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 (Active and Inactive)

Code	Description	Count	Percent
W	Closed	45557	45.3%
R	Removed	23865	23.7%
C	Currently In Use	19911	19.8%
E	Exempt From State Law	7337	7.3%
UR	Unregulated Removed	1641	1.6%
T	Temporarily Out of Use	1132	1.1%
P	Permanently Closed in Place	868	0.9%
TRANS	Transferred	140	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 4Frequency of Substance Types

Fuel Type	Count
Gasoline	50490
Diesel	19782
Heating_Oil	14476
Used_Oil	5625
Kerosene	3704
Unknown	2798
Hazardous	1134

[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 (Potential Defaults)

Date	Count
1974-01-01	581
1981-12-01	580
1985-12-01	568
1980-01-01	564
1979-12-01	559
1980-12-01	541
1970-01-01	531
1987-12-01	513
1974-12-01	511
1975-01-01	510

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

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

Attribute Code	Type Description	Count
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

Attribute Code	Type Description	Count
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

Attribute Code	Type Description	Count
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

Attribute Code	Type Description	Count
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

Attribute Code	Type Description	Count
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 and CORROSION PROTECTION

Attribute Code	Type Description	Count
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

Attribute Code	Type Description	Count
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

Attribute Code	Type Description	Count
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

Attribute Code	Type Description	Count
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

Attribute Code	Type Description	Count
8N	NO	1397
8Y	YES	11677

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

Table 27

Table 28Component Universe: FIRE MARSHAL PERMIT

Attribute Code	Type Description	Count
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

Attribute Code	Type Description	Count
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

Attribute Code	Type Description	Count
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

Attribute Code	Type Description	Count
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

Attribute Code	Type Description	Count
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

Attribute Code	Type Description	Count
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

Attribute Code	Type Description	Count
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

Attribute Code	Type Description	Count
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

Attribute Code	Type Description	Count
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

Attribute Code	Type Description	Count
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

Attribute Code	Type Description	Count
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

Attribute Code	Type Description	Count
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

Attribute Code	Type Description	Count
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

Attribute Code	Type Description	Count

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

Table 55

Table 56Component Universe: TANK UPGRADE

Attribute Code	Type Description	Count
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 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.

