Data Dictionary & Variable Construction Reference

Repo A: PA UST Combined Datasets

Analytical Engine

January 9, 2026

This document serves as the definitive technical reference for the raw datasets in Repo A (Facility/Tank Master Database). It characterizes the distributions of key variables—including tank status, substance types, and component attributes—and identifies specific data quality issues such as default installation dates. The frequency tables presented herein are intended to guide the "hotcoding" of binary variables and the treatment of missing data for downstream econometric analysis.

Table of contents

1 Facility & Tank Characteristics

This section details the primary grouping variables derived from the harmonized Active (PADEP) and Inactive (SSRS) datasets.

1.1 Tank Status

[Description: Explanation of how Active vs. Inactive status is defined.]

Table 1

Table 2Tank Status Distribution

STATUS_CODE	Tank_Status_Meaning	N	Pct
W	Closed	45557	44.8%
R	Removed	24173	23.8%
C	Currently In Use	20579	20.2%
E	Exempt From State Law	7339	7.2%
UR	Unregulated Removed	1641	1.6%
T	Temporarily Out of Use	1269	1.2%
P	Permanently Closed in Place	899	0.9%
TRANS	Transferred	142	0.1%
DC	2004 Data/Fee Cleanup	80	0.1%
UC	Unsubstantiated Claim	22	0.0%

[Table Note: Verify if 'Temporarily Out of Use' tanks should be treated as active for auction eligibility.]

1.2 Substance Profile

[Description: Breakdown of fuel types stored in the tanks, mapped from raw substance codes to boolean flags.]

Table 3

Table 4Fuel Types

Fuel	N
Gasoline	51104
Diesel	20170
Other_Substance	30427

[Table Note: These counts are based on the mapping of raw substance codes (e.g., “GAS”, “DIESL”) to consolidated categories.]

1.3 Installation Date Diagnostics

[Description: Assessment of data quality for DATE_INSTALLED.]

Table 5

Table 6Top 10 Most Frequent Installation Dates

DATE_INSTALLED	N
1974-01-01	587
1981-12-01	580
1985-12-01	572
1979-12-01	565
1980-01-01	564
1980-12-01	541
1970-01-01	531
1987-12-01	529
1983-12-01	526
1978-12-01	515

[Table Note: The dates listed above are likely system defaults (e.g., 01/01/1900) and should be treated as missing values.]

1.3.1 KNN Date Imputation Diagnostics

[Description: Comparison of reported installation dates vs. KNN-imputed dates for records with suspect default values. KNN regression (k=5) uses capacity, substance type, and region as features.]

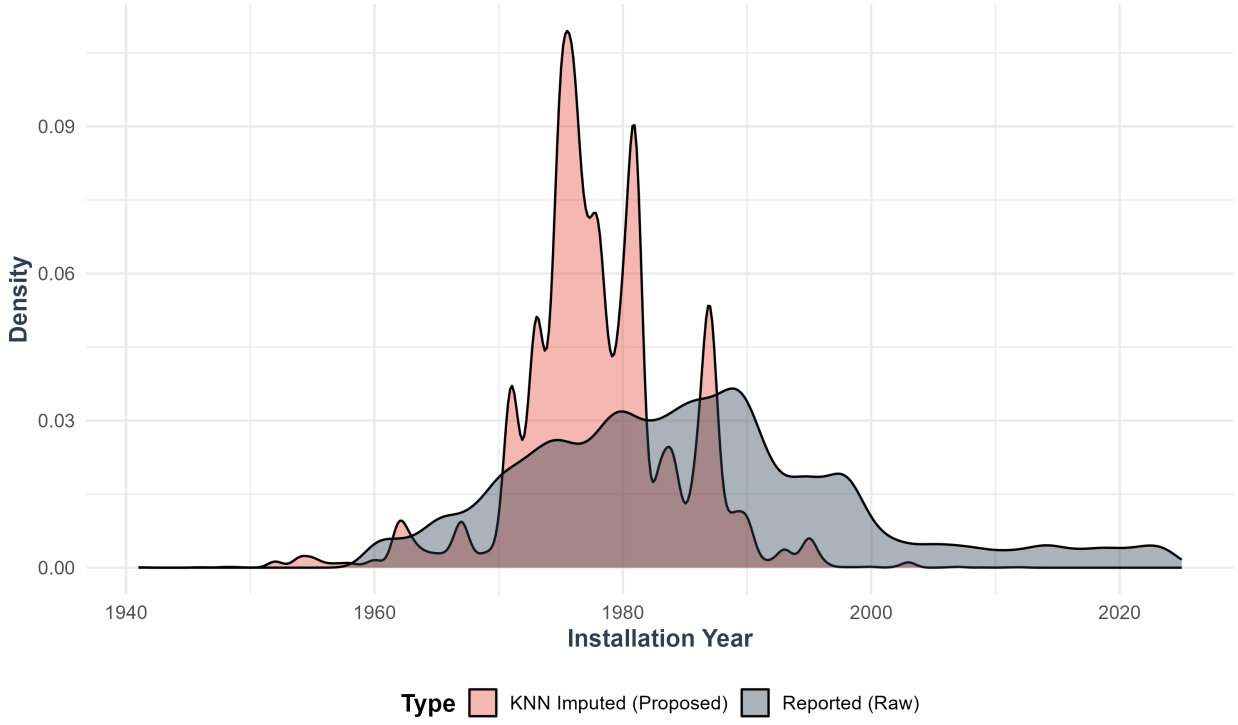

Figure 1: KNN Date Imputation Diagnostics: Reported vs. Imputed Installation Years

[Figure Note: Density comparison showing distributional plausibility of KNN-imputed dates relative to observed installation years. Imputed values should align with historical installation patterns.]

2 Component Universe

This section details the raw attributes available in the 'Compounds' table. Use these tables to determine which specific attributes to 'hotcode' into binary variables.

2.1 Tank Specifics

2.1.1 Release Detection

Table 7

Table 8Component Universe: TANK RELEASE DETECTION METHOD

COMPONENT_ATTRIBUTE_CODE	COMPONENT_TYPE	N
1299	OTHER	8
12A	MONTHLY INVENTORY CONTROL	3611
12B	ANNUAL TANK TIGHTNESS TESTING	2985
12C	TANK TIGHTNESS TESTING (EVERY 5 YEARS)	535
12D	STATISTICAL INVENTORY RECONCILIATION	4683
12E	AUTOMATIC TANK GAUGING	17834
12F	MANUAL TANK GAUGING (36 HRS)	407
12G	MANUAL TANK GAUGING (44 OR 58 HRS)	383
12H	INTERSTITIAL MONITORING (2 WALLS)	13753
12I	INTERSTITIAL MONITORING (LINER)	9
12J	GROUNDWATER MONITORING	79
12K	VAPOR MONITORING	58
12L	GROOVES MADE IN THE IMPERMEABLE PAD	2
12M	SLOTTED PIPE ABOVE THE IMPERMEABLE PAD	4
12N	NONE	2833
12O	EXEMPT	953

[Table Note: Identify primary detection methods.]

2.1.2 Tank Construction

Table 9

Table 10Component Universe: TANK CONSTRUCTION

COMPONENT_ATTRIBUTE_CODE	COMPONENT_TYPE	N
188	OTHER (COMPLIANT)	15
199	OTHER	69
1A	UNPROTECTED STEEL (SINGLE WALL)	5368
1B	CATHODICALLY PROTECTED STEEL (GALVANIC)	7094
1C	CATHODICALLY PROTECTED STEEL (IMPRESSED CURRENT)	2284
1D	UNPROTECTED STEEL (DOUBLE WALL)	75
1E	FIBERGLASS (SINGLE WALL)	8647
1F	FIBERGLASS (DOUBLE WALL)	8272
1G	STEEL W/PLASTIC OR FIBERGLASS JACKET (DOUBLE WALL)	5378
1H	STEEL W/FRP COATING (ACT 100 OR EQUIVALENT) (SINGLE WALL)	961
1I	STEEL W/LINED INTERIOR	750
1J	CONCRETE	31
1K	BOTTOM MODIFICATION	1
1N	UNKNOWN	194
1O	CATHODICALLY PROTECTED DOUBLE WALL STEEL (GALVANIC)	3100
1P	CATHODICALLY PROTECTED STEEL WITH LINER	572
1V	STEEL W/PLASTIC OR FRP JACKET W/ ANODES (DOUBLE WALL)	445
1W	STEEL W/FRP COATING W/ ANODES (SINGLE WALL)	15

[Table Note: Distinguish between Steel, Fiberglass, and Composite.]

2.2 Piping Infrastructure

2.2.1 Underground Piping Construction

Table 11

Table 12Component Universe: UG PIPING CONSTRUCTION

COMPONENT_ATTRIBUTE_CODE	COMPONENT_TYPE	N
288	OTHER (COMPLIANT)	31
299	OTHER	39
2A	BARE STEEL	5870
2B	CATHODICALLY PROTECTED, METALLIC	2137
2C	COPPER	213
2D	FIBERGLASS	6331
2E	FLEXIBLE NON-METALLIC	488
2F	UNKNOWN	218
2G	NONE	793
2H	MODIFICATION OF PIPING	15
2I	Double wall, metallic primary	490
2J	Double wall, rigid (FRP) primary	1623
2K	Double wall, flexible primary	2756
2L	TRENCH LINER	16
2M	JACKETED	15

[Table Note: Key variable for leak risk (e.g., Single Wall vs Double Wall, Galvanized vs Fiberglass).]

2.2.2 Underground Piping: Single Inner Wall

Table 13

Table 14Component Universe: UG SINGLE / INNER WALL PIPING

COMPONENT_ATTRIBUTE_CODE	COMPONENT_TYPE	N
2899	OTHER	34
28A	BARE STEEL	465
28B	CP PROTECTED	712
28C	COPPER	167
28D	FRP	8859
28E	FLEX	12507
28F	UNKNOWN	10
28G	NO DISPENSING PIPING	373
28I	STAINLESS STEEL	47

[Table Note: [Placeholder: Description of single inner wall piping characteristics.]]

2.2.3 Underground Piping: Outer Wall

Table 15

Table 16Component Universe: UG OUTER WALL PIPING

COMPONENT_ATTRIBUTE_CODE	COMPONENT_TYPE	N
2999	OTHER, OUTER	263
29A	BARE STEEL, OUTER	204
29B	CP PROTECTED, OUTER	51
29D	FRP, OUTER	3985
29E	FLEX, OUTER	12227
29F	UNKNOWN, OUTER	14
29I	POLY-ENCASED STAINLESS STEEL, OUTER	19
29N	NONE	6295

[Table Note: [Placeholder: Description of outer wall piping characteristics.]]

2.2.4 Aboveground Piping: Corrosion Protection

Table 17

Table 18Component Universe: AG PIPING CONSTRUCTION & CORROSION PROTECTION

COMPONENT_ATTRIBUTE_CODE	COMPONENT_TYPE	N
388	OTHER (COMPLIANT)	156
399	OTHER	21
3A	CARBON STEEL	615
3B	CATHODICALLY PROTECTED, METALLIC	72
3C	COPPER	30
3D	FIBERGLASS	40
3E	FLEXIBLE NON-METALLIC	29
3F	PVC	10
3G	NONE	448
3H	PIPING MODIFICATION	5
3I	DOUBLE WALL METALLIC PRIMARY	19
3J	DOUBLE WALL RIGID (FRP) PRIMARY	11
3K	DOUBLE WALL FLEXIBLE PRIMARY	8
3L	STAINLESS STEEL	9

[Table Note: [Placeholder: Description of aboveground piping corrosion protection methods.]]

