Error Detection in Prices

You are given a history of end of day prices for 100 securities. For the purpose of this problem, a security can be taken to be either a stock or a bond. Each security is uniquely indicated by the ID number.
They are valued in various currencies as indicated by the 'currency' field. Please use publicly available Foreign Exchange Rates if required (mention the exact sources used in the model documentation). In addition, you are given the dollar worth of every security held in a portfolio (**holdings.csv**) as of Aug 26, 2013.

You are also given a fresh set of prices coming into the pricing database for the next day i.e. August 27, 2013. You are to:

a) Finding all breaks - potentially erroneous prices in the fresh set of prices.
b) Are all the breaks equally important? How does one measure their relative importance?

In addition:

a) Are all historical prices accurate? Briefly describe the approach you will take to classify all historical prices as accurate or inaccurate. Also produce a file listing out all the historical prices that are potentially inaccurate (**output_history.csv)**
b) What happens when there are stock splits? There is precisely one security that has undergone a stock split, can you identify which one? Does this change what the other files should look like? If so how? (Provide answer as part of Model Documentation)
c) Types A, B and C of securities are different in a fundamental way. A represents stocks whereas B and C represent bonds. C is more likely to have "jumps" in prices compared to B. Can you devise a statistical test that will help you "classify" a new product into A, B or C given only a history of prices? Describe your approach. (Provide answer as part of Model Doc)
d) What happens when there are notable market wide moves in prices on any given new day? Briefly describe the approach you will take. You do not need to produce a file. Provide answer as part of the model documentation

Input

Filename	Fields	Available from
history.csv	ID, date, price, currency, type	9pm, Sep 28, 2013
holding.csv	ID, USD value held	9pm, Sep 28, 2013
prices_A.csv	ID, price on Aug-27-2013	9pm, Sep 28, 2013
prices_B.csv	ID, price on Aug-27-2013	6pm, Sep 30, 2013
prices_C.csv	ID, price on Aug-27-2013	6pm, Sep 30, 2013

Output

Zipped file to be submitted by **9pm, Sep 30, 2013** with the following contents:

Filename	Fields
ModelDocumentation	-
SourceCode	(zipped folder with all relevant files)
output_A.csv	ID, Min Price, Max Price, Error Rank
output_B.csv	ID, Min Price, Max Price, Error Rank
output_C.csv	ID, Min Price, Max Price, Error Rank

output_history.csv	ID, Date

- Min Price and Max Price will be double values in the range beyond which the price is considered an Error.

- Date in output_history.csv should be in the same format as available in history.csv.
- Error Rank should be between 0 & N, inclusive. 0 will indicate no error. 1 (lowest importance) through N (highest importance) will stand for the relative importance of the N errors identified.

- Partial submissions i.e., without one/more of the output_* files will be accepted. However, there should be at least one output_* file in the submission. ModelDocumentation & SourceCode are mandatory.