5.1 Capacity Evolution by Decade

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

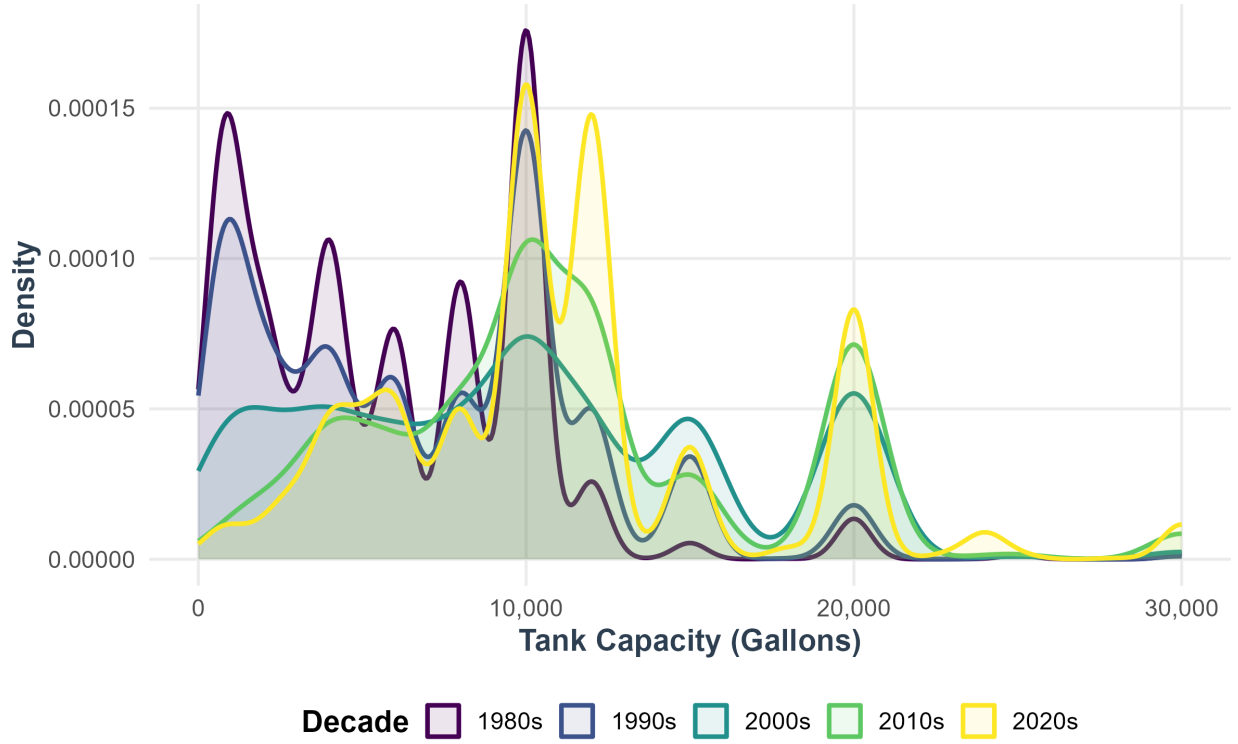

Figure 1: Evolution of Tank Capacity Distribution by Installation Decade

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

5.2 Fuel Mix Evolution

5.3 Fuel Mix Evolution

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

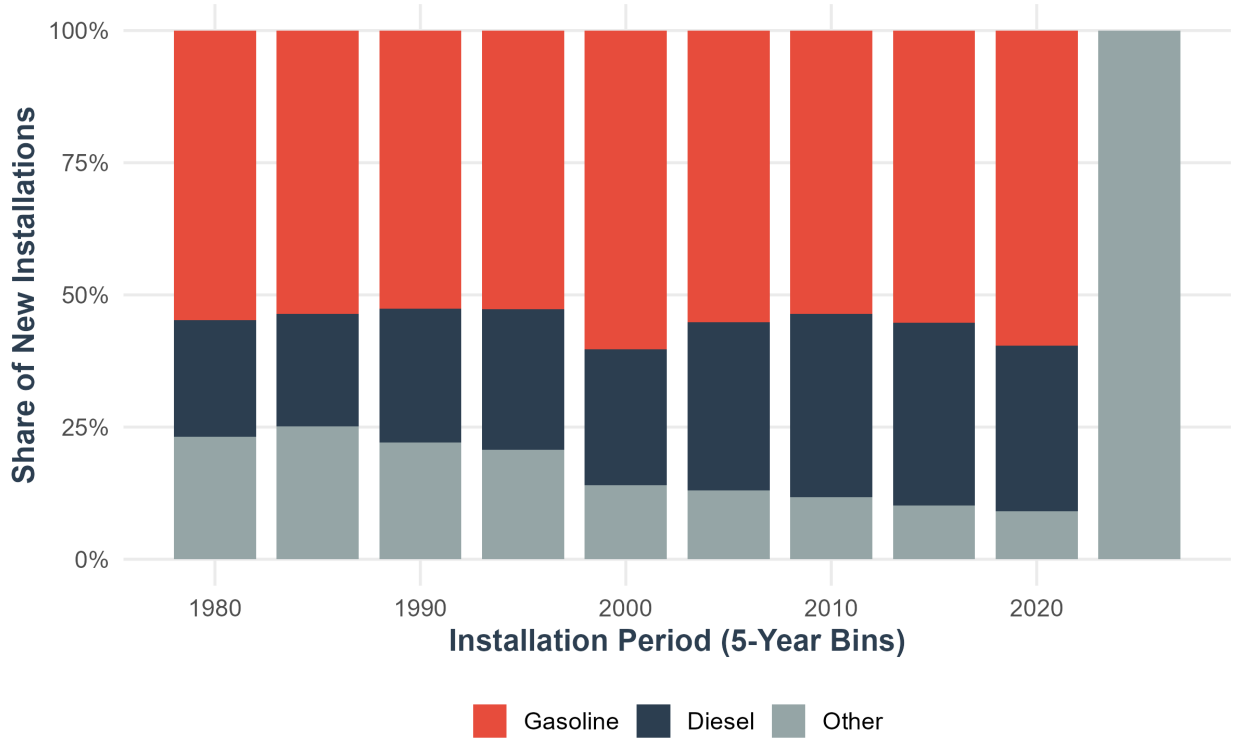

Figure 2: Evolution of Fuel Type Mix in New Installations (5-Year Periods)

