DATA SCIENCE CHALLENGE

Inventory Optimization & Supply Chain Analysis

Slooze Take-Home Challenge

Technical Documentation

Comprehensive Analysis Report

February 2026

Table of Contents

1. Executive Summary

This analysis presents a comprehensive inventory optimization system for Slooze, a retail wine and spirits company operating across multiple locations. Using 2.37 million purchase records, 1.05 million sales transactions, and inventory data from 2016, we implemented four core analytical frameworks to transform raw transactional data into actionable business intelligence.

Key Deliverables

- **ABC Analysis:** Classified 7,658 products by revenue contribution using the 80/20 rule
- **Demand Forecasting:** Built Facebook Prophet time-series models for top 5 A-Class products
- **Reorder Point Analysis:** Calculated optimal inventory triggers with 95% service level safety stock
- **Lead Time Analysis:** Evaluated 120 vendors across $321.9 million procurement spend

> **Critical Finding**
>
> 80% of revenue comes from only 19.6% of products (A-Class), yet none of these critical products use Premium suppliers. This creates a significant supply chain vulnerability that requires immediate attention.

Summary Metrics

Metric	Value
Total Revenue Analyzed	$33.1 Million
Unique Products	7,658
Vendors Evaluated	128
Store Locations	79

2. Problem Statement & Objectives

2.1 Business Context

Slooze manages millions of transactions across sales, purchases, and inventory records spanning 79+ store locations. Traditional spreadsheet-based analysis is inadequate for this data volume, creating risks of:

- Stockouts of high-revenue items leading to lost sales
- Excess inventory carrying costs tying up working capital
- Missed optimization opportunities from lack of data-driven insights
- Supplier inefficiencies going undetected

2.2 Core Objectives

Objective	Description	Success Metric
Inventory Optimization	Determine ideal stock levels by category	Reduced stockouts + carrying costs
Sales & Purchase Insights	Identify trends and supplier efficiency	Clear product/vendor segmentation
Process Improvement	Optimize procurement and stock control	Data-driven reorder triggers

Analytical Tasks Completed

1. ABC Analysis - Product classification by revenue contribution
2. Demand Forecasting - Time-series prediction using Prophet
3. Reorder Point Analysis - Inventory trigger calculations with safety stock
4. Lead Time Analysis - Supplier performance evaluation

3. Dataset Overview

Data Sources (6 Files)

File Name	Size	Records	Purpose
SalesFINAL12312016.csv	127.86 MB	1,048,575	Sales transactions
PurchasesFINAL12312016.csv	401.75 MB	2,372,474	Purchase orders
BegInvFINAL12312016.csv	19.31 MB	206,529	Beginning inventory
EndInvFINAL12312016.csv	21.00 MB	224,489	Ending inventory
InvoicePurchases12312016.csv	591 KB	5,543	Invoice records
2017PurchasePricesDec.csv	1.16 MB	12,261	Price reference

Data Quality Summary

Dataset Characteristics

- **Date Range:** January 1 - December 31, 2016 (with February anomaly)
- **Geography:** 79-81 store locations across multiple cities
- **Products:** ~7,658 unique SKUs
- **Vendors:** 128 significant suppliers (>10 purchase orders)
- **Total Records:** 3.87 million rows analyzed

4. Methodology

4.1 Phase 1: Data Exploration & Cleaning

The dataset was automatically downloaded using KaggleHub (61.9 MB compressed). All 6 CSV files were loaded into pandas DataFrames with appropriate data type optimization for memory efficiency.

Data Cleaning Steps

1. Standardized date formats (PODate, ReceivingDate, SalesDate)
2. Calculated derived metrics (LeadTime_Days, Revenue)
3. Removed invalid lead times (negative or zero values)
4. Converted 2,372,474 purchase records to datetime format
5. Stripped whitespace from string columns
6. Validated referential integrity across datasets

4.2 Phase 2: ABC Analysis

Classify inventory by revenue contribution to prioritize management attention and allocate resources effectively.

Classification Criteria

Class	Criteria	Priority
A-Class	Top 80% cumulative revenue	High
B-Class	80-95% cumulative revenue	Medium
C-Class	Bottom 5% cumulative revenue	Low

4.3 Phase 2.2: Demand Forecasting

Implemented Facebook Prophet models with daily and weekly seasonality, multiplicative seasonality mode, and 14-day forecast horizon. MAE and RMSE calculated for model validation.

4.4 Phase 2.3: Reorder Point Analysis

ROP = (Average Daily Demand x Lead Time) + Safety Stock

Where Safety Stock = Z x Standard Deviation x sqrt(Lead Time)

Parameter	Value	Description
Service Level	95%	Z-score = 1.65 (industry standard)
Lead Time	7.3-7.6 days	Vendor-specific average
Safety Stock	Variable	Based on demand variability

4.5 Phase 2.4: Lead Time Analysis

Classification	Lead Time	Std Dev	Risk
Premium	<= 7.7 days	<= 2.2 days	Low
Fast but Variable	<= 7.7 days	> 2.2 days	Medium
Slow but Steady	> 7.7 days	<= 2.2 days	Low-Medium
High Risk	> 7.7 days	> 2.2 days	High

