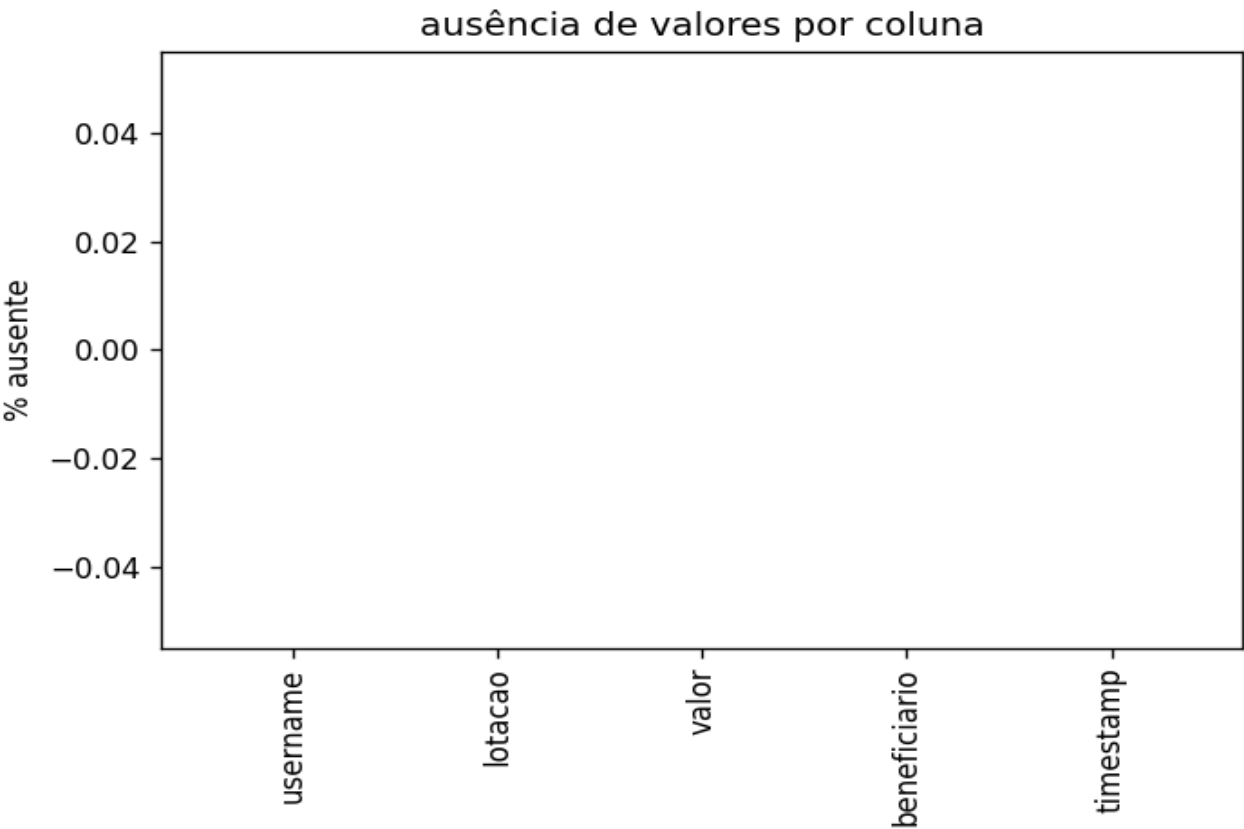


Relatório de Análise — 061025-1602

Arquivo analisado: **input.csv**
Linhas: 50000 Colunas: 5
Registros duplicados: 0

Ausência de valores por coluna



Coluna: username

Campo	Valor
Tipo	object
Nulos	0
Distintos (não vazios)	73
Frequências	mais frequente: 'NFDFX' (4994, 9.99%) menos frequente: 'ROBSO' (1, 0.00%)
Comprimentos	min=5, q1=5.0, mediana=5.0, q3=5.0, max=5
Entropia	5.9868
Top 10	tabela abaixo

