SCC Inventory Consulting, LLC

White Paper

Optimizing Inventory for On-Time Delivery

The Case for Data Cleansing

Steve Cimorelli, CFPIM **April 20, 2011**

Optimizing inventory investment to achieve desired levels of on-time delivery or material availability is among the greatest challenges for inventory managers. One thing is certain: the quality of data going into an inventory optimization process has a direct and profound impact on the quality of the resulting output. Sanitizing data on the front end helps to ensure that the optimization process is in fact optimized.

One method of sanitizing data is to determine the presence of "statistical outliers" in the data, then eliminating or replacing those data points. This process eliminates non-representative data that result from outright errors (bad data), that are accurate but unlikely to recur, or values that are deemed to be so large that they should simply be ignored. For example, a major recall or sales event may create a spike in demand that is not expected to recur. If the spike is allowed to remain in the data stream, safety stock investment may be elevated beyond what is actually needed to optimize on-time delivery and inventory levels.

Table 1 below shows sample data for a subset of part numbers supplied from a warehouse. The demand values for days 1 through 5 are shown; however, the full table includes a year of historical usage by day for each part for a total of 250 working days. (Obtaining this data is generally quite straightforward, either by extracting it directly from inventory databases or generating it from transaction history records.)

Part Number	Total Demand	Avg Day	Std Dev	Outlier Threshold	Outlier? (Yes/No)	Max Day	Day1	Day2	Day3	Day4	Day5
PN123	24,506	102	109	539	Yes	1,300	107	198	234	0	198
PN234	302	1	2	8	Yes	13	0	0	3	0	1
PN345	853	4	5	23	Yes	32	7	13	0	8	0
PN456	11,131	46	35	186	No	175	0	99	84	47	30
PN567	2,499	10	8	44	No	25	0	22	0	14	6
PN678	27,549	115	79	433	No	250	0	223	249	139	56

Table 1 – Daily Usage History & Statistics by Item

To determine the statistical outlier threshold for each part, several calculations are needed:

- Mean: The average of all daily usage values for the item.
- Standard Deviation (SD): The standard deviation of all daily usage values for the item.
- Sigma Level (N): A user-defined number of standard deviations. (A good starting point is 4 Sigma, which covers 99.997% of a normal distribution. In other words, data points greater than the 4 Sigma level have only a 0.003% probability of recurring.)
- Outlier Threshold = Mean + (N x SD).

Applying the concept, any daily usage value for an item which is greater than N-sigma above the mean is *suspect* and should be examined, replaced or eliminated. Manually examining all potential outliers is not always practical. An alternate approach is to apply an algorithm which replaces outliers with a calculated value such as the average, 4th largest daily value, or other appropriate value. An automated technique such as this can save considerable analysis time and provide repeatable results.

Fully sanitizing the data requires an iterative process. In the first iteration, all daily usage values exceeding the threshold are replaced. This of course alters the mean and standard deviation calculations, yielding a new threshold. A second and third iteration follow, replacing outliers and recalculating mean and standard deviation. Three iterations are almost always sufficient to complete the analysis.

As an example, consider the full data stream for Part Number PN123, shown graphically in Figure 1 below. Note the presence of three distinct spikes around days 10, 100 and 175.

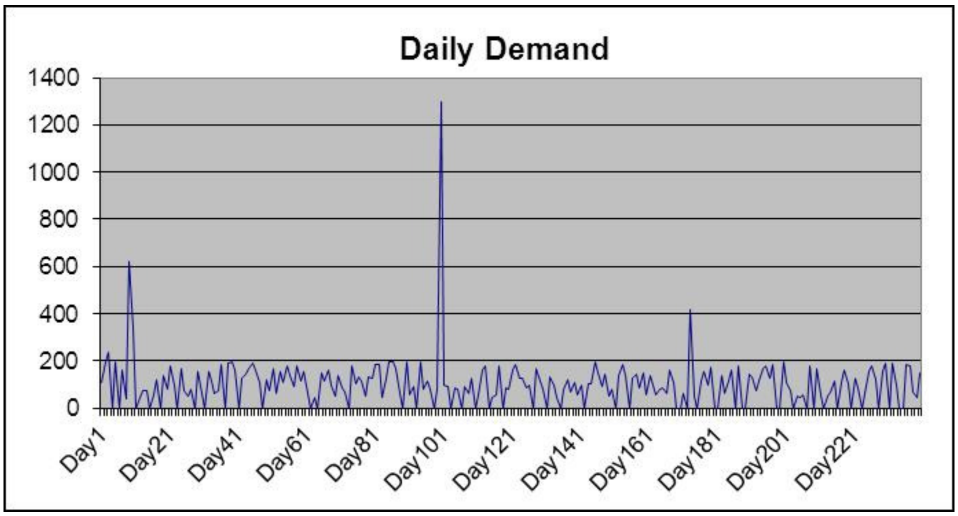

Figure 1 – Daily Usage for PN123

Applying the three-iteration model to PN123, replacing the outliers with a value of 200, yields the revised statistics shown in Table 2 below.

Part Number	Total Demand	Avg Day	Std Dev	Outlier Threshold	Outlier? (Yes/No)	Max Day	Day1	Day2	Day3	Day4	Day5
PN123	22,766	95	67	361	No	336	107	198	234	0	198

Table 2 – Revised Daily Usage History & Statistics for PN123

Note that the standard deviation value has dropped from 109 before removing outliers, to 67.

Finally, to demonstrate the power of this approach on inventory optimization, consider the resulting safety stock calculation using the before-and-after data streams. (Note that inventory optimization models use far more sophisticated techniques; however, the safety stock calculation shown here provides a clear indication of the potential improvement.)

Safety Stock Formula: SS = SF x SD x Sqrt(LT)

- SS = Safety stock quantity
- SF = Safety Factor (number of standard deviations to the right of mean in a normal distribution that are required to provide a defined service level; readily available in Excel using the NORMSINV function)
- SD = Standard deviation of the item's daily demand values
- Sqrt(LT) = Square Root of the item's replenishment lead time

Our sample item has a cost of $10/part, replenishment lead time of 40 days, and a service factor of 2.05 (yielding an availability level of 98%).

- Original SS = 2.05 x 109 x sqrt(40) = 1,413 pieces ($ 14,130)
- Revised SS = 2.05 x 67 x sqrt(40) = 869 pieces ($ 8,690)
- Net savings = $ 14,130 – 8,690 = $ 5,440 (38% savings!)

Conclusions

Using this technique on all items in a warehouse can significantly reduce safety stock levels required to provide a desired level of service. The process can be applied equally well to a finished goods warehouse where demand represents sales to customers, as to a raw material warehouse where demand represents consumption by a manufacturing process. Cleansing the data stream on the front end of an inventory optimization process can dramatically reduce inventory investment while maintaining or even improving delivery performance, boosting inventory turns and improving your bottom line.

SCC Inventory provides low-cost highly-customizable solutions to a variety of inventory management challenges, including data cleansing. Please visit our website's Publications page for more information.

About the Author

Steve Cimorelli is President of SCC Inventory Consulting, LLC. He is a Certified Fellow in Production and Inventory Management (CFPIM), a published author and educator. Steve holds a BS in Engineering with more than 25 years of experience in aerospace, industrial equipment and commercial manufacturing, distribution and supply chain management. He can be reached at 321-269-3407 or by email at Steve.Cimorelli@SCCInventory.com.