2.2.5 Piping Release Detection

Table 19

Table 20Component Universe: PIPE RELEASE DETECTION METHOD

COMPONENT_ATTRIBUTE_CODE	COMPONENT_TYPE	N
5A	AUTOMATIC LINE LEAK DETECTOR	15109
5B	ANNUAL LINE TIGHTNESS TESTING (PRESSURE)	12413
5C	LINE TIGHTNESS TEST - 3 YEARS (SUCTION)	1646
5D	INTERSTITIAL MONITORING	10888
5E	GROUNDWATER MONITORING	63
5F	VAPOR MONITORING	33
5G	VISUAL INSPECTION	62
5H	NONE	2796
5I	EXEMPT	11903
5J	STATISTICAL INVENTORY RECONCILIATION	2558
5K	ELECTRONIC LINE LEAK DETECTOR	4762
5L	INTERSTITIAL MONITORING W/CONTINUOUS ALARM/SHUT OFF	7686

[Table Note: [Placeholder: Description of piping release detection methods.]]

2.2.6 Line Leak Detectors

Table 21

Table 22Component Universe: LINE LEAK DETECTOR SHUTS OFF PUMP

COMPONENT_ATTRIBUTE_CODE	COMPONENT_TYPE	N
23N	NO	17043
23Y	YES	10425

[Table Note: [Placeholder: Description of line leak detector functionality.]]

2.2.7 Secondary Containment

Table 23

Table 24Component Universe: UST TOTAL SECONDARILY CONTAINED

COMPONENT_ATTRIBUTE_CODE	COMPONENT_TYPE	N
18N	NO	20287
18Y	YES	12244

[Table Note: [Placeholder: Description of secondary containment systems.]]

3 Regulatory & Compliance

3.1 Certificates & Permits

Table 25

Table 26Component Universe: REGISTRATION CERTIFICATE

COMPONENT_ATTRIBUTE_CODE	COMPONENT_TYPE	N
8N	NO	1397
8Y	YES	11677

[Table Note: [Placeholder: Description of registration certificate requirements.]]

Table 27

Table 28Component Universe: FIRE MARSHAL PERMIT

COMPONENT_ATTRIBUTE_CODE	COMPONENT_TYPE	N
9A	ISSUED PRIOR TO AUGUST 5, 1989	3316
9B	ISSUED ON OR AFTER AUGUST 5, 1989	1288
9C	NO PERMIT OBTAINED	3709
9D	TANKS NOT REGULATED BY FIRE MARSHAL	268

[Table Note: [Placeholder: Description of fire marshal permit requirements.]]

3.2 Prevention Systems (Spill & Overfill)

Table 29

Table 30Component Universe: SPILL PREVENTION

COMPONENT_ATTRIBUTE_CODE	COMPONENT_TYPE	N
6D	DOUBLE WALL SPILL PREV	2564
6E	EXEMPT	512
6N	NO	5543
6S	SINGLE WALL SPILL PREV	4680
6Y	YES	29689

[Table Note: [Placeholder: Description of spill prevention systems.]]

Table 31

Table 32Component Universe: OVERFILL PREVENTION

COMPONENT_ATTRIBUTE_CODE	COMPONENT_TYPE	N
7A	OVERFILL ALARM	9490
7B	BALL FLOAT VALVE	2598
7E	EXEMPT	885
7N	NO	6183
7S	DROP TUBE SHUTOFF DEVICE	24814
7Y	YES	3550

[Table Note: Presence of these systems often correlates with lower premiums.]

3.3 Vapor Recovery Systems

Table 33

Table 34Component Universe: VAPOR RECOVERY

COMPONENT_ATTRIBUTE_CODE	COMPONENT_TYPE	N
11A	STAGE I INSTALLED	3693
11B	STAGE II INSTALLED	89
11C	STAGE I AND STAGE II INSTALLED	774
11D	NONE	4494

[Table Note: [Placeholder: Description of vapor recovery systems.]]

Table 35

Table 36Component Universe: STAGE I VAPOR RECOVERY

COMPONENT_ATTRIBUTE_CODE	COMPONENT_TYPE	N
19A	COAX	8099
19B	2 POINT	13101
19N	NONE OR INCOMPLETE	12382
2I	Double wall, metallic primary	1

[Table Note: [Placeholder: Description of Stage I vapor recovery.]]

Table 37

Table 38Component Universe: STAGE II VAPOR RECOVERY

COMPONENT_ATTRIBUTE_CODE	COMPONENT_TYPE	N
20A	COMPLETE BALANCE SYSTEM	623
20B	COMPLETE ASSIST SYSTEM	2256
20C	UG PIPING ONLY	6248
20D	DECOMMISSIONED	2359
20N	NONE	21258

[Table Note: Stage II recovery is largely phased out; check if this indicates older infrastructure.]

3.4 Containment & Sumps

Table 39

Table 40Component Universe: TANK-TOP CONTAINMENT SUMPS

COMPONENT_ATTRIBUTE_CODE	COMPONENT_TYPE	N
21A	AT ALL PENETRATIONS	18857
21N	NONE	7554
21S	AT SOME PENETRATIONS	1127

[Table Note: [Placeholder: Description of tank top containment sumps.]]

Table 41

Table 42Component Universe: UNDER-DISPENSER CONTAINMENT

COMPONENT_ATTRIBUTE_CODE	COMPONENT_TYPE	N
22A	AT ALL DISPENSERS	18566
22N	NONE	8667
22S	AT SOME DISPENSERS	253

[Table Note: ‘UDC’ (Under Dispenser Containment) is a critical modern safety feature.]

4 Miscellaneous Components

4.1 Flexible Connectors

Table 43

Table 44Component Universe: Piping Flexible Connectors

COMPONENT_ATTRIBUTE_CODE	COMPONENT_TYPE	N
88	Other (Compliant)	225
99	Other (Noncompliant)	4
PFLXA	Unprotected Metallic Components (incl wrapped or coated)	478
PFLXB	Cathodically Protected, Metallic	2465
PFLXC	Flexible Coupling w/ Protected Metallic Ends	53
PFLXD	Completely Inside Containment Sump, Secondary Pipe or Liner	4816
PFLXE	Completely Jacketed w/ Sealed Boot	1596
PFLXF	Not in Contact w/ Ground	976
PFLXX	None	197
UNK	Unknown	336

[Table Note: [Placeholder: Description of flexible connector types.]]

Table 45

Table 46Component Universe: FLEX - TANK END

COMPONENT_ATTRIBUTE_CODE	COMPONENT_TYPE	N
2699	OTHER	343
26A	UNPROTECTED METALLIC COMPONENTS (INCL WRAPPED OR COATED)	51
26B	CATHODICALLY PROTECTED, METALLIC	1676
26F	UNKNOWN	338
26I	COMPLETELY INSIDE CONTAINMENT SUMP, SECONDARY PIPE OR LINER	17133
26M	COMPLETELY JACKETED W/ SEALED BOOT	876
26N	NOT IN CONTACT W/ GROUND	1355
26X	NONE	476

[Table Note: [Placeholder: Description of flexible connectors at tank end.]]

Table 47

Table 48Component Universe: FLEX - DISPENSER END

COMPONENT_ATTRIBUTE_CODE	COMPONENT_TYPE	N
2799	OTHER	190
27A	UNPROTECTED METALLIC COMPONENTS (INCL WRAPPED OR COATED)	27
27B	CATHODICALLY PROTECTED, METALLIC	1788
27F	UNKNOWN	34
27I	COMPLETELY INSIDE CONTAINMENT SUMP, SECONDARY PIPE OR LINER	16280
27M	COMPLETELY JACKETED W/ SEALED BOOT	1926
27N	NOT IN CONTACT W/ GROUND	1344
27X	NONE	763

[Table Note: [Placeholder: Description of flexible connectors at dispenser end.]]

4.2 Other Hardware

Table 49

Table 50Component Universe: PUMP/DELIVERY SYSTEM

COMPONENT_ATTRIBUTE_CODE	COMPONENT_TYPE	N
4A	SUCTION: CHECK VALVE AT PUMP	11641
4B	SUCTION: CHECK VALVE AT TANK	3899
4C	PRESSURE	25845
4D	GRAVITY FED	201
4E	NONE	1110

[Table Note: [Placeholder: Description of pump delivery system types.]]

Table 51

Table 52Component Universe: EMERGENCY GENERATOR

COMPONENT_ATTRIBUTE_CODE	COMPONENT_TYPE	N
25N	NO - EMER GEN	21543
25Y	YES - EMER GEN	678

[Table Note: [Placeholder: Description of emergency generator installations.]]

4.3 Unclassified / Other

Table 53

Table 54Component Universe: NA

COMPONENT_ATTRIBUTE_CODE	COMPONENT_TYPE	N

[Table Note: [Placeholder: Description of unclassified or N/A component entries.]]

Table 55

Table 56Component Universe: TANK UPGRADE

COMPONENT_ATTRIBUTE_CODE	COMPONENT_TYPE	N
10A	TANK WAS RETROFITTED WITH CATHODIC PROTECTION	1689
10B	TANK WAS RETROFITTED WITH LINING	859
10C	TANK WAS RETROFITTED WITH RIGID BLADDER (EX. PHOENIX SYS)	4

[Table Note: [Placeholder: Description of tank upgrade activities.]]

5 Owner & Business Intelligence

This section characterizes facility ownership structure, business model classification, and market segmentation derived from the harmonized facility linkage table.

5.1 Business Model Distribution

[Description: Classification of facilities by business category based on owner sector and facility count logic.]

Table 57

Table 58Facility Counts by Business Category

business_category	N
Unknown/Unclassified	58226
Private Firm - Non-motor fuel seller	10646
Retail Gas (Branded Commercial)	10223
Retail Gas - Single Proprietor	9258
Retail Gas - Multi-property Not Branded	6746
Non-Retail: Fleet Fuel Facility	3341
Publicly Owned	3261

[Table Note: Business categories derived from owner name pattern matching and facility count thresholds. “Publicly Owned” includes government and municipal entities; “Retail Gas (Branded Commercial)” includes major chains.]

5.2 Owner Fleet Size Distribution

[Description: Distribution of facilities and tanks across owner size classes.]

Table 59

Table 60Distribution by Owner Fleet Size

Owner_Size_Class	Facilities	Tanks	Pct_Tanks
Unknown/Unlinked	22863	58226	57.3%
Single-Site Owner (Mom & Pop)	3799	18631	18.3%
Large Fleet/Corporate (50+)	2049	10746	10.6%
Small Fleet (2-9)	1510	7913	7.8%
Medium Fleet (10-49)	1114	6185	6.1%

[Table Note: Owner size classification based on facility count per owner: Single-Site (1), Small Fleet (2-9), Medium Fleet (10-49), Large Fleet/Corporate (50+).]

5.3 Owner Sector Breakdown

[Description: Top 25 owner sectors by tank count, derived from owner name pattern classification.]

Table 61

Table 62Top 25 Owner Sectors by Tank Count

final_owner_sector	Facilities	Tanks
Unknown	22863	58226
Private Commercial/Other	3098	16004
Real Estate/Property Mgmt	1703	9437
Major Chain (Sheetz)	313	2126
Major Chain (7-Eleven)	182	1377
Local Govt/Muni	395	1324
State Govt/Agency	189	1011
Major Chain (Sunoco)	153	969
Utility/Energy	179	869
Major Chain (Wawa)	184	843
Major Chain (Speedway)	90	838
Major Chain (Turkey Hill/EG)	215	831
Education/School	168	809
Trucking/Logistics	171	806
Major Chain (United Refining)	140	802
Major Chain (GetGo/Giant)	199	695
Auto Dealership/Repair	128	694
Construction/Development	140	630
Major Chain (Rutters)	81	518
Major Chain (Other Fuel Brand)	81	467
Recreation/Hospitality	89	360
Agriculture	89	360
Utility/Telecom	166	342
Healthcare	58	262
Major Chain (Country Fair)	53	230

[Table Note: Sector classification uses regex pattern matching on owner names to identify major chains, government entities, utilities, and commercial sectors.]

5.4 Facility Operational Status

[Description: Aggregate facility-level operational status based on tank closure patterns.]

