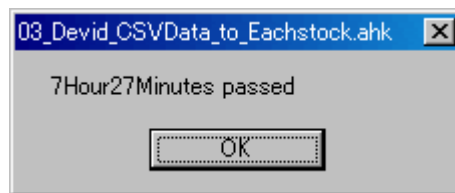
06. Divide_file ボタンをクリックする。



07. Dividing from "データ内容" のメッセージが出て CSV ファイルの分割が実行される。



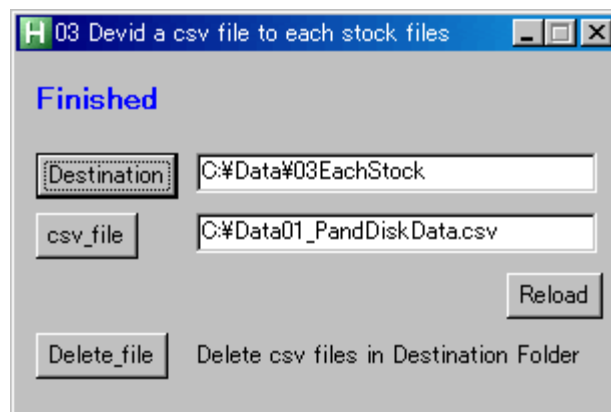
08. 変換・作成が終了すると、作業にかかった時間が表示される。OK ボタンをクリックする。



09. Finished のメッセージが表示される。

終了する場合はウィンドウ右上のクローズボタンをクリックする。

続けて作業する場合は Reload ボタンをクリックする。



10. Reload ボタンをクリックした場合には、02. からの手順を繰り返すことになる。

11. コード毎の CSV ファイルが作成されたら作業は終了となる。

(画像では英語 Divide が Devid と間違っています。)

2000 年以降のデータだけを使う方法

1. 「データ倉庫」からのダウンロード

<http://www.geocities.co.jp/WallStreet-Stock/9256/data.html>

このページから圧縮されたデータをダウンロードするのだが、データ数が多いため、一括してダウンロードする Fuerza というソフトを使うのがお勧め。

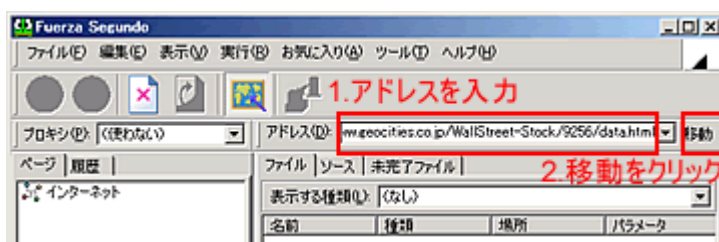
<http://www1.bbiq.jp/jylki/>

上記の作者のページから Furza をダウンロードし、適当なフォルダに解凍する。

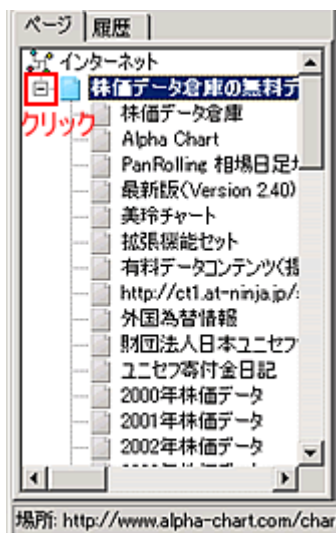
Isabel.exe のアイコンをダブルクリックして Fuerza を起動する。



アドレスの入力フォームに上記のデータ倉庫のアドレスを入力し、フォーム横の移動ボタンをクリックする。

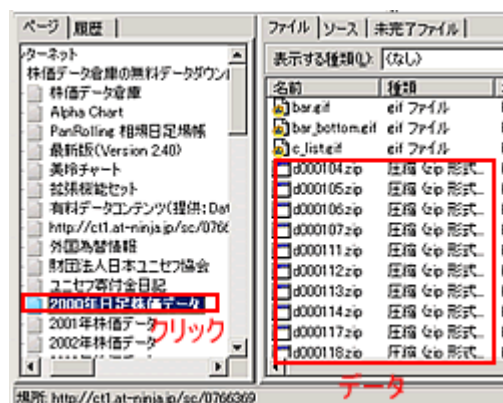


”インターネット”の下に「株価データ倉庫の無料データ」のファイルが現れるので、その左横の＋をクリック。すると入力したアドレスのページにリンクしたファイル類が表示される。

