Overview

Dataset info

Number of variables	16
Number of observations	62239
Missing cells	0 (0.0%)
Duplicate rows	0 (0.0%)
Total size in memory	7.6 MiB
Average record size in memory	128.0 B

Variables types

Numeric	5
Categorical	11
Boolean	0
Date	0
URL	0
Text (Unique)	0
Rejected	0
Unsupported	0

Warnings

Amendment_Value is highly skewed ($\gamma_1 = 195.9911812$)	`Skewed`
Amendment_Value has 50207 (80.7%) zeros	`Zeros`
Commodity has a high cardinality: 1525 distinct values	`Warning`
Country_of_Origin has a high cardinality: 137 distinct values	`Warning`
Entity has a high cardinality: 63 distinct values	`Warning`
Exemption is highly skewed ($\gamma_1 = 43.60094747$)	`Skewed`
Exemption has 62174 (99.9%) zeros	`Zeros`
Limited_Tendering_Reason has 52172 (83.8%) zeros	`Zeros`
Original_Value is highly skewed ($\gamma_1 = 102.7278833$)	`Skewed`

Variables

Amendment_…
Numeric

Distinct count	10306	**Mean**	71769.73839	
Unique (%)	16.6%	**Minimum**	-16522116	
Missing (%)	0.0%	**Maximum**	1157080504	
Missing (n)	0	**Zeros (%)**	80.7%	
Infinite (%)	0.0%			
Infinite (n)	0			

Toggle details

Commodity
Categorical
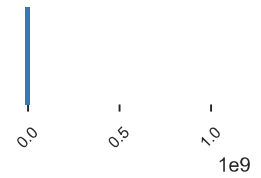
Distinct count	1525
Unique (%)	2.5%
Missing (%)	0.0%
Missing (n)	0

Toggle details

Common values	Composition

Value	Count	Frequency (%)
Liquid Propellants and Fuels, Petrole…	4042	6.5%
Informatics Professional Services	2964	4.8%
Middle Distillate Fuels	1497	2.4%
Drugs and Biologicals	1476	2.4%
Commercial Security Guard and Rel…	1334	2.1%
Office Furniture	1319	2.1%
Structural Steel Work	1318	2.1%
ADP Software	1215	2.0%
Laboratory Equipment and Supplies	1172	1.9%
Passenger Motor Vehicles	1029	1.7%
Other values (1515)	44873	72.1%

Commodity_R…
Categorical

Distinct count	3	
Unique (%)	< 0.1%	
Missing (%)	0.0%	
Missing (n)	0	

Unknown	44533
Yes	16371
No	1335

Toggle details

Commodity_T…
Categorical

Distinct count	3	
Unique (%)	< 0.1%	
Missing (%)	0.0%	
Missing (n)	0	

Services	29964
Goods	26770
Construction	5505

Toggle details

Country_of_O…
Categorical

Distinct count	137
Unique (%)	0.2%
Missing (%)	0.0%
Missing (n)	0

Toggle details

Common values   Composition

Value	Count	Frequency (%)
CA	57454	92.3%
US	1849	3.0%
GB	413	0.7%
DE	280	0.4%
SN	230	0.4%
FR	131	0.2%
IT	92	0.1%
RO	71	0.1%
GR	69	0.1%
ES	64	0.1%

Value	Count	Frequency (%)
Other values (127)	1586	2.5%
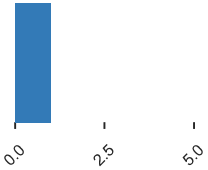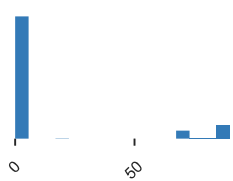
Entities_Rule
Categorical

Distinct count	3	Unknown		52917
Unique (%)	< 0.1%	Yes		9228
Missing (%)	0.0%	No	94	
Missing (n)	0			

Toggle details

Entity
Categorical

Distinct count	63
Unique (%)	0.1%
Missing (%)	0.0%
Missing (n)	0

Toggle details

Common values	Composition

Value	Count	Frequency (%)
DND	32368	52.0%
CSC	10167	16.3%
DFO	3818	6.1%
CBSA	2023	3.3%
GAC	1948	3.1%
CRA	1580	2.5%
HC	1401	2.3%
PHAC	1001	1.6%
ESDC	579	0.9%
VAC	544	0.9%
Other values (53)	6810	10.9%

