Procurement Risk & Spending Efficiency Analytics of Public Contracts

Sector:
Governance & Public Finance

Dataset Source:
Procurement Services Data - https://data.cityofchicago.org/Administration-Finance/Contracts-Contract-PDF-Present/kzv2-52bx/about_data

Tool Used:
Google Sheets

Course:
Data Visualization & Analytics (DVA)

Team Members:

Asad Ali - 2401010108
S. Harsith Priyan - 2401010398
Bhumi Rastogi - 2401010129
Siddhant Giri - 2401010455
Joshit Dutta - 2401010203
Himani Pinjani - 2401020115

Faculty:

Mrs. Aayushi Vashishth
Mr. Satyaki Das
Mr. Archit Raj

Executive Summary

Problem: The City of Chicago manages a $31.5 Billion procurement portfolio. However, the system faces significant challenges in spending concentration, administrative bottlenecks, and vendor dependency. Without centralized visibility, these inefficiencies lead to potential waste of taxpayer funds and operational delays.

Approach: We utilized a raw dataset of 10,000 contracts to build a dynamic Google Sheets dashboard. Our analysis focused on data cleaning, feature engineering (calculating approval delays and contract durations), and risk segmentation.

Key Insights

1. **Vendor Risk:** The top 5 vendors control nearly 30% of total spend, creating a "Too Big to Fail" (if a vendor fails, it poses a significant risk to the sector) dependency.
2. **Operational Lag:** The average contract takes **1180 days** to approve, indicating a severe administrative process failure.
3. **Spending Surge:** Procurement spending has increased sharply since 2019, while the number of contracts has remained stable, suggesting larger deal sizes.
4. **Spending is Concentrated:** A small subset of departments accounts for the major share of total contract spending, indicating budget concentration rather than uniform distribution.
5. **Approval Delays Vary Significantly Across Departments:** Certain departments show higher average approval delay, highlighting administrative inefficiencies localized to specific units, not system-wide.

Recommendations: We recommend implementing an "Express Lane" for low-value approvals, breaking down mega-contracts to diversify the vendor base, and placing strict caps on "Depends Upon Requirements" (blanket) contracts. Reduce dependency on a single vendor.

Sector & Business Context

Sector Overview: Public procurement is the process by which governments purchase goods and services. It requires a balance between **transparency** (fair bidding), **efficiency** (speed), and **cost-effectiveness**.

Current Challenges: Governments often suffer from "Vendor Lock-in," where incumbent suppliers dominate. Additionally, bureaucratic red tape leads to long lead times, and a lack of real-time data prevents agile decision-making.

Why This Problem: We chose this problem because efficient procurement directly impacts city services—from infrastructure to healthcare—making it a high-impact area for analytics.

Problem Statement

Is the City's procurement spending efficient and stable, or are there hidden risks in vendor concentration, approval bottlenecks, and contract volatility?

This project aims to evaluate whether the City's procurement spending is efficient, stable, and operationally sustainable, and to identify potential systemic risks embedded within the procurement process.
The analysis specifically investigates:

- Vendor concentration risk, where a small number of vendors dominate total spending
- Administrative approval delays indicate operational inefficiencies
- Contract spending volatility over time, reflecting financial instability or policy shifts

Project Objectives

Formal Problem Definition: To assess procurement efficiency and stability using **spending distribution patterns, approval timelines, and vendor dependency metrics**, and determine whether hidden structural risks exist in the procurement system.

Project Scope: Analysis of 10,000 contracts from 1990–2026, focusing on Departmental Spend, Vendor Distribution, and Time-to-Approval metrics.

Success Criteria:

- Identification of the Top 10 Departments and Vendors by total spending
- Quantification of the average administrative approval delay
- Detection of long-term spending trends and concentration risks
- Development of a dynamic, interactive dashboard enabling filtering by:
    - **Department**
    - **Year**
    - **Key performance indicators (KPIs)**

Data Description

Source: City of Chicago Data Portal – Administration & Finance
Raw Dataset Size: 10,000 records across 20 attributes

Cleaned Dataset Size: 10,000 records across 28 attributes

Key Variables:

- **Purchase Order (Contract)** Number – Unique identifier for each procurement contract
- **Award Amount** – Monetary value allocated to the contract
- **Start Date and Approval Date** – Temporal indicators used to measure approval delays and contract timelines
- **Vendor Name** – Entity receiving procurement funds

These variables enable analysis of financial distribution, administrative efficiency, and vendor concentration within the procurement system.