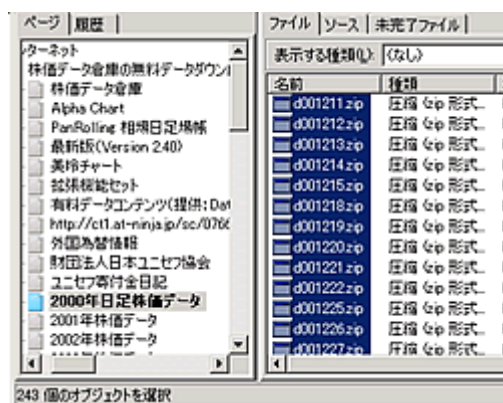


ここに 2000 年株価データ～2008 年株価データがあるのがわかる。

このうちほしい年のアイコンをクリックすると、右側に日にちごとの圧縮データが一覧で表示される。



Zip 形式の圧縮ファイルをすべて選ぶ。



ダウンロード用のボタンをクリックし、適当なフォルダを保存先として指定する。



ダウンロード中は次のウィンドウが表示される。

ダウンロードが完了したら閉じるのボタンをクリックする。



ほしい年すべてのファイルを1つのフォルダにダウンロードする。

ダウンロードが終わったら Fuerza を終了する。

ダウンロードされたファイルは圧縮ファイルなので解凍する必要がある。

今回は 7-zip という圧縮解凍ソフトを使い解凍するが、別に他のソフトでも構いません。

圧縮ファイルをすべて選択し、右クリックして現れたメニューから「7-zip」を選び「現在のフォルダに解凍」を選択すればよい。



解凍時に上書きしていいか？というメッセージが現れたら「はい」を選び、～は解凍できなかったというメッセージが現れても気にする必要はない。

解凍終了直後、圧縮ファイルが選択された状態にあるので、そのまま削除(Del キー)を行うと圧縮ファイルは削除される。

これでデータのダウンロードは終了。

2. データ倉庫のデータからコード毎の CSV ファイルを作成する

自作したプログラムは 04_Divide_DataSouko_to_Eachstock.exe と名づけた。

04_Divide_DataSouko_to_Eachstock.zip

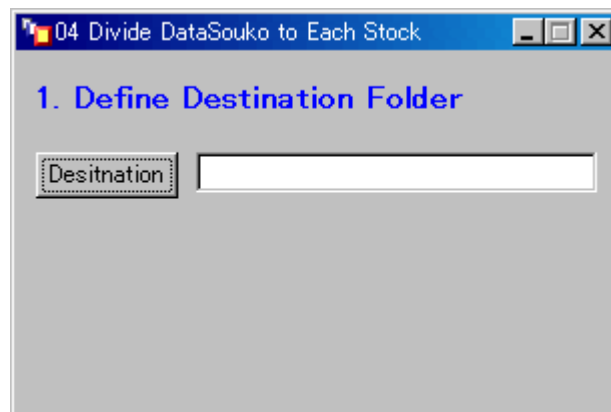
Zip 形式で圧縮されているので解凍する。

お決まりの文句ですが、このプログラムは私個人用に作ったものであるため、使用する人は自己責任でお使いください。不具合がある可能性もあり、動作を保障するものではありません。

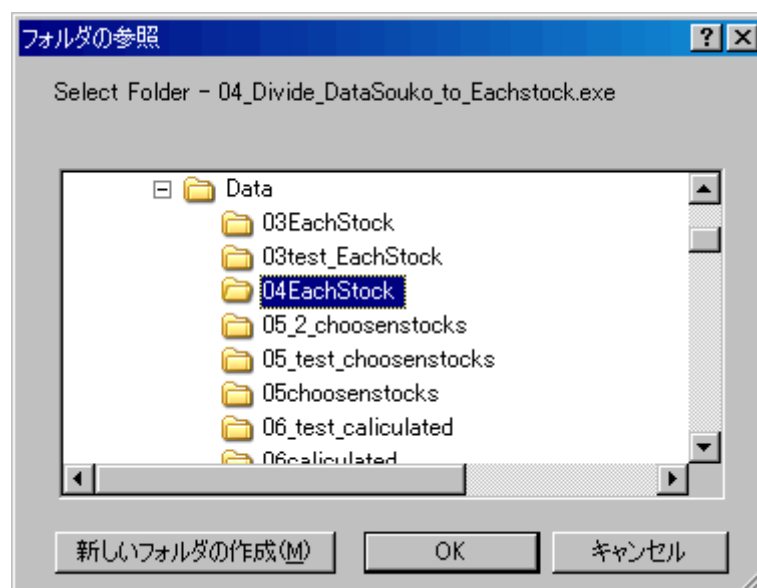
01. 04_Divide_DataSouko_to_Eachstock.exe のアイコンをダブルクリックして起動する。