Table 63

Table 64Facility Operational Status

facility_status	N	Pct
Fully Closed	24428	78.0%
Mixed Status	4057	12.9%
Fully Active	2850	9.1%

[Table Note: “Fully Active” = all tanks in use; “Fully Closed” = all tanks closed; “Mixed Status” = some active/some closed.]

5.5 Owner Size vs Business Model Cross-Tabulation

[Description: Cross-tabulation of owner fleet size against business category to identify structural patterns.]

Table 65

Table 66Facility Counts: Owner Size vs Business Model

Owner_Size_Class	Non-Retail: Fleet Fuel Facility	Private Firm - Non-motor fuel seller	Publicly Owned	Retail Gas (Branded Commercial)	Retail Gas - Multi-property Not Branded	Retail Gas - Single Proprietor	Unknown/Unclassified
Large Fleet/Corporate (50+)	156	3	151	1667	72	0	0
Medium Fleet (10-49)	151	403	68	84	408	0	0
Single-Site Owner (Mom & Pop)	303	1166	429	41	0	1860	0
Small Fleet (2-9)	174	412	151	15	758	0	0
Unknown/Unlinked	0	0	0	0	0	0	22863

[Table Note: Cell values represent unique facility counts. Useful for identifying which business models are dominated by small vs. large operators.]

5.6 Closure Rates by Business Category

[Description: Tank closure rates stratified by business category to identify differential attrition patterns.]

Table 67

Table 68Tank Closure Rates by Business Category

business_category	Total_Tanks	Closed_Tanks	Active_Tanks	Closure_Rate
Unknown/Unclassified	58226	58226	0	100.0%
Private Firm - Non-motor fuel seller	10646	5152	5494	48.4%
Retail Gas (Branded Commercial)	10223	3372	6851	33.0%
Retail Gas - Single Proprietor	9258	5155	4103	55.7%
Retail Gas - Multi-property Not Branded	6746	3573	3173	53.0%
Non-Retail: Fleet Fuel Facility	3341	2127	1214	63.7%
Publicly Owned	3261	2248	1013	68.9%

[Table Note: Higher closure rates may indicate market exit, consolidation, or infrastructure modernization patterns specific to certain business types.]

6 Temporal Evolution & Trends

This section presents visualizations of how tank characteristics, fuel types, and facility attributes have evolved over time. These temporal patterns inform understanding of regulatory compliance trends and infrastructure modernization.

6.1 Capacity Distribution

[Description: Overall distribution of tank capacities across the fleet.]

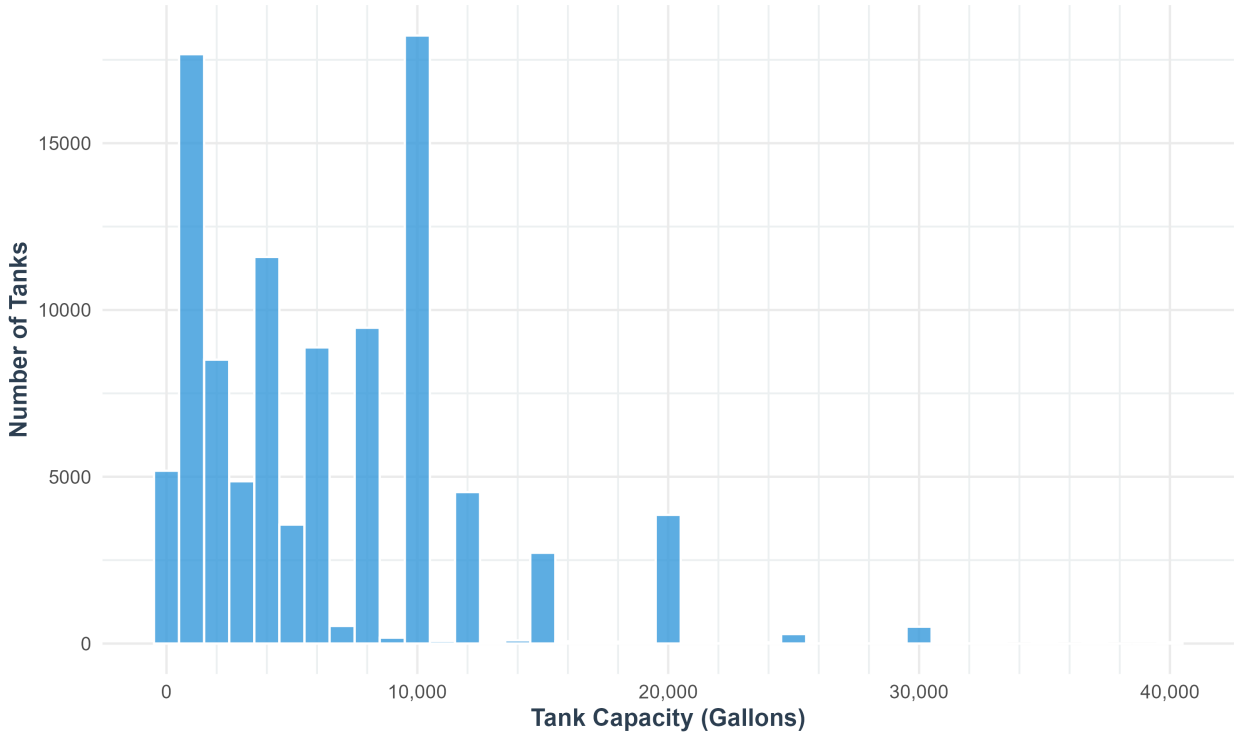

Figure 2: Distribution of Tank Capacities (Gallons)

[Figure Note: Modal peaks at standard capacities (e.g., 10,000, 12,000, 15,000 gallons) indicate industry standardization. Long right tail represents commercial/industrial facilities.]

6.2 Capacity Evolution by Decade

[Description: Distribution of tank capacities across installation decades, showing trends in tank sizing over time.]

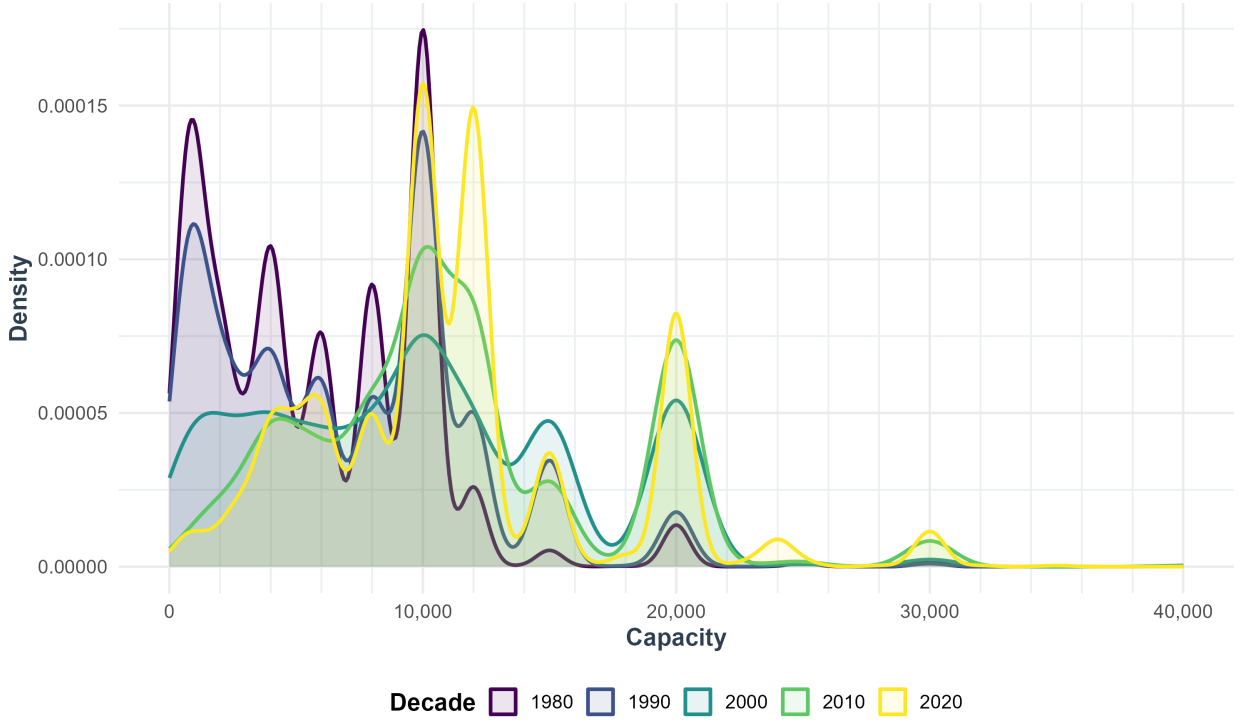

Figure 3: Evolution of Tank Capacity Distribution by Installation Decade

[Figure Note: [Placeholder: Interpretation of capacity trends—have tanks gotten larger over time? Implications for replacement costs and risk assessment.]]

6.3 Fuel Mix Evolution

[Description: Temporal shift in the proportion of gasoline, diesel, and other fuel types in new tank installations.]

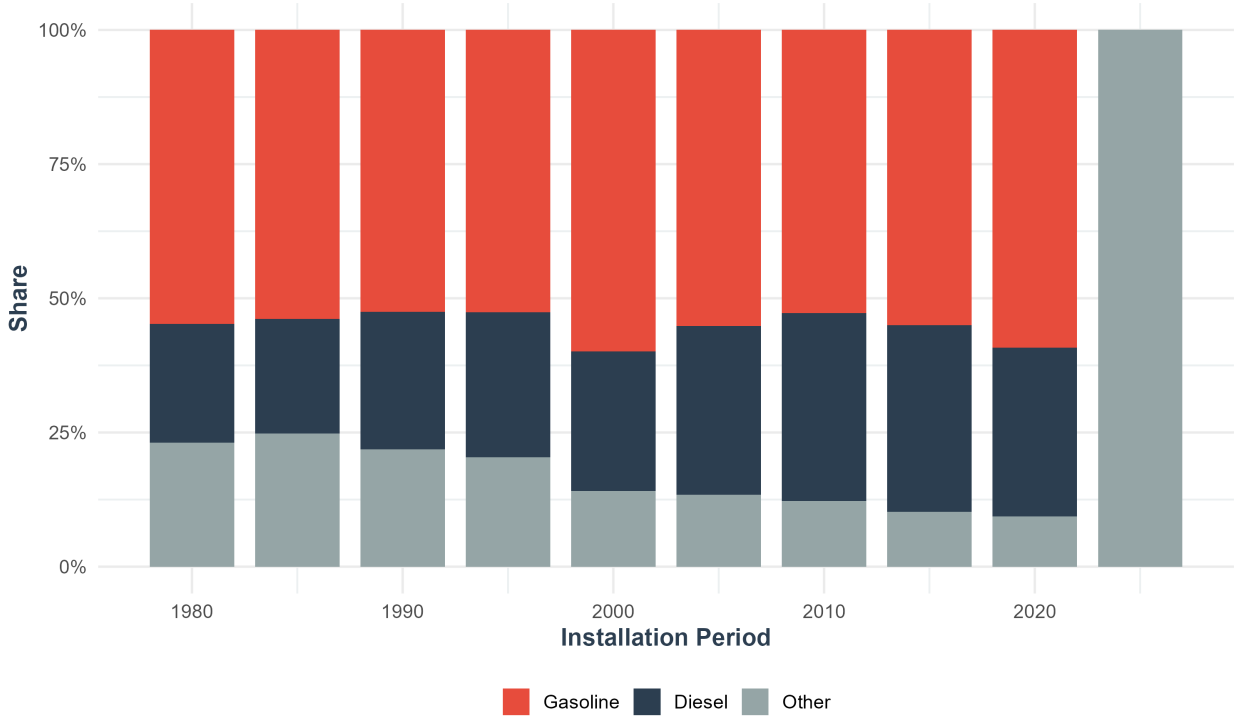

Figure 4: Evolution of Fuel Type Mix in New Installations (5-Year Periods)

[Figure Note: [Placeholder: Analysis of fuel mix shifts—decline in gasoline share? Growth in diesel? Implications for risk profiles and insurance premiums.]]

