The attached is a six month times series of some of stocks on russel 2000 in a csv

1. Load the data csv into table it should look like this

A picture containing graphical user interface

Description automatically generated

**2- Create a new table holding the data in this structure Row Date, Columns ticker, cell value= daily\_return sorted by date as below**

A picture containing text, window

Description automatically generated

**3- PreProcess the data**

1. **For all single instances of empty or 0’ in the returns (forward fill the data make the value = prior dates value) and write dates above and below the observation to a log table with the following column structure**

**-ticker,Prior Date, Later Date, Method**

**Method set method to value ‘FF’**

1. **For all contiguous instances of zero greater then 1 you must linear interpolate (find a function or write it yourself) the values fill them in, and write the dates above and below the instances to the log table mentioned above setting the Method to value ‘Interpolate”**

Table

Description automatically generated with medium confidence

**4- Given the new clean table, write a stored procedure that takes a column name for example ZYXI, a start and end date and a dollar amount. Calculate the date and amount of most that investment would have been worth, calculate the date and amount the least that investment would have been worth during that date period.**

**Formula from i start to i end Amount(i-1)\*(1+dailyReturn(i))**

**ZYXI example starting with 100 dollars**

Graphical user interface, application, table, Excel

Description automatically generated

**5 -Change the stored procedure to only show tickers starting with the letter A or B or C to people who login with admin rights who try to execute it.**

**6- Load test and fine tune the process , think about design:**

**Create a CSV containing 1 million records from a random sample using table you created in task 1, simulating dates**

**For steps 1 – 4: using the new CSV, Analyze how your work performs, show the query plan, the trace. What will you do to fine tune this?**

**Assuming I wanted to speed this up, without touching the query and it was azure, what azure components could I use any hardware specific ideas?**

**Assuming we wanted to capture the entire universe of securities 200K public companies to do this entire process in Sql Server daily, what would that system look like and how would you insure we could scale, how would we preserve data both in terms redundancy and transactions.**