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

5.4 Tank Lifespan Distribution

5.5 Tank Lifespan Distribution

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

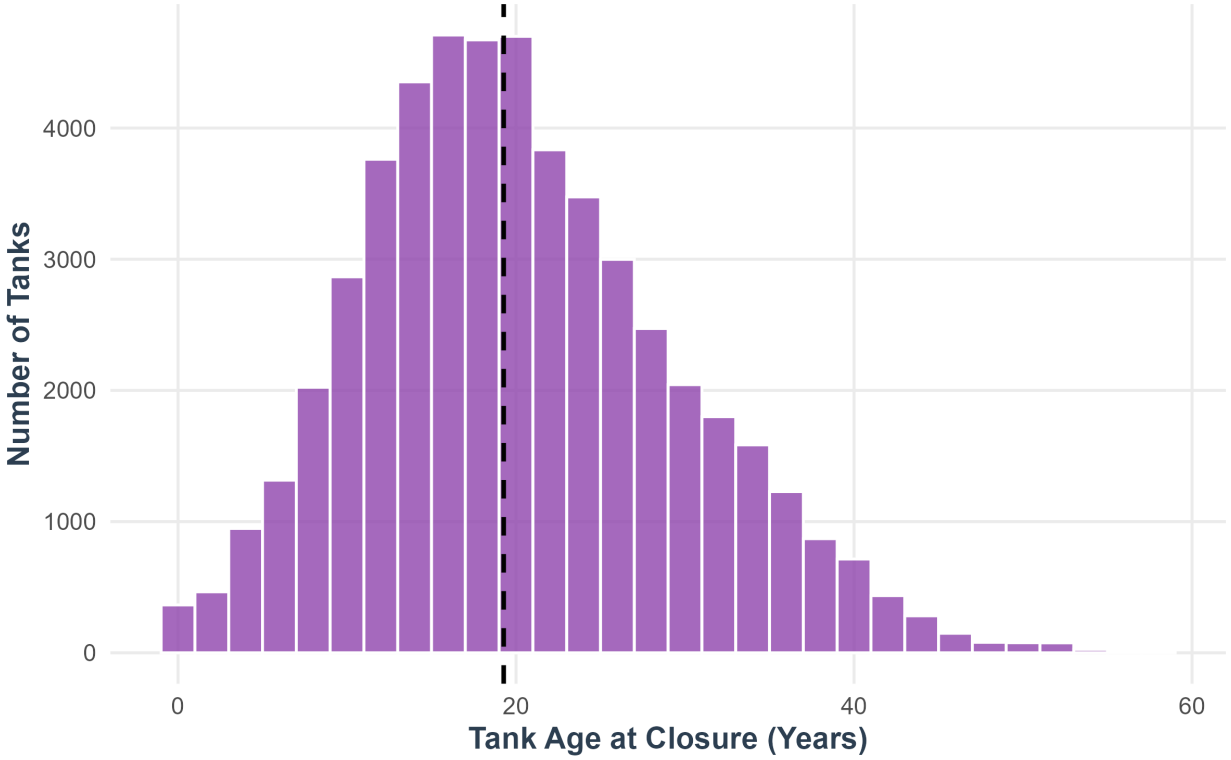

Figure 3: Distribution of Tank Age at Closure

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

5.6 Material Evolution

5.7 Material Evolution

[Description: Temporal trends in tank construction materials (Steel vs. Fiberglass/Composite) over installation periods.]