6.4 Tank Lifespan Distribution

[Description: Distribution of tank ages at closure, showing typical service life and identifying outliers.]

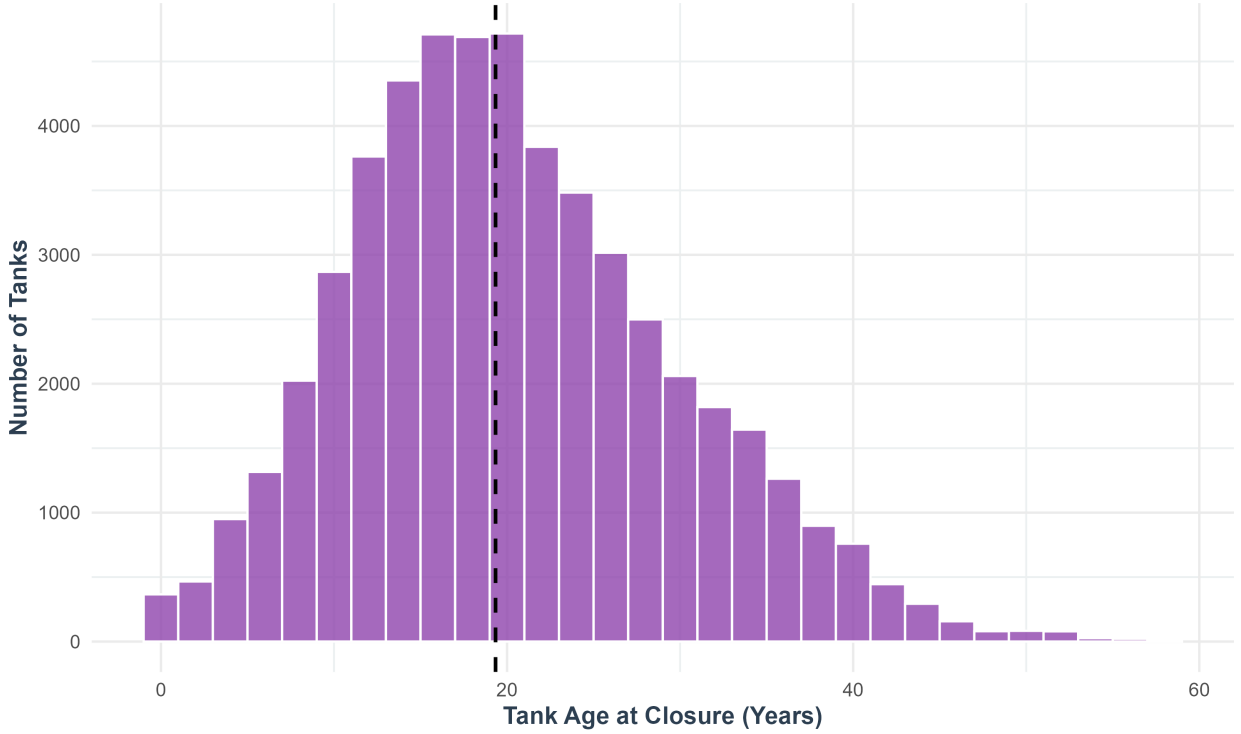

Figure 5: Distribution of Tank Age at Closure

[Figure Note: [Placeholder: Interpretation of lifespan patterns—median age, typical range, and implications for remaining useful life of active tanks.]]

7 Facility-Level Intelligence

This section examines facility-level aggregations, including size distributions, survival patterns, and relationships between facility age and complexity.

7.1 Facility Size Evolution by Decade

[Description: Distribution of facility sizes (number of tanks per facility) across facility vintage decades.]

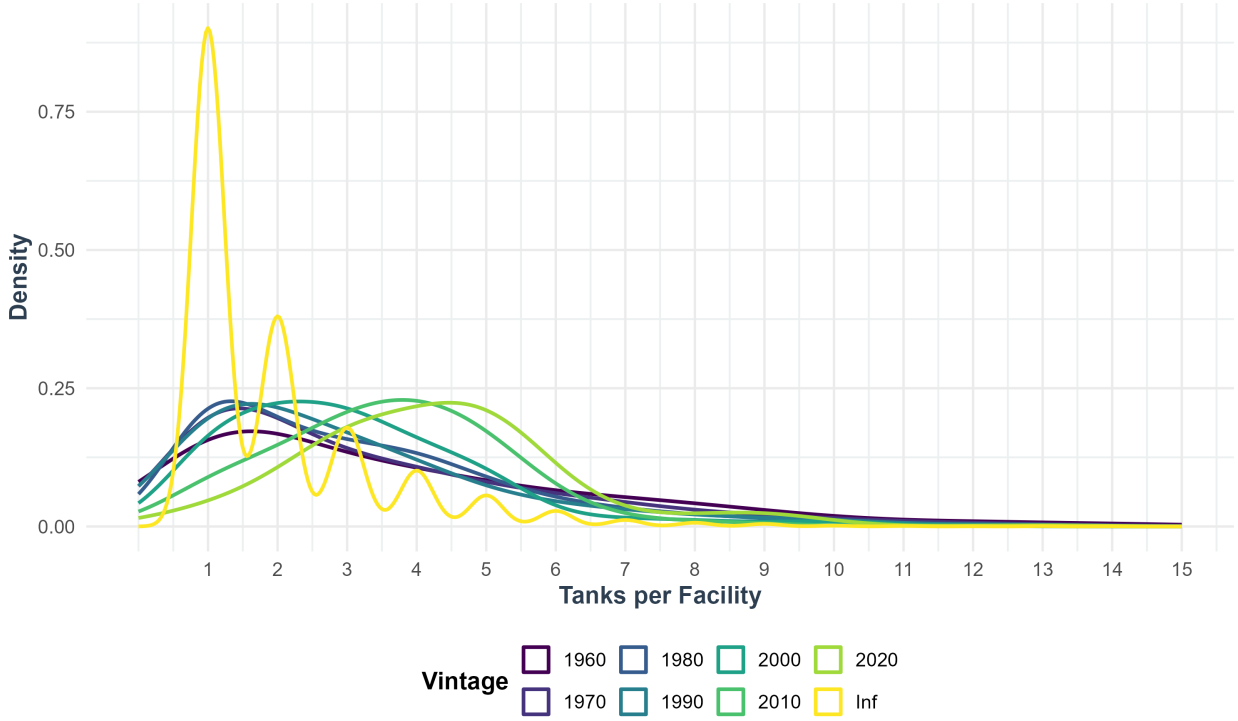

Figure 6: Evolution of Facility Size Distribution by Vintage Decade

[Figure Note: [Placeholder: Interpretation of facility size trends—are newer facilities larger or smaller? Implications for operational complexity and risk concentration.]]

7.2 Facility Status by Vintage

[Description: Proportion of facilities in different status categories (Fully Active, Fully Closed, Mixed Status) by vintage decade.]

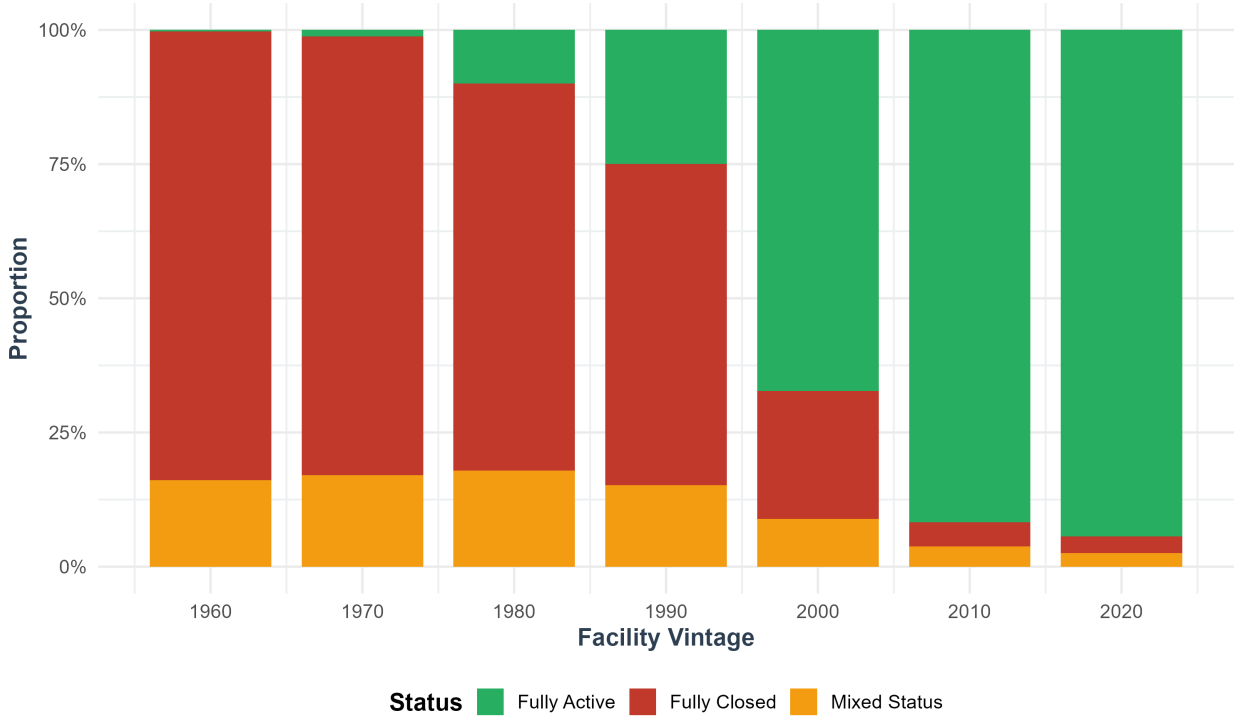

Figure 7: Facility Survival Status by Vintage Decade

[Figure Note: [Placeholder: Analysis of facility survival patterns—do older facilities show higher closure rates? Implications for portfolio risk assessment.]]

7.3 Facility Age vs. Size Relationship

[Description: Scatter plot examining the relationship between facility age and number of tanks, with smoothed trend line.]