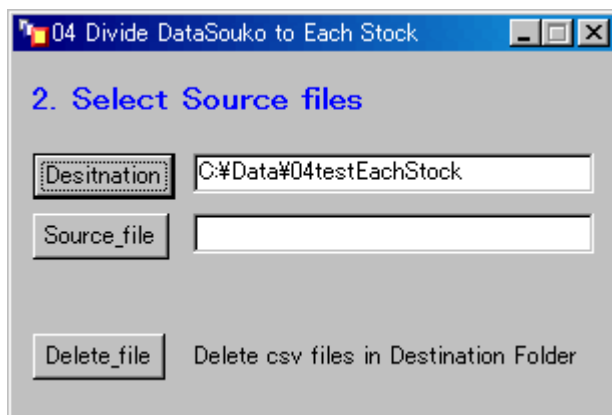
02. 次のウィンドウが現れるので Destination のボタンをクリックする。



03. 作成される CSV ファイルを保存するフォルダ(ディレクトリ)を指定する。OK ボタンをクリックする。



04. Source_file のボタンをクリックする。

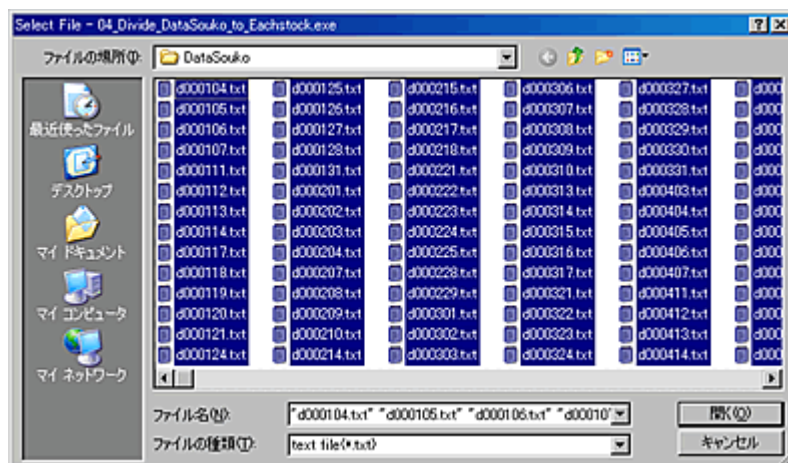


05. ダウンロードと解凍を済ませたデータ倉庫のデータを選択する。

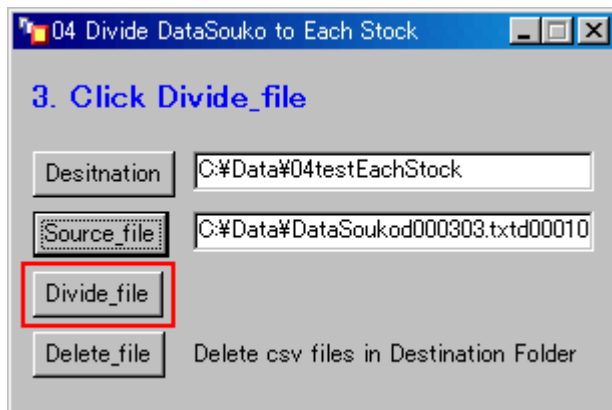
そして「開く(O)」ボタンをクリックする。

メモリーもしくは字数制限が原因と思われる理由で、データをすべて選ぶとエラーが出る場合があるので、選ぶデータを2つか3つに分けて2～9の作業を繰り返し行う必要がある。

