DATA ANALYTICS

Out of stock prediction And Buffer Stock Estimation

Introduction:

Out Of Stock: A stock-out, or out-of-stock (OOS) event occurs when an item is not available for sale as intended.If a retailer intends an item for sale but there is no physical salable unit available on shelf is OOS situation.The OOS events begins when the final salable unit is removed from the shelf and ends when its presence is replenished.
Buffer Stock : Buffer stock corresponds to level of extra stock to be maintained in order to avoid stockout due to sudden variation in demand or supply of items. Buffer Stock helps to reduce the negative effects (stock-out costs) of an unusually large usage of stock. Buffer stock schemes seek to stabilize the market price of products by buying up supplies of the product when stocks are plentiful and selling stocks of the product onto the market when supplies are low.

Problem Statement :

1. **Identifying Out Of Stock situations** - In a retail chain consisting of 'n' number of stores and 'x' number of products, possible out-of-stock situations are to be identified using real world data.
2. **Buffer Calculation Logic** - The best algorithm is to be found out for the calculation of buffer/safety stock in order to reduce the lost sales opportunity occurring due to the above mentioned out-of-stock situations.

Challenges :

A variety number of challenges were faced while preparing this report. For the calculation of buffer stock,service level , demand and lead time are to be considered for a more accurate prediction.Since lead time data were not provided its effect is less in our calculation. - was not defined which led us to having to omit some parts related to it from our calculations.

Approach Used:

Buffer stock calculation

Buffer will only be used when demand exceeds stock. So for a certain value of buffer stock we can calculate a buffer error.

We define,

Buffer error = ∑(Sales-Stock-Buffer_Stock)^2

SS=Z*σd

BS = Buffer Stock,

where Z = Safety Factor and σ = Standard Deviation of Demand

Our major task is to minimize this error.

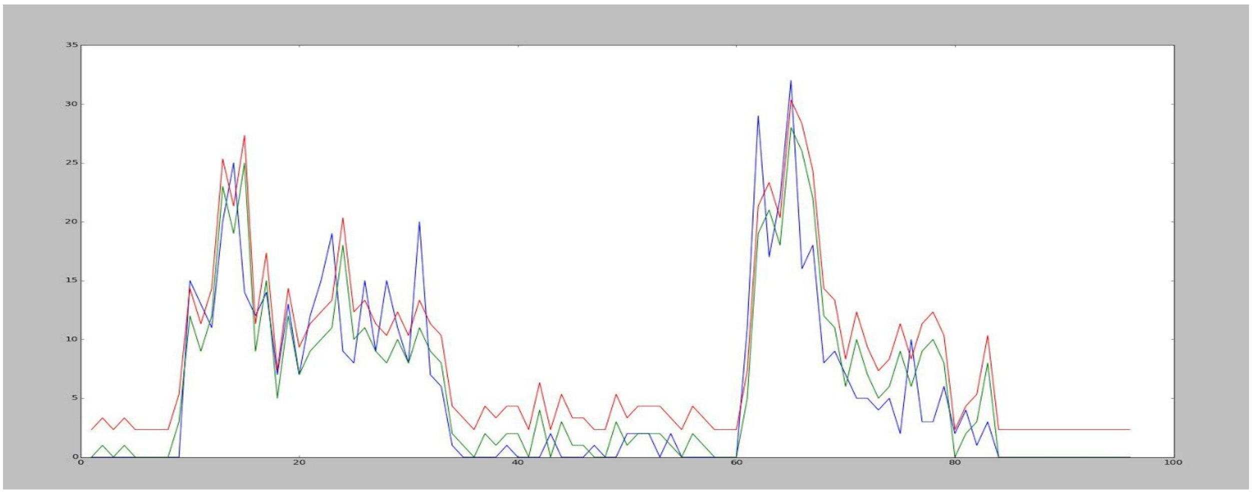
The buffer stock is calculated using the following formula. We assume the predicted sales to be the demand.

- In the scenario when our sales exceed the predicted sales, we need to take buffer stock in account to satisfy the customers demand.
- We define the Buffer error function as

$$Cost\ Function\ f(x)\ :\ \sum(PS - S - BS)^2$$

We need to find the safety factor as a function of time such that the value of the cost function is minimum. Optimization of the cost function was done using the gradient descent method.

ss = matrix form of above formula



In the above graph, the blue line represents sales data, the green line represents predicted sales data and the red line represents the sum of predicted sales and buffer that we calculated. This graph pertains to a particular product ID - 13245 of city A.