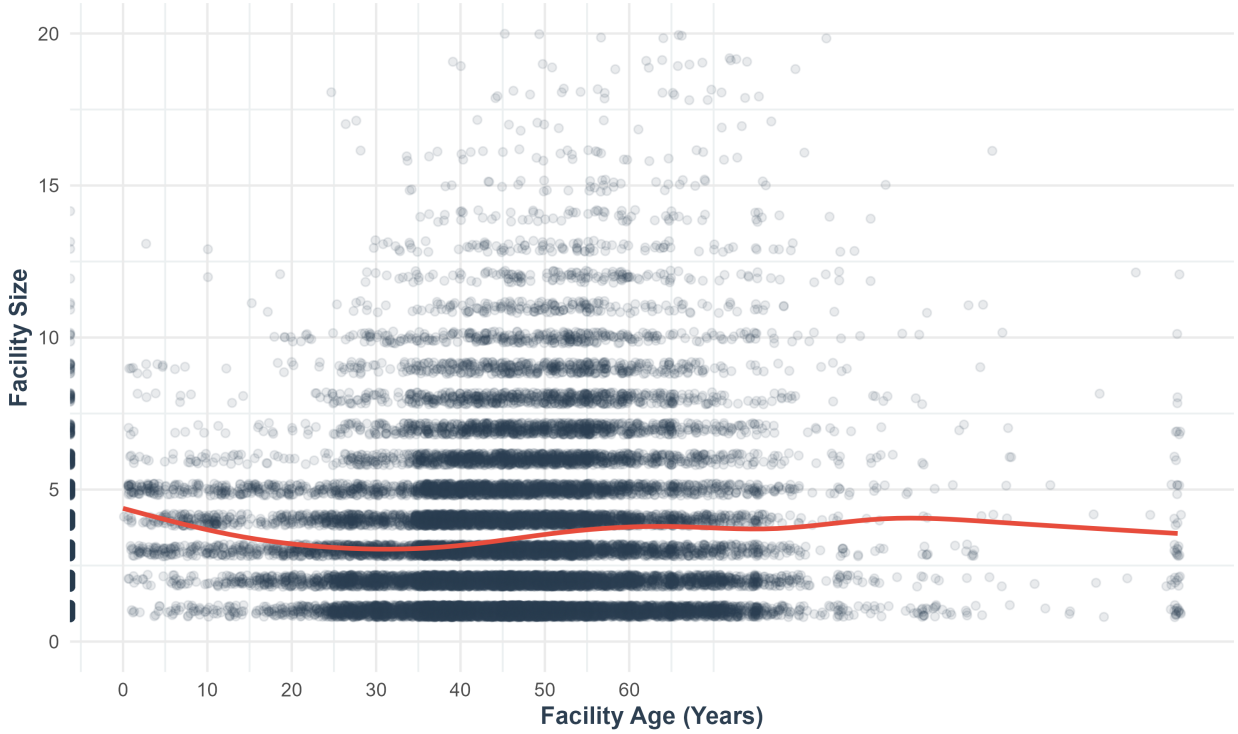

Figure 8: Relationship Between Facility Age and Size

[Figure Note: [Placeholder: Interpretation of age-size relationship—do older facilities tend to be larger? Implications for modernization costs and operational risk.]]

8 Fleet Risk & Infrastructure Analytics

This section examines construction-based risk tiers and capacity standardization trends to characterize infrastructure modernization patterns.

8.1 Fleet Risk Tier Transition

[Description: Temporal evolution of tank construction risk tiers based on wall construction (single vs. double) and material type.]

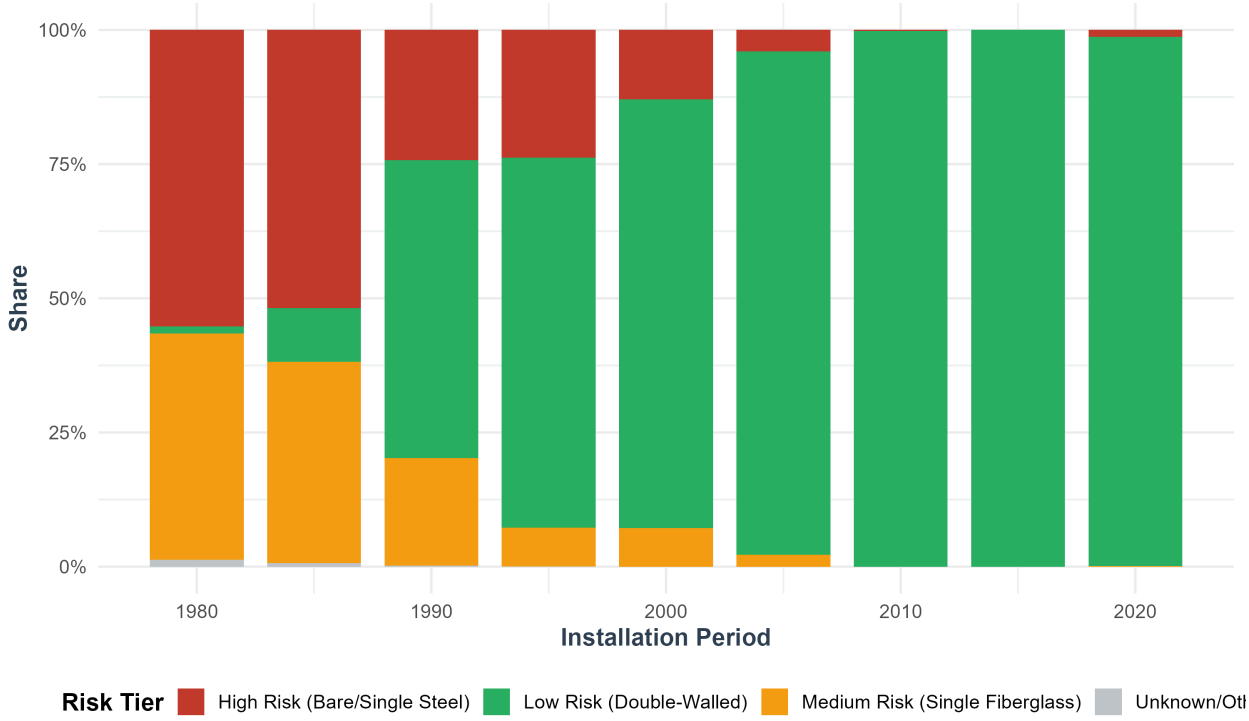

Figure 9: Fleet Risk Tier Transition by Installation Period

[Figure Note: Risk tiers derived from TANK CONSTRUCTION component: "High Risk" = bare/single-wall steel; "Medium Risk" = single-wall fiberglass; "Low Risk" = double-walled or jacketed tanks. Transition toward lower-risk construction reflects regulatory evolution (e.g., 1998 EPA deadline).]

8.2 Capacity Standardization Trends

[Description: Installation frequency of standard tank capacities over time, showing market convergence on common sizes.]

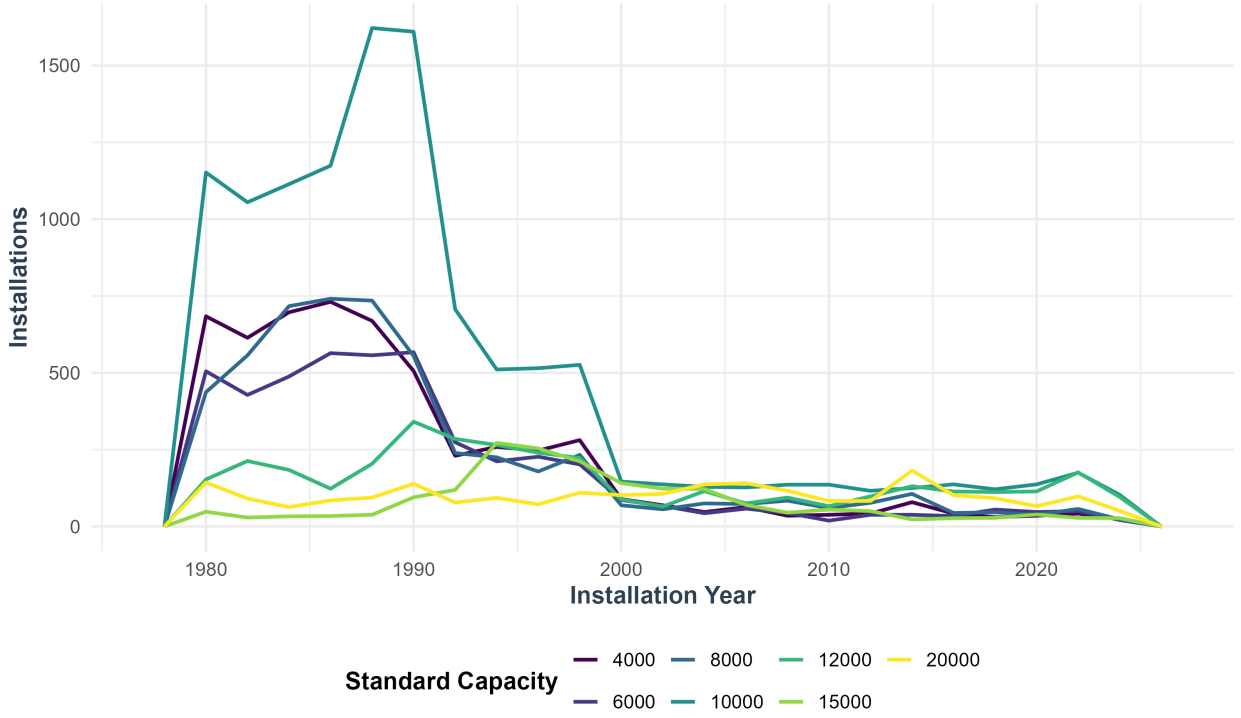

Figure 10: Installation Trends for Standard Tank Capacities (1980+)

[Figure Note: Standard capacities (4000, 6000, 8000, 10000, 12000, 15000, 20000 gallons) shown. Trends indicate market preferences and potential cost efficiencies from standardization.]

9 Owner & Market Structure Analytics

This section presents visualizations of ownership patterns, market concentration, and temporal evolution of business categories.

9.1 Owner Size Distribution

[Description: Distribution of tank counts across owner size classes.]

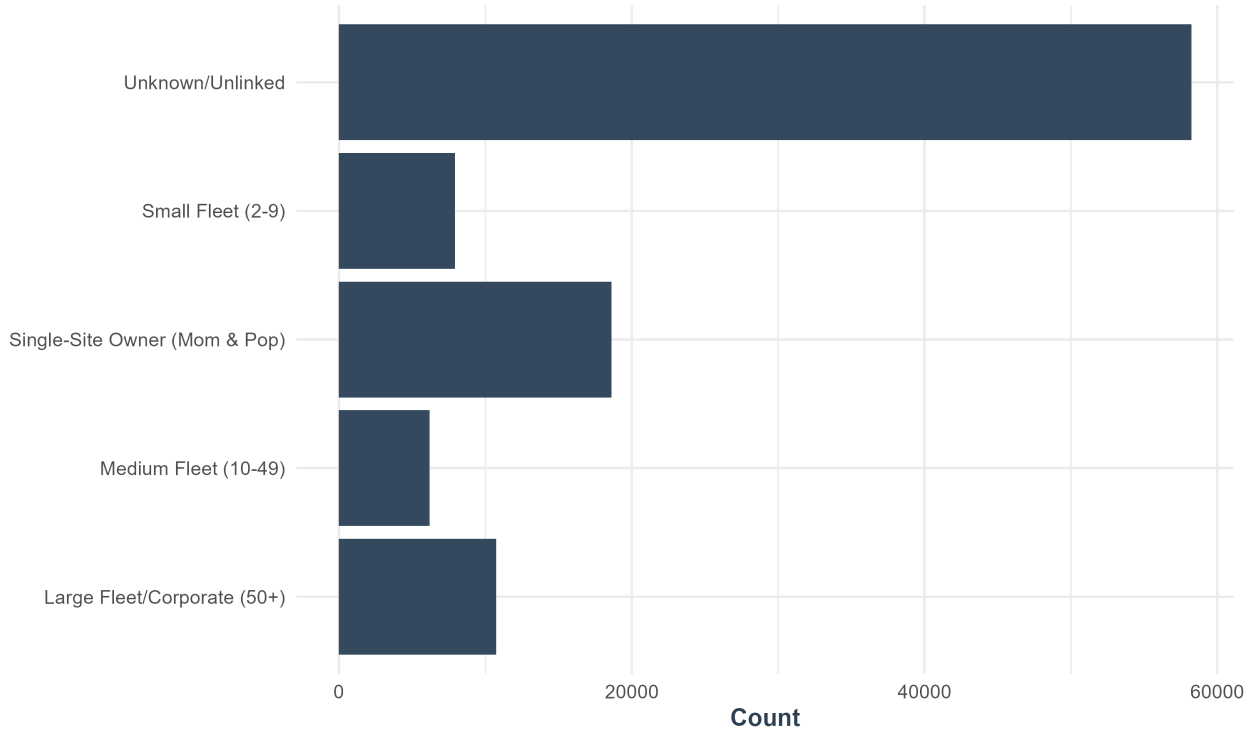

Figure 11: Tank Count Distribution by Owner Size Class

[Figure Note: Horizontal bar chart showing relative market share by owner fleet size. Single-site operators vs. corporate fleet concentration.]

9.2 Mom & Pop Sector Breakdown

[Description: Sector distribution within single-site ("Mom & Pop") owners, showing business type diversity.]