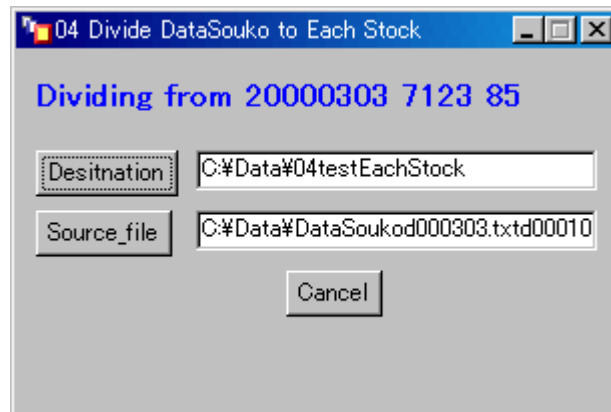
また、データを選ぶ順序は時系列順にしなければならない。



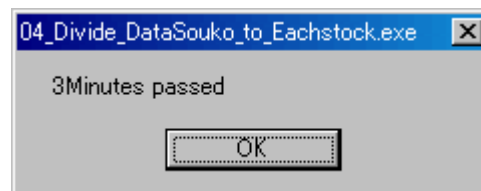
06. Divide file ボタンをクリックする。



07. Dividing from データ内容 のメッセージが出て CSV ファイルの分割が実行される。



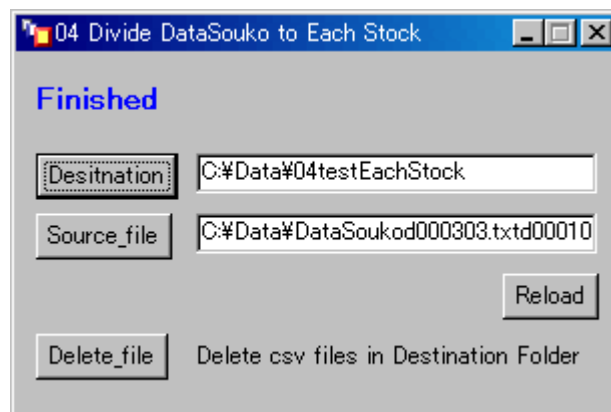
08. 変換・作成が終了すると、作業にかかった時間が表示される。OK ボタンをクリックする。



09. Finished のメッセージが表示される。

終了する場合はウィンドウ右上のクローズボタンをクリックする。

続けて作業する場合は Reload ボタンをクリックする。



10. Reload ボタンをクリックした場合には、02. からの手順を繰り返すことになる。

11. コード毎の CSV ファイルが作成されたら作業は終了となる。

新しいデータを追加する方法

1. 「データ倉庫」からのデータのダウンロード

<http://www.geocities.co.jp/WallStreet-Stock/9256/data.html>

このページから圧縮されたデータをダウンロードするのだが、データ数が多い場合は一括してダウンロードする Fuerza というソフトを使うのがお勧め。

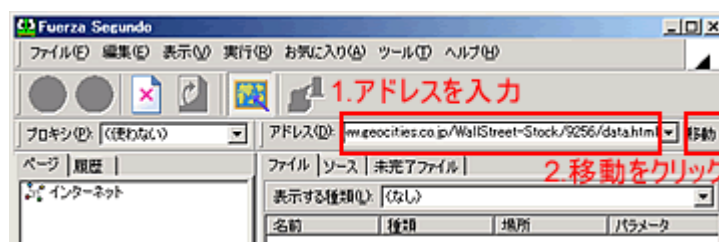
<http://www1.bbq.jp/jylki/>

上記の作者のページから Furza をダウンロードし、適当なフォルダに解凍する。

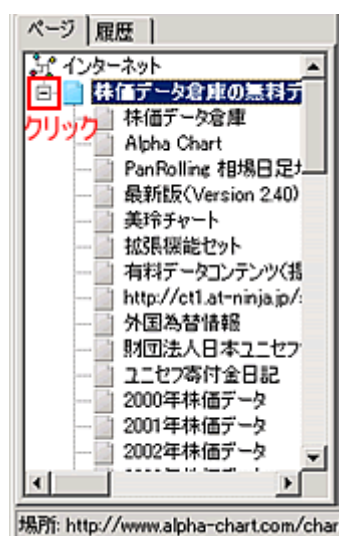
Isabel.exe のアイコンをダブルクリックして Fuerza を起動する。