Data Cleaning & Preparation

Missing Value Treatment

The dataset contained null values primarily in the **Start Date** and **End Date** fields.
These were handled using **logical imputation strategies**:

- Missing **Start Dates** were inferred from the **Approval Date** where appropriate.
- Missing **End Dates** were either derived from available temporal information or flagged and excluded from **time-series and duration-based analysis** to prevent distortion.

This ensured temporal metrics remained analytically valid.

Data Standardization

Vendor names exhibited **inconsistent capitalization and formatting** (e.g., *“Abest Scale”*, *“ABEST SCALE INC”*).

To ensure accurate aggregation and vendor-level analysis:

- All vendor names were converted to a **uniform uppercase format**.
- Naming inconsistencies were standardized to prevent **duplicate vendor identities** in pivot tables and KPIs.

Feature Engineering

Several analytical features were derived to support KPI computation and dashboard insights:

- **Approval_Delay_Days**
  - Calculated as: **Approval Date − Start Date**
  - Used to evaluate **administrative processing efficiency**.
- **Contract_Duration_Days**
  - Calculated as: **End Date − Start Date**
  - Used to analyze **operational contract timelines**.
- **Is_High_Value**
  - Contracts with **Award Amount > $3M** were flagged as **“Large”**, enabling value-tier segmentation.
- **Is_Negative_Modification**
  - Contracts where **modifications reduced the total awarded value** were flagged to identify **financial risk or renegotiation events**.

Cleaning Logs

Cleaning Logs		
Column Name	**Issue Found**	**Action Taken**
Purchase Order Description	16 missing, inconsistent whitespace	Filled blanks with "Not Provided", trimmed

Start Date	657 missing, mixed formats	Created Start_Date_Clean using DATEVALUE; flagged blanks as ""
End Date	654 missing,	Created End_Date_Clean; for standard contracts, blank may be "Not Applicable."
Approval Date	0 but appears after Start	Date
Department	212 missing	Trimmed and filled blanks with "Unknown."
Vendor Name	0 missing, inconsistent casing	Trimmed + UPPER to standardize
Vendor ID	0 missing, inconsistent whitespace	Trimmed
Address 1	150 missing, inconsistent whitespace	Trim + fill blanks with "Address Not Provided."
Address 2	9028 missing	HIGH NULL RATE → Remove
City	16 missing, spelling variants	Trim + UPPER; blanks → "Unknown City"
State	7 missing, mixed full names/abbrev	Converted to 2-letter uppercase codes using mapping; blanks → "Unknown State."
Zip	104 missing, numeric formatting issue	Format column as Plain Text and convert to 5-digit text
Award Amount	0 missing, contains zeros & negatives	Kept raw; created flags Is_Zero_Award, Is_Negative_Modification, Is_High_Value
Procurement Type	5482 missing	Filled blanks with "Not Specified."
Start Date Year	Missing in raw	source derived from Start_Date_Clean
Contract_Duration_Days	Some contracts are missing an end date	Calculated as End - Start when both are present
Approval_Delay_Days	Initially negative due to the wrong formula	Corrected to Approval - Start
Is_Zero_Award	Derived	Flag for depends-upon (award=0)
Is_Negative_Modification	Derived	Flag for award_amount < 0
Is_High_Value	Derived	Flag for award >= 3000000
Award_Category	Derived	Segment the award into Depends/Negative/Small/Medium/Large
Vendor_Key	Derived	Normalized key combining Vendor ID and Name
Revision_Count_Per_Contract	Derived	Count of rows per Purchase Order to show the number of revisions

KPI Framework

The following Key Performance Indicators were constructed:

1. **Total Procurement Spend**
    a. SUM(Award Amount)
2. **Total Number of Contracts**
    a. Count(Contract ID)
3. **Total Negative Modifications**
    a. SUM(Is_Negative_Modification)
4. **Average Approval Delay**
    a. Average(Start Date – Approval Date)
5. **Average Contract Value**
    a. Average(Award Amount)

These KPIs measure financial exposure, efficiency, and risk.