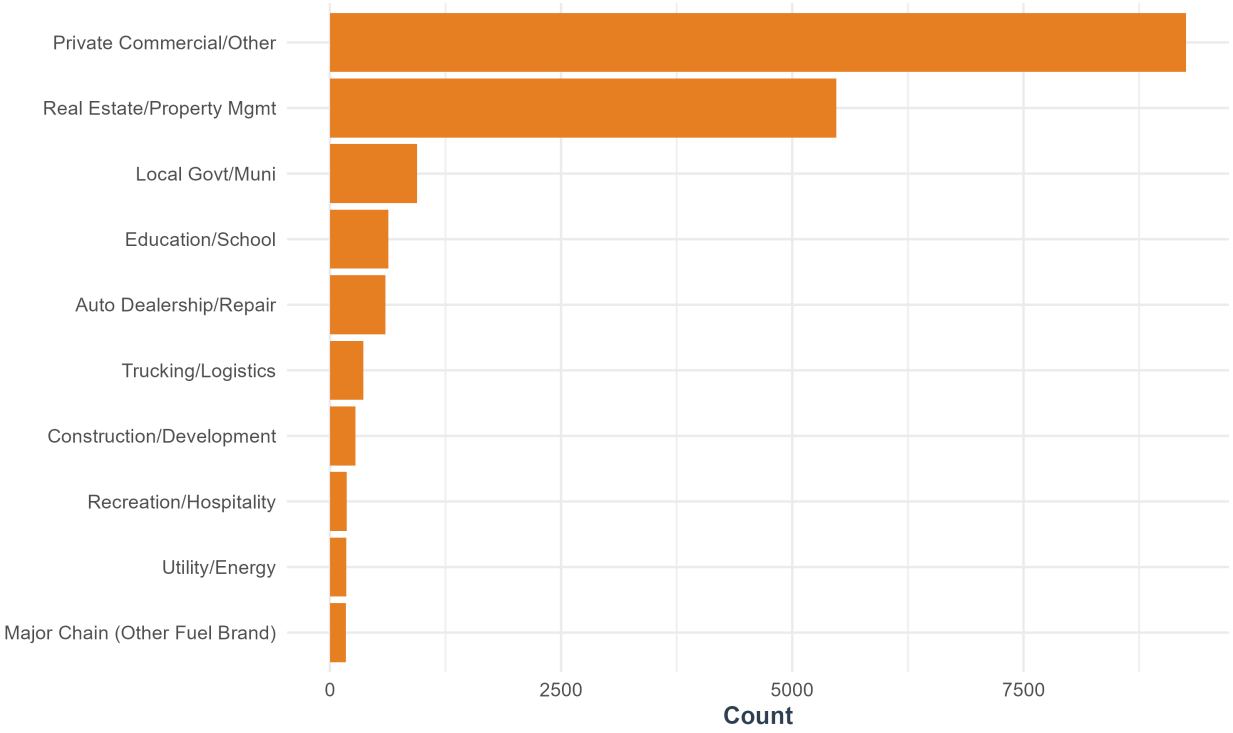

Figure 12: Top 10 Sectors Among Single-Site Owners

[Figure Note: Identifies which business sectors dominate the single-site owner segment—critical for understanding small operator vulnerability to remediation costs.]

9.3 Major Chains Market Share

[Description: Tank counts for major retail chains operating in Pennsylvania.]

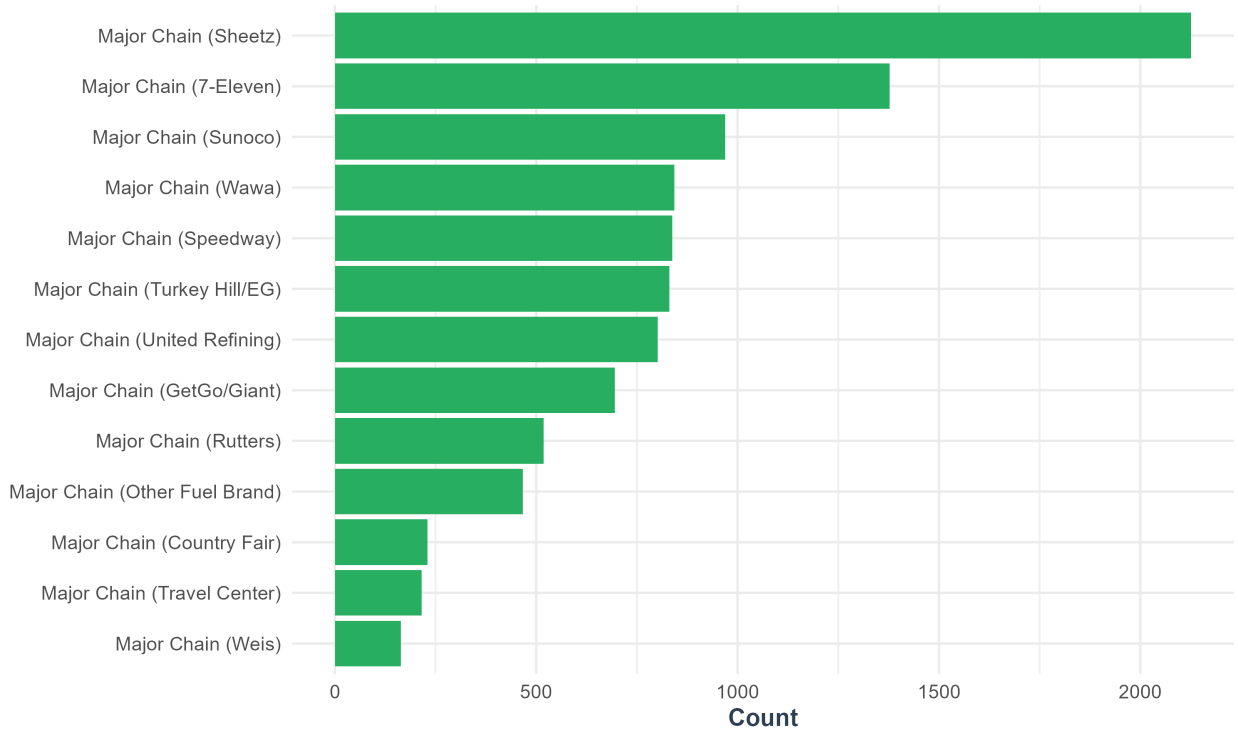

Figure 13: Major Chains Market Share by Tank Count

[Figure Note: Identifies dominant retail fuel chains (Sheetz, Wawa, GetGo, etc.) and their relative infrastructure footprints in Pennsylvania.]

9.4 Owner Size Evolution

[Description: Temporal evolution of installation share by owner size class.]

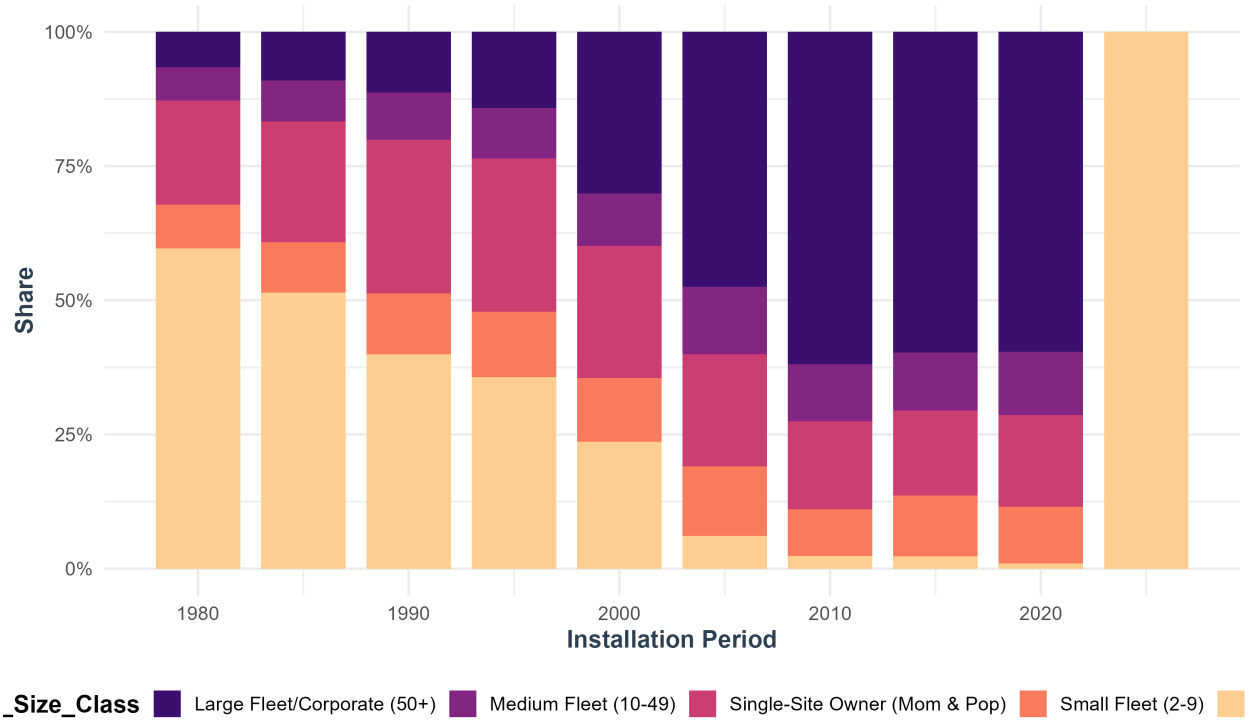

Figure 14: Evolution of Tank Installations by Owner Size Class (5-Year Periods)

[Figure Note: Stacked area showing market consolidation trends—increasing corporate share vs. declining single-site operator share over time.]

9.5 Business Category Evolution

[Description: Temporal evolution of installation share by business category classification.]

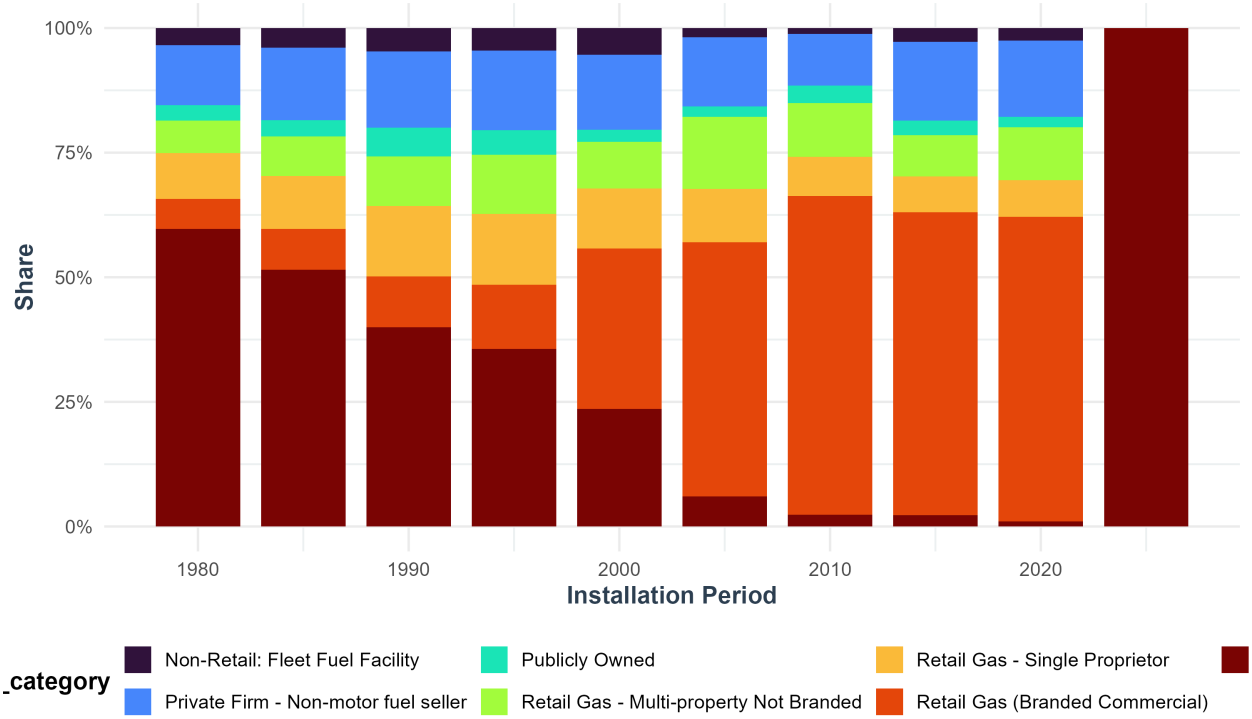

Figure 15: Evolution of Tank Installations by Business Category (5-Year Periods)

[Figure Note: Shows structural shifts in who is installing tanks—retail gas, fleet operations, public sector, etc. Useful for understanding changing composition of USTIF portfolio.]