アドレスの入力フォームに上記のデータ倉庫のアドレスを入力し、フォーム横の移動ボタンをクリックする。

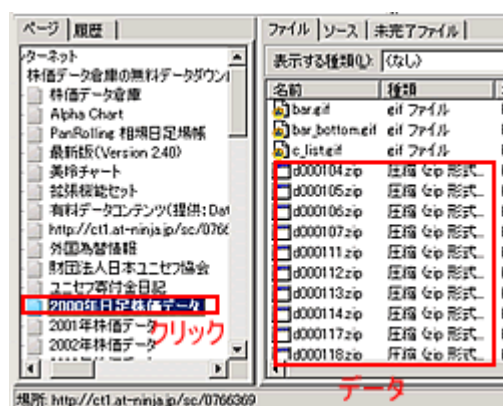


”インターネット”の下に「株価データ倉庫の無料データ」のファイルが現れるので、その左横の＋をクリック。すると入力したアドレスのページにリンクしたファイル類が表示される。

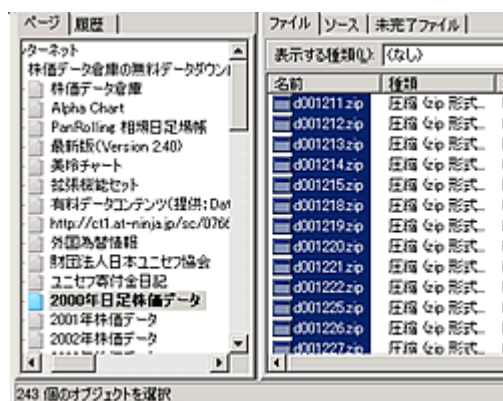


ここに 2000 年株価データ～2008 年株価データがあるのがわかる。

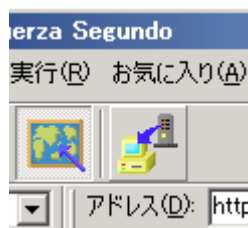
このうちほしい年のアイコンをクリックすると、右側に日にちごとの圧縮データが一覧で表示される。