Exploratory Data Analysis

Yearly Trend Analysis: Analyzed procurement spend over time to identify growth trends and fluctuations.

Vendor Analysis: Identified top vendors by total spend to assess concentration risk.

Department-Wise Spending: Compared total contract values across departments.

Contract Duration by Value Category: Indicates whether high-value contracts run longer, helping evaluate operational planning efficiency.

Revision Frequency Analysis: Analyzed contracts with multiple revisions to identify potential governance inefficiencies.

Negative Modification Analysis: Studied contracts with negative award amounts to detect scope reductions or corrections.

Department-wise Distribution of Contracts by Value Tier: Most contracts are small in value, but financial risk is concentrated in a few large contracts

Advanced Analysis

Risk Segmentation

The contracts were grouped into risk categories to understand hidden financial issues.

- **Zero-award contracts:** Many contracts show an award value of $0. These are mainly master agreements used for future purchases and do not represent actual spending.
- **Negative modifications:** Around **80 contracts** have negative award values, indicating scope reduction, penalties, or corrections. These may point to planning or performance issues.

This shows that total spending alone cannot fully explain procurement stability.

Root Cause Insights

When comparing **department spending** with **approval delay**, we found that:

- High-spending departments such as **Finance and Aviation** also have **longer approval times**.
- This suggests that administrative complexity and approval bottlenecks increase with contract size.

Key Interpretation

Overall, the analysis indicates that:

- Financial risk is concentrated in a small number of large or revised contracts.
- Delays are mainly department-specific, not system-wide.
- Monitoring risk indicators and approval efficiency is important for better procurement management.

Dashboard Design

The final dashboard presents key insights from the procurement dataset through clear visualizations and summary metrics.

Main components include:

- KPI summary cards showing total spending, number of contracts, negative modifications, average approval delay, and average contract value
- Year-wise spending trend using a line chart
- Top vendors by spending displayed as a bar chart
- Department-wise expenditure comparison
- Contract type distribution chart
- Revision frequency distribution
- Approval delay comparison across departments

Interactive slicers allow users to filter the dashboard by:

- Start Date Year
- Department
- Procurement Type
- Award Category

This interactive design enables **quick analysis and real-time decision support** for understanding spending patterns, vendor dependency, and administrative efficiency.

KEY INSIGHTS

- **Budget Concentration:** The Departments of Finance and Aviation together control more than 35% of the total procurement budget, indicating high spending concentration.
- **Vendor Dependency:** A single vendor, Blue Cross Blue Shield, receives around 15% of total city spending (over $5B), creating potential dependency risk.
- **Process Inefficiency:** The average approval delay is 1180 days (more than two years), highlighting serious administrative inefficiency.
- **Contract Structure Risk:** About 25% of contracts are “Depends Upon Requirements”, meaning spending is not fixed and may lead to uncertain financial commitments.

RECOMMENDATIONS

- **Diversify Vendor Base:** Large contracts should be divided into smaller competitive packages to reduce reliance on a single vendor and improve pricing efficiency.
- **Faster Approval Mechanism:** Introduce an automated fast-track approval system for low-value standard contracts (e.g., below $50K) to reduce backlog and administrative delay.
- **Control Blanket Contracts:** Limit “Depends Upon Requirements” contracts to a fixed share of the total budget (such as 20%) to prevent uncontrolled future spending.

IMPACT ESTIMATION

If these recommendations are implemented:

- **Financial Savings:** Even a **10% increase in competition** for major contracts could generate **approximately $1.5 billion in savings** (around 5% of total spend).
- **Operational Efficiency:** Reducing approval delay from **1180 days to about 400 days** would save significant administrative time and enable faster project execution.

- **Risk Reduction:** Monitoring **negative modification patterns** can help detect failing projects early and prevent future financial losses.

LIMITATIONS

- **Data Granularity:** The analysis is limited to contract-level totals; we cannot analyze unit-level pricing inefficiencies.
- **Historical Inconsistencies:** Data entered before 2010 has higher rates of missing fields compared to recent data.
- **External Factors:** The analysis does not account for external economic factors (inflation), which may have driven the spending increase.