10 Closure Dynamics

This section examines tank closure patterns over time, stratified by facility type.

10.1 Tank Closures by Facility Type Timeline

[Description: Annual tank closures by business category, showing differential exit/modernization patterns.]

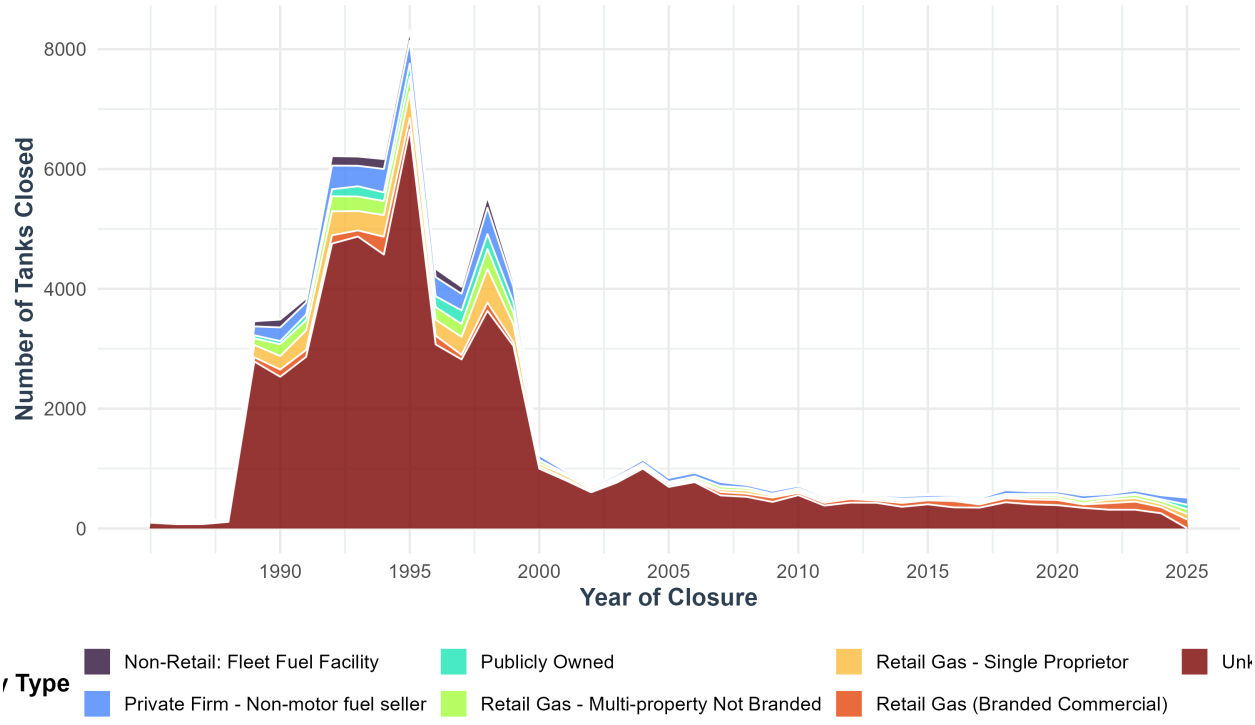

Figure 16: Tank Closures by Facility Type (1985-2025)

[Figure Note: Stacked area chart showing closure volume by business category over time. Peaks may correspond to regulatory deadlines (e.g., 1998 EPA compliance), market consolidation waves, or economic shocks. Useful for understanding historical claim volume drivers.]

11 USTIF Claims & Contracts Analysis

This section provides a summary of the USTIF claims and remediation contracts, adjusted for inflation to **2024 Real Dollars**.

11.1 Claims Overview

[Description: Operational status and financial severity of claims, adjusted for inflation.]

Table 69

Table 70USTIF Claims: Status & Real Costs

claim_status	Count	Avg_Real_Cost	Total_Real_Paid	Share
Closed Eligible	4407	$361,543	$1,593,320,744	56.6%
Closed Withdrawn	1264	$1,502	$1,898,367	16.2%
Closed Denied	1157	$5,086	$5,884,614	14.8%
Open Eligible	627	$607,121	$380,664,923	8.0%
Closed Post Remedial Care	217	$748,635	$162,453,746	2.8%
Open Pending	97	$1,058	$102,618	1.2%
Open Post Remedial Care	18	$1,176,190	$21,171,415	0.2%
Open Appealed	5	$16,303	$81,516	0.1%

[Table Note: Claims categorized by operational status with real dollar costs adjusted to 2024 base year.]

11.2 Financial Severity

[Description: Distribution of real claim costs (log scale) showing the prevalence of high-severity events.]

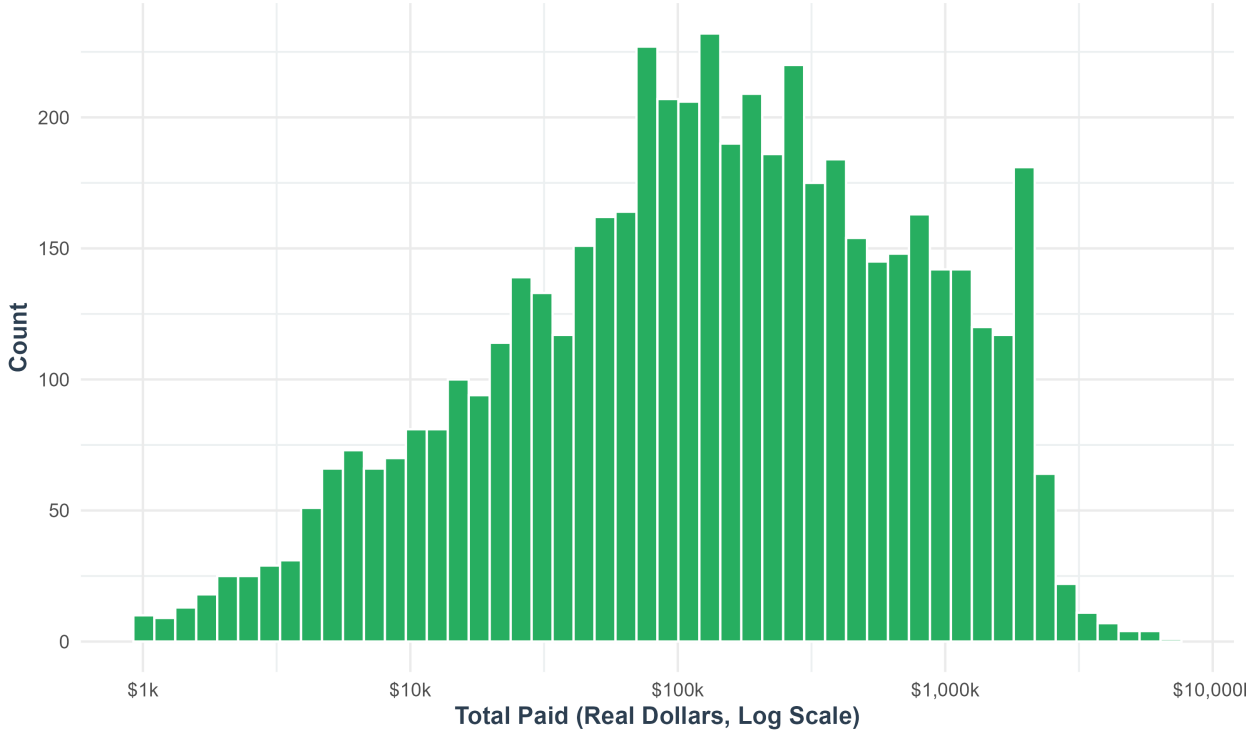

Figure 17: Distribution of Real Claim Severities (2024 Dollars)

[Figure Note: Log-scale distribution reveals long right tail characteristic of environmental remediation costs.]

11.3 Contracts & Auctions

[Description: Analysis of remediation contracts, comparing Bid-to-Result mechanisms against traditional models.]

Table 71

Table 72Contracts: Real Value by Mechanism

auction_type	Contracts	Total_Real_Value	Median_Real_Value
Other/Unknown	398	$110,541,797	$173,405
Bid-to-Result	83	$42,289,577	$438,017
Scope of Work	177	$28,650,874	$109,601

[Table Note: Contract mechanism breakdown showing PFP auction vs. T&M allocation patterns.]

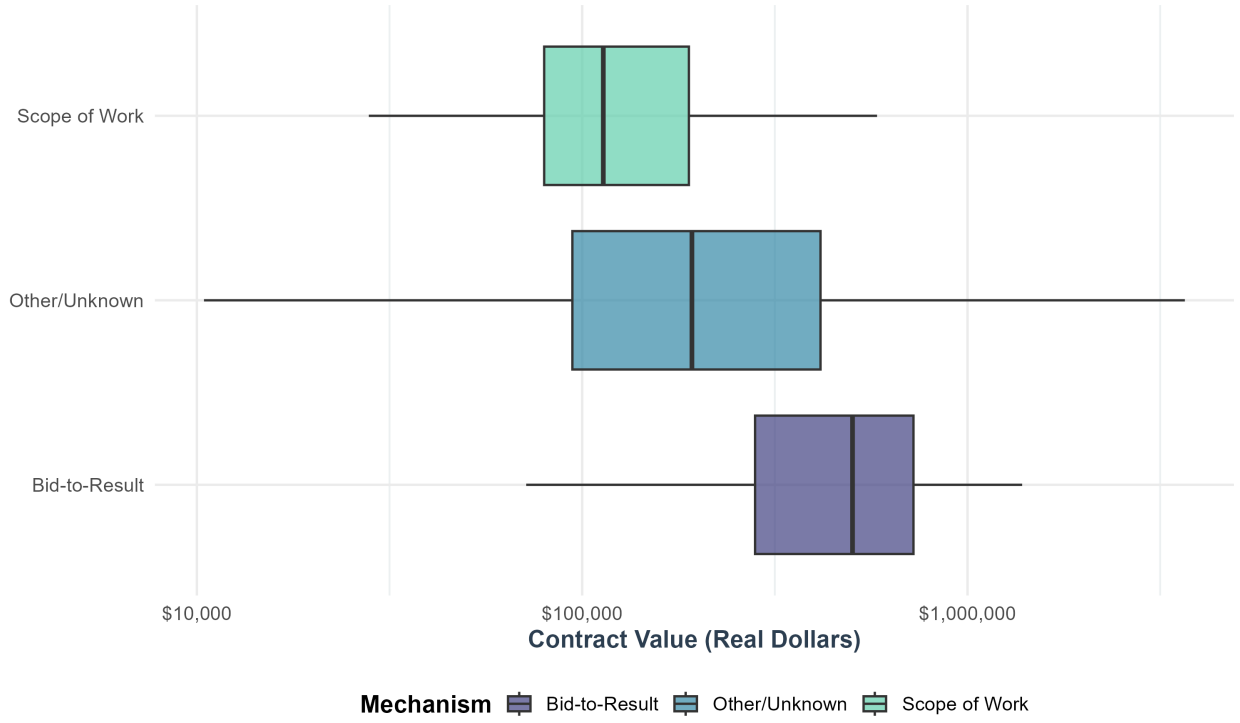

Figure 18: Contract Value Distribution by Mechanism (Real 2024 Dollars)

[Figure Note: Boxplot comparison of contract values by mechanism type. Caution: raw comparisons subject to selection bias per "naive fallacy" identified in research design.]