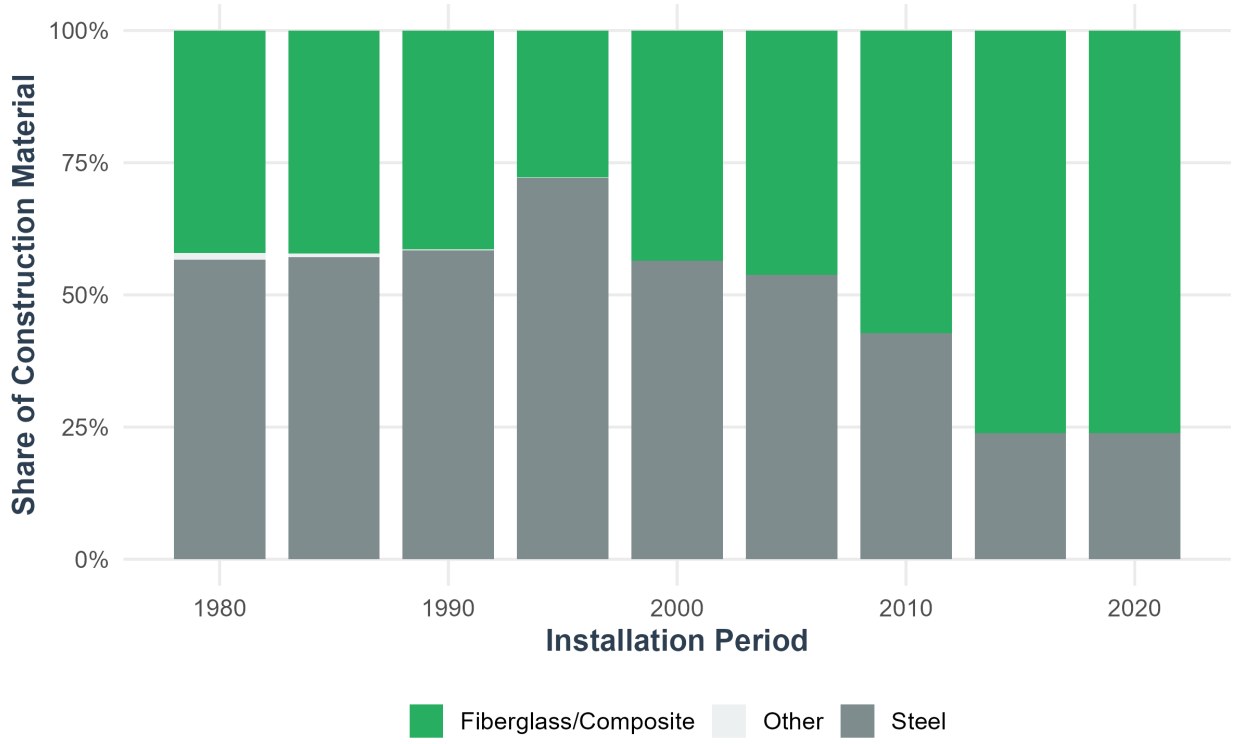

Figure 4: Evolution of Tank Construction Materials by Installation Period

[Figure Note: [Placeholder: Analysis of material shift—transition from steel to fiberglass/composite? Regulatory drivers? Cost and durability implications.]]