FUTURE SCOPE

- **Predictive Modeling:** Build a machine learning model to predict "Approval Delay" based on contract type and department.
- **Text Analysis:** Use NLP on the "Description" column to automatically categorize miscellaneous contracts.
- **Live Data Pipeline:** Connect the dashboard directly to the City's API for real-time daily updates.

Conclusion

This project successfully transformed raw and fragmented contract data into a meaningful strategic asset.
By identifying the **$31.5 billion procurement scale**, the **1180-day approval inefficiency**, and the **risks of vendor concentration**, the analysis provides actionable insights for improving public procurement management in **Chicago**.

The developed dashboard functions not only as a reporting interface but also as a **decision-support tool** that can enable more transparent, efficient, and data-driven governance in the future.

Appendix

Data Dictionary				
Column Name	**Data Type**	**Description**	**Example Value**	**Notes / Transformation**
Row ID	Number	Unique row identifier	1	no change
Purchase Order Description	Text	Short description of contract/PO	Annual office supplies	Trim whitespace; fill missing with "Not Provided."
Purchase Order (Contract) Number	Text	Primary contract identifier	CNTR-2021-001	Keep it as a unique key
Revision Number	Number	Revision index for a contract (0 = initial)	2	Used to compute revision frequency
Specification Number	Text	Reference spec number (if any)	SPEC-45	Preserve as text
Contract Type	Text	Type of contract (Blanket/Standard/COMPTROLLER-OTHER)	Blanket	Standardize casing and typos
Start Date	Date	Contract effective date	2021-03-01	Standardize to date;
End Date	Date	Contract end date	2022-03-01	Standardize to date;
Approval Date	Date	Date when the contract was approved in FMPS	2021-03-10	Standardize to date;
Department	Text	The department that issued the contract	Finance	Trim + PROPER; fill missing "Unknown."
Vendor Name	Text	Supplier name	Acme Supplies Ltd.	Trim + PROPER
Vendor ID	Text	Supplier identifier	V-10023	Trim; combine with Vendor Name in Vendor_Key
Address 1	Text	Primary vendor address line	123 Main St	Trim; replace blanks with "Address Not Provided."
City	Text	City of the vendor	Chicago	Trim + PROPER; fill "Unknown Cit.y"
State	Text	State of vendor	IL	Standardize to 2-letter uppercase; fill "Unknown State."
Zip	Text	Postal code	60601	Format as 5-digit text; keep leading zeros

Award Amount	Number	Monetary amount awarded for the row	150000	Keep numeric; do NOT remove negatives/zeros
Procurement Type	Text	Procurement sub-type	Open Bid	Fill missing with "Not Specified."
Start Date Year	Number	Year extracted from Start_Date_Clean	2021	Derived
Contract_Duration_Days	Number	Days between start and end (empty if no both dates)	365	Derived
Approval_Delay_Days	Number	Days between start and approval (Approval_Date - Start_Date)	9	Derived
Is_Zero_Award	Number	1 if Award Amount = 0, else 0	1	Derived
Is_Negative_Modification	Number	1 if Award Amount < 0, else 0	1	Derived
Is_High_Value	Number	1 if Award Amount >= threshold (e.g., 100000), else 0	1	Derived
Award_Category	Text	Segment contract by value / special cases	Large	Derived
Vendor_Key	Text	Normalized vendor key (ID|Name)	V-10023|ACME SUPPLIES LTD.	Derived
Revision_Count_Per_Contract	Number	Number of rows for the same contract (raw count of revisions)	3	Derived

Contribution Matrix

Team Member	Dataset & Sourcing	Cleaning	KPI & Analysis	Dashboard	Report Writing	PPT	Overall Role
Asad Ali		✓	✓	✓	✓	✓	Analysis Lead
S. Harsith Priyan	✓	✓	✓	✓	✓		Project Lead
Joshit Dutta		✓				✓	PPT & Quality Lead
Siddhant Giri	✓				✓		Strategy Lead
Bhumi Rastogi	✓		✓	✓			Data Lead
Himani Pinjani			✓	✓			Dashboard Lead

Declaration: We confirm that the above contribution details are accurate and verifiable through version history and submitted artifacts.

Team Signature Block:

S. Harsith Priyan

Asad Ali

Joshit Dutta

Bhumi Rastogi

Siddhant Giri

Himani Pinjani