Code quality, unit test and documentation will be evaluated.

Solve it with Spark + Scala/Python, put it in a public GIT repository of your choice and then share it with me, I ask you to add a readme file in GIT explaining how to run the test of the solution.

Here is the information to perform the test

There are 4 data files that can be downloaded on the links below, this files is:
orders (orders.json) https://favo-data-test.s3.amazonaws.com/TestFolder/orders.json
sellers (sellers.json) https://favo-data-test.s3.amazonaws.com/TestFolder/sellers.json
buyers (buyers.json) https://favo-data-test.s3.amazonaws.com/TestFolder/buyers.json
products (products.json) https://favo-data-test.s3.amazonaws.com/TestFolder/products.json

Given the above information generate the following data

1) Generate an enriched file that contains all unified information.

The next items can be made using the dataset enriched in item 1, if item 1 has not been resolved, use the separate files

2) Sales growth per seller weekly (find the difference between the current week analyzed and last week, for every week of the year). The solution should be a dataset per seller and per week (use week in format ISO 8601).

3) Find the percentage that each of the sellers has contributed to the total sales since always.

4) There are orders that do not have the information on who made the purchase, prepare a table that has who could be the possible buyer of each of these orders. Explain the logic used to arrive at this information. I'm waiting for the solution!

Good luck!!!