Project Summary: Bar Inventory Forecasting System

1 Core Business Problem and Its Importance

The national hotel chain faces significant inventory management challenges across its bar operations, including frequent stockouts of high-demand beverages, leading to lost sales and poor customer experiences, and overstocking of slow-moving items, resulting in increased holding costs and wastage. These inefficiencies directly impact profitability and operational performance. The core objective is to develop a data-driven forecasting and inventory recommendation system to predict item-level demand for beverages (e.g., rum, vodka, wine) and recommend optimal stock levels. This matters because accurate inventory management ensures product availability, enhances customer satisfaction, reduces waste, and improves profit margins. For a hotel chain with multiple bars, optimizing inventory across locations can lead to significant cost savings and operational efficiency, directly supporting business goals.

2 Assumptions and Their Rationale

Several assumptions were made in developing the forecasting system:

- **Stable Demand Patterns**: Historical consumption patterns from 2023–2024 are assumed to remain consistent into 2025. This is necessary for time-series forecasting but may not hold if external factors (e.g., market trends, promotions) change significantly.
- **Zero-Filled Missing Dates**: Days with no recorded consumption are assumed to have zero consumption. This simplifies the time-series analysis but may underestimate demand for sporadically sold items.
- **Fixed ARIMA Parameters**: The ARIMA model uses a fixed order of (1, 1, 1) for all items, assuming it adequately captures demand trends. This reduces computational complexity but may not be optimal for all time series.
- **Uniform Safety Buffer**: A 30% safety buffer is applied to all forecasted consumption to account for demand variability. This assumes consistent variability across items, which may oversimplify real-world fluctuations.

These assumptions were made to streamline model development and ensure feasibility with the available data, but they highlight the need for further validation and refinement.

3 Model Choice and Rationale

The system uses an ARIMA (AutoRegressive Integrated Moving Average) model for time-series forecasting of daily beverage consumption. ARIMA was chosen because:

- **Time-Series Suitability**: The dataset consists of daily consumption data, making ARIMA a natural fit for capturing temporal trends and seasonality.
- **Simplicity and Interpretability**: ARIMA is computationally efficient and provides interpretable forecasts, which is valuable for bar managers making inventory decisions.
- **Robustness**: ARIMA can handle non-stationary data through differencing, suitable for the fluctuating demand patterns in bar sales.

Other models, such as Prophet or machine learning models (e.g., XGBoost), were not chosen because:

- **Prophet**: While effective for seasonality, it requires additional configuration for daily data and external regressors (e.g., holidays), which were not readily available.
- **Machine Learning Models**: Models like XGBoost or LSTM require extensive feature engineering (e.g., temporal features, external variables) and larger datasets, which could complicate implementation given the dataset's scope.

4 System Performance and Potential Improvements

The system generates daily forecasts for each bar, alcohol type, and brand, with recommended stock levels (forecast $\times$ 1.3) for June 19, 2025. For example, Anderson's Bar forecasts 79.53 ml for Captain Morgan rum (stock: 103.39 ml) but only 3.68 ml for Budweiser beer (stock: 4.79 ml), indicating significant demand variation. However, performance evaluation is limited due to the absence of validation metrics (e.g., MAE, RMSE) or test set comparisons, making it unclear how accurate the forecasts are.

Potential Improvements:

- **Model Optimization**: Implement grid search to optimize ARIMA parameters (p, d, q) or use SARIMA to capture seasonality (e.g., weekly or monthly patterns).
- **Performance Evaluation**: Split the data into training and testing sets to compute error metrics and validate forecast reliability.
- **Feature Engineering**: Incorporate planned features (e.g., Day of the Week, Season) to enhance model accuracy and capture demand drivers like weekend spikes.
- **Data Validation**: Address anomalies (e.g., near-zero opening balances like 8.526513e-14 ml) to improve data quality.
- **Alternative Models**: Test Prophet or machine learning models to compare performance, especially for items with complex demand patterns.
- **Dynamic Safety Buffers**: Use item-specific buffers based on historical demand variability instead of a fixed 30% buffer.

5 Real-World Application in a Hotel

In a real hotel, the system would integrate with the bar's inventory management software to provide daily or weekly stock recommendations. Bar managers would receive a report (e.g., a table or dashboard) listing forecasted consumption and stock levels for each beverage (e.g., "Stock 149.31 ml of Smirnoff vodka at Smith's Bar for June 19, 2025"). This would guide purchasing decisions, ensuring high-demand items like Grey Goose vodka are adequately stocked while minimizing excess inventory for low-demand items like Budweiser beer. The system could be automated to run nightly, pulling updated sales data and generating forecasts for the next day or week. Managers would use these insights to place orders, adjust stock levels, and plan promotions, improving operational efficiency and customer satisfaction.

6 Scalability and Production Monitoring

Scalability Challenges:

- **Data Volume**: At scale (e.g., hundreds of bars), processing large datasets could slow down ARIMA fitting. Parallel processing or cloud-based solutions could address this.

- **Model Maintenance**: Fixed ARIMA parameters may become less accurate as demand patterns evolve. Regular retraining or automated parameter tuning would be needed.
- **External Factors**: The system does not account for events (e.g., holidays, local festivals), which could cause demand spikes at scale. Integrating external data sources would be critical.

Production Monitoring:

- **Forecast Accuracy**: Track MAE or RMSE to monitor model performance over time, retraining when errors exceed a threshold.
- **Stockout/Overstock Incidents**: Monitor instances of stockouts or excess inventory to assess the effectiveness of the 30% safety buffer.
- **Data Quality**: Track anomalies (e.g., inconsistent balances, missing data) to ensure reliable inputs.
- **User Adoption**: Measure how often managers follow recommendations and correlate with sales and cost metrics to evaluate impact.

This system provides a strong foundation for data-driven inventory management but requires validation, feature enhancements, and scalability considerations for robust real-world deployment.