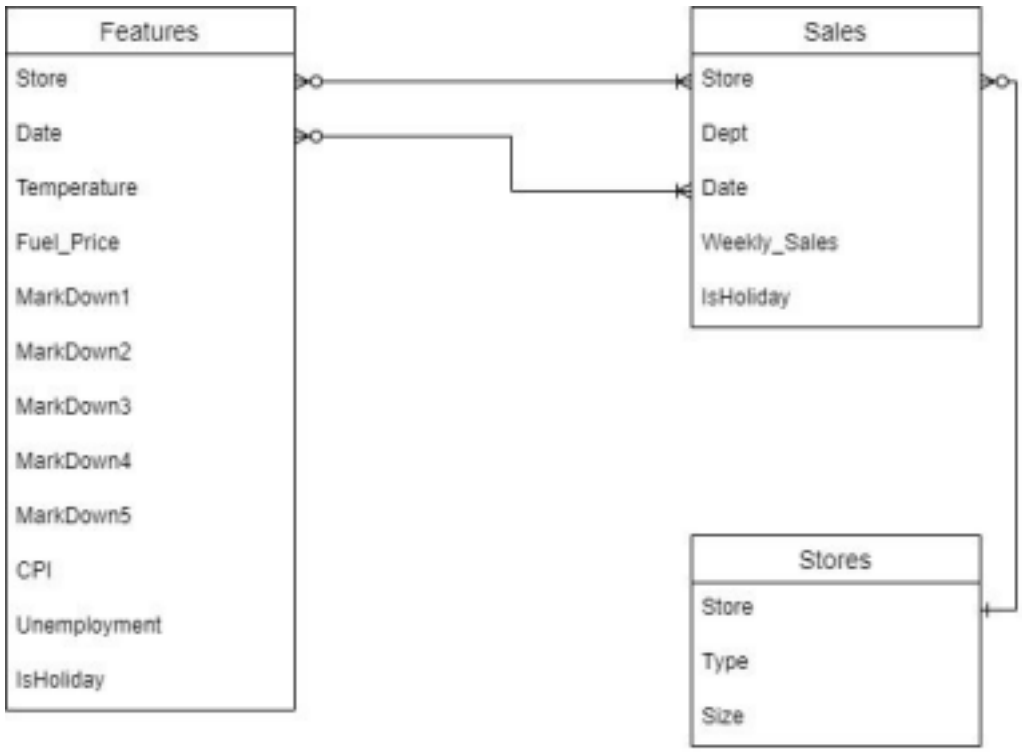
Danske Bank

Forecasting Homework

Today is 2012-12-10. You are a Data Scientist at a large US publicly traded retailer. Before the end of the year the Chief Financial Officer (CFO) needs to put out a financial forecast for the coming calendar year (2013) to comply with regulations, and to inform shareholders of expected business performance. You have been asked to provide these forecasts. You have 1 week as people are going on Christmas holiday on 2012-12-17, and the CFO wants some time to look over the forecasts and make sure they are ok.

You have been provided with 3 files; features.csv, sales.csv, stores.csv, which contain data about retail sales.

The structure of these files is as follows



Features - various features related to a store
Store - store number
Date - sales week date
Temperature - average temperature in the region
Fuel_price - cost of fuel in the region
Markdowns - promotion discounts, only available after November 11 2010 and may not exist for every row

CPI - consumer price index
Unemployment - unemployment rate
isHoliday - whether the week is considered a holiday or not

Sales - sales data covering 2010-01-10 to 2012-12-10
Store - store number
Dept - department number
Date - the week
Weekly_sales - sales for a given department in a store
isHoliday - whether the week is considered a holiday or not

Stores
Store - store number
Type - store type
Size - the maximum number of items the store can hold

Tasks

- Analyse data, identify key drivers (features) for sales success (mandatory task).
- Prepare suggestions to CFO on how to improve sales for the next year (mandatory task).
- Predict next year's (2012-12-10 to 2013-12-10) weekly sales (task for bonus points).

You may use any technology that solves the task above, however we recommend submitting your code in a jupyter notebook. Please send us a link to your GIT repository with the code. If you will be using Excel, it is fine to send it as an attachment to the letter.

The purpose of this task is to emulate a real world scenario in which you encounter constraints (like time limits), unfamiliar tasks or problems, and to evaluate your ability to deliver value under these circumstances.

You will be evaluated on

- Your approach when dealing with a new task (formulating the task into a technical problem, learning wherever necessary to make sure you deliver on the objective) under time constraints.
- Analytical skills when evaluating information to make decisions, identify important metrics.
- Communication skills when presenting the whole solution on how to improve sales to your stakeholders.