Out-Of-Stock Identification

Mean and Standard Deviation of Sales Approach

- ### Methodology

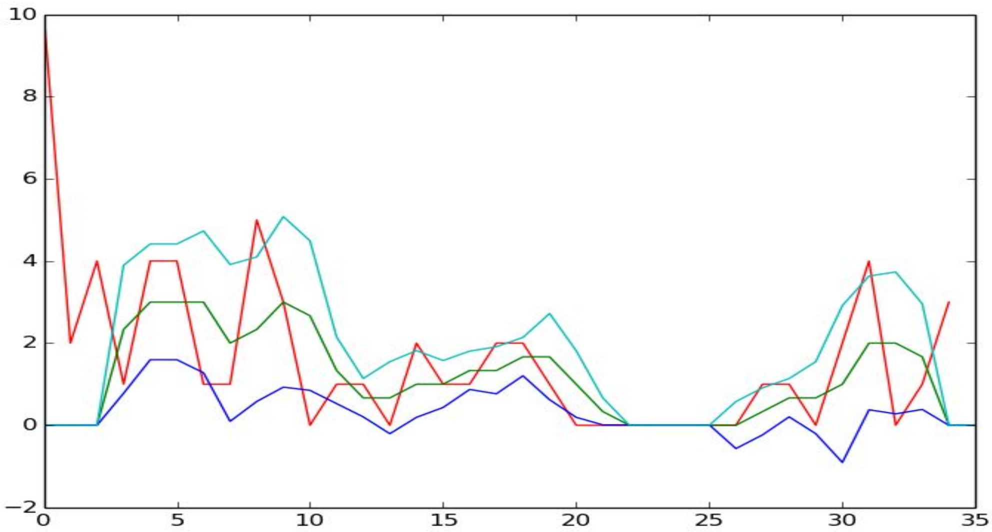
On the Sales vs Time graph, we plot a Line curve of (mean +– Standard Dev of sales). We get a lower bound and upper bound as per that.

Firstly we were doing it over all the data provided(Total Time). There this model was useless as it did not take into account the seasonal factors and gave a very bad result.

Then we calculated the same on a time segmented basis, like the mean of previous three sales and standard deviation calculated accordingly. This covered our data strongly.

Now if the point-of-sales data is less than the lower bound at that point we can term it as out of stock(OOS) and if the sales is more than the upper-bound at that point we can term it as an unusual sale. It could be due to some seasonal demand which can further be helpful for buffer stock calculation.As the demand rises unusually it is more likely that our stocks and buffer both get depleted before our previous calculated safe-zone.

So, Here we see how this approach not only solves the problem of OOS situation prediction but also warns us about possible increase of Sales opportunity.



Here:

- The Red Line is actual point of sales data.
- The Green Line is each time segment's average of point of sales.
- The Blue Line is average - standard deviation.
- The Green Line is average + standard deviation.

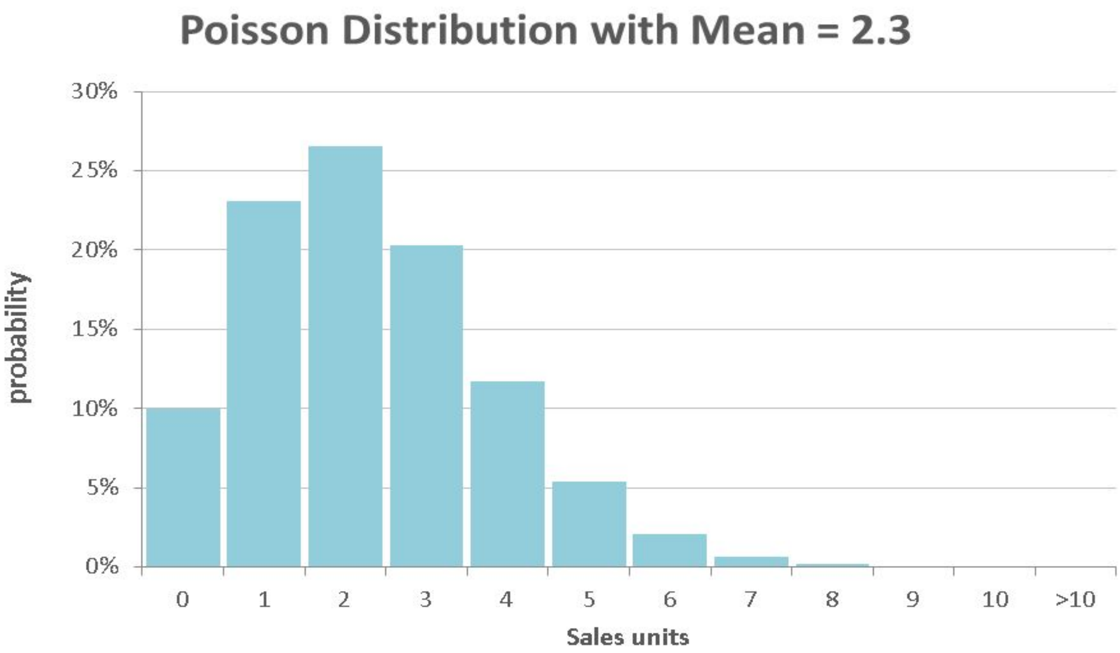
We can see how beautifully the model covers the data, predicting Out of Stock intelligently.