6 Facility-Level Intelligence

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

6.1 Facility Size Evolution by Decade

6.2 Facility Size Evolution by Decade

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

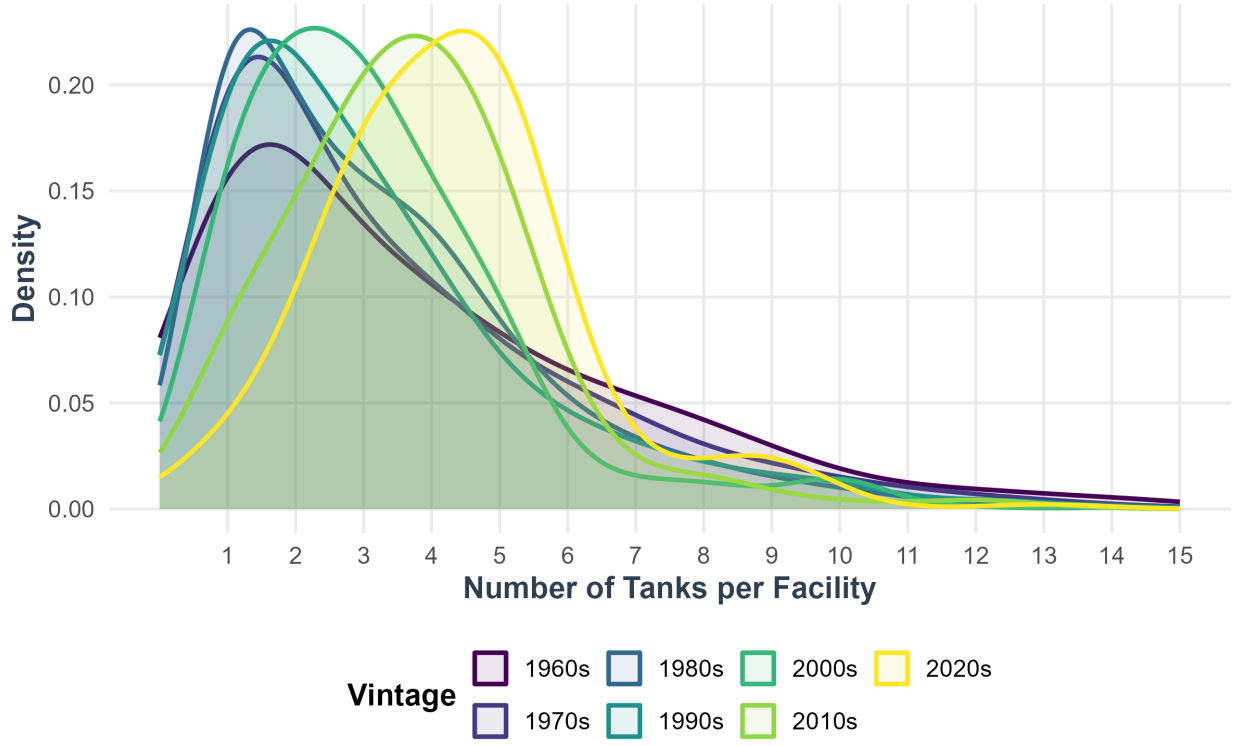

Figure 5: Evolution of Facility Size Distribution by Vintage Decade

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

6.3 Facility Status by Vintage

6.4 Facility Status by Vintage

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

