Paul Whittingslow

Prof. Daniel Parada

Data Science Tools 1

14 August 2025

Mapping Defense Procurement and Logistics Risk: An Exploratory Analysis of FPDS and BTS Indicators Through 30 September 2024

The United States Department of Defense (DoD) relies on a sprawling and internationally dispersed network of suppliers to furnish platforms, subsystems, and parts at the tempo required for readiness. Over the past several years, geopolitical volatility, port congestion, and sanctions enforcement have complicated the movement of goods, creating new uncertainty for program offices that manage schedules and budgets. This paper presents an exploratory data analysis that joins openly available federal procurement records with transportation performance indicators to characterize where procurement activity may intersect with logistics friction. The central question is straightforward: how do time‑varying signals of port performance and sanctions exposure relate to year‑over‑year and month‑to‑month patterns in defense contracting activity, and what do those relationships imply for acquisition planning? By bounding the analytical window from the earliest available observations to 30 September 2024, we ensure that the most recent fiscal year is represented without partial fourth‑quarter spillover that might distort the comparison.

The project follows a research design common in operations research and public‑sector analytics: define a measurable proxy for potential disruption, standardize and enrich the core operational records, and test simple associations before proposing targeted follow‑ups. Contracting activity is observed using the Federal Procurement Data System (FPDS), which reports award‑level details such as vendor name, obligated value, place of performance, and North American Industry Classification System (NAICS) codes (U.S. General Services Administration). To approximate the state of goods movement, we incorporate monthly metrics from the Bureau of Transportation Statistics (BTS) Port Performance program, including average turnaround time and percent on‑time (Bureau of Transportation Statistics). Finally, we build a conservative sanctions signal using the OpenSanctions consolidated watchlist (OpenSanctions). These three sources allow us to create a repeatable, logistics‑aware view of contracting that remains tractable for a Tools‑1‑level study.

The dataset used for this paper is intentionally compact but structured to scale. FPDS records are cleaned to unify column names, coerce dates, and standardize currency fields. Vendor identities are normalized by uppercasing, removing punctuation, and stripping corporate suffixes to reduce aliasing between common variants (for example, “LOCKHEED MARTIN CORPORATION” and “Lockheed Martin Corp”). NAICS codes are trimmed to a six‑digit canonical form, and a two‑digit sector is derived for coarse industry analysis. Place‑of‑performance state names are mapped to USPS codes to support a left join with BTS by (state, year‑month). Temporal features—year, month, and year‑month—are extracted from action dates and used to produce yearly bar charts and monthly time‑series through 30 September 2024. The sanctions indicator is assembled by normalizing names on both sides of the comparison and applying a high‑threshold fuzzy match to avoid false positives. Together, these steps create a tidy, analysis‑ready table without relying on black‑box modeling. (U.S. General Services Administration; Bureau of Transportation Statistics; OpenSanctions).

Exploratory analysis proceeds along three tracks aligned to the project rubric. First, temporal behavior: the monthly series show clear variability across the 2021–2024 window, while the yearly bars—capped at 2024‑09‑30—make year‑over‑year changes interpretable at a glance. Second, vendor concentration: after normalization, a small set of vendor groups accounts for the majority of contract counts, which suggests sensitivity if one of these firms faces a material disruption. Third, logistics context: joining BTS metrics to the same year‑month granularity enables a qualitative overlay between procurement tempo and congestion proxies. Months that exhibit elevated turnaround times often coincide with shifts in place‑of‑performance or a change in the mix of industry sectors—signals that are operationally relevant, even if not causal. These patterns inform where to deepen analysis in future iterations. (Bureau of Transportation Statistics).

Two additional observations follow from the cleaned feature set. Obligated values are markedly right‑skewed—most awards are small, while a minority of contracts carry very high dollar figures. Applying a log transform (or simply reporting medians alongside totals) improves interpretability for stakeholders. And in the sanctions overlay, even when no prime vendor is directly flagged, retaining the normalized name and the similarity score in the dataset makes subsequent auditing far easier. Should the watchlist change, only the matching step and the threshold need to be re‑run; dashboards and reports update without rewriting the pipeline (OpenSanctions).