Poisson Distribution Approach

- **Methodology**

This is what we want to get to, a rule like. "If you should have sold more than X (and you've sold nothing) call an off-shelf alert."

Let's assume we have sold nothing now for product X at the target store for 3 days and that prior to this point our average sales over a 3 day period is 2.3 units. This is what a Poisson distribution looks like for average demand of 2.3 units:



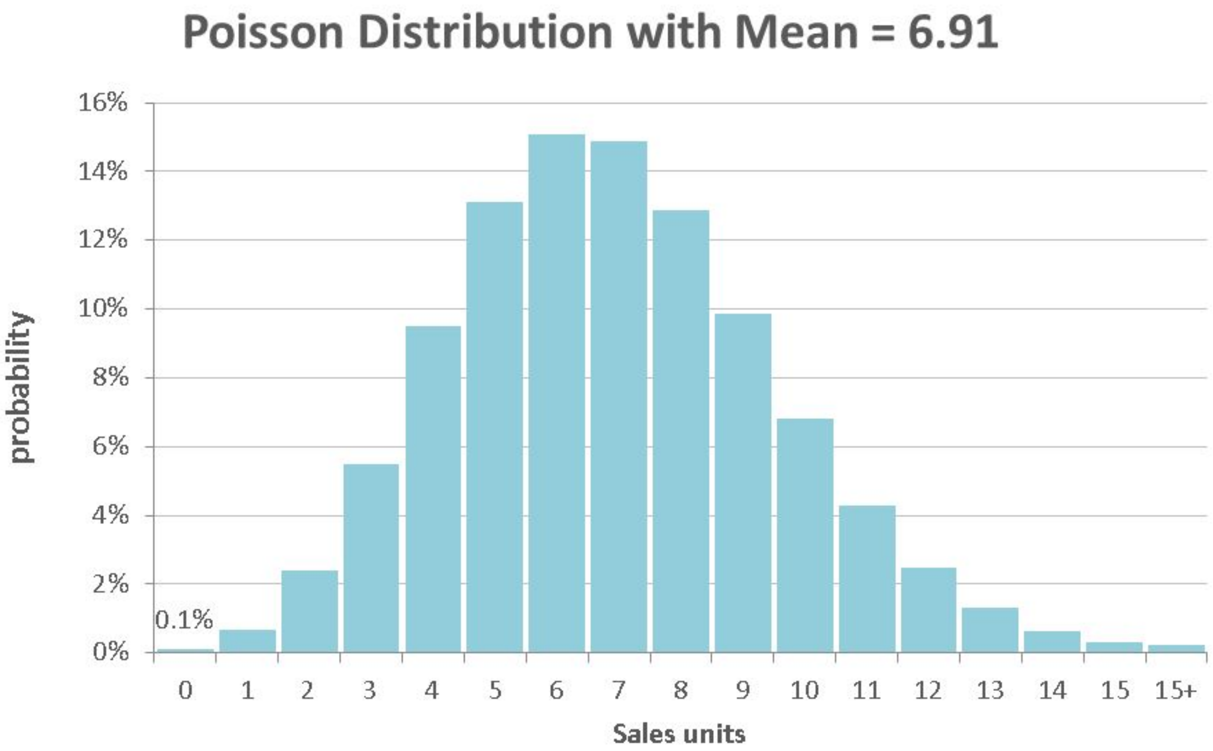
It shows the probability of actually getting sales of 1,2,3, thru 10 units in the 3 day period. As we can see there is a little over 25% chance of selling exactly 2 and about at 20% chance of selling 3. It's also possible (but very unlikely) that you could sell 8, 9, 10 or even more than 10 units.

Now, if this distribution really is a good representation of reality, how odd is it that we actually sold nothing at all in the most recent 3 day period? This tells us that seeing no sales is going to happen about 10% of the time just based on random chance. We probably don't want to call out an off-shelf alert with such a high chance that nothing is wrong.So, We wait.

When, a few days later, we reach the point at which you should have sold 4.6 units, and have still sold none, the probability of actually selling nothing through random chance is now just 1%.

Ignoring errors in data, your estimation of average sales or your assumptions (perhaps it's not a Poisson) you will be wrong about 1 in 100 times.

To get more surety we wait a little longer. At the point that we should have sold 6.91 units, there is only 0.1% chance that the zero sales you are seeing is due to random chance: far more likely in fact that there really is some issue inhibiting sales at the shelf.