Valor	Contagem
NFDFX	4994
AFPEX	643
AJZQY	643
BFFLZ	643
BZSHR	643
CCVSF	643

CKWGY	643
CWFEE	643
ACOWL	643
DAUDP	643

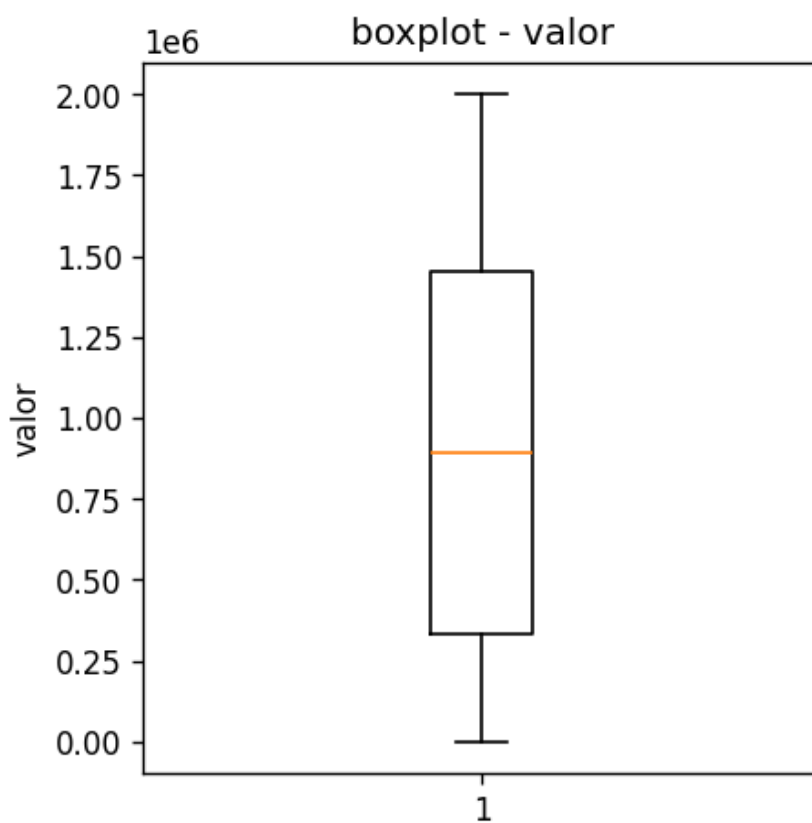
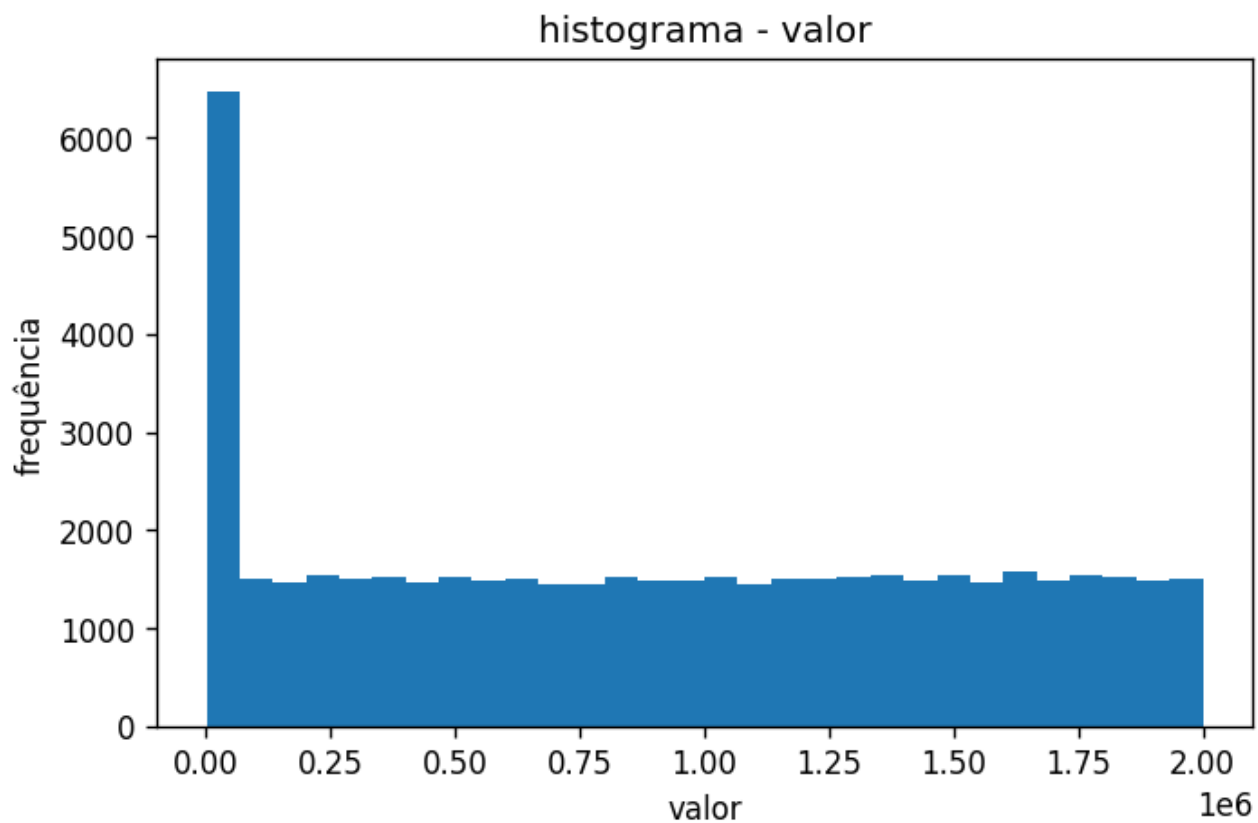
Coluna: lotacao