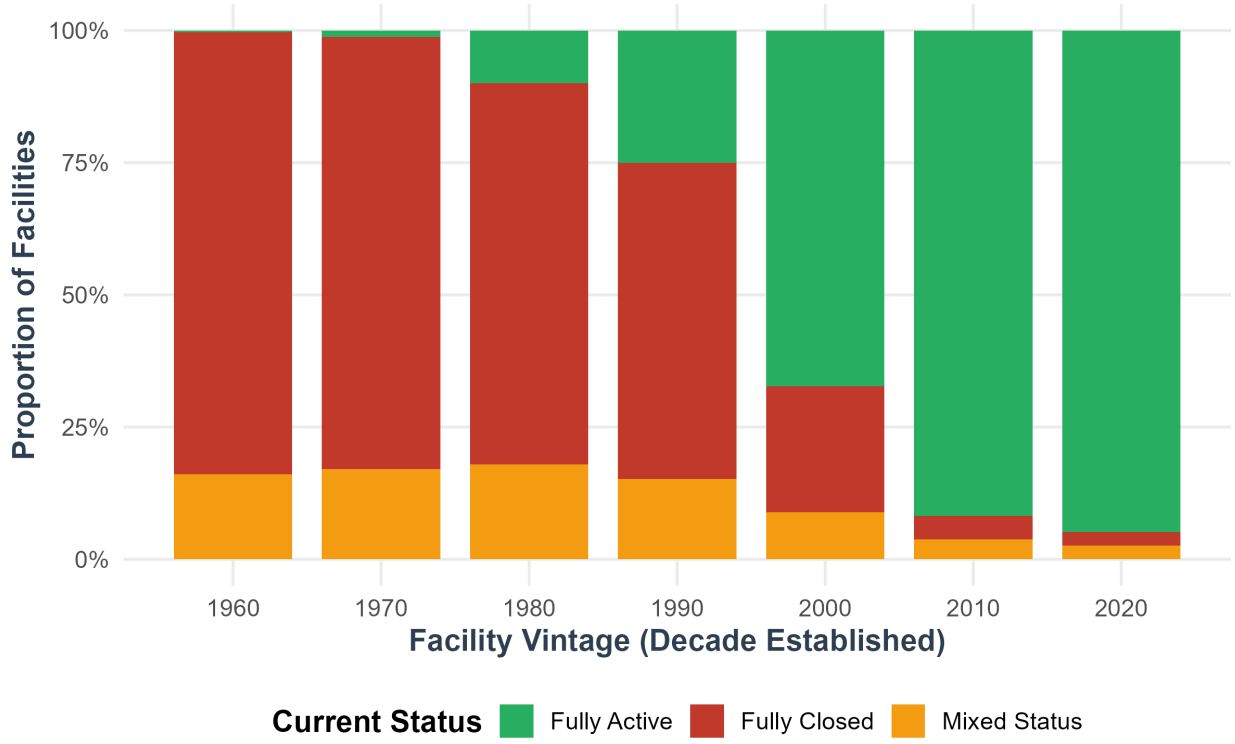

Figure 6: Facility Survival Status by Vintage Decade

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

6.5 Facility Age vs. Size Relationship

6.6 Facility Age vs. Size Relationship

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

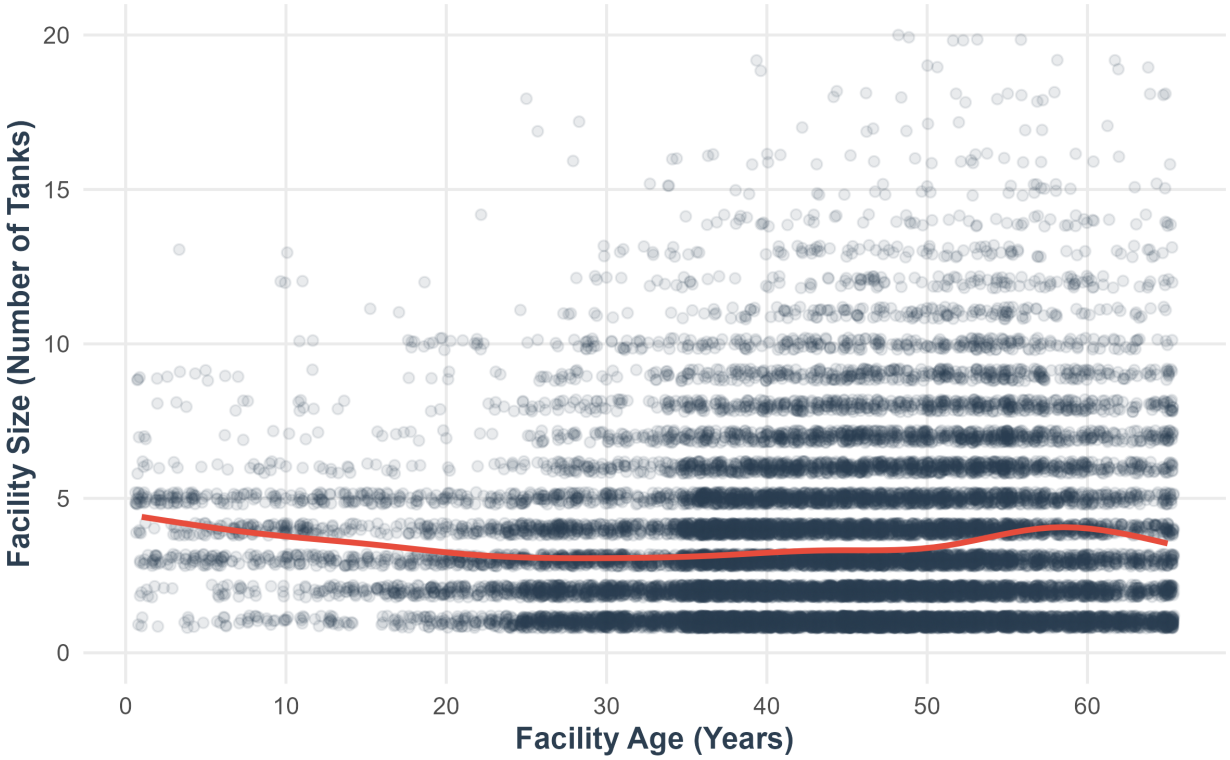

Figure 7: Relationship Between Facility Age and Size

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

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