Waiting helped us gain accuracy but it also cost us in lost sales. Combining this approach with time segment average we built a model and termed two threshold values as per experimental data:

- Harsh: Will generate more off-self alerts, on even a little of sales dwindling.
- Easy: Will generate less off-self alerts, only in extreme cases.

- ## Our Model

We took the average sales of previous three weeks into account then as per its poisson distribution we calculated the probability of getting an out of stock.

1. If it came less than the threshold then it is surely out of stock.
2. If it is not less than the threshold but is near it we term as not out of stock but add the lost sales number in check for the next point of sales data and so on.
3. If it is far greater than the threshold, we term it as not out of stock.

Then to make the model more realistic we added a smooth normalizing function depending on number of consecutive out of stock days.

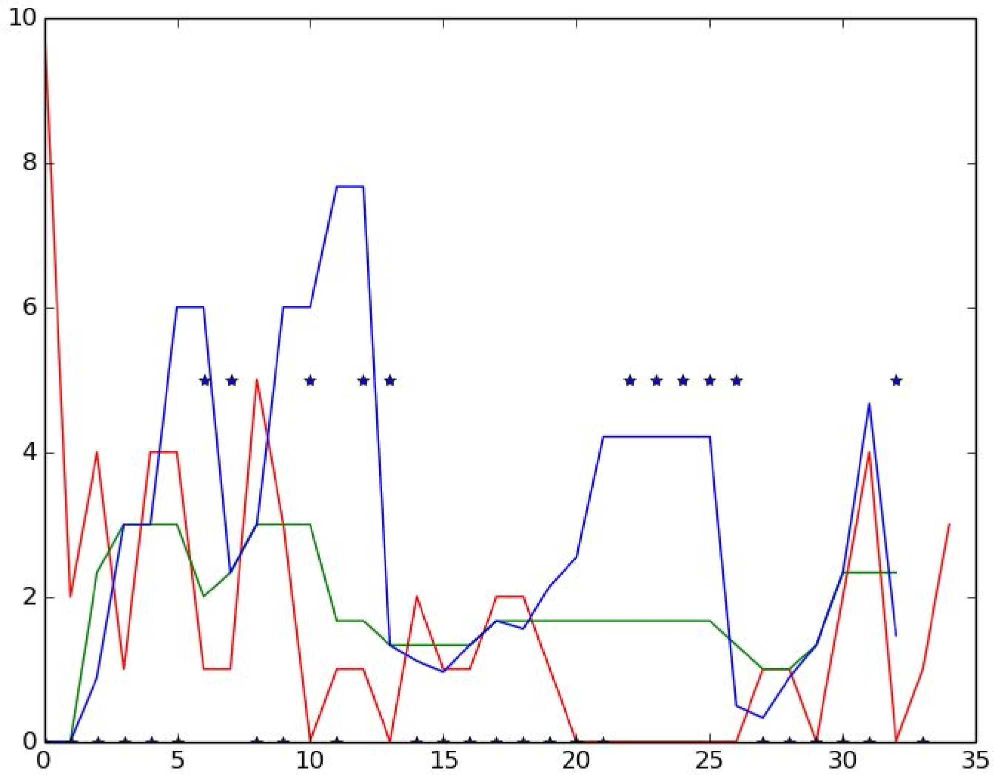
[e^(-x/100)]

where, x is the number of consecutive out of stock days.

Because more the number of days with out of stock increase more the mood of customer will turn away from the product.

So this model took into account the seasonal factor and mood of customers both.

- **Sample Graphs**



Here:

1. The Red Line is point of sales data.
2. The Green Line is calculated average as per rules.
3. The Blue LIne is the dynamic window i.e. the amount of sales expected that day.
4. A star marks an OOS event.

RESULT

For the buffer stock estimation, if we implement the above method for calculating the buffer we find that we are out of stock by 13618 units which is 11.74 % of total demand, while without buffer we are out of stock by 21234 units which is 18.29 % of total demand.This means that there is 32.9 % improvement after implementing the buffer strategy for the given data. Here we preferred Calculating Out Of Stock based on Poisson Ratio distribution as it strictly predicts Out of Stock, even on a little variation and it predicted 38.06% Out of Stock situation while the Standard Deviation Model Predicted only 12.35% Out of Stock situations.

DISCUSSION

A more precise model could have been presented if data pertaining to the frequency of renewal of goods would have been given, as mentioned in the challenges above.And also if some other little factors like cost of storage and profit per unit were given, then we could give optimized buffer for maximum profit.However we do realize that collecting such data not be very practical in a real - world setting, and we hope that our model is able to closely approximate the underlying distributions and mathematically give a good prediction with close accuracy.

Furthermore Calculating Out of Stock has been a little vague due to lack of train data and some of the products not being fast moving.