F4SC end-of-semester project

Instructor: Bertrand Haas

Due: 23 January 2026

1 Generalities

- All forecasts have to use TSE coded by yourself.
- We recommend you use the Walmart data posted in Moodle, but you can use any other dataset posted there.
- In the data you are using, reserve 4 weeks or 30 days (according to the data granularity) as testing data for the project evaluation (pretend you did not know the actuals).
- What you should submit: Your project notebook (Quatro or Jupyter notebooks), well documented, that is, with markdown text and plots when where necessary, and that I can run on my laptop. These should be the same notebooks you will be presenting Friday Jan. 30th.

2 Part I

The goal is to compare forecasting at the smallest unit level (usually item) versus forecasting at an aggregate level (like at the store level or department level) and then de-aggregate to the item level.

Here forecasts are horizon-1 forecasts. That is: You are at the end of today, or of this week, and you know the actuals until and including then, and you forecast for tomorrow, or next week.

2.1 How to proceed

You will have to use method 1 and method 2 to get 2 different set of forecasts at the item level. Then you will have to analyze these forecasts in terms of accuracy to determine which set performed best.

2.1.1 Method 1: Forecast at the item level.

Optimize your forecast at the smallest unit level (item level) with TSE on the training set. Call this set of forecasts X_1. This is part F_1 in the diagram below.

2.1.2 Method 2: Forecast at the aggregated level, then de-aggregate at the unit level (see pic).

1. Find meaningful aggregated groups, using TS analysis (specially seasonality), at the store or department level. This is part A_2 in the diagram below.
2. On such aggregated groups perform the forecast, say Y, with the best accuracy you can get . This is part F_2 in the diagram.
3. Choose a way to de-aggregate the Y-forecasts to item level forecasts. Call this set of forecasts X_2. This is part D_2 in the diagram.

2.1.3 Analysis of method 1 vs method 2

Analyze the differences between X_1 and X_2 in terms of accuracy (choose a few accuracy KPIs and explain why you chose what).

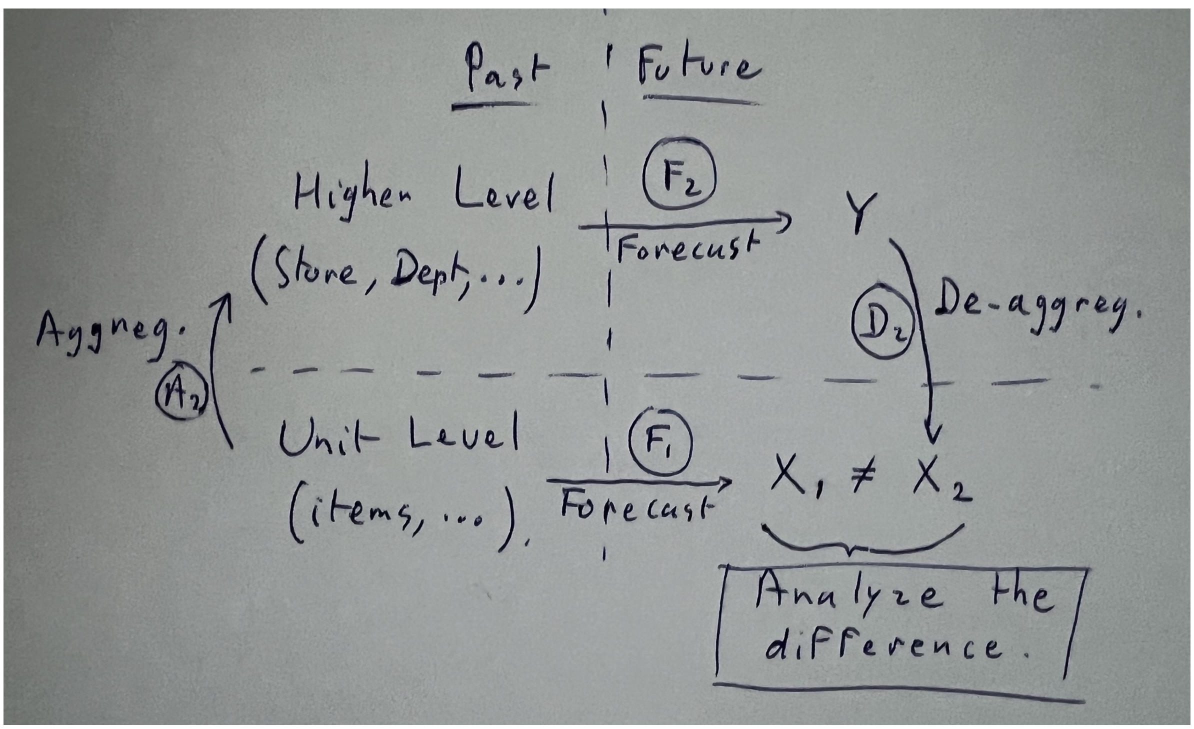

Figure 1: Direct Forecast vs Aggregate/Forecast/De-aggregate

3 Part II

Extract one store from your data, with enough data, and with some clear seasonality. We assume you get all your items from one supplier.

The goal is to maximize your chances of getting a rebate laid out in a contract between your store and your supplier, as explained below.

Your data should be weekly data (if you chose daily data, just aggregate it to weekly). Check your data to determine what is the first day of the week, Sunday (usually preferred in North America) or Monday (usually preferred in Europe).

3.1 The contract

In the contract you are bound to send every week at the beginning of the first day of the week, a forecast to your supplier on how much total number of items you are going to order from them in 2 weeks time. Notice that the first forecast is for the current week, so it is "horizon 0" (as opposed to horizon 1 in part I), the next forecast is for next week, horizon 1, and the forecast for next-next week that you send to the supplier is the horizon-2 forecast. We assume that at the beginning of the week you already know the actuals up until and including the week before.

In the contract, at the end of the month the supplier computes some "accuracy KPI" from the weekly forecasts of that month. Notice that a week is counted in the month if its first day is in the month, regardless of whether the rest of the week is in the month or not. For example suppose that in your data weeks start on Sunday. March 31st 2025 is a Sunday. So, the whole week, including the first 6 days of April, are counted for March, not April.

3.1.1 The contract's accuracy KPI

Now, the accuracy KPI for the rebate is the sum of the forecasts for the month, minus the sum of the actuals, all that divided by the sum of the actuals. If we write f_i for the forecast of week i of the month of interest, and a_i for the corresponding actual. let's write ρ_m for the rebate KPI for the month m of interest. Then:

$$\rho_m = \frac{\sum_i f_i - \sum_i a_i}{\sum_i a_i}$$

Where the sum is over all 4 or 5 weeks of month m If ρ_m is between $-X\%$ and $+X\%$ (X to be determined), then the month is selected for rebate. If ρ_m is outside of this range, then it is not selected for rebate. At the end of the year the store receives a rebate over only the months selected for rebate.

3.2 How to proceed

1. Optimize your TES algorithm to get the best horizon-2 forecasts.

2. Determine the bracket $\pm X\%$ for which half the months in the past year are selected for rebate, and half are not.

3. Devise a strategy to modify the forecasts you send to the suppliers in order to improve your chances to get the rebate.

    a. For how many more months would you get a rebate that way for the past year using the same $\pm X\%$ bracket?
    b. How much accuracy (with proper accuracy KPI, like MAPE, MedAPE, WAPE, ...) do you lose on weekly forecasts by modifying them that way (despite higher chances to get the rebates)?