Campo	Valor
Tipo	object
Nulos	0
Distintos (não vazios)	72
Frequências	mais frequente: 'FREQUENTE' (4994, 9.99%) menos frequente: 'FRAUDE' (5, 0.01%)
Comprimentos	min=5, q1=5.0, mediana=5.0, q3=5.0, max=9
Entropia	5.9866
Top 10	tabela abaixo

Valor	Contagem
FREQUENTE	4994
GBYKJ	644
AAVDJ	643
ANRGJ	643
BOGTK	643
BYIUM	643
ANSTD	643
BIXHV	643
DEFIK	643
COCXY	643

Coluna: valor

Campo	Valor
Tipo	float
Nulos	0
Distintos (não vazios)	49767
Frequências	mais frequente: '867.0' (3, 0.01%) menos frequente: '500.01' (1, 0.00%)
Comprimentos	min=5, q1=9.0, mediana=9.0, q3=10.0, max=10
Estatísticas	min=500.01, q1=334032.8325, mediana=894146.25, q3=1451355.56, max=1999973.98 média=903092.1257208, desvio padrão=625536.3338248847, iqr=1117322.7275 outliers(IQR)=0, outliers(MAD)=0, skew=0.06711471238249822, curtose(excesso)=-1.27400
Histograma	ver imagem abaixo
Boxplot	ver imagem abaixo
Benford	Erro - 'Power_divergenceResult' object has no attribute 'cdf'



Coluna: beneficiario

Campo	Valor
Tipo	object
Nulos	0
Distintos (não vazios)	204

Tipo semântico	cnpj
Frequências	mais frequente: '74.028.960/9599-16' (2559, 5.12%) menos frequente: '28.276.922/5855-18'
Comprimentos	min=18, q1=18.0, mediana=18.0, q3=18.0, max=18
Entropia	7.4469
Top 10	tabela abaixo

Valor	Contagem
74.028.960/9599-16	2559
65.495.584/0804-13	2435
74.394.889/5457-59	255
47.238.592/2640-48	255
86.714.800/8947-86	254
02.419.822/1821-46	254
92.892.395/4359-21	253
78.065.047/1978-38	253
07.417.078/2965-00	253
12.663.157/1313-61	253

Coluna: timestamp

Campo	Valor
Tipo	datetime
Nulos	0
Distintos (não vazios)	47007
Frequências	mais frequente: '2025-03-28 05:38:00' (4, 0.01%) menos frequente: '2025-01-01 00:01:00' (1)
Comprimentos	min=16, q1=16.0, mediana=16.0, q3=16.0, max=16
Data/hora	formato detectado: %d/%m/%Y %H:%M intervalo: 2025-01-01 00:01:00 → 2025-09-30 22:26:00 dias únicos: 273, média por dia: 183.15

Regras sugeridas

FDs

Determinante	Implicado/Chave	Cobertura
username	lotacao	100%

CFDs (aproximadas)

Determinante	Implicado	Cobertura
lotacao	username	98.61%
valor	username	99.99%
valor	lotacao	99.99%
valor	beneficiario	99.79%
valor	timestamp	99.55%

Denial constraints

Regra (DC)	Violações
valor>=0	0