Estimated_Tra...
Categorical

Distinct count	3		Yes	38634
Unique (%)	< 0.1%		Unknown	17967
			No	5638
Missing (%)	0.0%			
Missing (n)	0			

Toggle details

Exemption
Numeric

Distinct count	6	**Mean**	0.0025707353	
		Minimum	0	
Unique (%)	< 0.1%	**Maximum**	6	
		Zeros (%)	99.9%	
Missing (%)	0.0%			
Missing (n)	0			
Infinite (%)	0.0%		0.0 2.5 5.0	
Infinite (n)	0			

Toggle details

Exemptions_R...
Categorical

Distinct count	2		Unknown	51967
Unique (%)	< 0.1%		Yes	10272
Missing (%)	0.0%			
Missing (n)	0			

Toggle details

Limited_Tende...
Numeric

Distinct count	16	**Mean**	12.44785424	
		Minimum	0	
Unique (%)	< 0.1%	**Maximum**	90	
		Zeros (%)	83.8%	
Missing (%)	0.0%			
Missing (n)	0		0 50	
Infinite (%)	0.0%			
Infinite (n)	0			

Toggle details

Limited_Tende…
Categorical

Distinct count	2	Unknown	53222
		Yes	9017
Unique (%)	< 0.1%		
Missing (%)	0.0%		
Missing (n)	0		

Toggle details

Original_Value
Numeric

Distinct count	45758	**Mean**	347253.1686
Unique (%)	73.5%	**Minimum**	0
Missing (%)	0.0%	**Maximum**	1772776794
Missing (n)	0	**Zeros (%)**	0.4%
Infinite (%)	0.0%		
Infinite (n)	0		

Toggle details

Statistics	Histogram	Common values	Extreme values

Quantile statistics

Minimum	0
5-th percentile	10298.003
Q1	15451.485
Median	26266
Q3	76979.33
95-th percentile	557736.925
Maximum	1772776794
Range	1772776794
Interquartile range	61527.845

Descriptive statistics

Standard deviation	12542809.19
Coef of variation	36.12007123
Kurtosis	12162.61784
Mean	347253.1686
MAD	545430.4892
Skewness	102.7278833
Sum	2.161268996e+10
Variance	1.573220623e+14
Memory size	486.4 KiB

thresholds
Categorical

Distinct count	3	Yes	34703
		Unknown	22436
Unique (%)	< 0.1%	No	5100
Missing (%)	0.0%		
Missing (n)	0		

Toggle details

Unnamed_0
Numeric

Distinct count	62239	**Mean**	31119
Unique (%)	100.0%	**Minimum**	0
		Maximum	62238
Missing (%)	0.0%	**Zeros (%)**	< 0.1%
Missing (n)	0		
Infinite (%)	0.0%		
Infinite (n)	0		
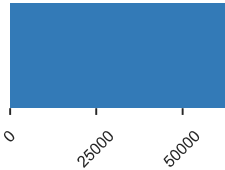
Toggle details

Year
Categorical

Distinct count	3	2018	32681
Unique (%)	< 0.1%	2017	21907
		2019	7651
Missing (%)	0.0%		
Missing (n)	0		