From a methodological standpoint, the emphasis in this Tools‑1 submission is on data cleaning, feature engineering, and communication rather than formal causal inference. The pipeline deliberately favors transparent transformations over opaque models: date coercion with explicit error handling; numeric standardization for obligated dollars; vendor normalization by rule; NAICS cleaning and sector extraction; and a documented state mapping table. The codebase is organized with a small utilities module, a cleaning notebook that can install its runtime dependencies into the active Jupyter kernel, and two analysis notebooks—one for static figures and one for interactive views. This structure prioritizes reproducibility for graders and collaborators while keeping the door open for scaling to larger extracts or additional logistics indicators.

The findings are intentionally framed as decision aids rather than definitive causal claims. First, contract volumes exhibit seasonal and episodic variability, the latter plausibly associated with exogenous shocks to shipment flows. Second, vendor concentration is non‑trivial; program managers should track dependency ratios for top vendors and identify alternate qualified sources where feasible. Third, while BTS metrics are not contract‑specific, they provide a practical context for interpreting procurement rhythms—planning buffers may merit adjustment during historically constrained months. Last, the sanctions framework, though conservative, is a low‑cost early‑warning surface that can be maintained as lists evolve. Together, these insights anchor concrete next steps for acquisition and logistics teams. (Bureau of Transportation Statistics; OpenSanctions).

Several limitations deserve emphasis. FPDS contracts describe obligations and attributes at the time of award; they do not, on their own, measure production progress or delivery timeliness. Place‑of‑performance is a useful proxy for geographic exposure but may not capture upstream tiers or multi‑site execution. BTS indicators average conditions at the state‑month level and therefore cannot distinguish which specific shipments served which awards. And sanctions matching at the prime level does not automatically illuminate exposure to sanctioned sub‑suppliers. Each of these gaps points to clear follow‑ups: incorporate delivery milestone data where available; enrich with port‑level or lane‑level movement data; and expand entity resolution to approved supplier lists beyond the prime.

Even with those caveats, the analysis yields actionable implications for defense acquisition. Maintain a rolling view of vendor dependency and identify categories where dual‑sourcing is practical. Use BTS indicators to inform schedule risk registers, particularly for programs that depend on coastal supply chains during periods of known congestion. Adopt normalized vendor identities as a common key across contracting, quality, and logistics systems to reduce friction in cross‑functional analysis. And treat sanctions proximity as a governance metric that is periodically refreshed and adjudicated—supported by saved match scores and audit trails (OpenSanctions).

In sum, consolidating FPDS procurement records, BTS logistics indicators, and a sanctions watchlist into a clean, feature‑rich table enables a logistics‑aware read of contracting activity through 30 September 2024. The approach is simple by design—anchored in data cleaning, transparent engineering, and clear visuals—so that program offices can adopt and extend it with modest effort. Future work should add delivery timeliness and port‑level movement detail, evaluate vendor dependency ratios over longer horizons, and test whether congestion proxies anticipate measurable schedule slippage. As a starting point for resilient acquisition planning, however, the present analysis already elevates the right questions and provides a reproducible foundation for deeper study (U.S. General Services Administration; Bureau of Transportation Statistics; OpenSanctions).

Works Cited

Bureau of Transportation Statistics. Port Performance Freight Statistics Program. U.S. Department of Transportation, 2023, www.bts.gov/port-performance. Accessed 14 Aug. 2025.

OpenSanctions. “Sanctions Database.” OpenSanctions, 2023, www.opensanctions.org. Accessed 14 Aug. 2025.

U.S. General Services Administration. Federal Procurement Data System (FPDS). 2023, www.fpds.gov. Accessed 14 Aug. 2025.