12 Appendix: Variable Inventory

[Description: Complete frequency dictionary for categorical variables in the Claims and Contracts datasets.]

12.1 Claims Variables

12.1.1 Claim Status

Table 73

Table 74Variable Inventory: $claim_status$ (Claims Data)

Value	N	Share
Closed Eligible	4407	56.6%
Closed Withdrawn	1264	16.2%
Closed Denied	1157	14.8%
Open Eligible	627	8.0%
Closed Post Remedial Care	217	2.8%
Open Pending	97	1.2%
Open Post Remedial Care	18	0.2%
Open Appealed	5	0.1%

[Table Note: Distribution of claim operational status (open/closed).]

12.1.2 DEP Region

Table 75

Table 76Variable Inventory: dep_region (Claims Data)

Value	N	Share
PADEP Southeast Regional Office	2018	25.9%
PADEP Southwest Regional Office	1512	19.4%
PADEP Southcentral Regional Office	1433	18.4%
PADEP Northeast Regional Office	1333	17.1%
PADEP Northwest Regional Office	907	11.6%
PADEP Northcentral Regional Office	589	7.6%

[Table Note: Geographic distribution of claims across PA DEP regional offices.]

12.1.3 County

Table 77

Table 78Variable Inventory: county (Claims Data)

Value	N	Share
Allegheny	687	8.8%
Montgomery	544	7.0%
Philadelphia	502	6.4%
Bucks	395	5.1%
Delaware	303	3.9%
Luzerne	285	3.7%
Chester	274	3.5%
Westmoreland	254	3.3%
Lancaster	254	3.3%
Erie	218	2.8%
York	212	2.7%
Berks	190	2.4%
Lehigh	188	2.4%
Dauphin	186	2.4%
Northampton	179	2.3%
Monroe	177	2.3%
Washington	171	2.2%
Lackawanna	161	2.1%
Schuylkill	133	1.7%
Beaver	113	1.5%
Mercer	106	1.4%
Fayette	105	1.3%
Butler	104	1.3%
Cumberland	104	1.3%
Blair	100	1.3%
Lycoming	95	1.2%
Clearfield	81	1.0%
Cambria	80	1.0%
Indiana	72	0.9%
Franklin	72	0.9%
Lebanon	70	0.9%
Adams	69	0.9%
Centre	68	0.9%
Somerset	64	0.8%
Northumberland	62	0.8%
Venango	61	0.8%
Bedford	61	0.8%
Columbia	58	0.7%
Armstrong	57	0.7%
Carbon	57	0.7%
Crawford	52	0.7%
Susquehanna	51	0.7%
Lawrence	46	0.6%

[Table Note: County-level claim frequency distribution.]

12.1.4 Location Description

Table 79

Table 80Variable Inventory: $location_desc$ (Claims Data)

Value	N	Share
Commercial	7292	93.6%
Local Government	251	3.2%
Private	237	3.0%
State Government	11	0.1%
	1	0.0%

[Table Note: Facility location type classification.]

12.1.5 Products

Table 81

Table 82Variable Inventory: products (Claims Data)

Value	N	Share
Unleaded Gasoline	4881	62.6%
Diesel	807	10.4%
Unleaded Gasoline, Diesel	553	7.1%
Heating Oil	329	4.2%
Unleaded Gasoline, Kerosene, Diesel	190	2.4%
Unleaded Gasoline, Other	175	2.2%
Unleaded Gasoline, Kerosene	169	2.2%
Other	142	1.8%
Kerosene	96	1.2%
Unleaded Gasoline, Heating Oil, Diesel	73	0.9%
Unleaded Gasoline, Other, Diesel	69	0.9%
Unleaded Gasoline, Kerosene, Heating Oil, Diesel	46	0.6%
Unleaded Gasoline, Heating Oil	42	0.5%
Unleaded Gasoline, Other, Heating Oil	41	0.5%
Unleaded Gasoline, Other, Heating Oil, Diesel	33	0.4%
Unleaded Gasoline, Other, Kerosene, Diesel	25	0.3%
Heating Oil, Diesel	16	0.2%
Other, Diesel	16	0.2%
Unleaded Gasoline, Other, Kerosene	15	0.2%
Kerosene, Diesel	14	0.2%
Unleaded Gasoline, Other, Kerosene, Heating Oil, Diesel	11	0.1%
Unleaded Gasoline, Kerosene, Heating Oil	9	0.1%
Other, Heating Oil	7	0.1%
Other, Heating Oil, Diesel	6	0.1%
Unleaded Gasoline, Undetermined	5	0.1%
Unleaded Gasoline, Undetermined, Kerosene, Diesel	5	0.1%
Unleaded Gasoline, Other, Kerosene, Heating Oil	4	0.1%
Unleaded Gasoline, Undetermined, Other	3	0.0%
Kerosene, Heating Oil, Diesel	2	0.0%
Unleaded Gasoline, Undetermined, Other, Heating Oil	2	0.0%
Kerosene, Heating Oil	2	0.0%
Other, Kerosene, Heating Oil	1	0.0%
Unleaded Gasoline, Undetermined, Other, Kerosene	1	0.0%
Unleaded Gasoline, Undetermined, Diesel	1	0.0%
Undetermined, Other, Diesel	1	0.0%

[Table Note: Substance/product types associated with claims.]

12.1.6 Is Closed

Table 83

Table 84Variable Inventory: is_closed (Claims Data)

Value	N	Share
TRUE	7045	90.4%
FALSE	747	9.6%

[Table Note: Binary indicator for claim closure status.]

12.1.7 Is Open

Table 85

Table 86Variable Inventory: is_open (Claims Data)

Value	N	Share
FALSE	7045	90.4%
TRUE	747	9.6%

[Table Note: Binary indicator for claim open status.]

12.2 Contracts Variables

12.2.1 Adjuster

Table 87

Table 88Variable Inventory: adjuster (Contracts Data)

Value	N	Share
Marion,Shane	83	12.6%
Mackewicz,Bonnie	81	12.3%
Aubel,Tracy	69	10.5%
Smith,Bethany	66	10.0%
Bilder,Jack	57	8.7%
Ferro,James	54	8.2%
Hawk,Gerald	48	7.3%
Headdings,Kyle	41	6.2%
Cramer,Jolene	34	5.2%
Melvin, CPCU,Linda M.	33	5.0%
Bollana,Debra	29	4.4%
Goodyear,Jennifer	23	3.5%
Condran,Patricia	14	2.1%
Moore,Ronald	12	1.8%
Kern,Beth	9	1.4%
Crabb,Linda	5	0.8%

[Table Note: ICF Claims Evaluator assignment distribution. Key variable for instrumental variable identification strategy.]

12.2.2 Consultant

Table 89

Table 90Variable Inventory: consultant (Contracts Data)

Value	N	Share
Letterle & Associates Inc,	128	19.5%
MEA, Inc.,	65	9.9%
Groundwater & Environmental Services, Inc.,	58	8.8%
Mountain Research LLC,	55	8.4%
Core Environmental Services, Inc.,	43	6.5%
Kleinfelder Eastern Merger Corporation,	20	3.0%
In-Site Group, Inc.,	18	2.7%
Converse Consultants,	17	2.6%
DMS Environmental Services LLC,	16	2.4%
B&B Diversified Enterprises Inc,	15	2.3%
Environmental Alliance,	14	2.1%
Liberty Environmental, Inc.,	14	2.1%
Austin James Associates,	13	2.0%
Environmental Remediation & Recovery, Inc.,	13	2.0%
Moody & Associates, Inc.,	9	1.4%
United Environmental Services, Inc.,	9	1.4%
American Environmental Assoc Inc,	8	1.2%
R.A.R. Engineering Group, Inc.,	8	1.2%
Chambers Environmental Group Inc,	8	1.2%
Pennsylvania Tectonics, Inc.,	7	1.1%
Cribbs & Associates, Inc.,	7	1.1%
Environmental Consulting, Inc.,	6	0.9%
Alternative Environmental Solutions Inc,	6	0.9%
Flynn Environmental, Inc.,	5	0.8%
P. Joseph Lehman, Inc.,	5	0.8%
Monridge Environmental, LLC DBA JK Environmental Services, LLC,	5	0.8%
Labella Associates DPC Labella Associates PC,	5	0.8%
EnviroTrac, LTD,	4	0.6%
Juniata Geosciences LLC,	4	0.6%
Alpha Geological Services Inc.,	4	0.6%
KU Resources Inc.,	3	0.5%
American Geosciences Inc,	3	0.5%
Synergy Environmental, Inc.,	3	0.5%
Compliance Environmental Services,	3	0.5%
Keystone Environmental Health and Safety Services,	3	0.5%
Environmental Compliance Services, Inc.,	2	0.3%
Langan Engineering & Environmental Services, Inc.,	2	0.3%
ATC Associates,	2	0.3%
,	2	0.3%
CP Environmental Group, Inc.,	2	0.3%
Onesky Engineering, Inc.,	2	0.3%
Storb Environmental, Inc.,	2	0.3%
ATC Group Services LLC,	2	0.3%

[Table Note: Environmental consulting firm assignment frequency.]

12.2.3 Brings to Closure

Table 91

Table 92Variable Inventory: $brings_to_closure$ (Contracts Data)

Value	N	Share
Yes	394	59.9%
No	264	40.1%

[Table Note: Contract scope classification regarding site closure responsibility.]

12.2.4 Contract Category

Table 93

Table 94Variable Inventory: $contract_category$ (Contracts Data)

Value	N	Share
Sole Source	409	62.2%
Competitively Bid	249	37.8%

[Table Note: High-level contract type categorization.]

12.2.5 Bid Type

Table 95

Table 96Variable Inventory: bid_type (Contracts Data)

Value	N	Share
	420	63.8%
Defined Scope of Work	177	26.9%
Bid to Result	61	9.3%

[Table Note: Procurement mechanism classification (competitive bid vs. negotiated).]

12.2.6 Contract Type Raw

Table 97

Table 98Variable Inventory: $contract_type_raw$ (Contracts Data)

Value	N	Share
Fixed Price	620	94.2%
Pay for Performance	24	3.6%
Time and Material	12	1.8%
	2	0.3%

[Table Note: Original contract type codes from source data.]

12.2.7 Auction Type

Table 99

Table 100Variable Inventory: $auction_type$ (Contracts Data)

Value	N	Share
Other/Unknown	398	60.5%
Scope of Work	177	26.9%
Bid-to-Result	83	12.6%

[Table Note: Auction mechanism subtype for PFP contracts.]

12.2.8 Is Bid to Result

Table 101

Table 102Variable Inventory: $is_bid_to_result$ (Contracts Data)

Value	N	Share
FALSE	575	87.4%
TRUE	83	12.6%

[Table Note: Binary indicator for Performance-Fixed-Price (PFP) auction contracts. Primary treatment variable for causal analysis.]

12.2.9 Is Scope of Work

Table 103

Table 104Variable Inventory: $is_scope_of_work$ (Contracts Data)

Value	N	Share
FALSE	481	73.1%
TRUE	177	26.9%

[Table Note: Binary indicator for traditional Time-and-Materials (T&M) contracts.]

12.2.10 Brings to Closure Flag

Table 105

Table 106Variable Inventory: $brings_to_closure_flag$ (Contracts Data)

Value	N	Share
TRUE	394	59.9%
FALSE	264	40.1%

[Table Note: Binary indicator for contracts responsible for achieving site closure.]