5. Key Results

ABC Analysis Results

Category	Products	% SKUs	Revenue	% Revenue
A-Class	1,502	19.6%	$26.51M	80.0%
B-Class	1,813	23.7%	$4.97M	15.0%
C-Class	4,343	56.7%	$1.66M	5.0%

Top 5 Revenue Products

Rank	Product	Revenue	% of Total
1	Captain Morgan Spiced Rum	$444,811	1.34%
2	Ketel One Vodka	$357,759	1.08%
3	Jack Daniels No 7 Black	$344,712	1.04%
4	Absolut 80 Proof	$288,135	0.87%
5	Tito's Handmade Vodka	$275,163	0.83%

Reorder Point Results

Product	ROP	Current	Status
Captain Morgan	5,676	16,769	Healthy
Ketel One	3,616	16,770	Healthy
Jack Daniels	2,620	15,047	Healthy
Absolut	2,978	12,268	Healthy
Tito's	2,811	14,018	Healthy

> **Inventory Status**
>
> All products are currently healthy with current stock levels 2-4x above reorder points. No immediate stockout risk detected.

Lead Time Analysis Results

Tier	Vendors	Spend	Risk
Premium	26 (21.7%)	$50.6M	Low
Fast but Variable	34 (28.3%)	$165.0M	Medium
Slow but Steady	35 (29.2%)	$82.9M	Low-Medium
High Risk	25 (20.8%)	$23.5M	High

Critical Supplier Risk

All top 5 revenue products use "Fast but Variable" or "Slow but Steady" vendors. None use Premium vendors. Vendor 3960 (Diageo) supplies 40% of A-Class revenue - significant concentration risk.

6. Assumptions & Limitations

Data Limitations

1. Temporal Scope: Only 60 days of reliable data (January 2016). February showed 90% sales drop.
2. Missing Costs: No carrying cost or ordering cost data prevented EOQ calculation.
3. Vendor Names: Some vendors only had ID numbers; names extracted from InvoicePurchases.

Analytical Assumptions

1. Service Level: 95% used for safety stock calculations (Z = 1.65, industry standard)
2. Lead Time Distribution: Assumed normal distribution for safety stock formula
3. Demand Stability: Used January 2016 averages (ignoring February anomaly)
4. Product Classification: ABC based on 2-month snapshot; annual data would be more robust

Technical Constraints

1. Prophet Forecasts: Limited by short time series (60 observations)
2. External Regressors: No weather, holidays, or promotions included
3. Vendor Classification: Thresholds based on median rather than business SLAs

7. Business Recommendations

Immediate Actions (0-30 Days)

1. Dual-Source Vendor 3960: 40% of A-Class revenue depends on single supplier (Diageo)
2. Renegotiate SLAs: $85M spent with vendors slower than 7.7 days; implement penalties
3. Increase Safety Stock: Apply 1.5x multiplier for 'Fast but Variable' vendors
4. Consolidate Orders: Top 3 vendors = 618K orders; negotiate volume discounts

Strategic Initiatives (30-90 Days)

1. Supplier Development: Move A-Class products to Premium vendors (currently 0/5)
2. Inventory Rationalization: Review 4,343 C-Class products for discontinuation
3. Data Infrastructure: Collect full year of data for seasonal forecasting

Risk Mitigation Priorities

Risk Factor	Exposure	Mitigation
Vendor Concentration	21.7% Premium tier	Diversify supplier base
Variable Suppliers	$165M (51.2%) spend	Increase safety stock
Lead Time Buffer	7.6-day avg, 2.2-day var	Maintain safety coverage

8. How to Run the Code

Prerequisites

- Python 3.8 or higher
- Google Colab (recommended) or Jupyter Notebook
- 16GB RAM (for large CSV processing)

Installation

```
pip install pandas numpy matplotlib seaborn plotly prophet scikit-learn kagglehub
```

Execution Steps

1. Download Dataset: Automatically via KaggleHub
2. Phase 1: Data Loading & Cleaning
3. Phase 2: ABC Analysis
4. Phase 2.2: Demand Forecasting
5. Phase 2.3: Reorder Point Analysis
6. Phase 2.4: Lead Time Analysis

Output Files

- ABC_Analysis_Results.csv
- Reorder_Point_Analysis.csv
- Vendor_Performance_Scorecard.csv
- AClass_Vendor_Analysis.csv
- Demand_Forecast_Detailed.csv

9. Conclusion

This analysis transformed Slooze's raw transactional data into actionable inventory intelligence. Through ABC classification, we identified that 20% of products drive 80% of revenue - yet these critical items rely on variable suppliers, creating significant supply chain vulnerability.

The reorder point system provides data-driven triggers for procurement, while vendor analysis reveals $85 million in spend with suboptimal suppliers. By implementing the dual-sourcing and safety stock recommendations, Slooze can protect high-revenue streams while optimizing working capital across the portfolio.

Key Deliverables Provided

- Automated data pipeline with KaggleHub integration
- Statistical product classification (ABC Analysis)
- Predictive demand models with uncertainty quantification
- Operational reorder triggers with safety stock buffers
- Strategic vendor scorecards with risk classification

> **Final Note**
>
> All code is modular, documented, and ready for production deployment. The analysis framework can be extended to include additional data sources and more sophisticated forecasting models as the data infrastructure matures.

--- End of Documentation ---