Zip 形式の圧縮ファイルで、追加したい新しいファイルだけを選ぶ。

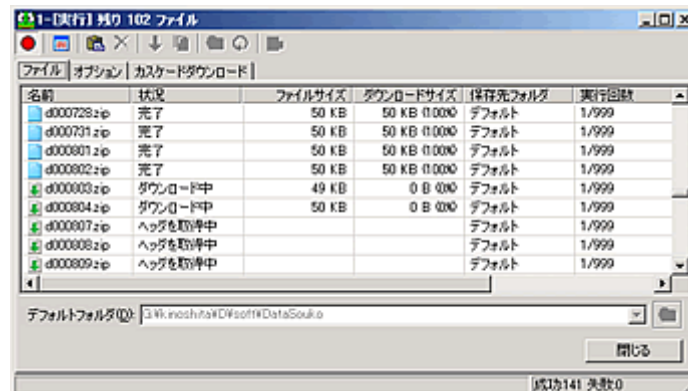


ダウンロード用のボタンをクリックし、適当なフォルダを保存先として指定する。



ダウンロード中は次のウィンドウが表示される。

ダウンロードが完了したら閉じるのボタンをクリックする。



ほしいファイルを1つのフォルダにダウンロードする。

ダウンロードが終わったら Fuerza を終了する。

ダウンロードされたファイルは圧縮ファイルなので解凍する必要がある。

今回は 7-zip という圧縮解凍ソフトを使い解凍するが、別に他のソフトでも構いません。

圧縮ファイルをすべて選択し、右クリックして現れたメニューから「7-zip」を選び「現在のフォルダに解凍」を選択すればよい。



解凍時に上書きしていいか？というメッセージが現れたら「はい」を選び、～は解凍できなかったというメッセージが現れても気にする必要はない。

解凍終了直後、圧縮ファイルが選択された状態にあるので、そのまま削除(Del キー)を行うと圧縮ファイルは削除される。

これでデータのダウンロードは終了。

2. データ倉庫のデータからコード毎の CSV ファイルを作成する

自作したプログラムは 04_Divide_DataSouko_to_Eachstock.exe と名づけた。

04_Divide_DataSouko_to_Eachstock.zip

Zip 形式で圧縮されているので解凍する。

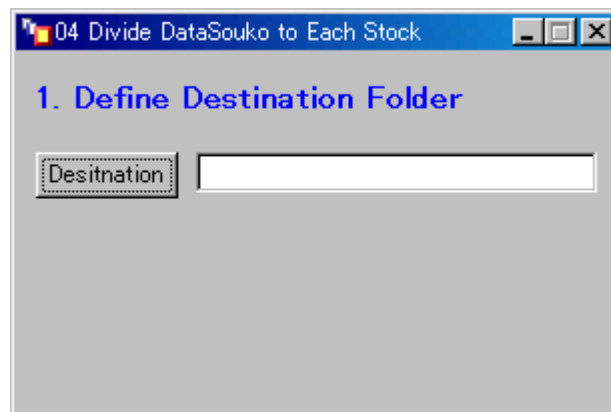
お決まりの文句ですが、このプログラムは私個人用に作ったものであるため、使用する人は自己責任でお使いください。

不具合がある可能性もあり、動作を保障するものではありません。

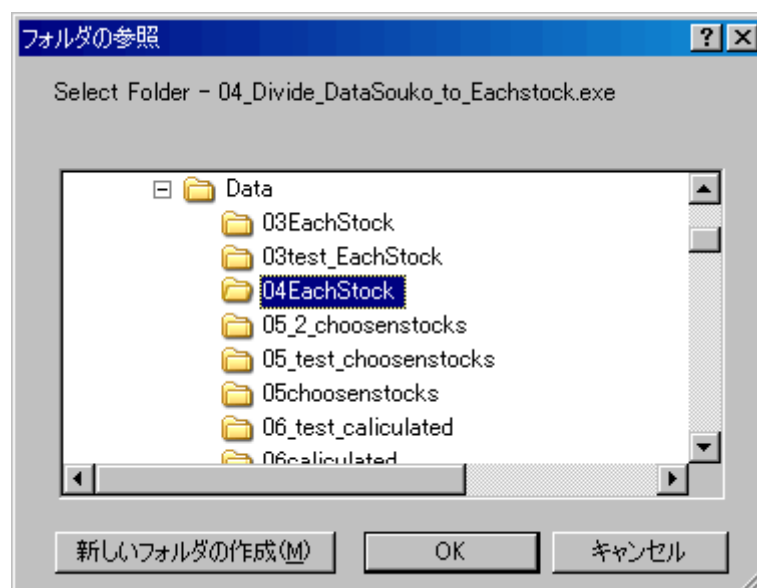
01. 04_Divide_DataSouko_to_Eachstock.exe のアイコンをダブルクリックして起動する。



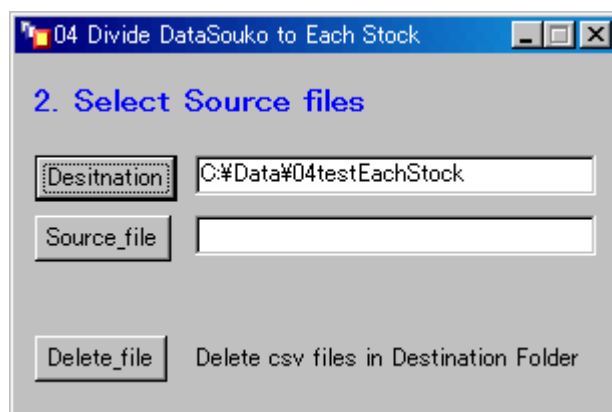
02. 次のウィンドウが現れるので Destination のボタンをクリックする。



