

U.S. Government Procurement Analytics for Strategic Supply Chain Decisions

Executive Summary

This project leverages real-time, U.S. based government procurement data to uncover actionable insights in vendor management, spend distribution, and procurement efficiency. By combining Excel (Power Query and Pivot analysis) with Power BI dashboards, the analysis provides transparency into top vendors, cross-agency overlaps, quarterly spend cycles, and vendor efficiency. The insights are designed to support strategic supply chain decisions, cost savings, and risk mitigation in public procurement.

Business Problem Statement

Government procurement involves large-scale vendor relationships across diverse U.S. agencies. However, without centralized insights, decision-makers face challenges in:

- Identifying top-spend vendors and their market dominance.
- Understanding agency–vendor overlaps, which can create dependencies and risks.
- Tracking spend trends over time to forecast demand and improve budgeting.
- Evaluating vendor efficiency in terms of spend vs. order volume.

Business Questions

1. Who are the top vendors by spend and what share of the total do they represent?
2. How concentrated is vendor spend (using HHI index)?
3. Which vendors are shared across multiple agencies, and where do overlaps occur?
4. What are the historical spend trends across agencies?
5. Which vendors deliver the most efficient procurement (high spend vs. lower order volume)?

Project Objective

To build a vendor intelligence framework using real-world U.S. procurement data, enabling transparency into vendor performance, agency dependencies, and procurement efficiency. The ultimate goal is to drive smarter contracting, cost savings, and risk mitigation in supply chain decisions.

Data Overview

The analysis was conducted using a U.S.-based public procurement dataset consisting of real-time purchase order records from multiple government agencies. In total, more than 18,000 purchase orders were analyzed to uncover vendor performance and spending patterns. Key fields chosen for the study included Vendor_Normalization, which standardized vendor names for consistency; Agency_Name, identifying the government agencies making the purchases; PO_Number, serving as unique identifiers for each transaction; PO_Total, representing the monetary value of each order; and Fiscal Year/Quarter, which enabled the analysis of procurement trends over time. The dataset was cleaned, transformed, and normalized in Excel (Power Query) to ensure accuracy, followed by advanced visualization in Power BI.

Methodology

The project followed a structured three-step methodology. First, data cleaning and normalization was carried out using Excel Power Query. This step focused on unifying vendor naming conventions, removing duplicate or mismatched purchase orders, and ensuring fiscal year and quarter alignment for accurate trend analysis. Second, pivot analysis was conducted in Excel to generate deeper insights. This included building a vendor leaderboard to rank top vendors by spend, calculating the Herfindahl–Hirschman Index (HHI) to measure vendor concentration and market share dominance, identifying cross-agency vendor dependencies through an Agency x Vendor overlap view, and analyzing quarterly and yearly agency spend trends. Finally, Power BI dashboards were developed to deliver impactful visual storytelling. These dashboards featured bar and donut charts to highlight top vendors by spend, a heatmap-style matrix to display agency–vendor overlaps, a scatterplot to assess vendor efficiency in terms of spend versus order count, and line and area charts to track quarterly spend trends across agencies.

Results & Interpretations

The analysis revealed several important insights. Vendor concentration was significant, with top vendors accounting for nearly 20% of total spend. Among these, “mcn build” and “fort myer construction” alone captured over \$3 billion in contracts, highlighting potential supplier dependency risks. Agency–vendor overlap analysis showed that vendors such as “Computer Aid” and “Gilbane” supply multiple agencies, suggesting opportunities for contract consolidation. Spend trends demonstrated a strong seasonal pattern, with recurring spikes in the fourth quarter, reflecting typical U.S. fiscal year-end budget cycles. Vendor efficiency analysis indicated that high-spend, low-order vendors such as “mcn build” and “fort myer” construction provide procurement efficiency, while vendors with a large number of small purchase orders highlight fragmented and potentially inefficient spending practices.

Recommendations

Based on these findings, several strategic recommendations were developed. Agencies should reduce dependency on their top two or three vendors by promoting competitive bidding and vendor diversification. Contract consolidation should be pursued where multiple agencies engage the same vendors, as this can provide cost leverage and reduce redundancy. To address the fiscal year-end spikes, budget smoothing practices should be implemented to distribute procurement more evenly across quarters. Purchase order processes should also be optimized by encouraging vendors with frequent small-value orders to move toward bundled procurement. Finally, agencies should adopt regular benchmarking practices, such as tracking HHI and spend concentration, to continuously monitor competitiveness and risk in vendor relationships.

Tools & Technologies Used

The project combined Excel and Power BI for both analytical depth and visual clarity. Excel PivotTables and Power Query were used for data preparation, vendor normalization, and HHI analysis, ensuring clean and structured data. Power BI dashboards were then leveraged for interactive visual storytelling, enabling agency–vendor overlap analysis, vendor efficiency evaluation, and trend forecasting. Additionally, the Herfindahl–Hirschman Index (HHI) was applied as a statistical tool to measure vendor concentration and market competitiveness.

Limitations

While the analysis was extensive, certain limitations remain. The dataset was restricted to publicly available procurement records, excluding private or undisclosed contracts that may impact overall spend analysis. Moreover, the dataset focused solely on financial and order metrics, meaning vendor performance factors such as quality, timeliness, or service level were not included. Finally, while vendor normalization was performed, manual adjustments may have introduced minor inconsistencies in naming conventions.

Future Enhancements

Looking ahead, the analysis can be enhanced in several ways. Integrating supplier performance metrics such as on-time delivery and quality would provide a more comprehensive view of vendor effectiveness. Predictive forecasting models could be developed to project future spend by vendor and agency, allowing for proactive planning. Establishing automated Power BI refresh pipelines with live procurement data would ensure real-time decision support. Additionally, machine learning techniques such as clustering could be employed to categorize vendors into high-risk versus efficient groups, supporting data-driven supply chain strategy and risk management.