Toggle details

Correlations

Pearson's r	Spearman's ρ	Kendall's τ	Phik (φ$_k$)	Cramér's V (φ$_c$)	Recoded

Year —

endment_Value | Exemption | dering_Reason | Original_Value | Unnamed_0 | Year

−1.00

Missing values

Count | Matrix
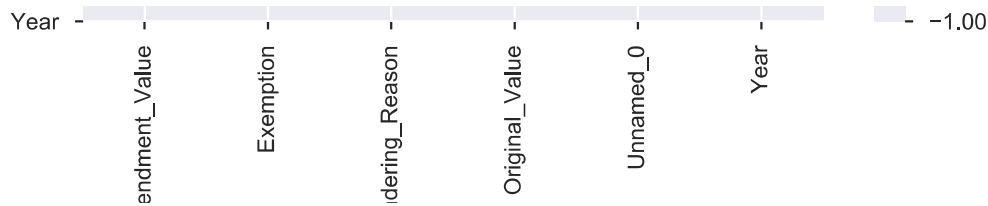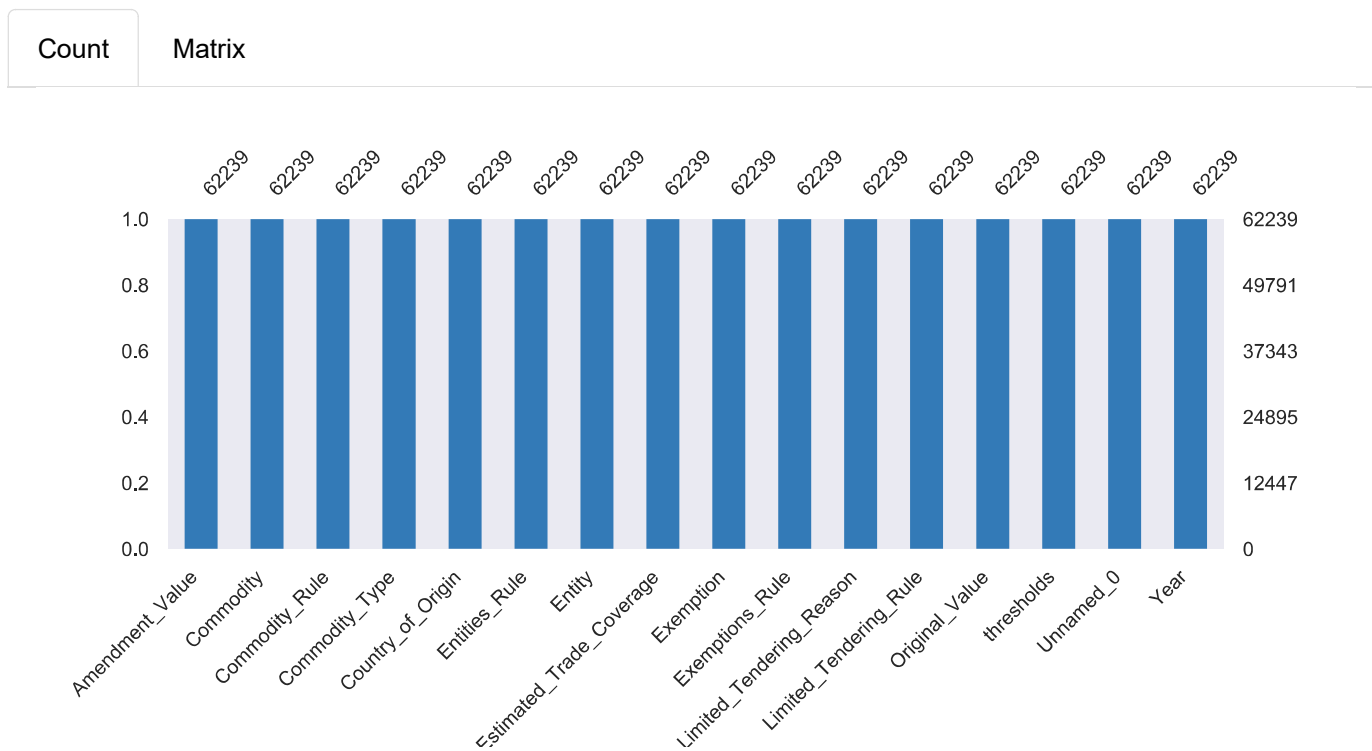
Sample

First rows

	Amendment_Value	Commodity	Co
0	5016.25	Consulting Services	Yes

	Amendment_Value	Commodity	Co
1	0.00	Maintenance and Repair Shop Specialized Equipment, Miscellaneous - Repair	Un
2	0.00	Outerwear	Un
3	0.00	Trash/Garbage Collection Services	Un
4	164553.65	Cleaning and Maintenance, Janitorial	Ye:
5	0.00	Scaffolding Equipment and Concrete Forms	Un
6	0.00	Roofing, Siding and Sheet Metal Installation Services	Un
7	88000.01	Professional Services / Program Research Analysis	No
8	0.00	Bars and Rods, Nonferrous Base Metal	Un
9	0.00	Laboratory Equipment and Supplies	Ye:

Last rows

	Amendment_Value	Commodity	Comm
62229	0.00	Communication Marketing Management	Yes
62230	0.00	Building/Installation Special Trade Services, Not Elsewhere Specified	Unknov
62231	0.00	Harassment (Investigative Services)	No
62232	0.00	Passenger Motor Vehicles	Unknov
62233	73563.00	Informatics Professional Services	Yes
62234	0.00	ADP Software	Unknov
62235	0.00	Communication Writing and Editing Services	No
62236	264882.77	Transportation (R&D)	Yes
62237	0.00	Recreational and Gymnastic Equipment	Unknov
62238	-10723.70	Temporary Help Services - Professional and Technical	No

Report generated with pandas-profiling (https://github.com/pandas-profiling/pandas-profiling).