03. 追加先となる CSV ファイルの入っているフォルダ(ディレクトリ)を指定する。OK ボタンをクリックする。



04. Source_file のボタンをクリックする。

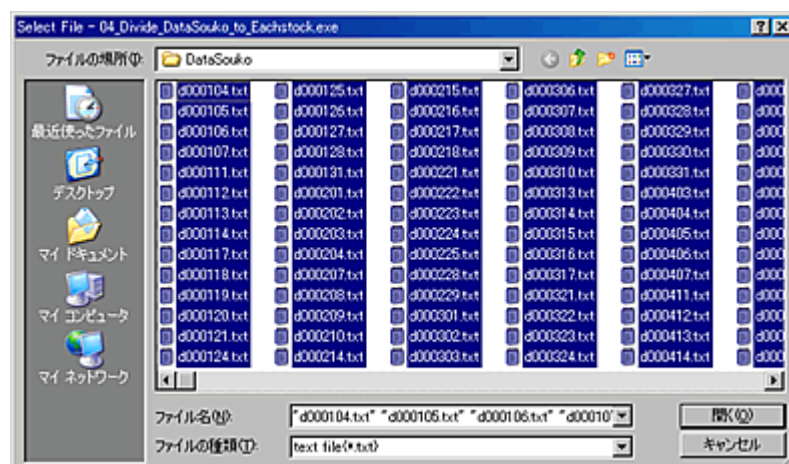


05. ダウンロードと解凍を済ませたデータ倉庫のデータを選択する。

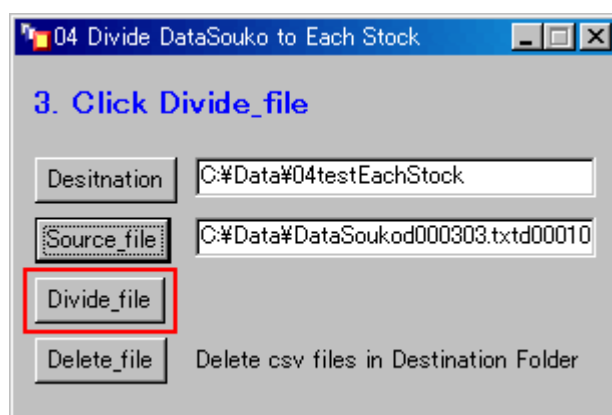
そして「開く(O)」ボタンをクリックする。

メモリーもしくは字数制限が原因と思われる理由で、データをすべて選ぶとエラーが出る場合があるので、選ぶデータが多い場合には2つか3つに分けて2～9の作業を繰り返し行う必要がある。

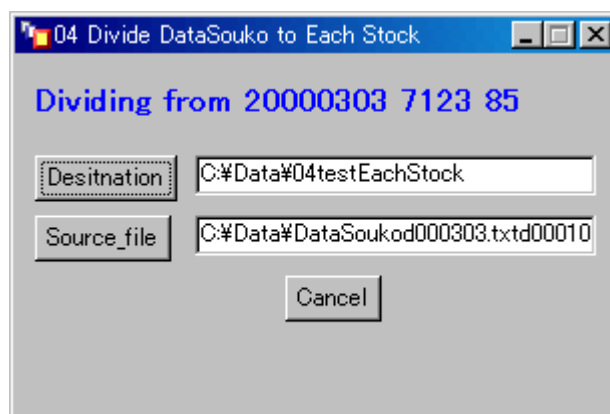
また、データを選ぶ順序は時系列順にしなければならない。



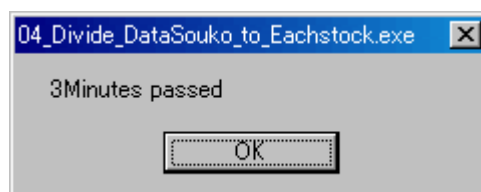
06. Divide_file ボタンをクリックする。



07. Dividing from データ内容 のメッセージが出て CSV ファイルの分割が実行される。



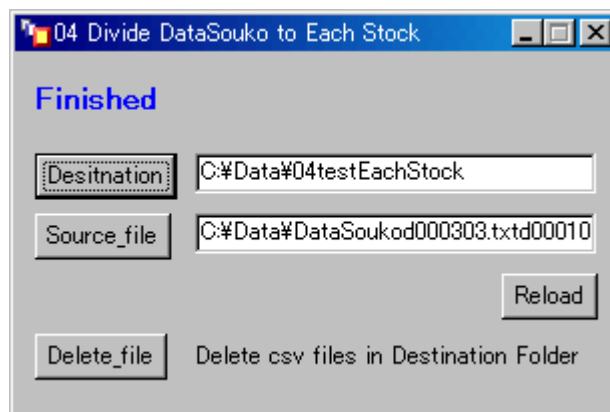
08. 変換・作成が終了すると、作業にかかった時間が表示される。OK ボタンをクリックする。



09. Finished のメッセージが表示される。

終了する場合はウィンドウ右上のクローズボタンをクリックする。

続けて作業する場合は Reload ボタンをクリックする。



10. Reload ボタンをクリックした場合には、02. からの手順を繰り返すことになる。

11. 銘柄コード毎の CSV ファイルにデータが追加され、作業は終了となる。

問合せ先(ウェブサイトのお問い合わせフォーム)

<http://t.12enjoy.net/contact/>